Karbosguide.com

Welcome to Michael Karbo's Online Service. Here you will find a modern online-magazine with more than 500 illustrated articles for the critical reader!

Use our menu to your left or the sitemap. You may also follow any of the links listed below. We hope that you appreciate our work!

- Start studying the design of a PC motherboard.
- Learn about harddisks and other drives.
- Learn about the PC I/O system.
- Learn about the PC video system.

News ... Sitemap

See our guestbook and add your comment.

Corrections

Misspellings, typos or other corrections. Please report!.

Karbo's Newsletter:

Sign up!

Pentium 4 and AthlonXP.

More than 50 photos of old cars free to download!

The MP3 article is re-written.

Editing photos with Photoshop

Cleaning Windows Me for temporary files ...

All modules 7 re-written

About Michael Karbo

NEW: German version.

Privacy politic

Software Guides

Dictionary

Photo Gallery

Search

1. About PC data

1a. About data

1b. Character tables

2. The PC system board

2a. Introduction

2b. Boot process, system bus

2c. I/O buses

2d. Chip sets

2e. On RAM

3. About CPUs

3a. An intro to CPUs

3b. CPU improvements

3c. 5th gener. CPUs

3d. Cooling and overclocking

3e. 6th gener. CPUs

4. Drives and other storage

4a. Drives

4b. Hard disks

4c. Optic storage media

4d. ZIP etc.

4e. Tape streamers

5. Expansion cards and interfaces

5a. Adapters

5b. EIDE, Ultra DMA, AGP

5c. SCSI, FireWire, USB

6. OSs and file systems

6a. File systems

6b. Windows 95

6c. BIOS, OS, hardware

6d. The Windows 98 page

7. Graphics and sound

7a. Display basics

7b. Graphics cards

7c. About sound cards

7d. Digital music MP3, MOD etc.

Main page

KarbosGuide.com

- Next page
- Previous page

Support our work

Karbosguide.com a tutorial used globally. It is a work made "con amore". We make no profit from it. However, just a little economical revenue would be great. It takes so much time to update and develop the site with high quality information.

You can help us in several ways:

- Click on the banners
- Register as a user and donate a few dollars.
- Tell your friends an partners about Karbosguide.com, share the URL.

- Next page
- Previous page

KarbosGuide.com. Module 1a.

About data

- Next page
- Previous page

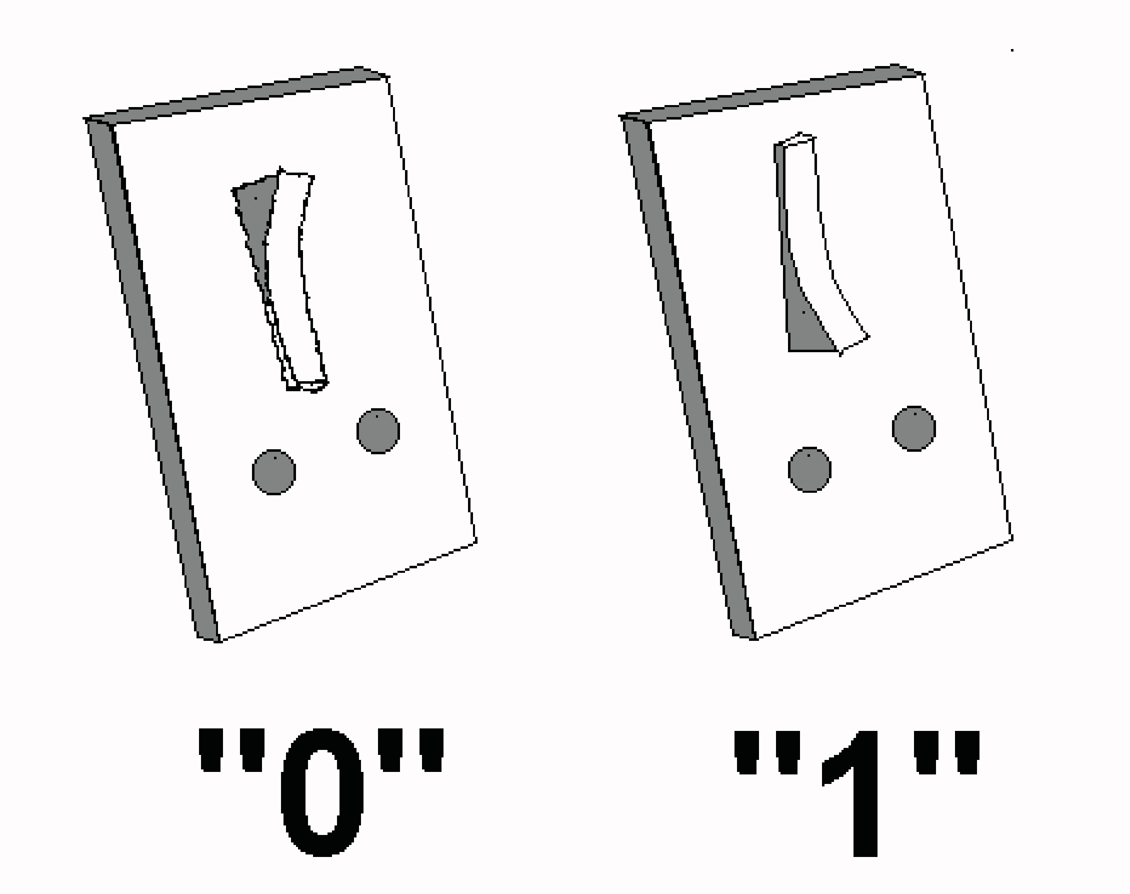
Our PCs are data processors. The PC's function is simple: to process data, and the processing is done electronically inside the CPU and between the other components. That sounds simple, but *what* is data, and *how* is it processed electronically in a PC? That is the subject of these pages.

Analog data

The signals, which we send each other to communicate, is data. Our daily data have many forms: sound, letters, numbers, and other characters (handwritten or printed), photos, graphics, film. All this data is in its nature *analog*, which means that it varies in type. In this form, the data-signals are unusable in a PC. The PC can only process concise, simple data formats. Such data can be processed very effectively.

Digital data

The PC is an electric unit. Therefore, it can only deal with data, which are associated with electricity. That is accomplished using electric switches, which are either off or on. You can compare with regular household switches. If the switch is off, the PC reads numeral 0. If it is on, it is read as numeral one. See the illustration below:



With our electric switches, we can write 0 or 1. We can now start our data processing!

The PC is filled with these switches (in the form of transistors). There are literally millions of those in the electronic components. Each represents either a 0 or a 1, so we can process data with millions of 0s and 1s.

Bits

[top]

Each 0 or 1 is called a *bit*. Bit is an abbreviation of the expression BInary digiT. It is called binary, since it is derived from the *binary number system:*

0	1 bit

1	1 bit
0110	4 bit
01101011	8 bit

The binary number system

[top]

The binary number system is made up of digits, just like our common decimal system (10 digit system). But, while the decimal system uses digits 0 through 9, the binary system only uses digits 0 and 1.

If you are interested in understanding the binary number system, then here is a brief course. See if you can follow the system. See how numbers are constructed in the binary system, using only 0s and 1s:

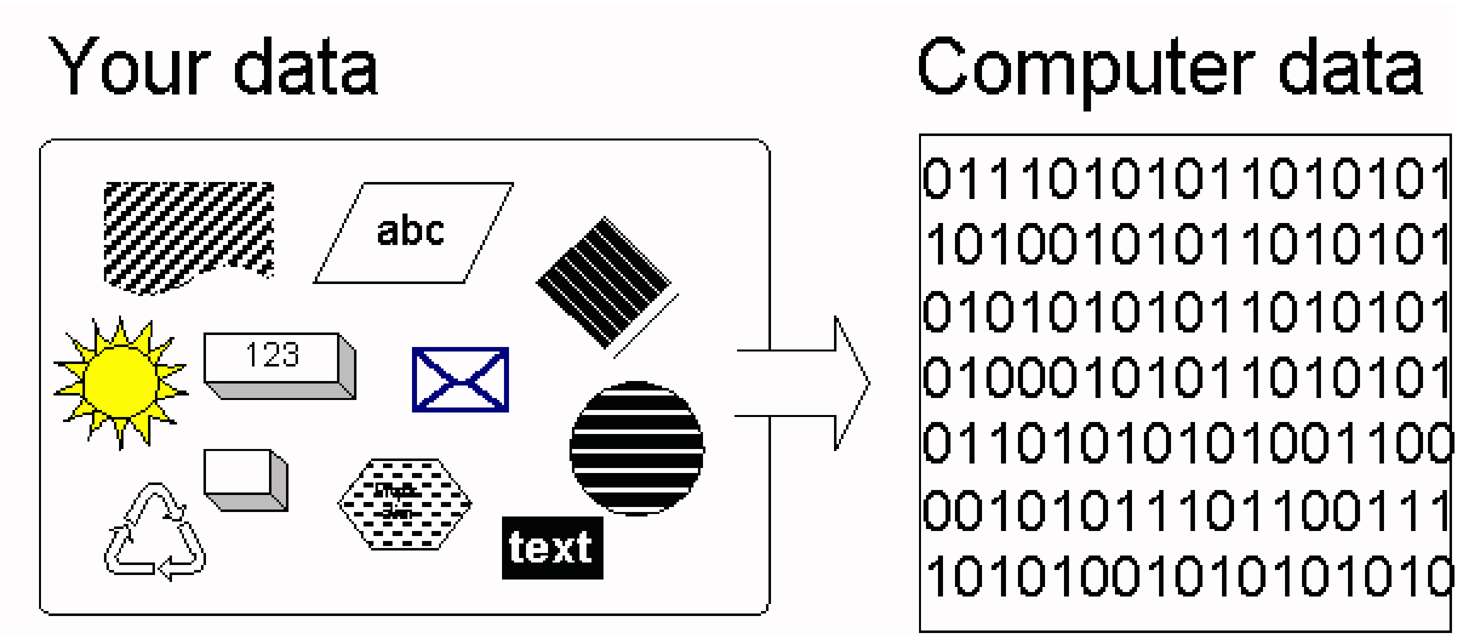
Numbers, as known in the decimal-system	Same numbers in binary system
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000

Digital data

[top]

We have seen that the PC appears capable of handling data, if it can receive them as 0s and 1s. This data format is called digital. If we can translate our daily data from their analog format to digital format, they will appear as chains of 0s and 1s, then the PC can handle them.

So, we must be able to digitize our data. Pour text, sounds, and pictures into a funnel, from where they emerge as 0s and 1s:



Let us see how this can be accomplished.

- Next page
- Previous page

Learn more [top]

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 1a2.

About Bytes

Contents:

- Introduction
- ASCII
- About text and code
- Data in files

- Next page
- Previous page

Introduction

The most basic data processing is word processing. Let us use that as an example. When we do word processing, we work at a keyboard similar to a typewriter. There are 101 keys, where we find the entire alphabet A, B, C, etc. We also find the digits from 0 to 9 and all the other characters we need:,.-;():_?!"#*%&etc..

All these characters must be digitized. They must be expressed in 0s and 1s. Bits are organized in groups of 8. A group of 8 bits is called a byte.

8 bits = 1 byte, that is the system. Then, what can we do with bytes? First, let us see how many different bytes we can construct. A byte is an 8 digit number. We link 0s and 1s in a pattern. How many different ones can we make? Here is one: 01110101, and here is another: 10010101.

We can calculate that you can make 2 x 2 x 2 x 2 x 2 x 2 x 2 x 2 different patterns, since each of the 8 bits can have 2 values.

- 2^8 (two in the power of eight) is 256. Then there are 256 different bytes!

Now we assign a byte to each letter and other characters. And since we have 256 patterns to choose from, there is plenty of room for all. Here you see some examples of the "translation:"

Character	Bit pattern	Byte number	Character	Bit pattern	Byte number
A	01000001	65	¼	10111100	188
B	01000010	66	.	00101110	46
C	01000011	67	:	00111010	58
a	01100001	97	$	00100100	36
b	01100010	98	\	01011100	92
o	01101111	111	~	01111110	126
p	01110000	112	1	00110001	49
q	01110001	113	2	00110010	50
r	01110010	114	9	00111001	57
x	01111000	120	©	10101001	169
y	01111001	121	>	00111110	62
z	01111010	122	‰	10001001	137

When you write the word "summer", you write 6 letters. If the computer has to process that word, it will be digitized to 6 bytes. In other words, the word summer occupies 6 bytes in the PC RAM, when you type it, and 6 bytes on the hard disk, if you save it.

ASCII

[top]

ASCII means American Standard Code for Information Interchange. It is an industry standard, which assigns letters, numbers, and other characters within the 256 slots available in the 8 bit code.

The ASCII table is divided in 3 sections:

- Non printable system codes between 0 and 31.

- "Lower ASCII" between 32 and 127. This part of the table originates from older, American systems, which worked on 7 bit character tables. Foreign letters, like Ø and Ü were not available then.

- "Higher ASCII" between 128 and 255. This part is programmable, in that you can exchange characters, based on which language you want to write in. Foreign letters are placed in this part.

Learn more about the ASCII table in Module 1b

An example

Let us imagine a stream of bits sent from the keyboard to the computer. When you type, streams of 8 bits are sent to the computer. Let us look at a series of bits:

001100010011001000110011

Bits are combined into bytes (each 8 bits). These 24 bits are interpreted as three bytes. Let us read them as bytes: 00110001, 00110010, and 00110011.

When we convert these byte binary numbers to decimal numbers, you will see that they read as 49, 50, and 51 in decimal numbers. To interpret these numbers, we have to look at the ASCII table. You will find that you have typed the numbers 1, 2, and 3.

About text and code

[top]

Now we have seen the PCs user data, which are always digitized. But there are many different kinds of data in the PC. You can differentiate between 2 fundamental types of data:

- Program code, which is data, that allows the PC to function.

- User data, like text, graphics, sound.

The fact is, that the CPU must have *instructions* to function. You can read more about this in the review of the CPU in module 3a. An instruction is a string of data, of 0s and 1s. The CPU is designed to recognize these instructions, which arrive together with the user input data to be processed.

The program code is thus a collection of instructions, which are executed one by one, when the program runs. Each time you click the mouse, or hit a key on the keyboard, instructions are sent from your software (program) to the CPU, telling it what to do next.

User data are those data, which tells the software how to respond. The letters, illustrations, home pages, etc., which you and I produce, are created with appropriate software.

Files

[top]

Both program code and user data are saved as *files on the* hard disk. Often, you can recognize the type of file by its suffix. Here are some examples:

Content	File name
Program code	START.EXE, WIN.COM, HELP.DLL, VMM32.VXD
User data	LETTER.DOC, HOUSE.BMP, INDEX.HTM

This is written as an introduction to naming files. The file name suffix determines how the PC will handle the file. You can read about this subject in some of my books, e.g. "DOS - teach yourself" (only available in Europe.

- Next page
- Previous page

Learn more [top]

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

Module 1b

Character tables

- Next page
- Previous page

The ASCII tables

Here you see the complete ASCII character table. First the part from ASCII-numbers 032 to 127:

ASCII-number	Common characters (in Windows)	Symbol	Wingdings
032			
033	!	!	!
034	"	∀	"
035	#	#	#
036	$	∃	$
037	%	%	%
038	&	&	&
039	'	∍	'
040	(	(	(
041	)	)	)
042	*	∗	*
043	+	+	+
044	,	,	,
045	-	−	-
046	.	.	.

047	/	/	/
048	0	0	0
049	1	1	1
050	2	2	2
051	3	3	3
052	4	4	4
053	5	5	5
054	6	6	6
055	7	7	7
056	8	8	8
057	9	9	9
058	:	:	:
059	;	;	;
060	<	<	<
061	=	=	=
062	>	>	>
063	?	?	?
064	@	≅	@
065	A	Α	A
066	B	Β	B
067	C	Χ	C
068	D	Δ	D
069	E	Ε	E
070	F	Φ	F
071	G	Γ	G
072	H	Η	H
073	I	Ι	I
074	J	ϑ	J
075	K	Κ	K
076	L	Λ	L
077	M	Μ	M

078	**N**	Ν	N
079	**O**	Ο	O
080	**P**	Π	P
081	**Q**	Θ	Q
082	**R**	Ρ	R
083	**S**	Σ	S
084	**T**	Τ	T
085	**U**	Υ	U
086	**V**	ς	V
087	**W**	Ω	W
088	**X**	Ξ	X
089	**Y**	Ψ	Y
090	**Z**	Ζ	Z
091	**[**	[	[
092	****	∴	\\
093	**]**	]	]
094	**^**	⊥	^
095	**_**	_	_
096	**`**	‾	`
097	**a**	α	a
098	**b**	β	b
099	**c**	χ	c
100	**d**	δ	d
101	**e**	ε	e
102	**f**	φ	f
103	**g**	γ	g
104	**h**	η	h
105	**i**	ι	i
106	**j**	ϕ	j
107	**k**	κ	k
108	**l**	λ	l

109	**m**	μ	m
110	**n**	ν	n
111	**o**	ο	o
112	**p**	π	p
113	**q**	θ	q
114	**r**	ρ	r
115	**s**	σ	s
116	**t**	τ	t
117	**u**	υ	u
118	**v**	ϖ	v
119	**w**	ω	w
120	**x**	ξ	x
122	**z**	ζ	z
123	**{**	{	{
124	**|**	|	|
125	**}**	}	}
126	**~**	~	~
127	**•**		•

Then the numbers from 0128 to 0255. Notice the leading zero.

ASCII-number	Common characters (in Windows)	Symbol	Wingdings
0128	**€**	&ευρο;	& e u r o ;
0129	**•**	•	•
0130	**‚**	‚	‚
0131	***ƒ***	ƒ	ƒ
0132	**„**	„	„
0133	**…**	…	…
0134	**†**	†	†
0135	**‡**	‡	‡

0136	ˆ	ˆ	ˆ
0137	‰	‰	‰
0138	Š	Š	Š
0139	‹	‹	‹
0140	Œ	Œ	Œ
0141	•	•	•
0142	Ž	Ž	Ž
0143	•	•	•
0144	•	•	•
0145	‘	‘	‘
0146	’	’	’
0147	“	“	“
0148	”	”	”
0149	•	•	•
0150	–	–	–
0151	—	—	—
0152	˜	˜	˜
0153	™	™	™
0154	š	š	š
0155	›	›	›
0156	œ	œ	œ
0157	•	•	•
0158	ž	ž	ž
0159	Ÿ	Ÿ	Ÿ
0160			
0161	¡	ϒ	¡
0162	¢	′	¢
0163	£	≤	£
0164	¤	⁄	¤
0165	¥	∞	¥
0166	¦	ƒ	¦

0167	**§**	♣	§
0168	**¨**	♦	¨
0169	**©**	♥	©
0170	**ª**	♠	ª
0171	**«**	↔	«
0172	**¬**	←	¬
0173	**-**	↑	-
0174	**®**	→	®
0175	**¯**	↓	¯
0176	**°**	°	°
0177	**±**	±	±
0178	**²**	″	²
0179	**³**	≥	³
0180	**´**	×	´
0181	**µ**	∝	µ
0182	**¶**	∂	¶
0183	**·**	•	·
0184	**¸**	÷	¸
0185	**¹**	≠	¹
0186	**º**	≡	º
0187	**»**	≈	»
0188	**¼**	…	¼
0189	**½**	⏐	½
0190	**¾**	⎯	¾
0191	**¿**	↵	¿
0192	**À**	ℵ	À
0193	**Á**	ℑ	Á
0194	**Â**	ℜ	Â
0195	**Ã**	℘	Ã
0196	**Ä**	⊗	Ä
0197	**Å**	⊕	Å

0198	Æ	∅	Æ
0199	Ç	∩	Ç
0200	È	∪	È
0201	É	⊃	É
0202	Ê	⊇	Ê
0203	Ë	⊄	Ë
0204	Ì	⊂	Ì
0314	Í	⊆	Í
0206	Î	∈	Î
0207	Ï	∉	Ï
0208	Ð	∠	Ð
0209	Ñ	∇	Ñ
0210	Ò	®	Ò
0211	Ó	©	Ó
0212	Ô	™	Ô
0213	Õ	∏	Õ
0214	Ö	√	Ö
0215	×	·	×
0216	Ø	¬	Ø
0217	Ù	∧	Ù
0218	Ú	∨	Ú
0219	Û	⇔	Û
0220	Ü	⇐	Ü
0221	Ý	⇑	Ý
0222	Þ	⇒	Þ
0223	ß	⇓	ß
0224	à	◊	à
0225	á	〈	á
0226	â	®	â
0227	ã	©	ã
0228	ä	™	ä

0229	**å**	∑	å
0230	**æ**	⎛	æ
0231	**ç**	⎜	ç
0232	**è**	⎝	è
0233	**é**	⎡	é
0234	**ê**	⎢	ê
0235	**ë**	⎣	ë
0236	**ì**	⎧	ì
0237	**í**	⎨	í
0238	**î**	⎩	î
0239	**ï**	⎪	ï
0240	**ð**		ð
0241	**ñ**	〉	ñ
0242	**ò**	∫	ò
0243	**ó**	⌠	ó
0244	**ô**	⎮	ô
0245	**õ**	⌡	õ
0246	**ö**	⎞	ö
0247	**÷**	⎟	÷
0248	**ø**	⎠	ø
0249	**ù**	⎤	ù
0250	**ú**	⎥	ú
0251	**û**	⎦	û
0252	**ü**	⎫	ü
0253	**ý**	⎬	ý
0254	**þ**	⎭	þ
0255	ÿ		ÿ

- Next page
- Previous page

Learn more [top]

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2a.1

The PC and its motherboard

The contents:

- Introduction to the PC
- The PC construction
- The motherboard (motherboard)
- POST and other ROM (BIOS etc.)

- Next page
- Previous page

Introduction to the PC

The technical term for a PC is *micro data processor* . That name is no longer in common use. However, it places the PC in the bottom of the computer hierarchy:

- Supercomputers and Mainframes are the largest computers - million dollar machines, which can occupy more than one room. An example is IBM model 390.

- Minicomputers are large powerful machines. They typically serve a network of simple terminals. IBM's AS/400 is an example of a minicomputer.

- Workstations are powerful user machines. They have the power to handle complex engineering applications. They use the UNIX or sometimes the NT operating system. Workstations can be equipped with powerful RISC processors like Digital Alpha or MIPS.

- The PCs are the Benjamins in this order: Small inexpensive, mass produced computers. They work on DOS, Windows , or similar operating systems. They are used for standard applications.

The point of this history is, that Benjamin has grown. He has actually been promoted to captain! Todays PCs are just as powerful as minicomputers and mainframes were not too many years ago. A powerful PC can easily keep up with the expensive workstations. How have we advanced this far?

The PC's success

[top]

The PC came out in 1981. In less than 20 years, it has totally changed our means of communicating. When the PC was introduced by IBM, it was just one of many different micro data processors. However, the PC caught on. In 5-7 years, it conquered the market. From being an IBM compatible PC, it became *the standard.*

If we look at early PCs, they are characterized by a number of features. Those were instrumental in creating the PC success.

- The PC was from the start standardized and had an open *architecture.*
- It was well documented and had great possibilities for expansion.
- It was inexpensive, simple and robust (definitely not advanced).

The PC started as IBM's baby. It was their design, built over an Intel processor (8088) and fitted to Microsoft's simple operating system MS-DOS.

Since the design was well documented, other companies entered the market. They could produce functionable copies (clones) of the central system software (BIOS). The central ISA bus was not patented. Slowly, a myriad of companies developed, manufacturing IBM compatible PCs and components for them.

The Clone was born. A clone is a copy of a machine. A machine, which can do precisely the same as the original (read *Big Blue* - IBM). Some of the components (for example the hard disk) may be identical to the original. However, the Clone has another name (Compaq, Olivetti, etc.), or it has no name at all. This is the case with "the real clones." Today, we differentiate between:

- Brand names, PCs from IBM, Compaq, AST, etc. Companies which are so big, so they develop their own hardware components.

- Clones, which are built from standard components. Anyone can make a clone.

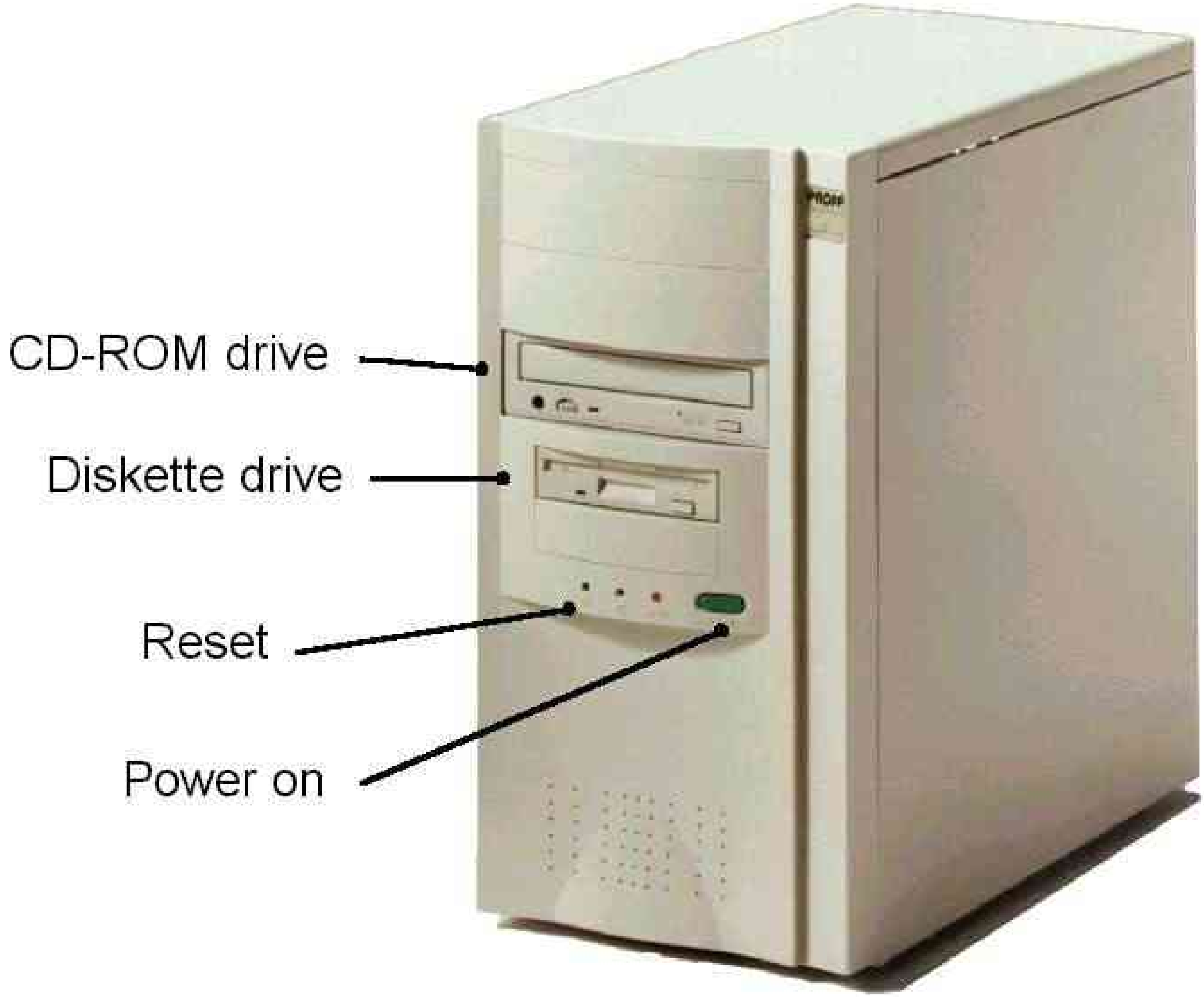
Since the basic technology is shared by *all* PCs, I will start with a review of that.

The PC construction

[top]

The PC consists of a *central unit* (referred to as the computer) and various peripherals. The computer is a box, which contains most of the working electronics. It is connected with cables to the peripherals.

On these pages, I will show you the computer and its components. Here is a picture of the computer:



Here is a list of the PC components. Read it and ask yourself what the words mean. Do you recognize all these components? They will be covered in the following pages.

Components in the central unit - the computer	Peripherals
The motherboard: CPU, RAM, cache, ROM chips with BIOS and start-up programs. Chip sets (controllers). Ports, buses and expansion slots. Drives: Hard disk(s), floppy drive(s), CD-ROM, etc. Expansion cards: Graphics card (video adapter), network controller, SCSI controller. Sound card, video and TV card. Internal modem and ISDN card.	Keyboard and mouse. Joystick Monitor Printer Scanner Loudspeakers External drives External tape station External modem

So, how are the components connected. What are their functions, and how are they tied together to form a PC? That is the subject of Click and Learn. So, please continue reading...

The von Neumann Model of the PC

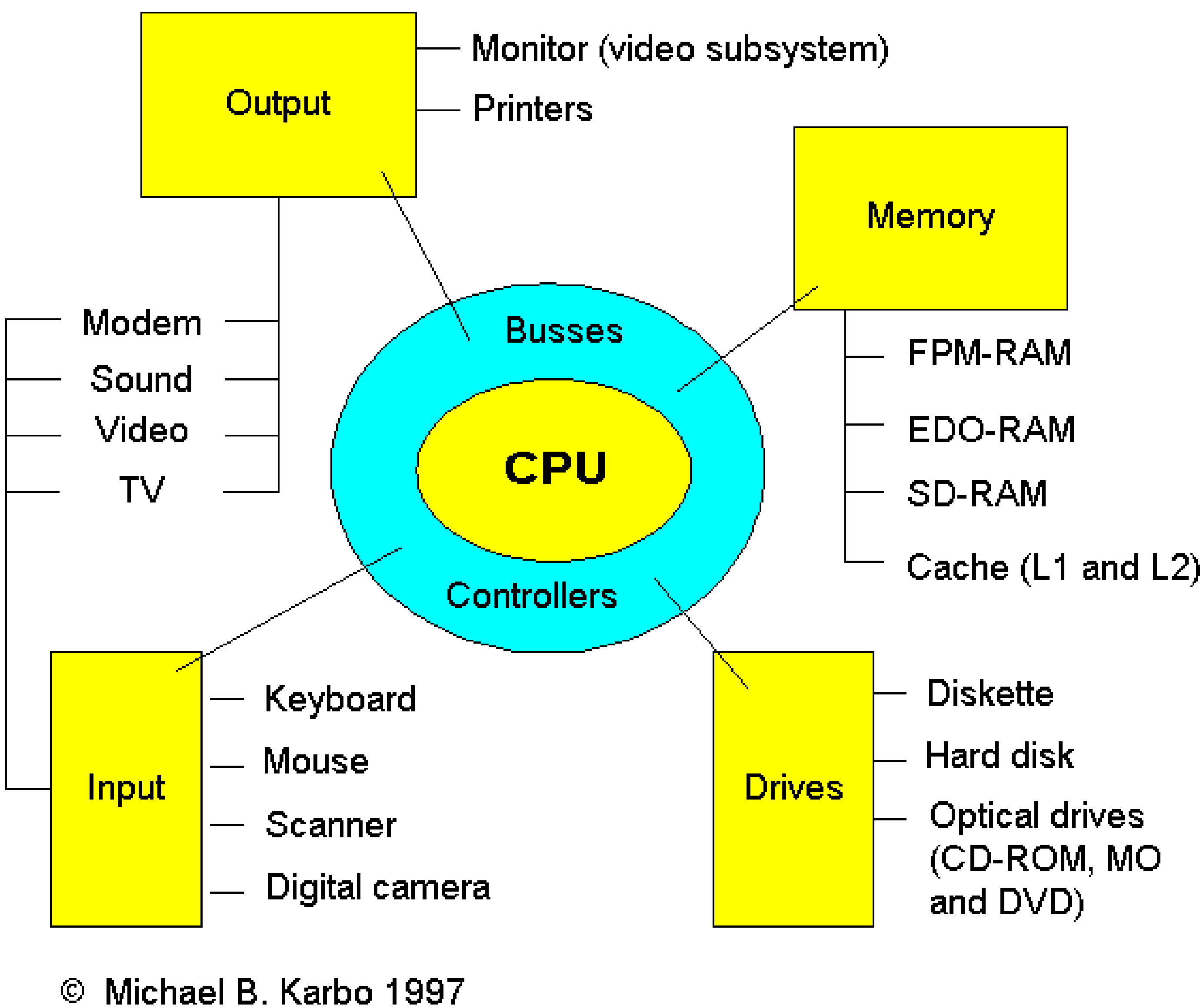
[top]

Computers have their roots 300 years back in history. Mathematicians and philosophers like Pascal, Leibnitz, Babbage and Boole made the foundation with their theoretical works. Only in the second half of this century was electronic science sufficiently developed to make practical use of their theories.

The modern PC has roots that go back to the USA in the 1940s. Among the many scientists, I like to remember John von Neumann (1903-57). He was a mathematician, born in Hungary. We can still use his computer design today. He broke computer hardware down in five primary parts:

- CPU
- Input
- Output
- Working memory
- Permanent memory

Actually, von Neumann was the first to design a computer with a working memory (what we today call RAM). If we apply his model to current PCs, it will look like this:



All these subjects will be covered.

Data exchange - the motherboard

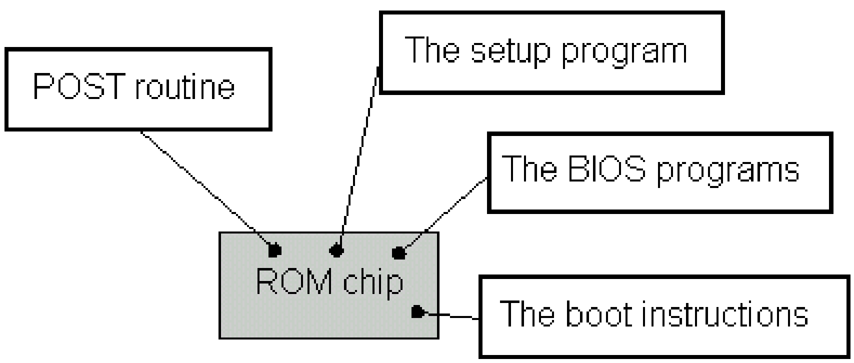
[top]

The ROM chips contain instructions, which are specific for that particular motherboard. Those programs and instructions will remain in the PC throughout its life; usually they are not altered.

Primarily the ROM code holds start-up instructions. In fact there are several different programs inside the start-up instructions, but for most users, they are all woven together. You can differentiate between:

- POST (Power On Self Test)
- The *Setup* instructions, which connect with the CMOS instructions
- BIOS instructions, which connect with the various hardware peripherals
- The *Boot* instructions, which call the operating system (DOS, OS/2, or Windows)

All these instructions are in ROM chips, and they are activated one by one during start-up. Let us look at each part.



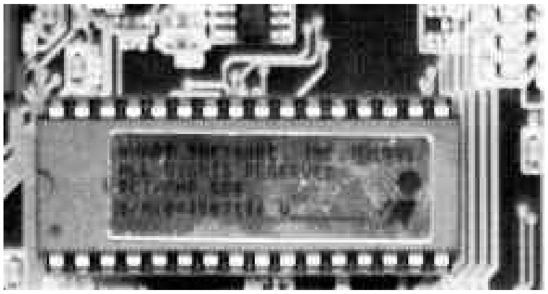
The suppliers of system software

[top]

All PCs have instructions in ROM chips on the motherboard. The ROM chips are supplied by specialty software manufacturers, who make BIOS chips. The primary suppliers are:

- Phoenix
- AMI (*American Megatrends*)
- Award

You can read the name of your BIOS chip during start-up. You can also see the chip on the system board. Here is a picture (slightly blurred) of an Award ROM chip:



Here is an AMI chip with BIOS and start-up instructions:

Let us look at the different components inside the ROM chip.

- Next page
- Previous page

Learn more **[top]**

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

I also recommend two books for further studies. Gunnar Forst: "PC Principles", from MIT is excellent. Also "The Winn L. Rosch Hardware Bible" from Brady covers the same subjects. Also "PC Intern" from Abacus is fine.

Links to BIOS information:

Mr BIOS FAQ

Click & Learn. Module 2a.2

The system software on the motherboard

The contents:

- The Setup program
- The POST
- The CMOS RAM
- Opening the Setup program

- Next page
- Previous page


Articles written
by Michael B. Karbo

The Setup programs

[top]

There are three elements in the start-up part of the ROM chip:

- The Initializing routine, which sets up the BIOS functions. The adapter ROM is integrated. A table covering all the BIOS programs is constructed. This is often called the *interrupt vectors*.
- The POST (the test programs)
- The *disk bootstrap loader*, which calls upon the operating system.

These programs are stored in the ROM chip, and they are activated one by one during the PC start-up.

The POST

Power On Self Test is the first instruction executed during start-up. It checks the PC components and that everything works. You can recognize it during the RAM test, which occurs as soon as you turn power on.

You may follow the checks being executed in this order, as the information are gathered:

1) Information about the graphics adapter
2) Information about the BIOS (name, version)
3) Information about the RAM (being counted)

As users, we have only limited ability to manipulate the POST instructions. But certain system boards enable the user to order a quick system check. Some enable the user to disable the RAM test, thereby shortening the duration of the POST. The duration of the POST can vary considerably in different PCs. On the IBM PC 300 computer, it is very slow. But you can disrupt it by pressing [Esc].

Error messages

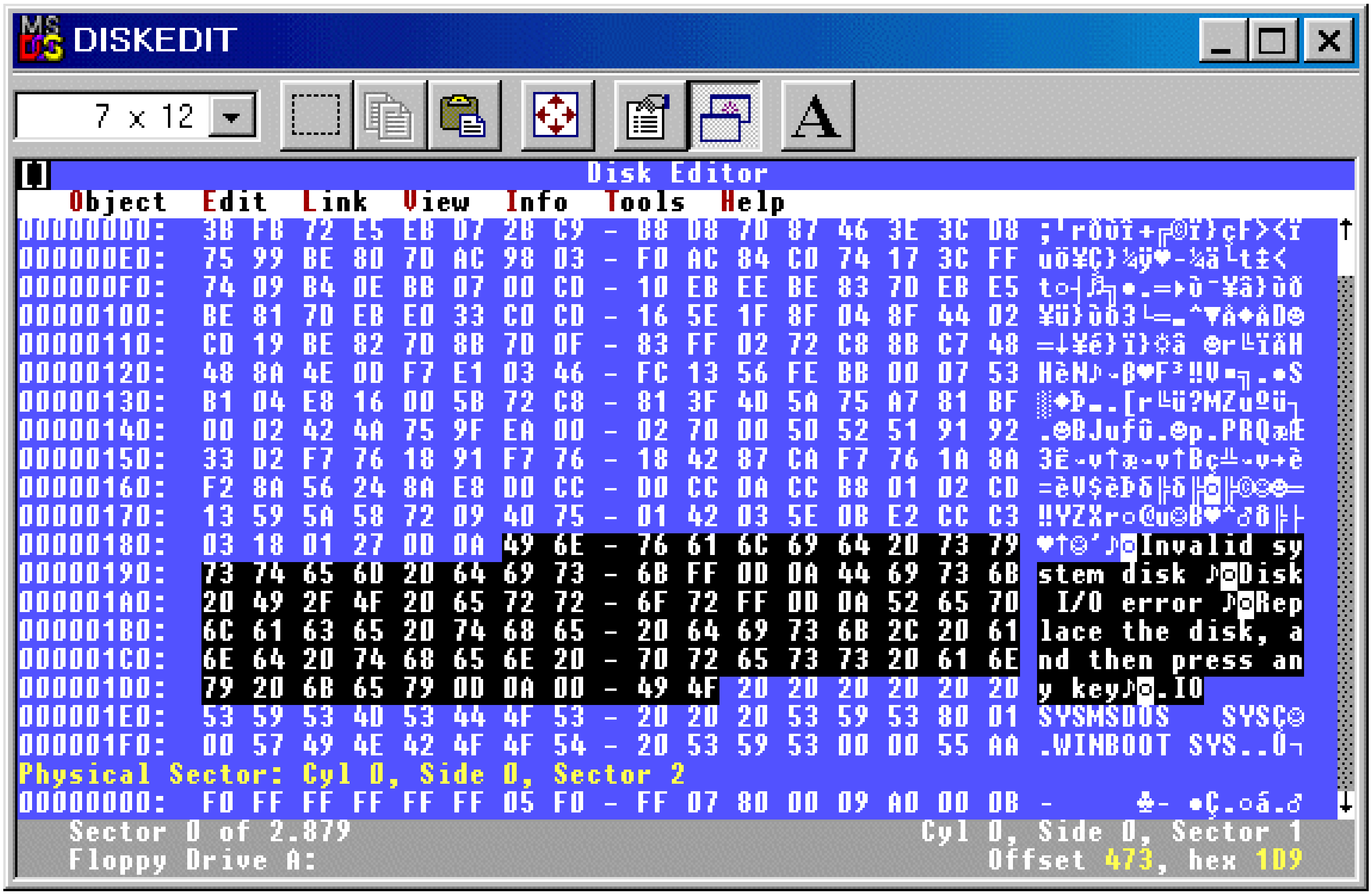
If POST detects errors in the system, it will write error messages on the screen. If the monitor is not ready, or if the error is in the video card, it will also sound a pattern of beeps (for example 3 short and one long) to identify the error to the user. If you want to know more of the beeps, you can find explanations on the Award, AMI and Phoenix web sites. For instance you will receive error messages if the keyboard is not connected or if something is wrong with the cabling to the floppy drive.

POST also reads those user data, which are found in the CMOS. This is discussed in the following chapter.

The bootstrap loader

The last part of the BIOS execution at start-up is the *bootstrap loader*. It is a tiny program, which only has one task: to find the bootsector on a disk (hard disk, floppy or another boot-drive).

The DOS Boot Record (DBR) also holds a *media descriptor* as well as information on the OS version. Please read module 6a4 on this issue. You can use DiskEdit (included in the "Norton Utilities") to read view the contents of the boot sector.



When the disk holds no boot strap routine, you get an error message like "Non-system disk, replace with system disk and press any key".

The bootstrap loader is the last step in BIOS execution during start-up. It hands over the control to the bootstrap routine found on the boot disk. The OS is being loaded.

CMOS RAM

[top]

CMOS stands for Complementary Metal Oxide Semiconductor. In PC's there is a small amount of memory in a special CMOS RAM chip. The data is maintained with electric power from a small battery.

CMOS is only a medium for storage. It could be used for any type of data. Here, it holds important system data, values to be used during the start process. These information take up maybe 100 or

200 bytes of data, and storage in the CMOS makes them instantly available to the POST and BIOS programs (loaded from ROM) during the start-up.

The values are regarding:

- Floppy and hard disk drives
- The keyboard
- The CPU, cache, chip set values, RAM type
- Date and time
- Much more ...

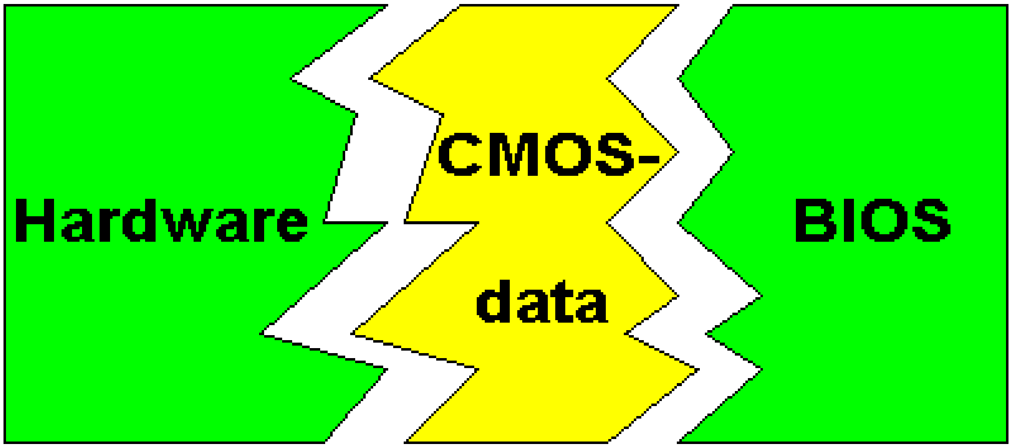
These data have to be set up correctly, and they are read during the start-up to make the PC operable.

Two types of data

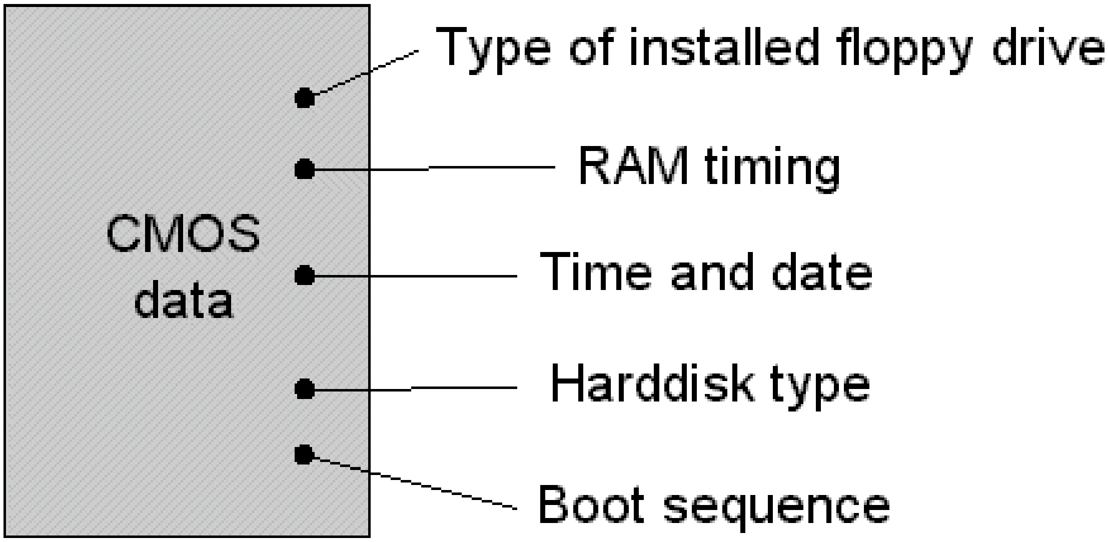
CMOS data can be divided in two groups:

- Data, which POST cannot find during the system test.
- Data, which contain user options.

For example, POST cannot by itself find sufficient information about the floppy drive(s). Floppy drives are so "dumb," that POST cannot read whether they are floppy drives or not, nor what type. About the same goes for IDE hard disks, while EIDE hard disks are a little more "intelligent," However, POST still needs assistance to identify them 100% correctly.



The same goes for RAM: POST can count how much RAM is in the PC. However, POST cannot always detect whether it is FPM, EDO or SD RAM. Since the CPU and BIOS reads data from RAM chips differently, depending on the RAM type, the type must be identified to setup the correct timing.



The configuration of CMOS data

The PC must be configured, be supplied with this information. That is done in the factory or store, where it is assembled. This information is stored in CMOS, where they stay. CMOS data only need to be updated, when different or additional hardware components are installed. This could be a different type hard disk or floppy disks or an new RAM type. Often the user can do this him/herself.

Other data in CMOS contain various *user options* . This is data, which you can write to CMOS. For example, you can adjust date and time, which the PC then adjusts every second. You can also choose between different system parameters. Maybe you want a short system check instead of a long one. Or if you want the PC to try to boot from hard disk C before trying floppy disk A, or vice versa. These options can be written to CMOS.

Many of the options are of no interest to the ordinary user. These are options, which regard controller chips on the motherboard, which *can* be configured in different ways. Ordinarily, there is no need to make such changes. The motherboard manufacturer has already selected the optimal configurations. They recommend in their manuals, that you do not change these *default* settings.

We can conclude, that CMOS data are essential system data, which are vital for operation of the PC. Their special feature is, that they are user adjustable. Adjustments to CMOS are made during start-up.

Opening the Setup program

[top]

You communicate with the BIOS programs and the CMOS memory through the so-called Setup program. This gives us a very simple user interface to configuring the PC with these vital data.

Typically you reach the Setup program by pressing [Delete] immediately after you power up the PC. That brings you to a choice of setup menus. You leave Setup by pressing [Esc], and choose "Y" to restart the PC with the new settings. Generally, you should not change these settings, unless you know precisely what you are doing.

Here you see the start menu of the American Megatrends BIOS Setup program, which has a kind of graphical user interface. You are supposed to use the mouse:

- Next page
- Previous page

Learn more **[top]**

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

I also recommend two books for further studies. Gunnar Forst: "PC Principals", from MIT is excellent. Also "The Winn L. Rosch Hardware Bible" from Brady covers the same subjects. Also "PC Intern" from Abacus is fine.

Links to BIOS information:

BIOS Guide

Mr BIOS FAQ

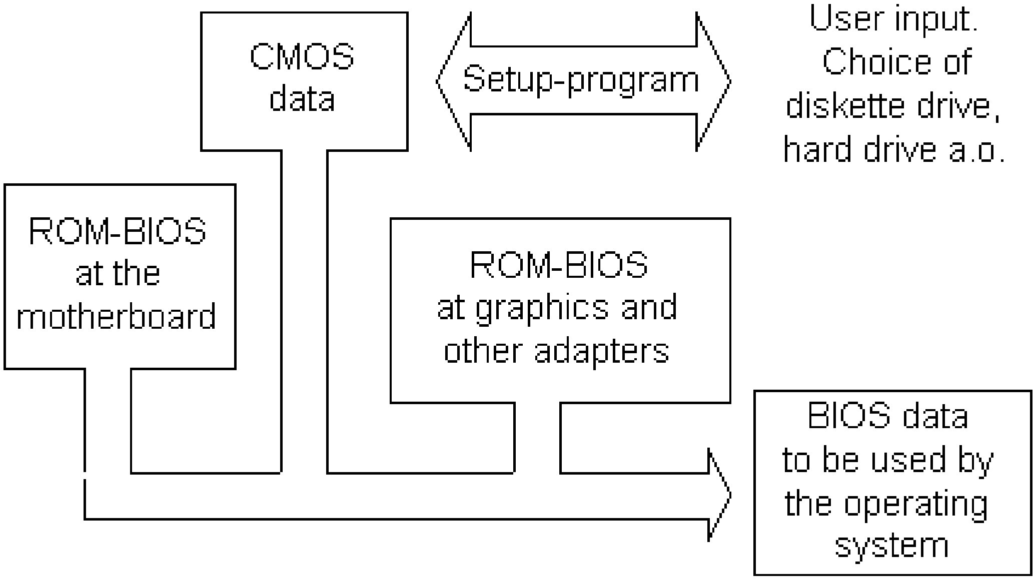
KarbosGuide.com. Module 2a.3

Using the system software of the motherboard

The contents:

- What use of Setup program?
- Modifying the boot sequence
- Images from the setup program

- Next page
- Previous page



What can I use the Setup program for? [top]

The Setup program can do many things for you. However, be careful. You should not change any values within the menus, unless you know what you are doing. Otherwise your PC may not function properly.

You *have* to enter Setup, if you install a different type or additional disk drive in your PC. Certain BIOSs will also need adjustment of its settings, if a CDROM drive is installed on one of the EIDE channels.

The Standard values

The standard values in the CMOS Setup are used to configure:

- The date and time.
- The keyboard.
- The display.
- The diskette drive.
- EIDE units number 1-4 (typically hard disks and CD-ROM-drive).

The values for date and time are stored in the CMOS RAM. You can always change them, from Setup or from DOS, Windows or any other OS.

The keyboard - obviously it has to be there. But it is possible to configure the PC to work *without* a keyboard. Otherwise the PC will error if the keyboard is missing.

The display is always VGA. From older times the Setup gives you options as EGA, CGA and MDA. You won't need them!

Diskette drive has to be selected. You can choose to have A: or B: or both. Each drive can be of five types or more. You probably have the 1.44 MB floppy drive. You choose among the options using [PgUp] and [PgDn]. Modern super floppies like Zip and LS120 are not to be installed as diskette drives, they are EIDE units.

The hard disk is the most important unit to install in this part of the Setup. With the modern motherboards and the EIDE drives you may experience an automatic configuration during the *Auto detect* . In other situations you have to run the auto detect yourself. With older drives, you have to enter all the CHS-values for the drive (number of cylinders, heads and sectors.

The BIOS Feature Setup

The Feature Setup is the next layer in the CMOS setup. Here you can choose among options like:

- Quick execution of POST (a good thing).
- Choice of boot device EIDE/SCSI. If you have both types of hard drives, which one is to be booted?
- The boot sequence.
- ....

Modifying the boot sequence

You can change the boot sequence from A:, C: to C:, A:. That means, that the PC will not try to boot from any diskette in the A drive. This will protect you from certain virus attacks from the boot sector. Also, the boot process will not be blocked by any diskette in the A drive. If you need to boot from A-drive (for example, if you want to install Windows 98), you have to enter Setup again, and change the boot sequence to A:, C:. That is no problem.

Power Management

You also use the Setup program to regulate the *power management* , which is the power saving features in the motherboard. For example, you can make the CPU shut down after one minute of no activity. There are plenty of settings available in this area. The power management functions found on the PC's motherboard will cooperate with the operating system. Especially Windows 98 is very good at using the power management.

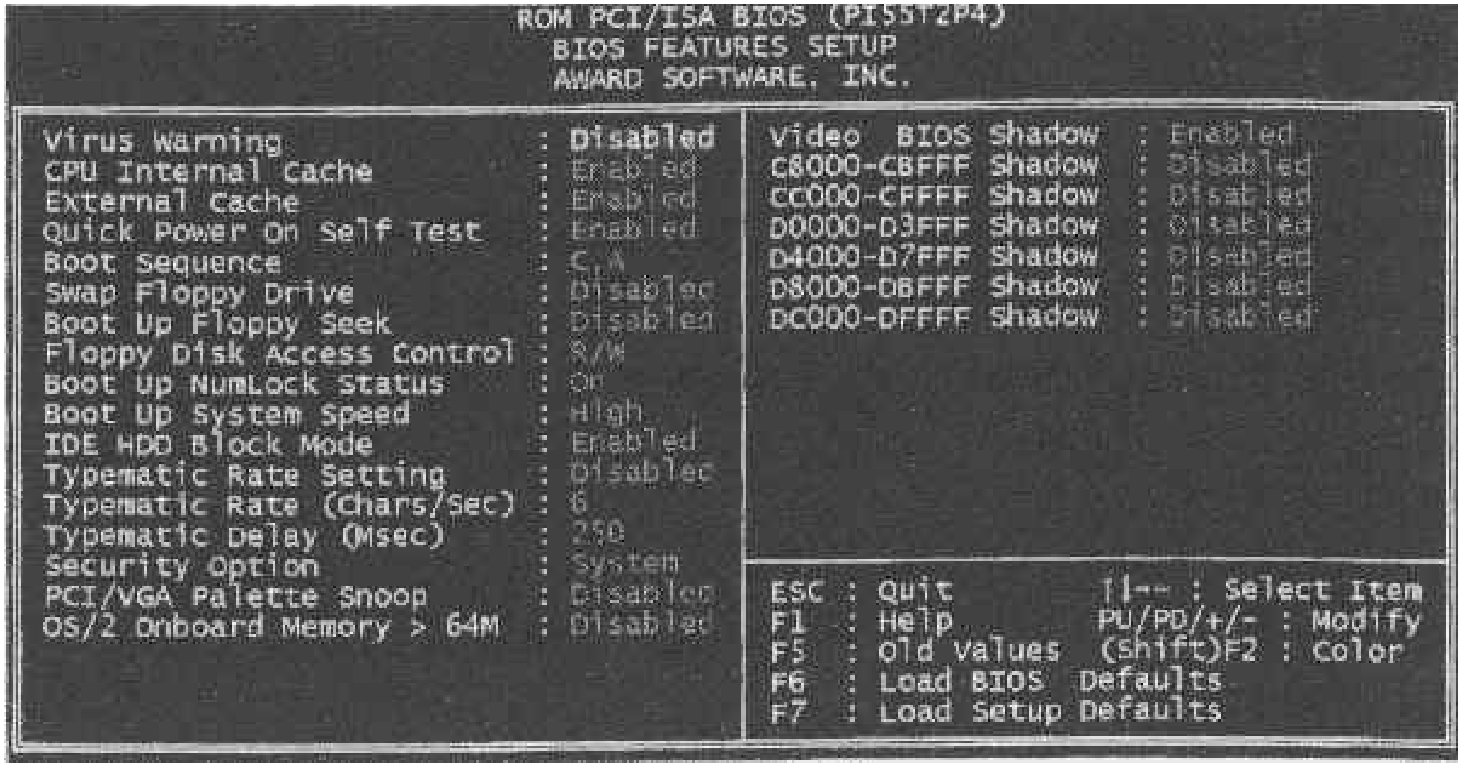
Password Protection

You can protect the Setup program with a password. This is used widely in schools, where the teachers do not want the little nerds to make changes in the setup. Please remember the password (write it down in the motherboard manual). If you forget it you have to remove the battery from the motherboard. Then all user input to the CMOS is erased - including the password.

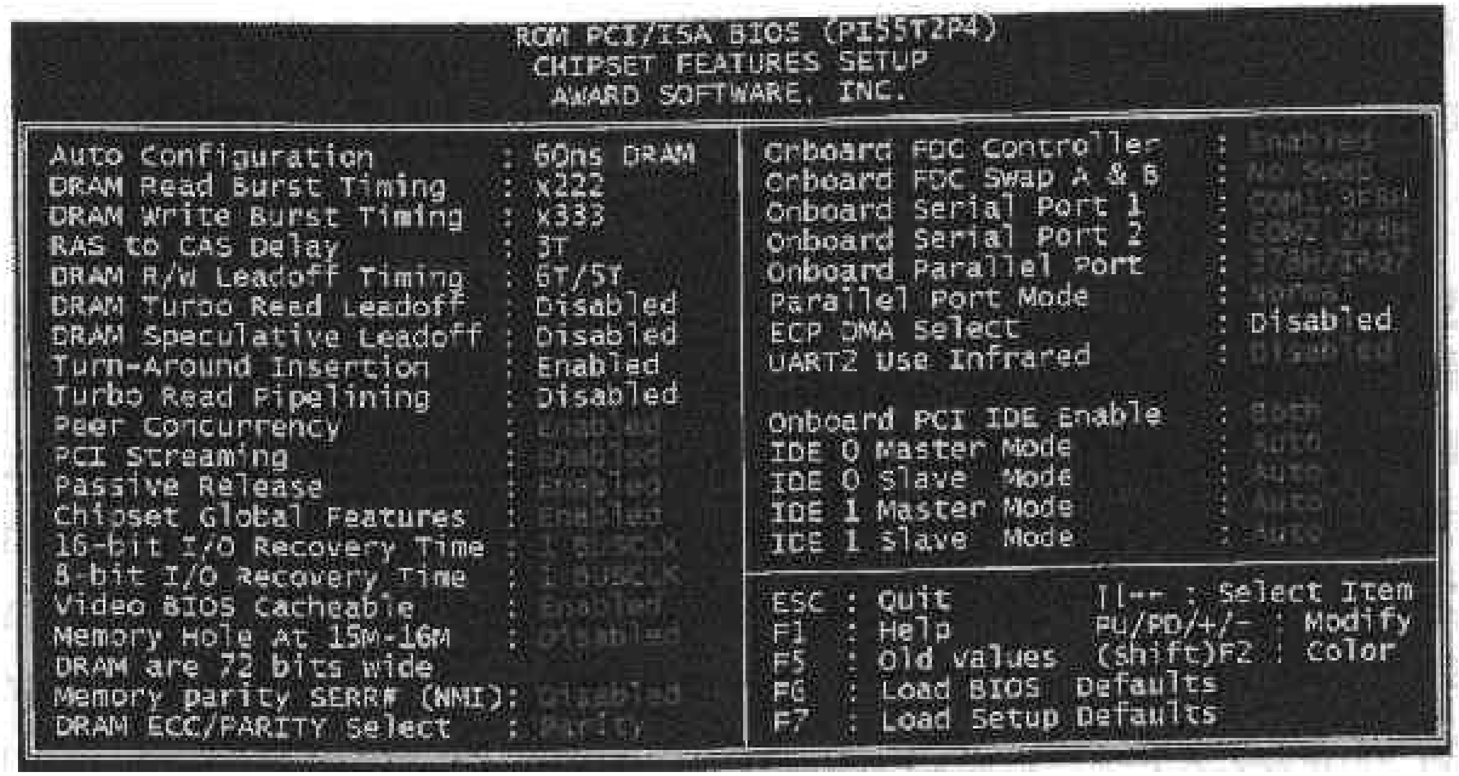
[top]

Images from the Setup program

Here is a scanned image from a Setup program. It belongs a very fine board from ASUS. Here you see the "BIOS Feature Setup," where you can select start-up choices:



Here we are in the special "Chip set Feature Setup." These choices relate to the chip sets and, most likely, need no changes:



- Next page
- Previous page

Learn more **[top]**

Module 2b. About the boot process and system bus

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

I also recommend two books for further studies. Gunnar Forst: "PC Principals", from MIT is excellent. Also "The Winn L. Rosch Hardware Bible" from Brady covers the same subjects. Also "PC Intern" from Abacus is fine.

Links to BIOS information:

BIOS Guide

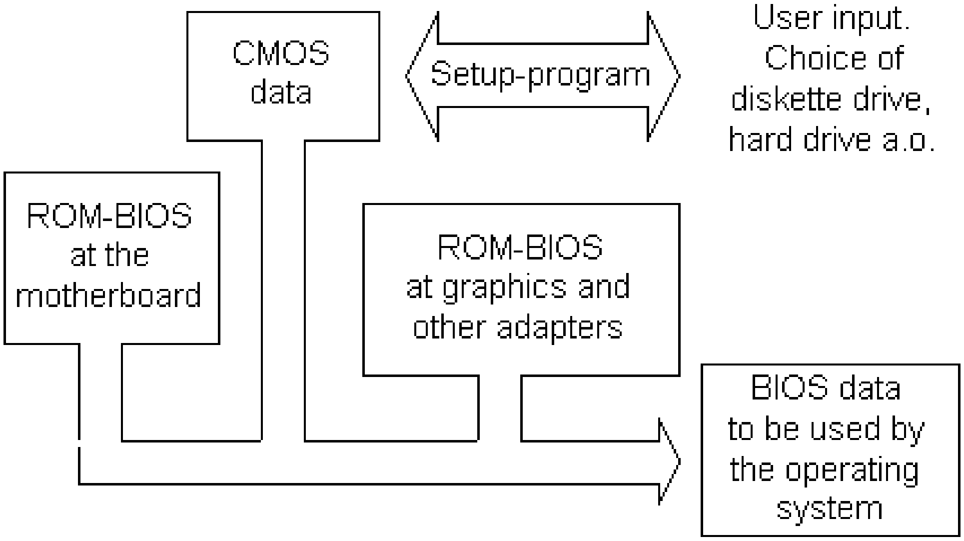
Mr BIOS FAQ

[Main page]	[Contact]	[Karbo's Dictionary]	[The Windows 98 pages]

KarbosGuide.com. Module 2a.4.

The system software of hardware

The contents:



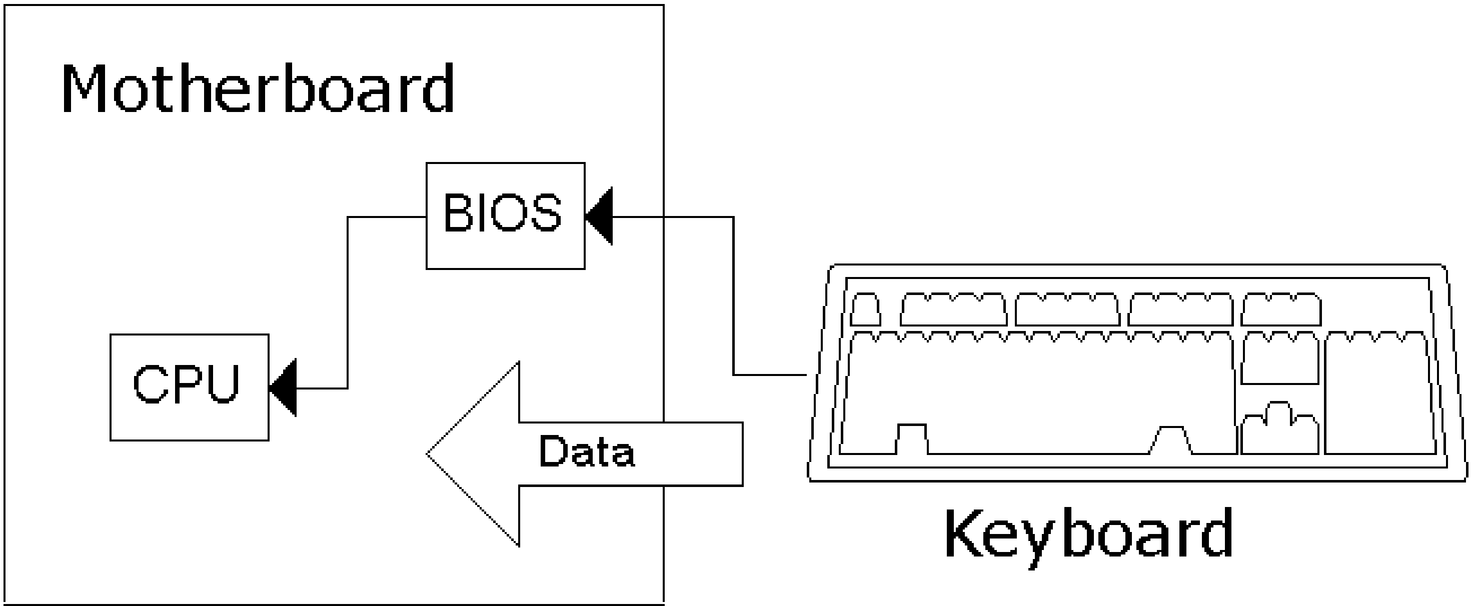
- Next page
- Previous page

The BIOS in adapter ROM

[top]

During the start-up process the BIOS programs are read from the ROM circuits. BIOS stands for *Basic Input Output System* and it is small program routines which controls specific hardware units.

For instance you have a BIOS routine which reads the keyboard:



The BIOS is a part of the modular design of the IBM Compatible PC. The OS and other programs access the hardware units by making requests to the BIOS routines.

BIOS typically occupies 64 KB, and the programs are stored in ROM chips on the motherboard.

The reserved areas

In the original PC design we only had 1 MB of RAM. This memory was adressed using hex numbers, so each byte had its own address going from 00000h to FFFFFh.

Important parts of the system software is mapped into this range, where we also find two *reserved* areas:

Hex address	Kilobytes	Occupied by
C0000-C8000	768-800	BIOS from the video card
F0000 - FFFFF	960-1024	BIOS from the Motherboard

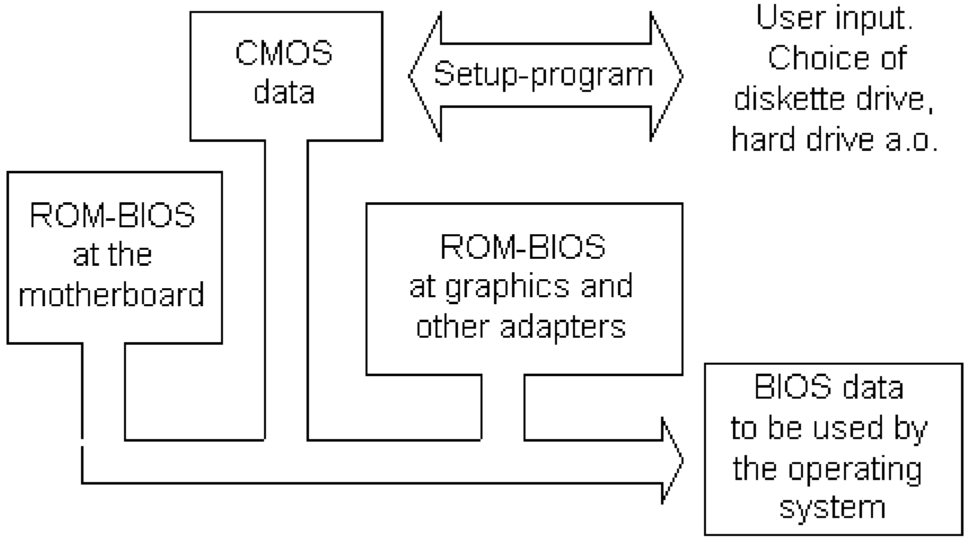
These two ranges are reserved for this special adapter ROM. Other adapters cannot *map* their BIOS routines into these addresses.

If it is setup to *shadowing* ("Shadow RAM" in the Setup utility), then this BIOS code is copied into RAM. If not, it has to be read directly from the ROM circuit. The last access is slower.

BIOS on many adapters

There are BIOS codes on many adapters (expansion cards). The adapters are external hardware, which are connected to and “integrated” with the motherboard during the hardware configuration and internalizing.

The adapters hold their own BIOS code making them functional. This BIOS must be included during the configuration. Therefore, the adapter ROM is read during start-up, and the program code is “woven” together with other BIOS programs and the CMOS data. It is all written into RAM, where it is ready for the operating system, as you can see here:



The BIOS routines are not always in use. They can be regarded as basic program layers in the PC, giving it a simple functionality.

Many programs routinely bypass BIOS. In that case, they "write direct to hardware", as we say. Windows contains program files, which can be written directly to all kinds of hardware - bypassing BIOS routines. One example is the COM ports. If you use the BIOS routines connected with them, you can transmit only at max. 9600 baud on the modem. That is insufficient. Therefore, Windows will assume control over the COM port.

BIOS update

BIOS programs can be *updated* . The modern motherboard has the BIOS instructions in *flash ROM,* which can be updated. You can get new BIOS software from your supplier or on the Internet, which can be read onto the motherboard. The loading is a special process, where you might need to change a *jumper switch* on the motherboard. Usually, you do not need to do this, but it is a nice available option.

ATX motherboards

[top]

The latest PC electronic standard is called ATX. It consists of a new type motherboard with a specific physical design like the traditional board (30.5 cm X 19 cm). However the board has been shifted 90 degrees for a better placing of the units.

The I/O connectors COM1, COM2 and LPT, keyboard, mouse and USB are mounted directly on the motherboard. The ATX board requires specifically designed chassises with an I/O access opening measuring 1¾ by 6¼ inch. ATX is designed by Intel, but has gained general acceptance.

The ATX motherboard is more ”intelligent” than the ordinary type. In a few years, it will be wide spread. It includes advanced control facilities, where the BIOS program continually checks the CPU temperature and voltages, the cooling fans RPM, etc. If over heating occurs, the PC will shut down automatically. The PC can also be turned on by for example modem signals, since the power supply is controlled by the motherboard. The on/off button will turn the PC "down" without turning it completely off.

If you want a PC designed for the future, the ATX layout is what you should go for.

- Next page
- Previous page

Learn more [top]

Module 2b. About the boot process and system bus

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

I also recommend two books for further studies. Gunnar Forst: "PC Principals", from MIT is excellent. Also "The Winn L. Rosch Hardware Bible" from Brady covers the same subjects. Also "PC Intern" from Abacus is fine.

Links to BIOS information:

BIOS Guide

Mr BIOS FAQ

[Main page] [Contact] [Karbo's Dictionary] [The Windows 98 pages]

KarbosGuide.com. Module 2b1.

About the System Bus

In this module, you can read about the following subjects, which add to our tour of the PC:

- The boot process
- Data on the motherboard

- Next page
- Previous page

The boot process

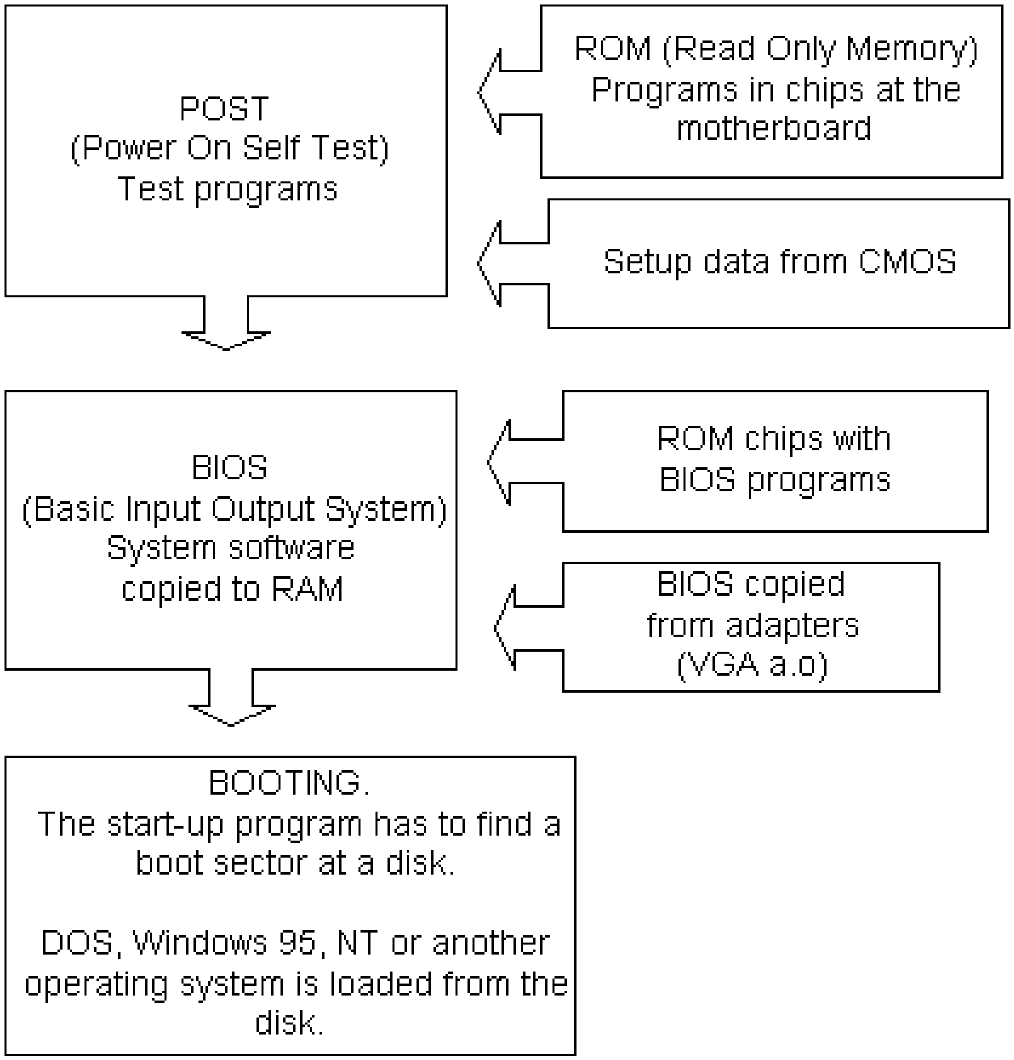
[top]

The last step in the PC start-up is reading the operating system. The start-up program is instructed to find the *Master Boot Record*. This is located in the very first sector on either hard disk (C) or floppy drive A. From the MBR it reads the boot-strap which points to the location of the startup files of the Operating System.

By default, the PC will look for a boot sector in floppy drive A. That is why the PC "drops dead" if there is a different diskette in A drive. If there is no diskette in A drive, the start-up program will search for the boot sector on hard drive C. When the boot sector is found, a small program segment (*boot-strap*) is read from there. The boot-strap then takes over control of the PC. The start-up program has done its job. Now DOS, Windows , or another operating system takes control.

Read more about boot sectors, etc. in module 6a, which deals with file systems.

Here is an illustration of the start-up process:



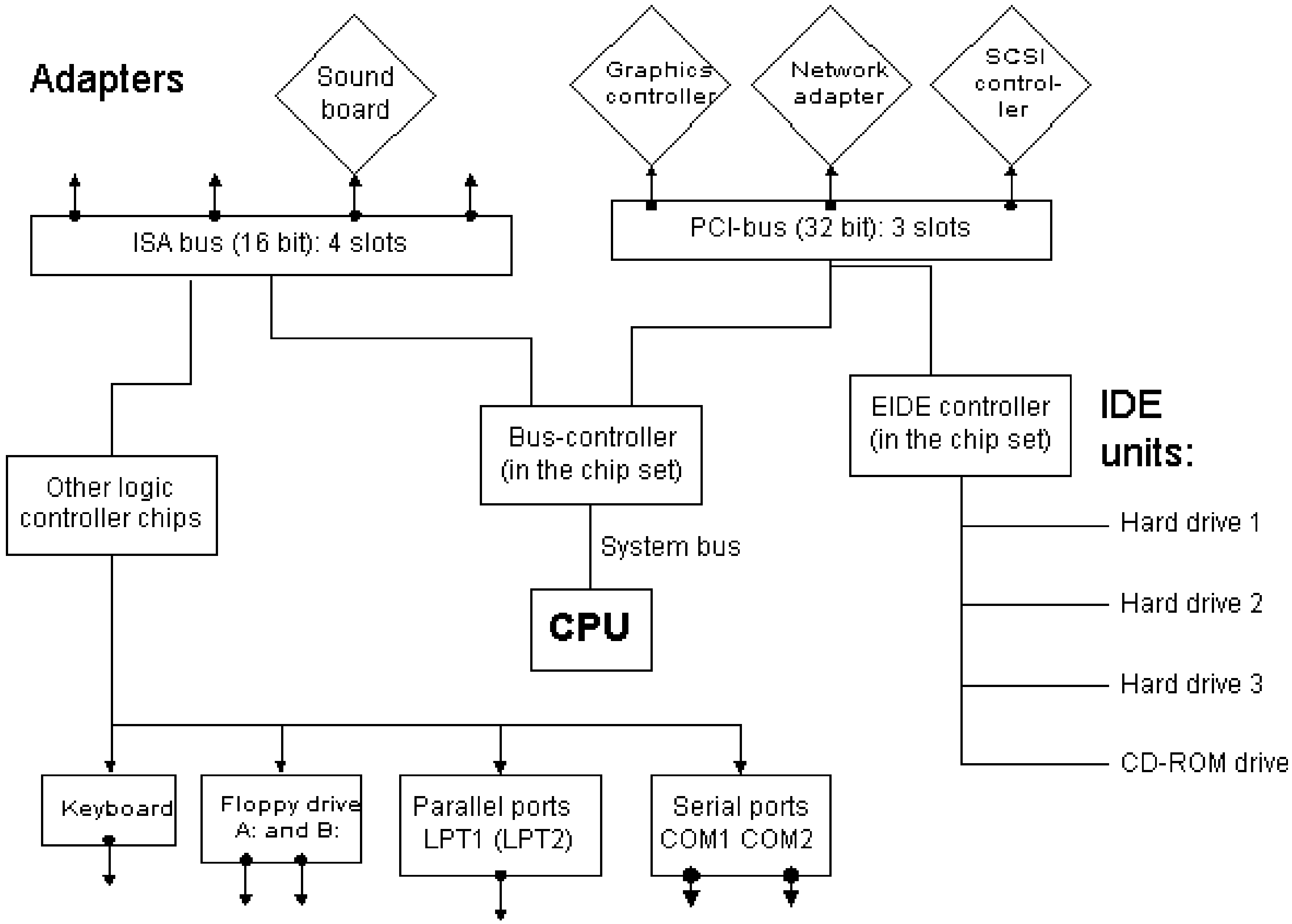
The data flow on the motherboard

[top]

On the motherboard, you will find the CPU, which is the "brain" of the PC and the *buses*. The buses are the nerve system of the motherboard. They connect the CPU to all the other components. There are at least three buses, which you can see below. You can read more about those on the following pages.

The buses are the PC's expressways. They are "wires" on the circuit board, which transmit data between different components.

One "wire" can move one bit at a time. In the following text, we start from a typical Pentium board. We will look at buses, chip sets and CPUs. Here is an illustration of some of the motherboard "logic." You can print it:



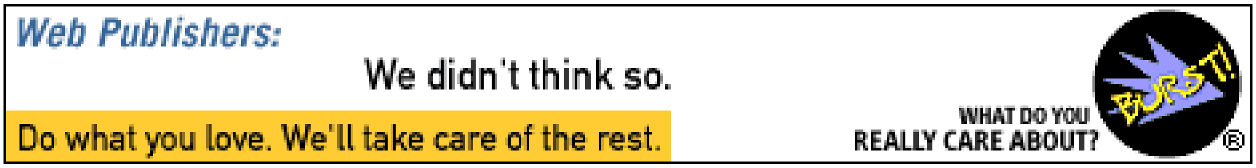
- Next page
- Previous page

Learn more **[top]**

Read more about the motherboards chip set in module 2d

Read more about RAM in module 2e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]



KarbosGuide.com. Module 2b2.

About the System Bus

In this module, you can read about the following subjects, which add to our tour of the PC:

- PC buses, an intro
- The system bus
- 66 MHz bus
- 100 MHz bus

- Next page
- Previous page

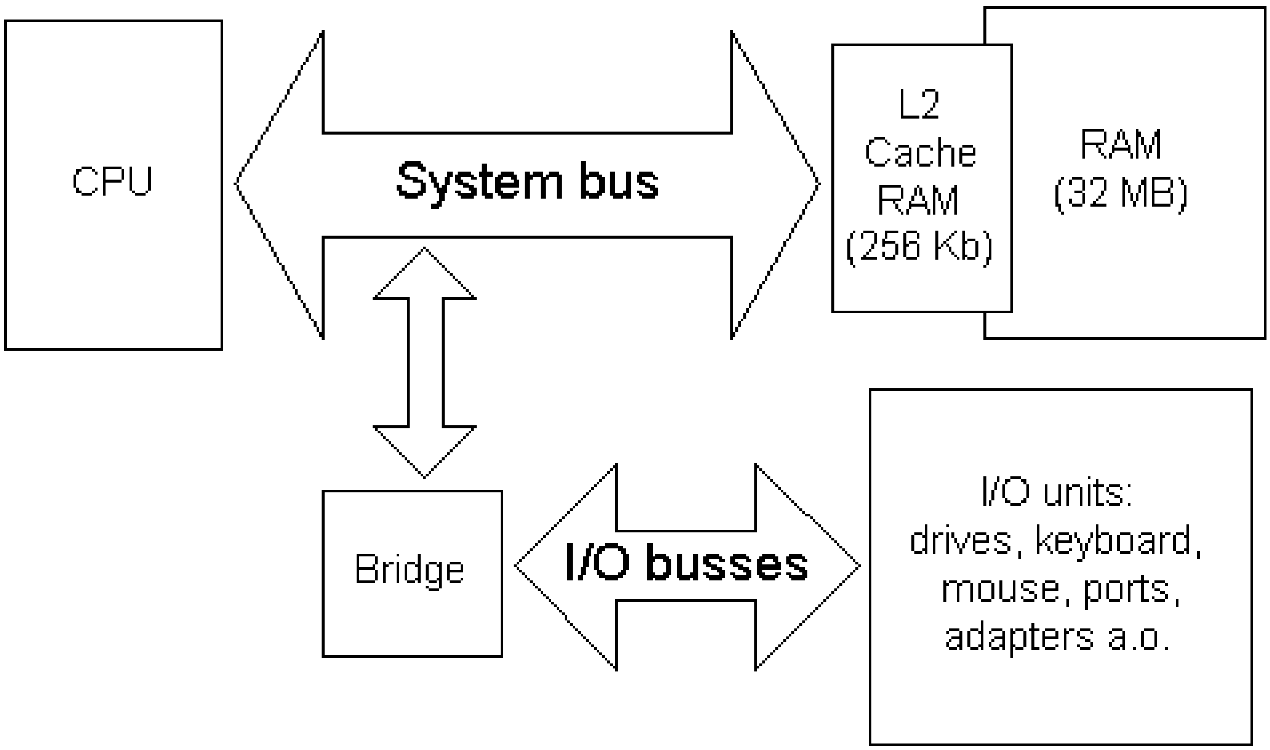
Introduction to the PC buses

[top]

The PC receives and sends its data from and to *buses*. They can be divided into:

- The system bus, which connects the CPU with RAM
- I/O buses, which connect the CPU with other components.

The point is, that the *system bus* is the central bus. Actually, it connects to the I/O buses, as you can see in this illustration. It is not completely correct, since the architecture is much more complex, but it shows the important point, that the I/O-buses usually derive from the system bus:



You see the central system bus, which connects the CPU with RAM. A *bridge* connects the I/O buses with the system bus and on to RAM. The bridge is part of the PC chip set, which will be covered in module 2c.

---

3 different I/O buses

[top]

The I/O buses move data. They connect all I/O devices with the CPU and RAM. I/O devices are those components, which can receive or send data (disk drives, monitor, keyboard, etc.). In a modern Pentium driven PC, there are two or three different I/O buses:

- The ISA bus, which is oldest, simplest, and slowest bus.
- The PCI bus, which is the fastest and most powerful bus.
- The USB bus, which is the newest bus. It may in the long run replace the ISA bus.

The three I/O buses will be described later. Here, we will take a closer look at the PC's fundamental bus, from which the others are branches from.

The system bus

[top]

The system bus connects the CPU with RAM and maybe a buffer memory (L2-cache). The system bus is the central bus. Other buses branch off from it.

The system bus is on the motherboard. It is designed to match a specific type of CPU. Processor technology determines dimensioning of the system bus. At the same time, it has taken much technological development to speed up "traffic" on the motherboard. The faster the system bus gets, the faster the remainder of the electronic components must be..

The following three tables show different CPUs and their system buses:

Older CPUs	System bus width	System bus speed
8088	8 bit	4.77 MHz
8086	16 bit	8 MHz
80286-12	16 bit	12 MHz
80386SX-16	16 bit	16 MHz
80386DX-25	32 bit	25 MHz

We see, that system bus speed follows the CPU's speed limitation. First at the fourth generation CPU 80486DX2-50 are doubled clock speeds utilized. That gives the CPU a higher *internal* clock frequency. The *external* clock frequency, used in the system bus, is only half of the internal frequency:

CPUs in the 80486 family	System bus width	System bus speed
80486SX-25	32 bit	25 MHz
80486DX-33	32 bit	33 MHz
80486DX2-50	32 bit	25 MHz
80486DX-50	32 bit	50 MHz

80486DX2-66	32 bit	33 MHz
80486DX4-100	32 bit	40 MHz
5X86-133	32 bit	33 MHz

66 MHz bus

[top]

For a long time all Pentium based computers ran at 60 or 66 MHz on the system bus, which is 64 bit wide:

CPUs in the Pentium family	System bus width	System bus speed
Intel P60	64 bit	60 MHz
Intel P100	64 bit	66 MHz
Cyrix 6X86 P133+	64 bit	55 MHz
AMD K5-133	64 bit	66 MHz
Intel P150	64 bit	60 MHz
Intel P166	64 bit	66 MHz
Cyrix 6X86 P166+	64 bit	66 MHz
Pentium Pro 200	64 bit	66 MHz
Cyrix 6X86 P200+	64 bit	75 MHz
Pentium II	64 bit	66 MHz

100 MHz bus

The speed of the system bus has increased in 1998. Using PC100 SDRAM a speed of 100 MHz is well proven and the use of RDRAM will give us much higher speeds.

However the rise from 66 MHz to 100 MHz has the greatest impact on Socket 7 CPUs and boards. In the Pentium-II modules 70-80% of the traffic is inside the SEC module, holding both L1 and L2 cache. And the module has its own speed independent of the system bus.

With the K6 the increase of system bus speed gives a vastly improved performance since the traffic between L1 and L2 cache crosses the system bus.

133 MHz

Intel's 820 and 815 chipsets to be used with Pentium III work with 133 MHz RAM as well as several VIA chipsets do.

In AMD's Athlon the system bus architecture was changed; it is not really a system bus any longer. Hence Athlon chipsets may work with many types of RAM.

Processor	Chip set	System bus speed	CPU speed
Intel Pentium II	82440BX 82440GX	100 MHz	350, 400, 450 MHz
AMD K6-2	Via MVP3ALi Aladdin V	100 MHz	250, 300, 400 MHz
Intel Pentium II Xeon	82450NX	100 MHz	450, 500 MHz
Intel Pentium III	i815 i820	133 MHz	600, 667 MHz and up
AMD Athlon	VIA KT133 and others	200 MHz	600 - 1000 MHz

With the 100 MHz bus, we dicovered that motherboards have to be well constructed with good power supply and many capacitors.

Newer buses

As mentioned under AMD Athlon, "system bus" is not that relevant a term looking at modern motherboards. The bus to RAM becomes separated from the other buses and this design opens up for better bandwidth between the CPU and the RAM.

Intels use of Rambus RAM working at 400 MHz as well as PC2100 RAM on non-Intel boards follows this trend.

The DDRAM operates with interfaces working at 200, 266 and 333 MHz.

- Next page
- Previous page

Learn more [top]

Read more about the motherboards chip set in module 2d

Read more about RAM in module 2e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

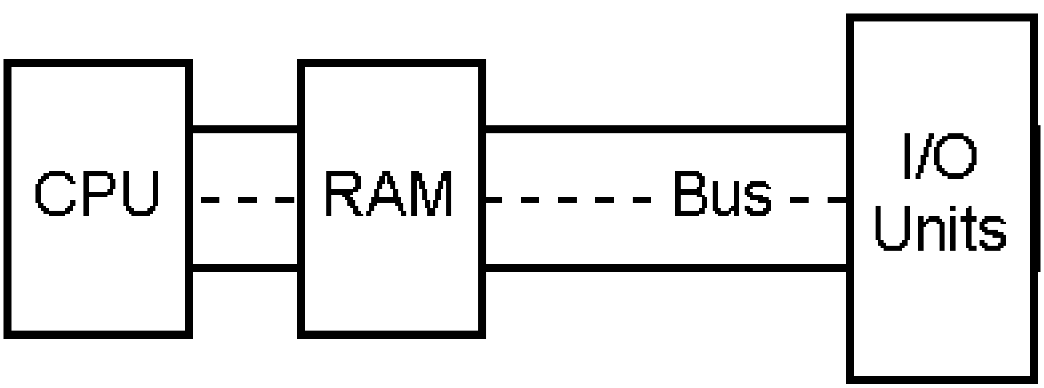
KarbosGuide.com. Module 2c.1

About the I/O buses

On these pages, you can read about the important system bus derivatives, the different I/O buses:

- Introduction to the I/O buses
- Technical and historical background for the I/O buses

- Next page
- Previous page



Introduction to the I/O buses

[top]

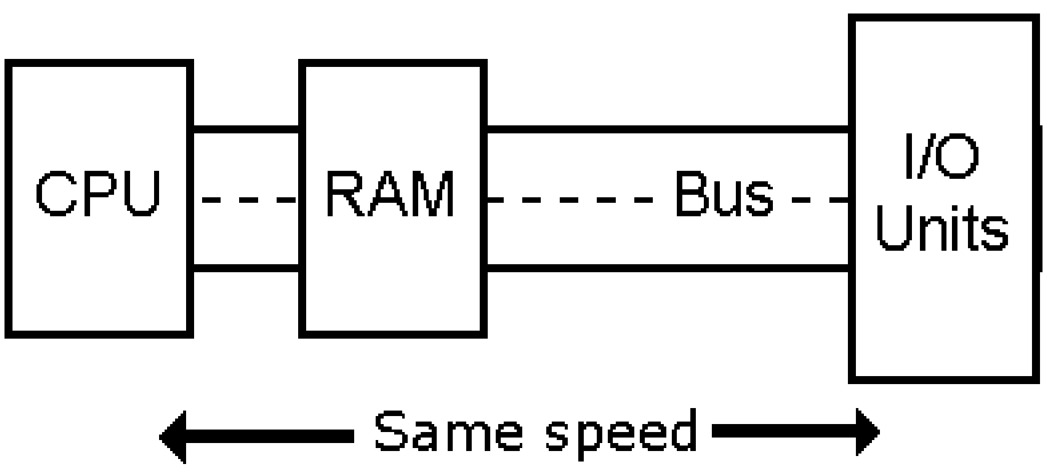
We have seen before, that the PC's buses are the fundamental data "highways" on the system board. The "first" bus is the *system bus*, which connects the CPU with RAM. In older designs it was a local bus. In newer designs this bus is called the front side bus (FSB).

The typical local bus has a speed and width depending on the type CPU installed on the motherboard. Typically, the system bus will be 64 bits wide and run at 66, 100 or 133 MHz. These high speeds create electrical noises and other problems. Therefore, the speed must be reduced for data reaching the expansion cards and other more peripheral components.

Very few expansion cards can operate at more than 40 MHz. Then the electronics shut down. The chips can just not react faster. Therefore, the PC has additional buses.

Originally only one bus

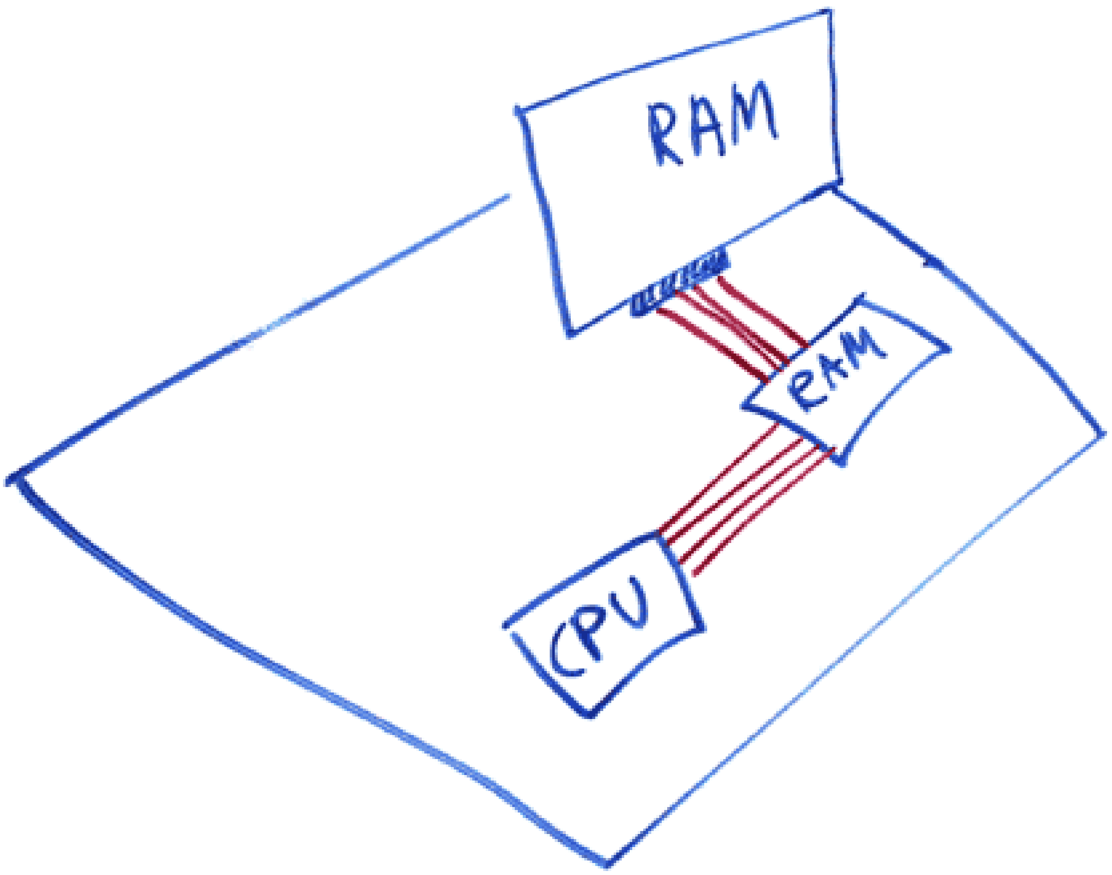
However, the first PCs had only one bus, which was common for the CPU, RAM and I/O components:



The older first and second generation CPUs ran at relatively low clock frequencies, and all system components could keep up with those speeds.

RAM on adapters

Among other things, that allowed additional RAM to be installed in expansion slots in the PC, by installing an adapter in a vacant expansion slot. An adapter, where RAM was mounted:



This setup would be unthinkable today. However it is truely a local bus. All units are united on one bus using the same clock.

First in 1987, Compaq figured out how to separate system bus from I/O bus, so they could run at different speeds. This multi-bus architecture has been industry standard ever since. Modern PCs also have more than one I/O bus.

What does an I/O bus do?

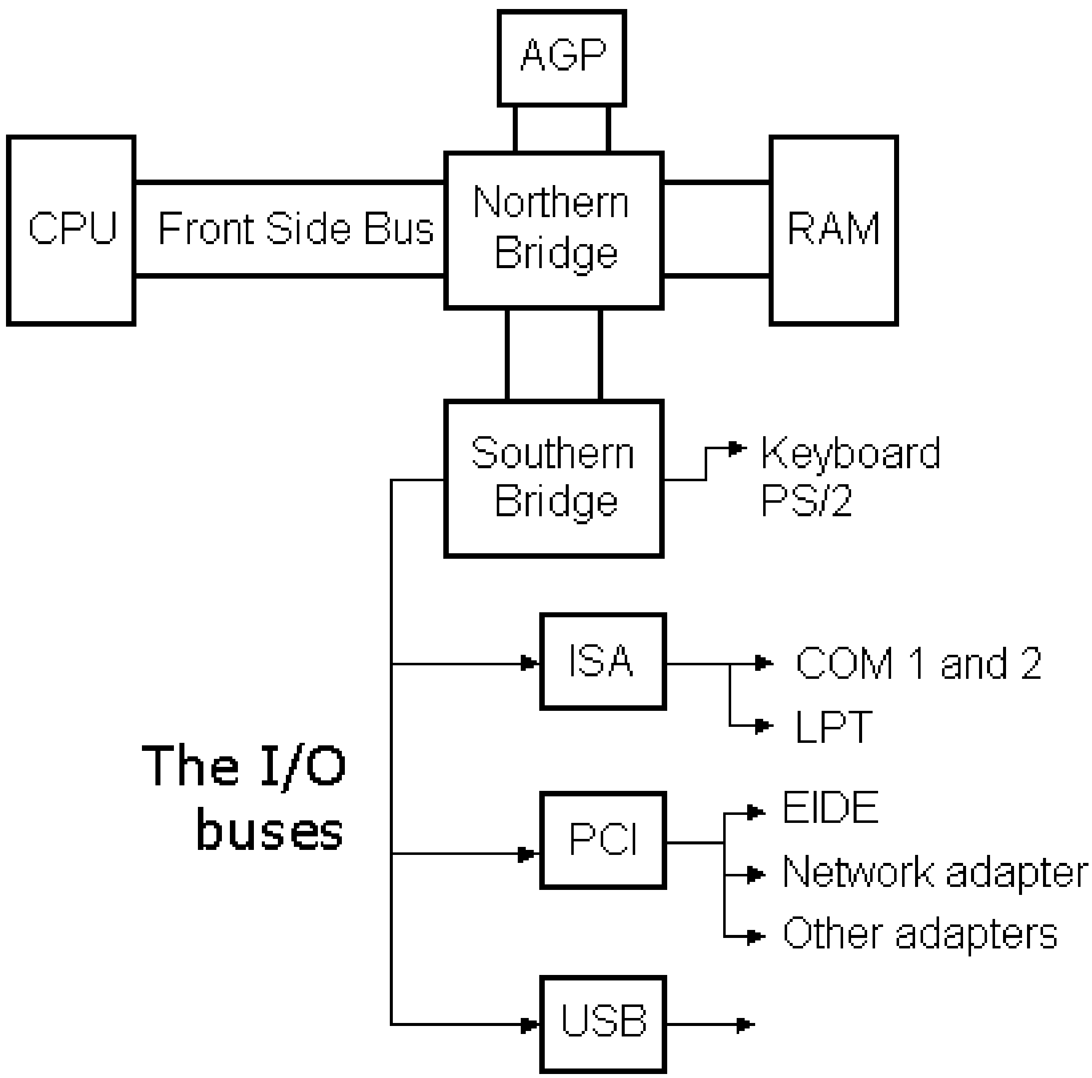
[top]

I/O buses connect the CPU to all other components, except RAM. Data are moved on the buses from one component to another, and data from other components to the CPU and RAM. The I/O buses differ from the system bus in speed. Their speed will always be lower than the system bus speed. Over the years, different I/O buses have been developed. On modern PCs, you will usually find four buses:

- The ISA bus, which is an old low speed bus, soon to be excluded from the PC design.
- The PCI bus, which is a new high speed bus.
- The USB bus (*Universal Serial Bus)*, which is a new low speed bus.
- The AGP bus which solely is used for the graphics card.

As mentioned earlier, I/O buses are really extensions to the system bus. On the motherboard, the system bus ends in a controller chip, which forms a bridge to the I/O buses.

All in all, the buses have had a very central placement in the PC's data exchange. Actually, all components except the CPU communicate with each other and with RAM via the different I/O buses. Here you see a demonstration of this logic:



The physical aspects of the I/O buses

[top]

Physically, the I/O bus consists of tracks on the printed circuit board. These tracks are used as:

- Data tracks, which each can move one bit at a time
- Address tracks, which identify where data should be sent to
- Other tracks for clock ticks, voltage, verification signals, etc.

When data are sent on the bus, they must be supplied with a *receiver*. Therefore, each device on the bus has an address. Similarly, the RAM is divided in sections, each having its address. Prior to sending data, a number is sent on the address track, to identify where the data should be sent to.

The bus width

The number of data tracks determine the data transfer capacity. The ISA bus is slow, partly because it only has 16 data tracks. The modern PCs send 32 bits per clock tick. On the ISA bus, 32 bits must be divided in two packages of 16 bits. This delays the data transfer. Another I/O bus concept is *wait states.*

Wait states

Wait states are small pauses. If an ISA adapter cannot keep up with the incoming data flow, its controller sends wait states to the CPU. Those are signals to the CPU to "hold on for a sec." A wait state is a wasted clock tick. The CPU skips a clock tick, when not occupied. Thus the old and slow ISA adapter can significantly reduce the operating speed of a modern computer.

Another aspect is the IRQ signals, which the components use to attract attention from the CPU. That and the concepts *DMA* and *bus mastering,* are described in module 5, which deals with adapters.

Technical and historical background for the I/O buses

[top]

In modern PCs you only find the PCI and ISA buses (besides USB, which we do not know much about yet). But, over the years, there have been other buses. Here is a diagram of the various I/O buses. Then comes a more detailed description of each of the buses:

Bus	Year	Bus width	Bus speed	Max. throughput (theoretical)
PC and XT	1980-82	8 bit	Synchronous with CPU: 4.77 - 6 MHz	4-6 MBps
ISA (AT) Simple bus.	1984	16 bit	Synchronous: 8-10 MHz	8 MBps
MCA. Advanced, intelligent bus by IBM.	1987	32 bit	Asynchronous: 10.33 MHz	40 MBps
EISA. Bus for servers.	1988	32 bit	Synchronous: max. 8 MHz	32 MBps
VL. High speed bus, used in 486s.	1993	32 bit	Synchronous: 33-50 MHz	100-160 MBps
PCI. Intelligent, advanced high speed bus.	1993	32 bit	Asynchronous: 33 MHz	132 MBps
USB. Modern, simple, and intelligent bus.	1996	Serial		1.2 MBps
FireWire (IEEE1394). High-speed I/O bus for storage, video etc.	1999	Serial		80 MBps
USB 2.0	2001	Serial		12-40 MBps

SCSI is another type of bus.

- Next page
- Previous page

Learn more **[top]**

5c about the modern I/O bus called USB.

Read module 5b about AGP and module 5c about Firewire.

Read more about chip sets on the motherboard in module 2d

Read about RAM in module 2e

Read Module 4b about hard disks.

Read Module 4c about optical media (CDROM and DVD).

Read Module 4d about super diskette and MO drives.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2c.2

About the ISA bus and other old PC buses

The contents:

- Introduction to the ISA bus
- MCA, Eisa and VLB buses

- Next page
- Previous page

Introduction to the ISA bus

Since about 1984, standard bus for PC I/O functions has been named ISA (*Industry Standard Architecture*). It is still used in all PCs to maintain backwards compatibility. In that way modern PCs can accept expansion cards of the old ISA type.

ISA was an improvement over the original IBM XT bus, which was only 8 bit wide. IBM's trademark is *AT bus*. Usually, it is just referred to as ISA bus.

ISA is 16 bit wide and runs at a maximum of 8 MHz. However, it requires 2-3 clock ticks to move 16 bits of data. The ISA bus works synchronous with the CPU. If the system bus is faster than 10 MHz, many expansion boards become flaky and the ISA clock frequency is reduced to a fraction of the system bus clock frequency.

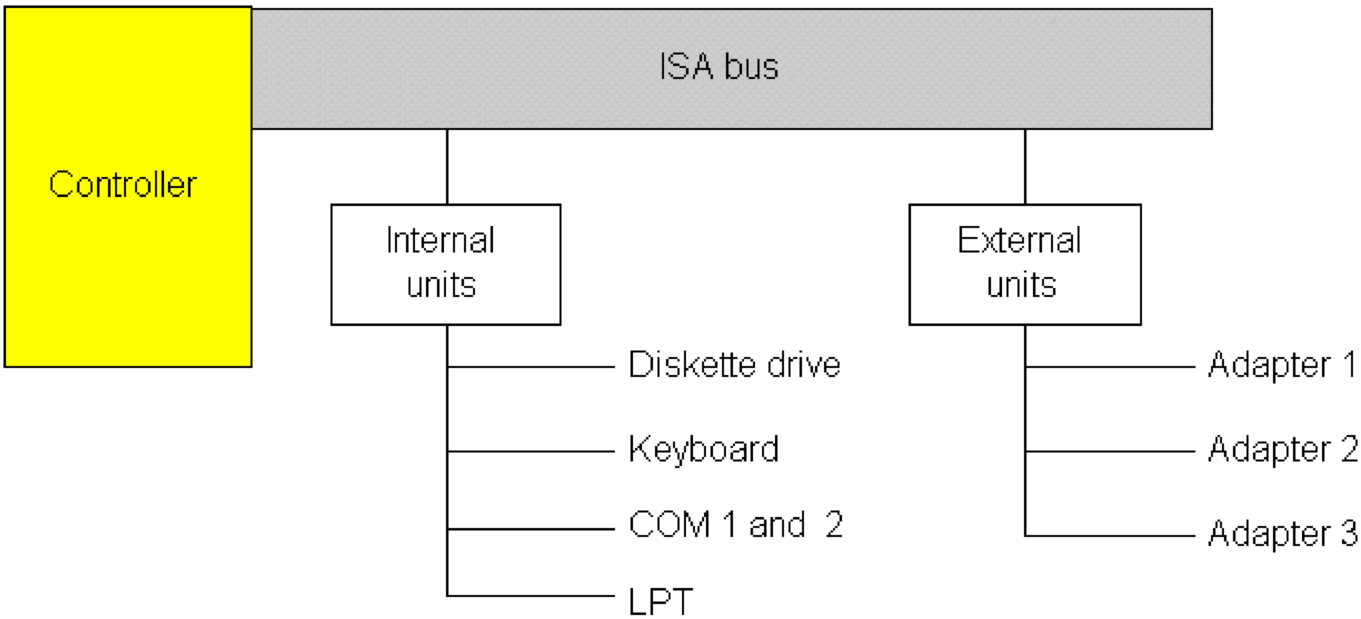
The ISA bus has an theoretical transmission capacity of about 8 MBps. However, the actual speed does not exceed 1-2 MBps, and it soon became too slow.

Two faces

The ISA bus has two "faces" in the modern PC:

- The internal ISA bus, which is used on the simple ports, like keyboard, diskette drive, serial and parallel ports.
- As external expansion bus, which can be connected with 16 bit ISA adapters.

ISA slots are today mostly used for the common 16 bit SoundBlaster compatible sound cards.



Problems

The problem with the ISA bus is twofold:

- It is narrow and slow.
- It has no intelligence.

The ISA bus cannot transfer enough bits at a time. It has a very limited bandwidth. Let us compare the bandwidths of ISA bus and the newer PCI bus:

Bus	Transmission time	Data volume per transmission
ISA	375 ns	16 bit
PCI	30 ns	32 bit

Clearly, there is a vast difference between the capacity of the two buses. The ISA bus uses a lot of time for every data transfer, and it only moves 16 bits in one operation.

The other problem with the ISA bus is the lack of intelligence. This means that the CPU has to control the data transfer across the bus. The CPU cannot start a new assignment, until the transfer is completed. You can observe that, when your PC communicates with the floppy drive, while the rest of the PC is waiting. Quite often the whole PC seems to be sleeping. That is the result of a slow and unintelligent ISA bus.

Problems with IRQs

The ISA bus can be a tease, when you install new expansion cards (for example a sound card). Many of these problems derive from the tuning of IRQ and DMA, which must be done manually on the old ISA bus.

Every component occupies a specific IRQ and possibly a DMA channel. That can create conflict with existing components. Read module 5 about expansion cards and these problems.

The ISA bus is out

As described, the ISA bus is quite outdated and should not be used in modern pcs. There is a good chance, that this "outdated legacy technology" (quoting Intel) will disappear completely.

The USB bus is the technology that will replace it. It has taken many years to get this working and accepted, but it works now.

Intel's chip set 810 was the first not to include ISA support.

MCA, EISA and VLB

[top]

In the 80s, a demand developed for buses more powerful than the ISA. IBM developed the MCA bus and Compaq and others responded with the EISA bus. None of those were particularly fast, and they never became particularly successful outside the server market.

MCA

IBM's top of the line bus from 1987 is named *Micro Channel Architecture*. The MCA bus was a masterpiece, unifying the best bus technology from the mainframe design with the demands from the PC. However, contrary to the ISA bus, MCA is patented, and IBM demanded high royalty fees, when other PC manufacturers wanted to use it. Thus the bus never became a great success, despite its advanced design. It ended up being a classic example of poor marketing strategy.

The MCA bus is 32 bit wide and "intelligent." The cards configure themselves with respect to IRQ. Thus, they can be installed without adjustments of jumper switches or other features. It works constantly at 10.33 MHz, asynchronous with the system bus.

The MCA bus is also relatively fast with transfer rates of up to 40 MBps in 32 bit mode at 10.33 MHz. MCA requires special adapters. There have never been too many adapters developed, since this bus is by and large used only in IBM's own PCs.

EISA

EISA is a bus from 1988-89. It is designed by the "Gang of Nine:" the companies AST, Compaq, Epson, Hewlett-Packard, NEC, Olivetti, Tandy, Wyse and Zenith. It came in response to IBM's patented MCA bus.

EISA is built on the ISA bus; the connector has the same dimensions and old ISA cards fit into the slots. To keep this compatibility, the EISA bus works at maximum 8 MHz. Like ISA, the bus bus is synchronous with the CPU at a clock frequency reduced to a fraction of the system bus clock frequency.

EISA is compatible with ISA in the sense that ISA adapters can be installed in EISA slots. The EISA adapters hold a second level of connectors in the button of the slot.

However, EISA is much more intelligent than ISA. It has bus mastering, divided interrupts and self configuration. It is 32 bit wide, and with it's *compressed transfers* and *BURST mode*gives a highly improved performance.

But, like the MCA, it did not have great success. The EISA bus is still used in some servers.

Vesa Local Bus

This Bus called VLB for short. It is an inexpensive and simple technology. This bus only achieved status as an interim phenomenon (in 1993-94). VLB was widely used on 486 motherboards, where the system bus runs at 33 MHz. VLB runs directly with the system bus. Therefore, data transfer is at CPU speed, synchronous and in width. The problem with VLB was compatibility. Adapters and system system boards would not always work together. Vesa is an organization with about 120 members, mostly monitor and graphics card manufacturers. Therefore, most VLB cards were video cards.

- Next page
- Previous page

Learn more **[top]**

Read module 5c about the modern I/O bus called USB.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire

Read more about chip sets on the motherboard in module 2d.

Read more about RAM in module 2e.

Read Module 4b about hard disks.

Read Module 4c about optical media (CDROM and DVD).

Read Module 4d about super diskette and MO drives.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2c.3

About the PCI bus

The contents:

- Introducing the PCI bus
- The internal and external face.
- The future design
- NGIO

- Next page
- Previous page

Introducing the PCI bus

[top]

The PCI is the high speed bus of the 1990s. PCI stands for *Peripheral Component Interconnect*. This bus is made by Intel. It is used today in all PCs and other computers for connecting adapters, such as network-controllers, graphics cards, sound cards etc.

Some graphics cards however use the AGP-bus, which is a separate bus only intended for graphics.

The PCI bus is the central I/O bus, which you find in all PCs!

A 32 bit bus

The PCI is actually 32 bit wide, but in practice it functions like a 64 bit bus. Running at 33 MHz, it has a maximum transmission capacity of 132 MBps.

According to the specifications - not in practice, it can have up to 8 units with a speed up to 200 MHz. The bus is processor independent. Therefore, it can be used with all 32 or 64 bit processors, and it is also found in other computers than PCs.

The PCI bus is compatible with the ISA bus in that it can react on ISA bus signals, create the same IRQs, etc.

Buffering and PnP

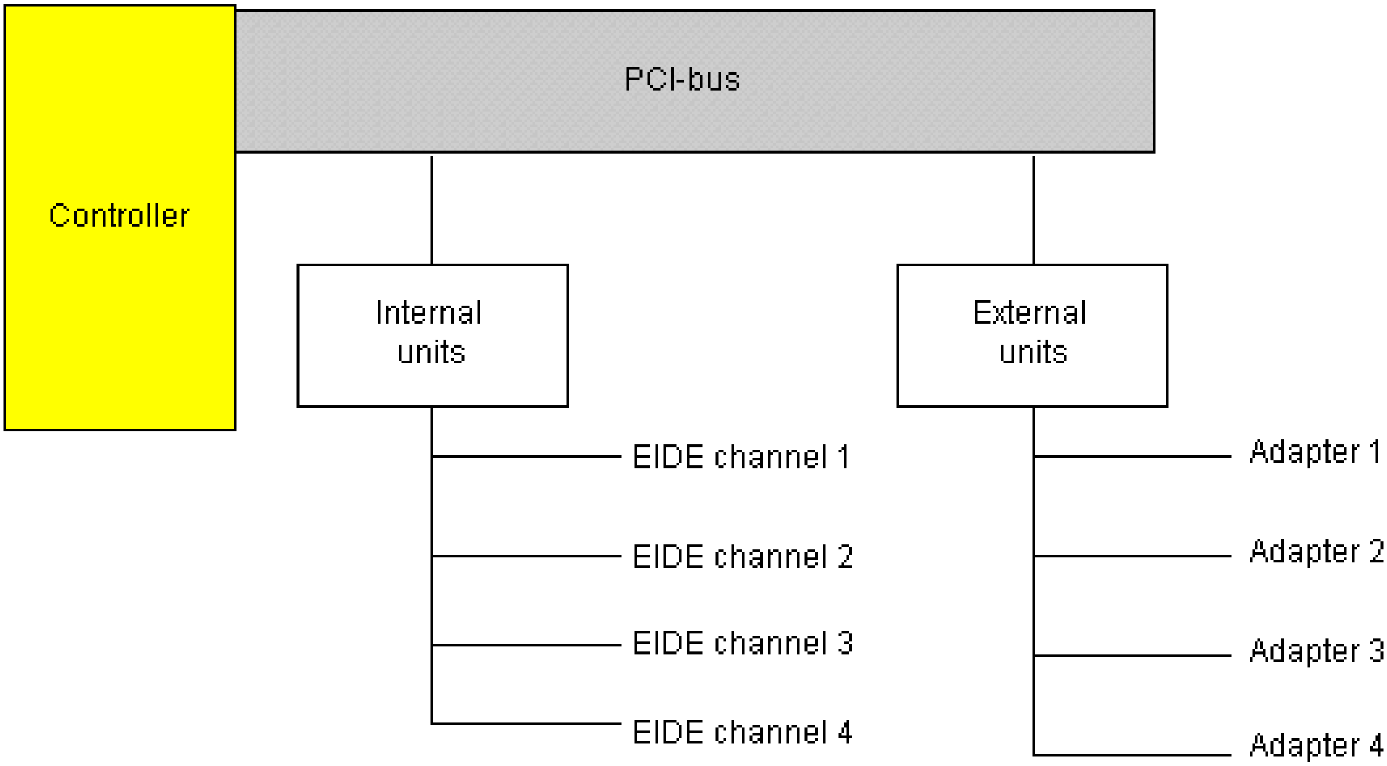
The PCI bus is *buffered* in relation to the CPU and the peripheral components. This means, that the CPU can deliver its data to the buffer, and then proceed with other tasks. The bus handles the further transmission in its own tempo. Conversely, the PCI adapters can also transmit data to the buffer, regardless of whether the CPU is free to process them. They are placed in a queue, until the system bus can forward them to the CPU. Under optimal conditions, the PCI bus transmits 32 bits per clock tick. Sometimes, it requires two clock ticks.

Because of this, the peripheral PCI units operate *asynchronous* . Therefore, the PCI (contrary to the VL bus) is not a local bus in a strict sense. Finally, the PCI bus is intelligent relative to the peripheral components, in that Plug and Play is included in the PCI specifications. All adapter cards for the PCI configure themselves. Plug and Play is abbreviated PnP.

PCI with two faces

On modern system boards, the PCI bus (like ISA) has two "faces:"

- Internal PCI bus, which connects to EIDE channels on the motherboard.
- The PCI expansion bus, which typically has 3-4 slots for PCI adapters.



The PCI bus is continuously being developed further. There is a PCI Special Interest Group, consisting of the most significant companies (Intel, IBM, Apple, and others), which coordinate and standardize the development.

Soon we shall see PCI with a higher bus speed (66 MHz) and greater width (64 bit). However alternative buses are also marketed. An example is the high speed AGP video bus (*Accelerated Graphics Port*) and the FireWire Bus. AGP is fundamentally a 66 MHz PCI bus (version 2.1) which has been enhanced with other technologies making it suitable for the graphics system.

PCI-X

Another new initiative is the so-called PCI-X (also called "Project One" and Future I/O). Companies like IBM, Mylex, 3COM, Adaptec, HP and Compaq want to launch a special high speed server version of the PCI bus. This new bus (also mentioned as PCIX) allows a bandwidth of up to 1 GB per second (with a 64 bit bus running at 133 MHz). Intel is not cooperating on this project, and neither is Dell. It is going to be interesting to follow.

Intel's NGIO (Next-Generation I/O)

NGIO server architecture is another initiative by the companies Dell Computer, Hitachi, NEC, Siemens, Sun Microsystems and Intel to produce a new architecture for I/O on servers. This is clearly an answer to the Project One mentioned above.

FIO to merge with NGIO

On August 31, 1999 seven of the leading companies (Compaq, Dell, Hewlett-Packard Company, IBM, Intel, Microsoft, Sun) announced the intent to merge the best ideas of the Future I/O

(FIO) and Next Generation I/O (NGIO). The new open *input/output architecture* will find use in servers. The bandwidth will be up to 6 GByte/sec.

The new standard NGIO will hardly go into production before 2001.

- Next page
- Previous page

Learn more [top]

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read Module 4b about hard disks.

Read Module 4c about optical media (CDROM and DVD).

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.01

On chip sets

Module 2d describes what chip sets are, and how they function on the motherboards. This module is subdivided into the following pages:

1: What is a chip set?
2: The first chip sets for the Pentium boards
3: Non-Intel chip sets (mostly for Super 7 boards)
4: Chip sets for Intel P6 processors
5: More chip sets for Intel P6 processors
6: Intel's i810 "Whitney"
7: Intel's i820 "Camino"
8: Intel's i815 "Solano"

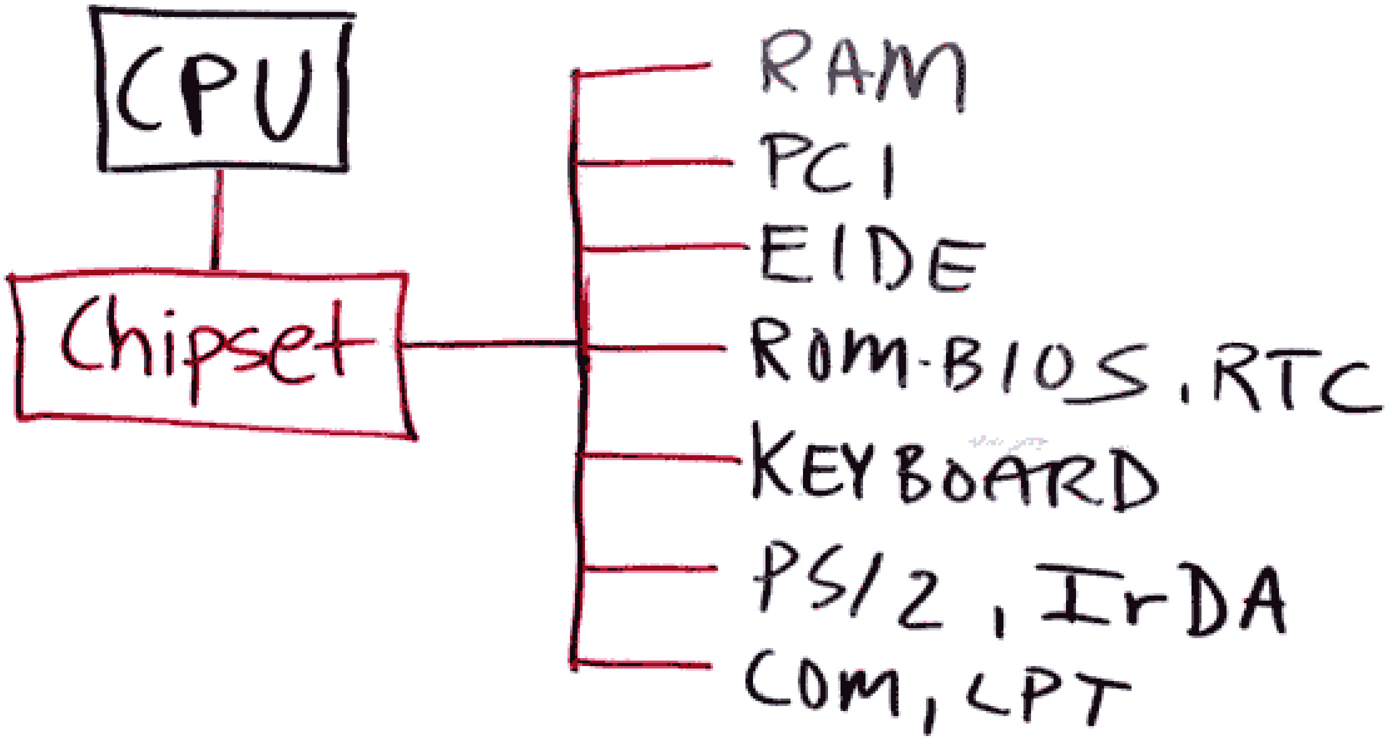
- Next page
- Previous page

I recommend that you read all the pages one by one. Just follow the links "Next page" to get through the textbook. I hope you find the information useful!

What is a chip set?

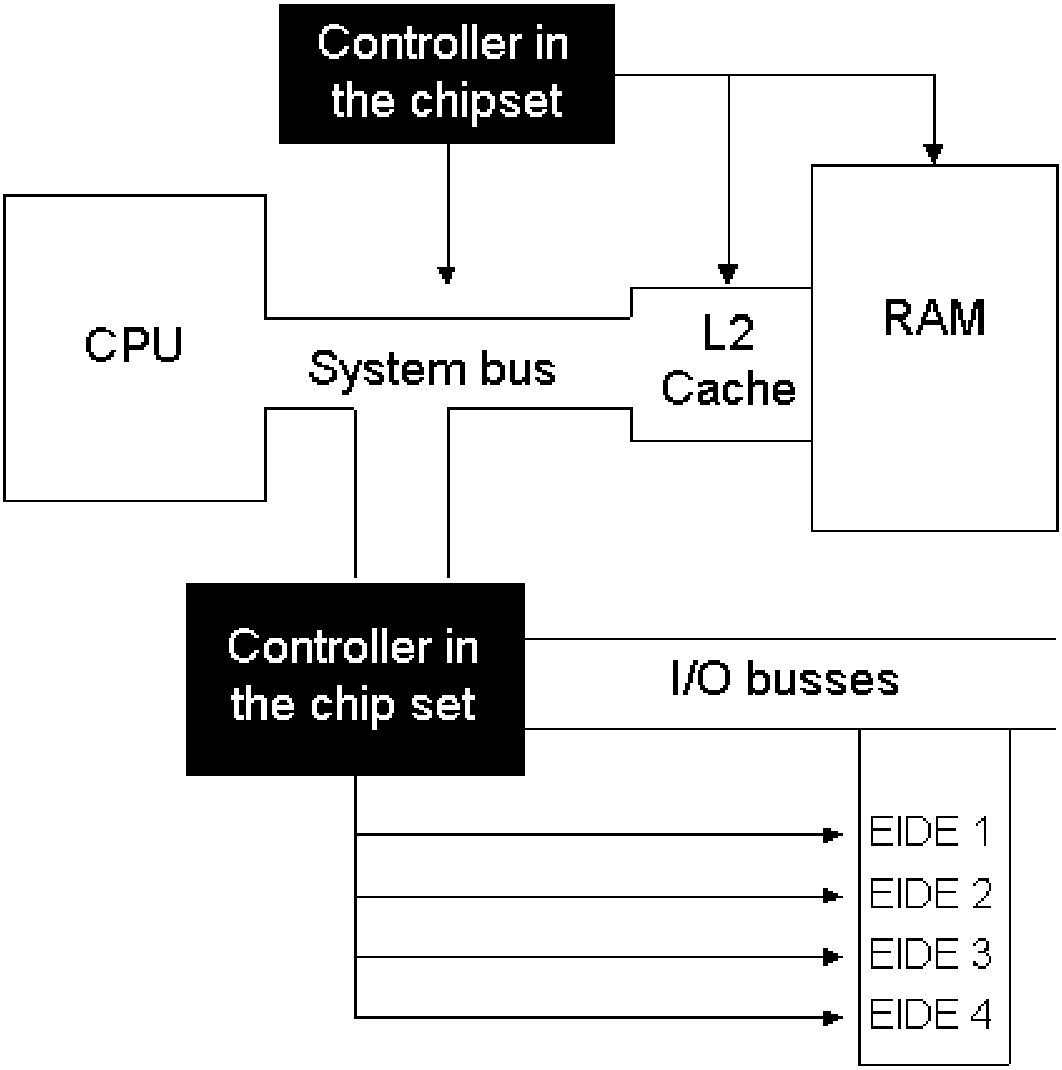
[top]

The chip set is very important to the modern PC and its performance. Many technologies meet on the motherboard and are "glued" together via these controllers, which we call the "chip set".



When we speak about buses and motherboards, we are also speaking about *chip sets*. The chip sets are a bunch of intelligent controller chips, which are on any motherboard.

The controllers are closely tied to the CPU, in that they control the buses around the CPU. Without the chip sets, neither RAM nor I/O buses could function together with the CPU:


Controller in the chipset
CPU
System bus
L2 Cache
RAM
Controller in the chip set
I/O busses
EIDE 1
EIDE 2
EIDE 3
EIDE 4

New technologies - new chip set

[top]

Therefore, the chip sets are quite central components on the motherboards. When new technological features are introduced (and this happens continuously) they are often accompanied by new chip sets. The new chip sets often enable:

- Higher speed on one or more buses
- Utilization of new facilities (new RAM types, new buses, improved EIDE, etc.)

The vendors

There are several suppliers of chip sets for the motherboard:

- Intel
- SIS
- Opti
- Via
- ALi

Intel has hitherto been the leader in supplying chip sets to the Pentium motherboard. Therefore, let us just mention their chip sets, which have astronomical names.

The Neptune chip set (82434NX) was introduced in June 1994. It replaced the Mercury set (82434LX). In both chip sets, there were problems with the PCI bus. In January 1995 Intel introduced the first Triton, where everything worked. This chip set supports some new features: it supports EDO RAM, and it offers bus master integrated EIDE control and NSP (*Native Signal Processing* - one of the many new creations, which was soon forgotten).

However, the following chip sets were of much higher quality, and within very few years they lead to several new generations of chip sets, each of them more powerful and offering great new features.

- Next page
- Previous page

Learn more **[top]**

If you want to read more about these and other chip sets, look for the excellent web site Toms Hardware Guide. Here, you will find all about these subjects.

Read more about RAM in module 2e

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.02

The first chip sets for the Pentium boards

The contents:

- Triton first and second
- A bridge to the I/O system
- USB and EIDE
- The differences between HX and VX
- The TX chip set

- Next page
- Previous page

Triton first and second

[top]

The interest in chip sets and their performance started in late 1995, when the Pentium processor became more popular. The Triton controllers were the first chip sets in this trend.

82430FX from late 1995 was Intel's next chip set and the first Triton. In February 1996 the second generation of Triton arrived. Two new chip sets were introduced: The 82430VX and 82430HX. The last (HX) was the fastest one.

VX and HX

The two sets were similar, yet different. 430HX consisted of two chips. It was designed for the more professional PCs. 430VX consisted of four chips, but the cost was slightly lower than HX. It was aimed at the home use PC market. Let us look at the contents of each chip set:

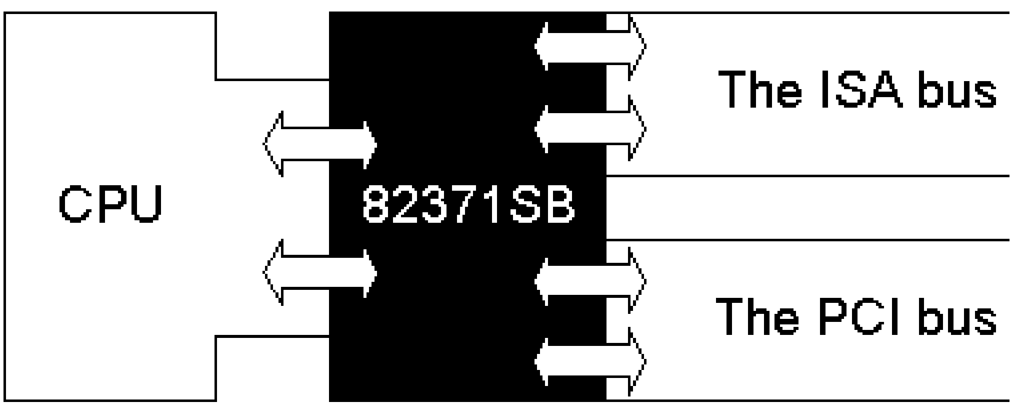
Chip set	Contents
82430HX	82439HX System Controller (TXC) + 82371SB PCI ISA IDE Accelerator
82430VX	82437VX System Controller (TVX) + two 82438VX Data Path Units (TDX) + 82371SB PCI ISA IDE Accelerator

Common to both chip sets is 82371SB, which is a "PCI ISA IDE accelerator chip". It is also called PIIX3, which some may recognize from the Windows 95 device driver, which comes with the ASUS T2P4 board.

A bridge to the I/O system

[top]

The chip makes a bridge between the CPU, ISA and PCI bus. The news was, that it permitted concurrent activity in all three locations, thus a new form of multitasking. This is significant for daily use. All data exchange to and from I/O units cross this intersection, which now has achieved greater width:

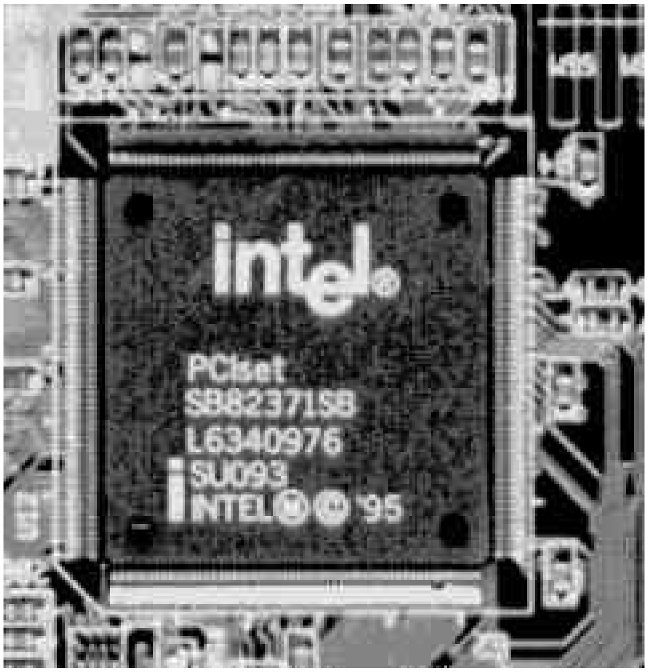


.

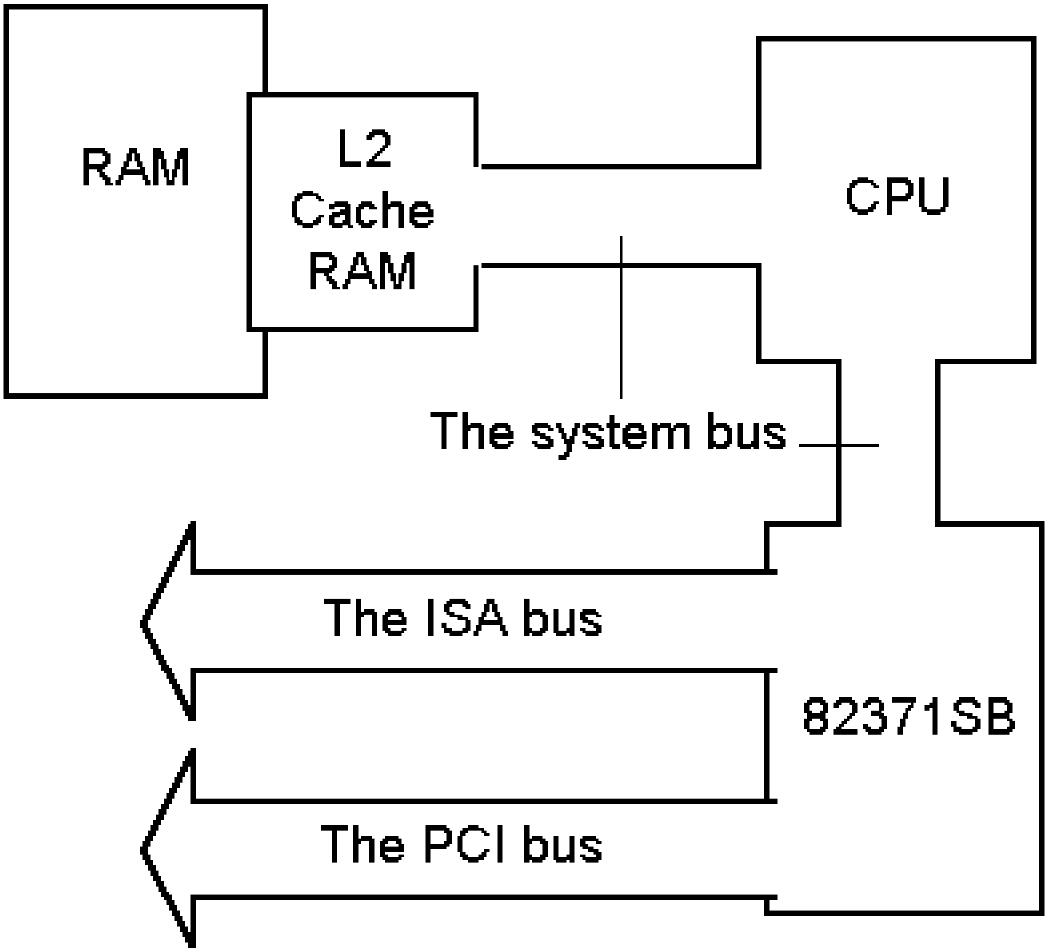
US2 and EIDE

[top]

New in the chip was also the host function for USB. It is the *Universal Serial Bus*, which was not much use at this time. Finally, the chip included a EIDE Bus Master control. In short that means, that EIDE components like hard disks, to some extent can deliver their data directly to RAM without taking up CPU time.



Above, you see the 82371SB chip and below, again, its placement relative to CPU and buses:



The differences between HX and VX

[top]

It was generally accepted, that the HX set yielded the best performance of the two chip sets described. But the VX set had two other facilities to offer: Capability for SMBA (*Shared Memory Buffer Architecture*). That means among other things, that you can integrate the video card on the motherboard with 1 or 2 MB standard RAM, from the working RAM. A technology, which is used only in the lowest cost PCs, and which soon was abandoned.

Also, the VX set also supported the fast RAM type SD-RAM. HX did not. The VX set could control up to 128 MB RAM, but it could not cache above 64 MB RAM.

HX controlled 512 MB RAM and was the only Intel Pentium chip set to cache above 64 MB RAM.

The VX and HX chip sets are both out. They were replaced by the TX chip set, which was the last Intel chip set for Socket 7 mounted CPUs. Today Ali and VIA produces chip sets for Socket 7 motherboards.

Intel TX chip set

[top]

The last chip set to Pentium processors were named 82430TX, which supports two new hot technologies:

- SD-RAM
- Ultra DMA

Ultra DMA was also called ATA/33, and it is a standard for harddisk interface, which permits EIDE hard disks to transfer at up to 33 MBps.

This improved EIDE standard is mostly marketed under the name Ultra DMA. Tests show that Ultra DMA results in a speed increase of 25-75 percent over the traditional EIDE PIO mode 4. Ultra DMA is the new EIDE standard and has been vastly enhanced since this chip set.

The controllers in the TX chip set

Chip set	Chips included
82430TX	82439TX System Controller (TXC) 82371AB PCI ISA IDE Accelerator

The TX set is an update and improvement of the VX set. Relative to this, the TX firstly supports SD RAM and Ultra DMA hard disks. Two important technologies. But the TX-set cannot cache above 64 MB RAM, and that was a problem. Please see this article on this subject.

Photos taken with Canon Powershot 600

The TX chip set was Intel's last and final set for Socket 7 motherboards. After that VIA and ALi took over and continued this work.

- Next page
- Previous page

Learn more **[top]**

Read more about RAM in module 2e

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5c2a.

About USB

The contents:

- [Next page]
- [Previous page]

- What is USB?
- After a slow start ..

On the following page:

- Next page: The USB hub
- Next module: IEEE1394 FireWire
- Device Bay

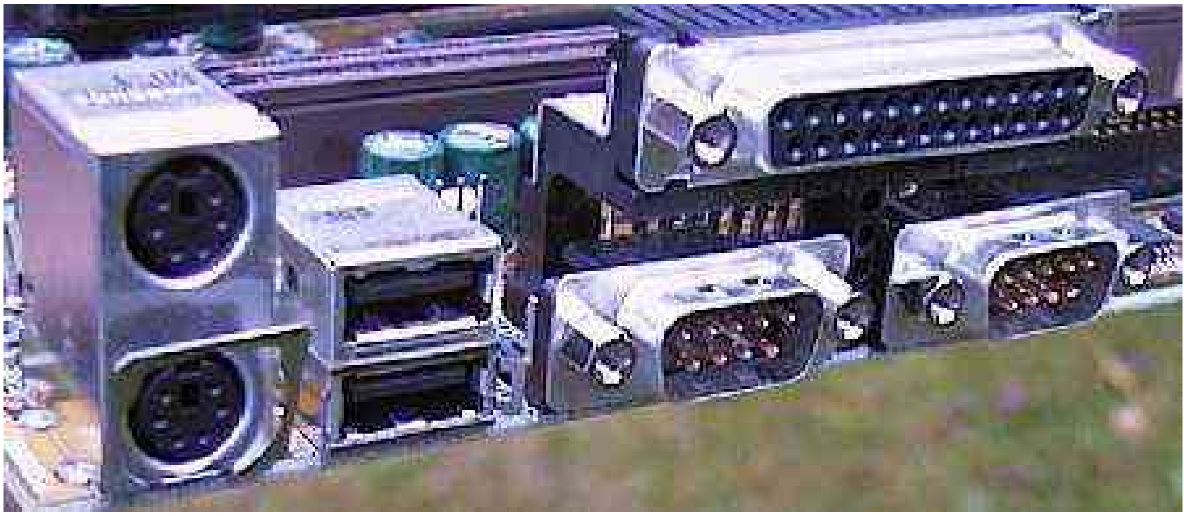
What is USB?

[top]

USB stands for *Universal Serial Bus*. It is a cheap and rather slow I/O bus, running at 12 Mbit/sec.

It can be compared to the FireWire bus, which however is a lot speedier.

USB is an open and royalty-free specification. Units can be plugged and unplugged on the fly very easily. Here you see the plugs, the two small ones, number two from the left:



There were problems with USB in the beginning, since many motherboard manufacturers produced their own versions of the port before it was fully standardized. Hence the nickname *Useless Serial Bus* .

USB is supported by Windows 95 OSR2.1, Windows 98/Me, Windows 2000 and Windows XP.

A success

USB has become a great succes. The bus simplifies PC design - giving us a simple and unified interface for a whole lot of PC units and devices like:

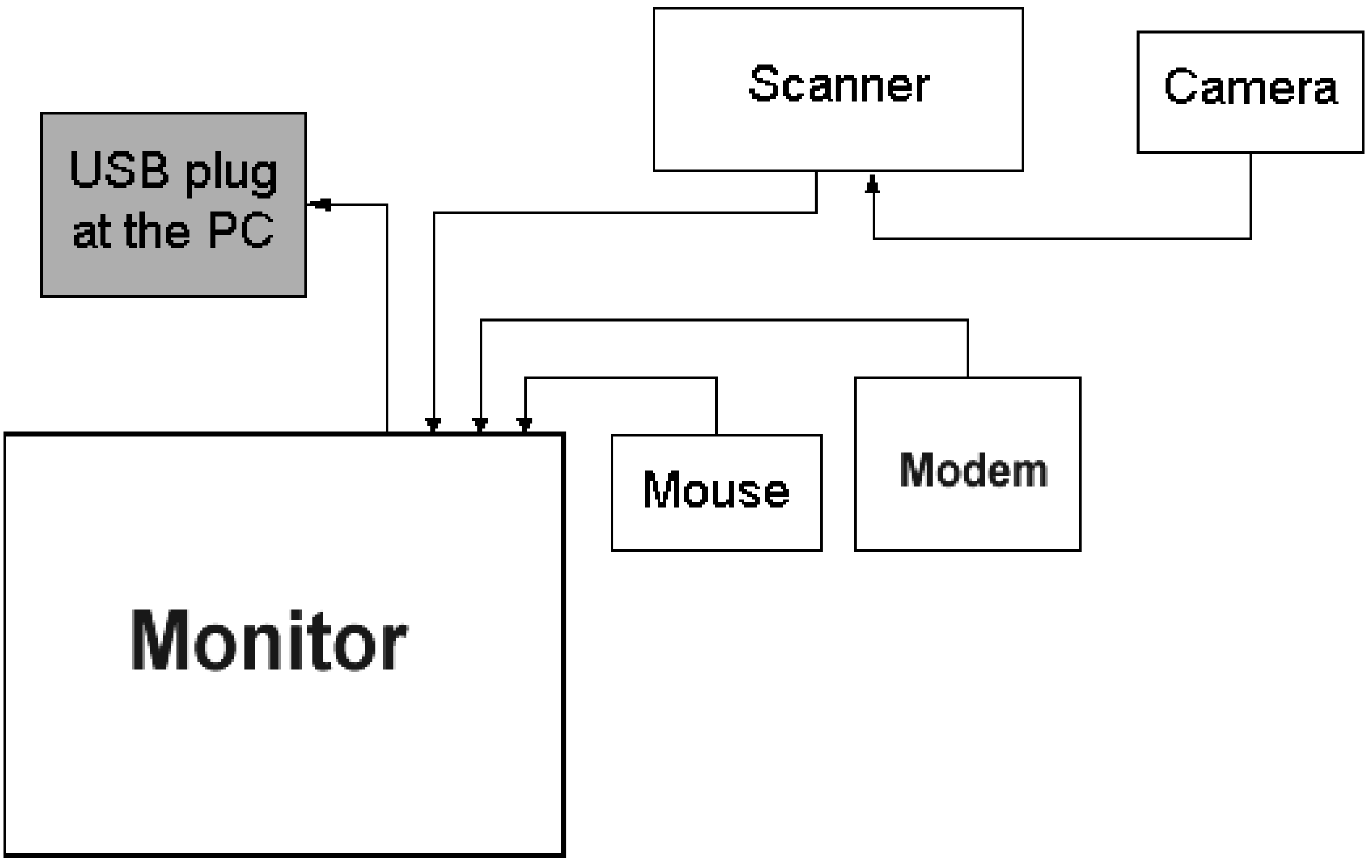
- Keyboard
- Mouse
- Loudspeakers, microphones, and other sound devices
- Printers
- Modems and ISDN adapters
- Scanners and cameras
- External drives like CD-RWs
- Card-readers and other adapters

All these units - and lots of others - will be connected using one single plug at the PC. USB holds up to 127 units in one long chain.

The keyboard may hold a hub, so other USB units are connected here (although it more often is the monitor to include a hub, as we shall see later):

Each unit may hold two USB connectors, so they all can be daisy chained.

This illustration is fiction - I never saw a setup like this, but it shows the intentions of the serial USB interface:



All units have a firmware identification code, that communicates with the OS (i.e. Windows). The unit must have a power feed (could be minimum 100 ma) to be recognized by the USB controller and Windows 98. If one unit fails this way, Windows shows an ! on yellow background to signalize that something has to be done. This could be to unplug other USB devices to increase the available power in the chain.

Many hardware manufacturers today produce their modems, cameras and scanners in versions with two-way interfaces. These devices connect either traditionally using a COM port or using the USB port.

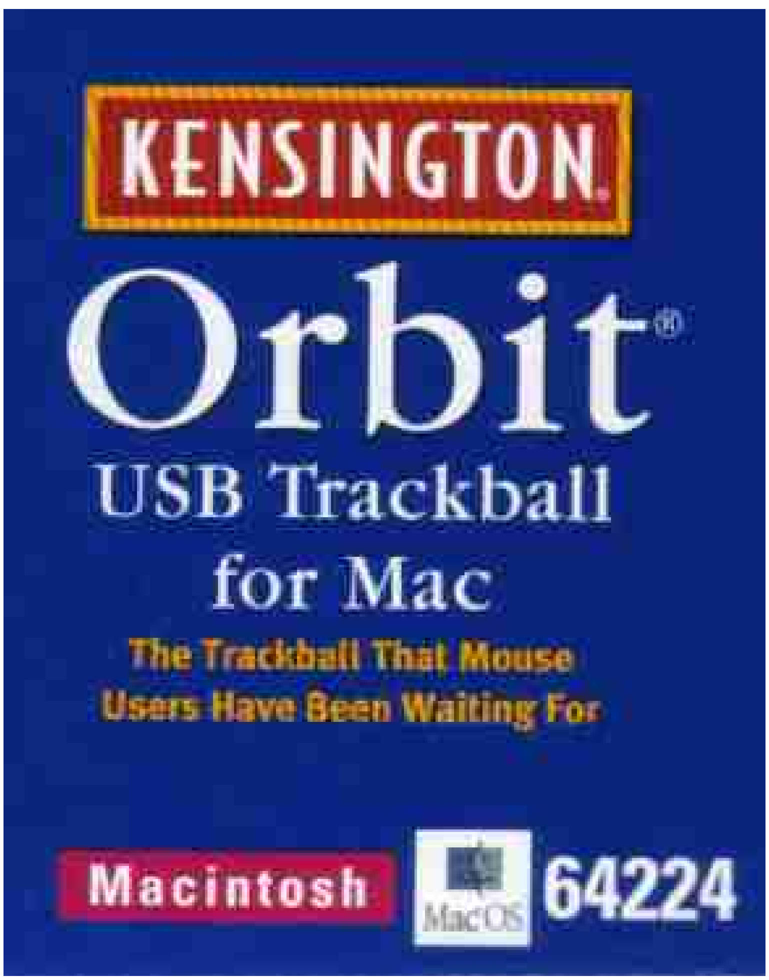
After a slow start ...

[top]

Personally I always believed, that USB a´´had to become a great renovation of the PC design. However, things were moving very slowly in the beginning.

Bigger companies like Swiss Logitech (producing the best mice and trackballs available, at least to my opinion) moved very slowly into USB. This probably has been due to serious concern over the correct technical implementations. The COM, PS/2, and LPT ports represent very well-known technology. Replacing them you have to be very certain of the consequences.

In 1999 and 2000 the USB products became available in large numbers. Many of them are being sold both to Mac and PC. My latest trackball, a Kensington Orbit is only a Mac-product, judging from the box:



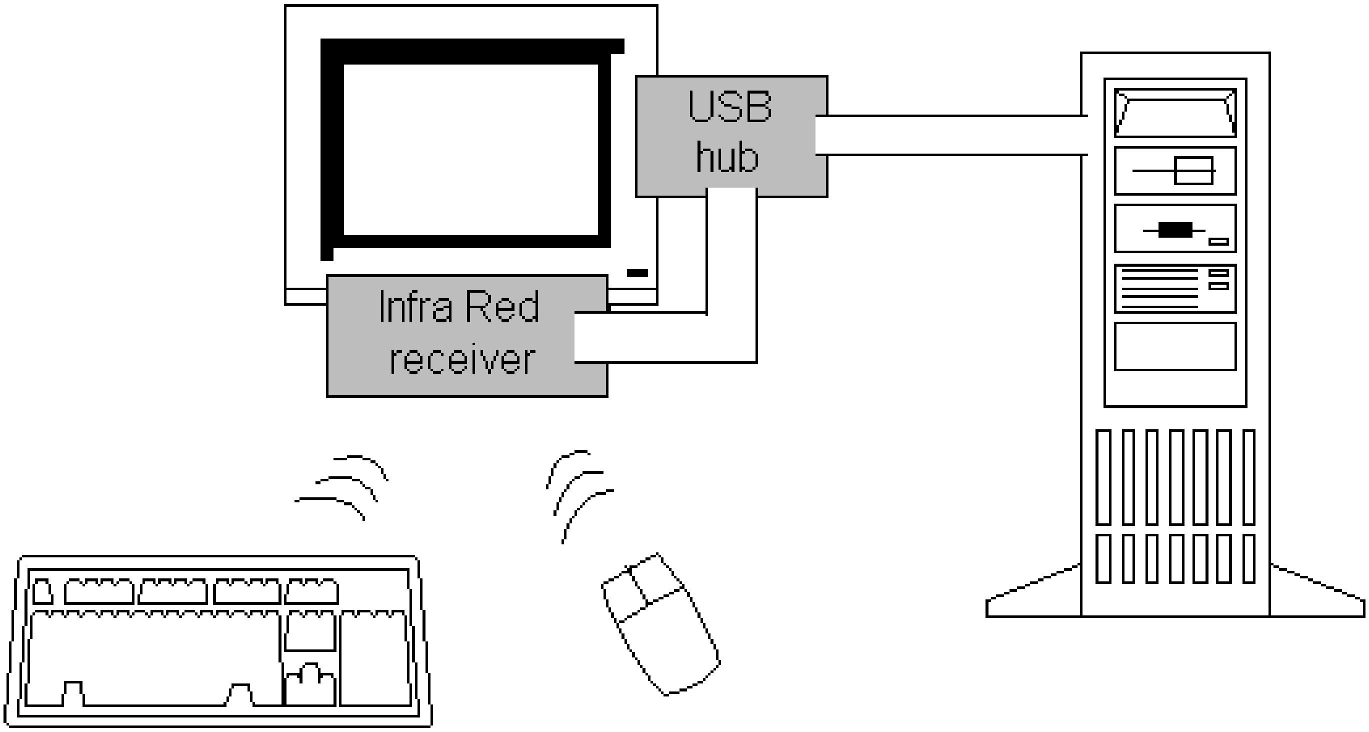
However, the trackball (which is very fine) works fine on any PC with USB. The Windows USB driver instantly recognizes the trackball.

The same goes for my great little tablet (Wacom Graphire):

Philips and Logitech - a private vision

[top]

If I were in charge, Philips should go further with USB. Already they build in a USB in their monitors. Why not bundle the monitor with a cordless set of keyboard & mouse. And place the infra-red receiver in the monitor using USB as interface? May I give this idea to Philips:



Philips even could buy Logitech as well. I think they would fit well together - two fine European vendors.

Links

You find technical specifications etc. in these sites:

Intel's USB site http://www.intel.com/design/usb

USB site: http://www.usb.org

- Next page
- Previous page

Learn more **[top]**

Next module (on FireWire and Device Bay)

Read Module 6a about file systems

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card

Read module 7c about sound cards, and 7d on digital sound and music

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5c3.

About FireWire IEEE1394

- IEEE1394 FireWire
- Device Bay

- Next page
- Previous page

FireWire is another interface connecting the PC to external units. It does not look very much like the SCSI we know, but is a further development being a serial high speed bus. It also a bit like USB in terms of hot-plugging and simple connections.

The interface IEEE1394 has a bandwidth of 400 Mbit per second, which is a lot better than USB and comparable to SCSI.

FireWire handles up to 63 units on the same bus. The units can be plugged and unplugged *hot* - meaning you do not have to power down the PC.

The Firewire was expected to replace:

- Parallel Centronics port (to some extent)
- IDE
- SCSI
- EIDE (later on)

However, Firewire so far has not become the real big thing for PCs. People with Macintosh computers soon found great use for FireWire, especially to connect high performance flatbed scanners.

Today Firewire is supported by Windows XP and it is gaining momentum. Among others it is used for:

- Connecting DV-cameras to video editing adapters
- High-end scanners
- Hot-plugged external harddisks from Maxtor with FireWire interface.

Here is the IEEE1394 port of a digital video camera (marked "DV" for Digital Video). It is very small:

FireWire for Macintosh

[top]

FireWire was developed by Lucent Technologies, but has ended by Apple, who claims a $1 per port royalty for use. IEEE1394 was conceived by Apple, who proposed it to IEEE, who approved it as a standard in 1995. Firewire is Apple's implementation of the IEEE1394 standard. Other companies have implemented their own versions of IEEE1394. Yes, they are not all the same! (Note: This is the purpose of the IEEE - to approve industry standards, not to make an implementation.)

Hence all major hardware companies have adapted FireWire in some way. Especially the entertainment electronic industry (Video/games/television) have great hopes with FireWire/IEEE1394. It will connect all types of digital electronics with the PC and this way open up for a much more modular design.

Since IEEE1394 is advanced and yet claimed to be cheap & simple, the communications protocol can handle a lot of other units including:

- Network controllers
- Hard disks, CDROM drives
- Printers

Two modes

The FireWire standard operates with two *modes.*

- *Asynchronous* as other buses. This means that operations across the bus are controlled using interrupt signals. The bus reports to the host when a task is fulfilled.

- *Iso-synchronous* . In this mode data is being transferred at a steady preset speed - continuously and without any supervision from the host. This opens up for data streaming useful for video or the multimedia presentation. The FireWire is a peer-to-peer interface. This means that data can be transferred between two units attached to the bus without supervision from the PC.

FireWire has a 64 bit address bus. Compared to SCSI each unit does *not* need a unique ID, they are dynamically configured "on the fly". Neither does the bus have to be *terminated*. All together a lot simpler than SCSI.

One of the problems with SCSI has been the limitation on distance between the units. FireWire can hold up to 16 units in the same "string" and there can be up to 4.5 meters between two units.

The first implementations of FireWire will connect it to the PCI bus using the new PIX6-controller, which will be a part of one of Intel's new chip sets. I think it will be at least 2-3 years before we really see this new technology in the market. But it will be worth waiting for it, it opens up for a new world of inter connectivity between TV, PC, video and all other types of electronic gear.

USB and FireWire - serial buses of the future

[top]

Soon high-end PC probably will hold both these serial I/O buses:

- USB version 2.0 for all low speed gear.

- FireWire for high speed I/O to Digital Video recorders, high-end scanners etc.

Both interfaces are about to be included in the south bridge of the chipsets. In 2002 it was difficult to whether USB 2.0 or Firewire was going to be the new high-speed I/O standard. Probably both of them will have place in the PC architecture the coming 1-2 years.

Device Bay

[top]

DeviceBay is another standard which follows IEEE1394 and USB. These busses can connect and disconnect units "on the fly", that is while the PC is operating. This ability to "hot plug" requires a new physical connection between the units. DeviceBay may be the answer to this. That is a standard for connection boxes that can hold hard disks, CDROM drives and similar units.

The mounting frame can be installed without tools and while the PC is running. With this one can imagine a storage unit filled with MP3 -files, which easily can be moved from a PC to the player in the car. Unfortunately, for a while it is only a futuristic dream.

If DeviceBay really gains widespread usage, it could be the end of loose ribbon cables in the PC cabinet. The whole PC can be made in modules, which all plug into the USB or the FireWire bus as DeviceBay units. The units can then freely be moved between the different computers and other electronic units in the home.

It is designed for hot plugging units like:

- Zip drives
- Tape streamers
- Modems
- Hard disks
- PC-card readers

These units will fit into a special bay connected to the USB and FireWire buses enabling the hot plug. A good thought; however not all good thoughts end up in good hardware.

Links

You find technical specifications etc. in these sites:

Apple about: FireWire

The IEEE has its own page at http://www.ieee.org

About DeviceBay: www.device-bay.org

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card

Read module 7c about sound cards, and 7d on digital sound and music

---

KarbosGuide.com. Module 6

About file systems: DOS formatting, FAT, etc.

The contents on this page:

- What are file systems?
- The file system and the OS
- Limitations in disk size

- Next page
- Previous page

We have seen before that the PC is a big data processor. We have also seen that data are bits and bytes, which are organized in files. One of the operating system's major tasks is to write these data to a disk. Hard, floppy, and zip disks must be formatted before we can save files on them. In these pages, we will review formatting, file systems, etc. We will start with a general view, then go in depth about FAT formatting, which is (still) the most common.

What are file systems?

As I wrote in module 4a, drives are storage media, which can hold a file system. When a disk is formatted in a drive, it becomes organized and prepared to receive data. When we format a disk, it receives a *file system* .

Formatting can be compared to starting a library. You must install the book shelves and the catalogue system before any books are put in place. Once the library is ready, bring on the books! Similarly with a disk. When we format it, we "burn in" a file system to make it ready to receive data (files).

We can format with any one of several different file systems:

FAT

File Allocation Table is the original, old 16 bit DOS system is probably used in 90% of all PC's. It is also called FAT16 contrary to:

FAT32

This is a new addition to FAT, which Microsoft introduced with Windows 95 B – the December -96 version (OSR2). The performance has been even improved with Windows 98.

HPFS

High Performance File System is from OS/2. It is an advanced 32 bit file system, which in all respects is far superior to FAT, except for possible usage. It can only be used with OS/2.

NTFS from Windows NT

A 32 bit file system like HPFS, but *not* compatible with it. NTFS can only be used in Windows NT/2000/XP. If it was available for use in Windows 95/98, it may be preferable to FAT and FAT32.

NetWare

NetWare is a server operating system from Novell. It has its own 32 bit file system. For that reason, the Novell server, contrary to NT or OS/2 servers, cannot be used as a work station. The file system is much faster than FAT, but it works only with Novell servers (typically file servers).

ISO 9660

This is for CDROMs and ISO 13346 for DVDs.

UDF

Universal Disk Format is for big capacity disks like DVD RAM. UDF is not directly supported by older versions of Windows , you need a driver.

UNIX

UNIX servers have their own filing system. Here the use of upper/lower case in file naming is significant. Read in the following pages about the concepts of these file systems.

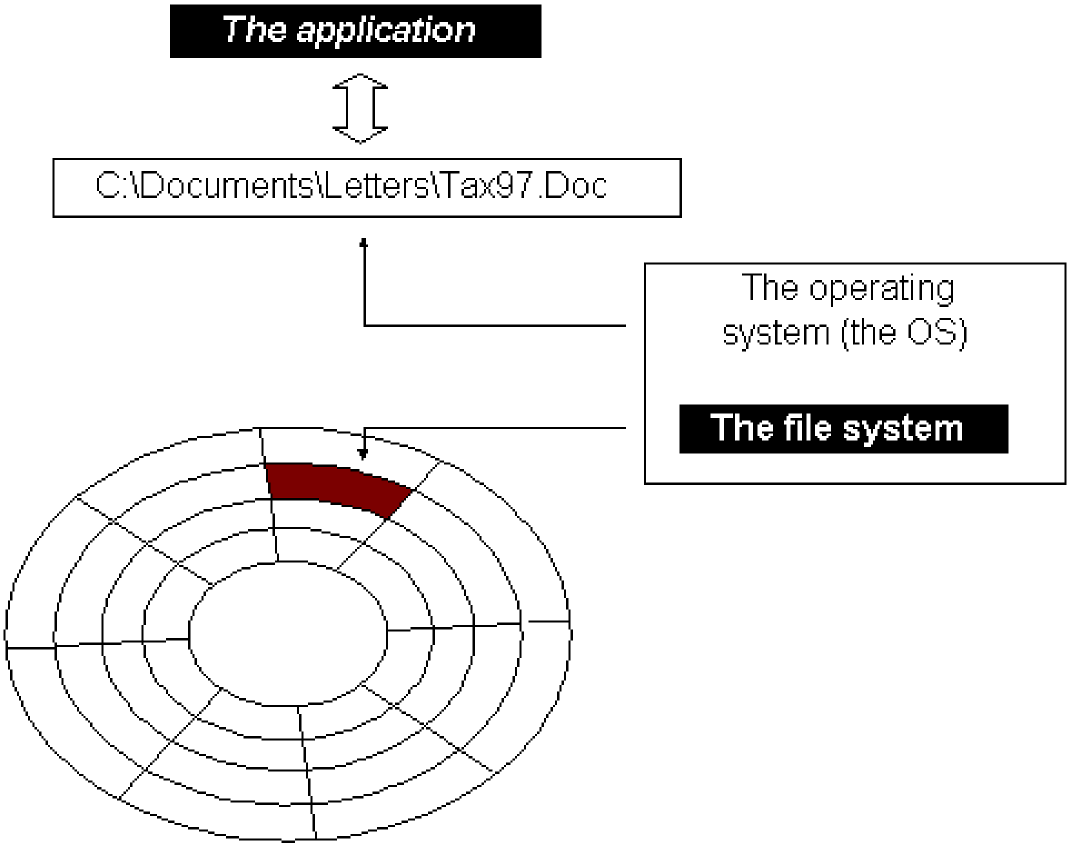
Relationship between file system and operating system

Top

We see that that the file system is an integral part of the operating system. An operating system can sometimes work with different file systems:

Operating system	File system(s)
DOS	FAT16
Windows 95/98	FAT16, FAT32
Windows NT	FAT16, NTFS
Windows 2000/XP	FAT16, FAT32, NTFS
OS/2	FAT16, HPFS
Novell NetWare	proprietary file system

The file system is actually the *interface* between operating system and drives. When the user software, such as MS Word, asks to read a file from the hard disk, the operating system (Windows 95/98 or NT) asks the file system (FAT or NTFS) to open the file:



The file system knows where files are saved. It finds and reads the relevant sectors and delivers the data to the operating system.

Limitations in disk size

Top

Over the years, the PC has suffered from a long list of irritating limitations. The hard disk industry has continuously developed hard disks with increasing capacity. However, the system software (BIOS, DOS, and FAT) has set its limitations:

- DOS versions below 3.0 could only handle hard disks up to 16 MB.
- Versions 3.0 to 3.32 could handle up to 32 MB.
- DOS 4.0 could handle up to 128 MB.
- DOS version 5.0 and the BIOS, which controls IDE drives, could only accept 1024 cylinders and disks up to 528 MB. This limit was broken with the EIDE standard.
- FAT16 can handle a maximum of 2 GB because of 16 bit calculations of the cluster size.
- FAT32 accepts disks up to 2048 GB. This standard will probably last another couple of years.

Let us return to the file system in next page.

- Next page
- Previous page

Learn more

Top

Read Windows tipswith a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6a.2

About file systems (continued)

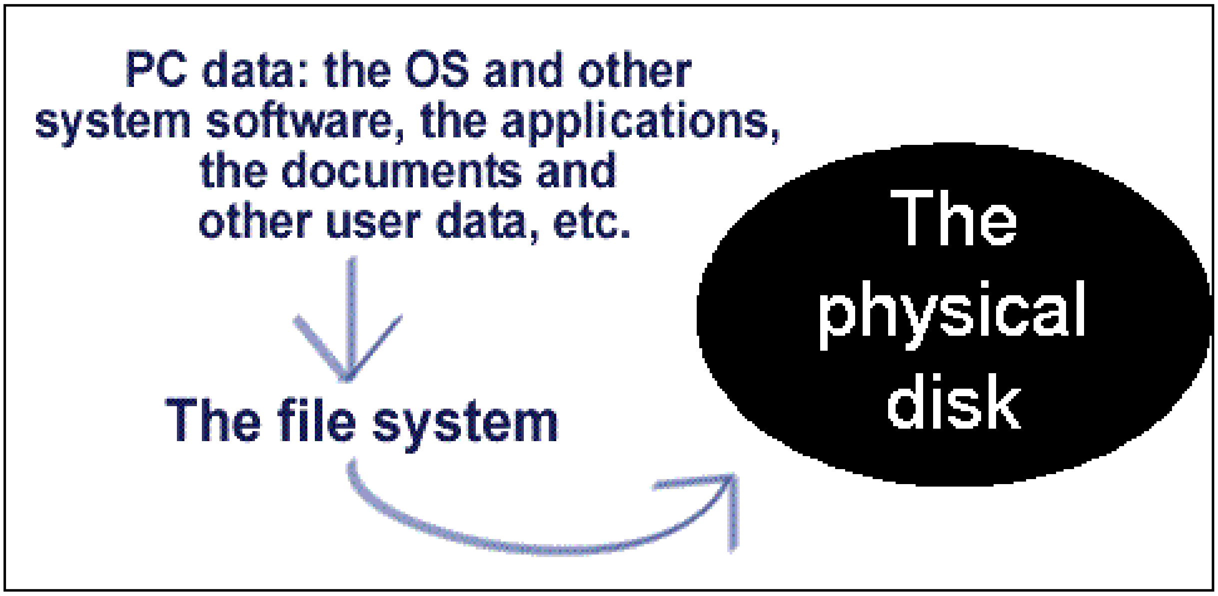
The contents:

- The formatted disk
- About sectors
- About clusters
- Small clusters with FAT32

- Next page
- Previous page

The formatted disk

We know that a disk must be formatted with a *file system* , before it can accept files to be saved:



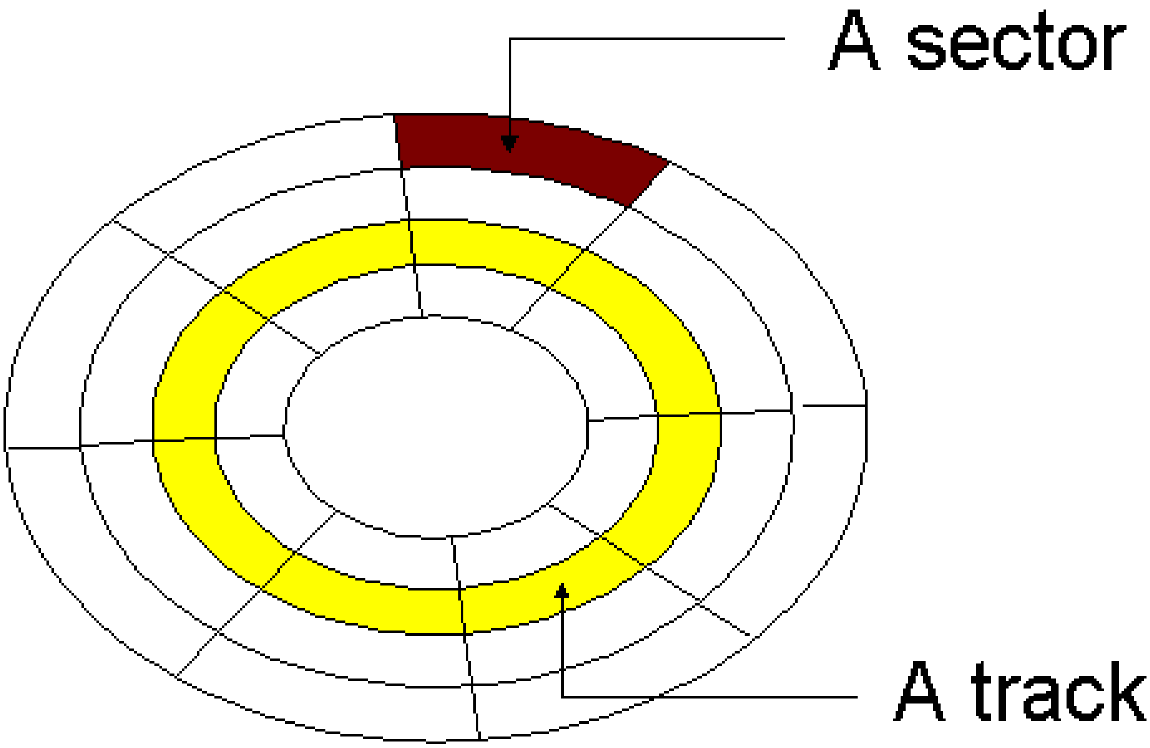
Now let us examine the disk formatting process. How does it work?

About sectors

Top

All disks are divided in 512 byte sectors. That is the standard size for the smallest disk unit. You could easily format with a different sector size, but that is not done. A sector is then the smallest disk unit and it holds 512 bytes of data.

Sectors are created when the circular disk is organized in concentric tracks. Each track is divided into sectors. Each sector can hold 512 bytes.



But, how are these sectors distributed? How are the files placed in the sectors? How do we handle a file larger than 512 bytes, which must occupy more than one sector? Who keeps track of what is in each sector?

This is a task for the file system. Below, we evaluate hard disks only and only FAT. Despite its age and flaws, it is still by far the most widely used file system. As for diskettes, read about diskette formatting .

About clusters

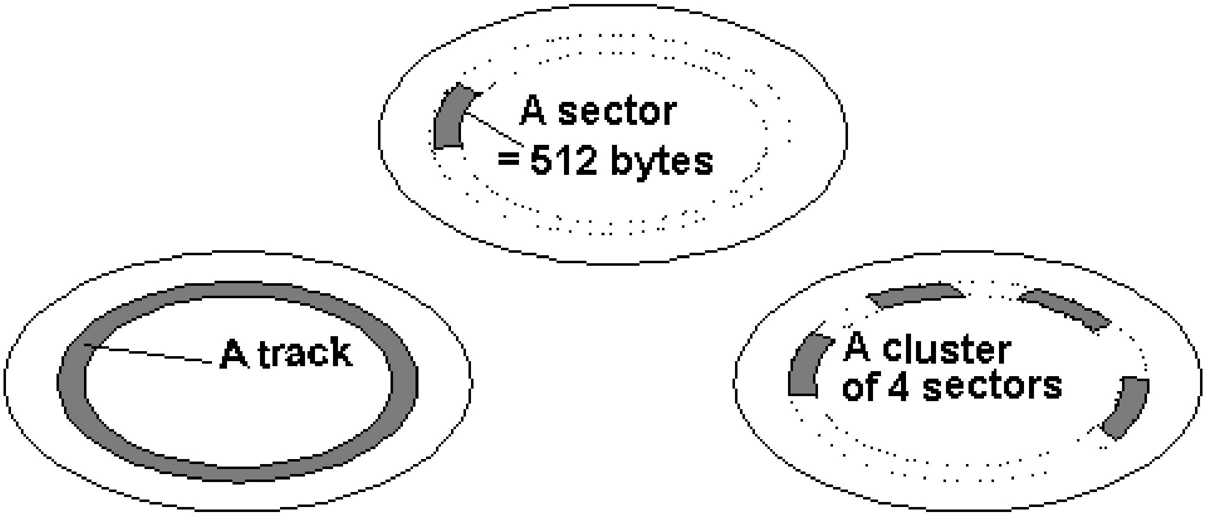
Top

To continue in the understanding of the file system, we must introduce a new concept - *clusters* .

Each sector holds 512 bytes and a sector is the smallest disk unit. However, often a sector is too small for DOS to handle. That is because DOS is a 16 bit operating system.

By design, DOS can only handle 2^{16} disk units at a time. A disk unit (my expression) is either a sector, or a cluster of sectors. Thus, DOS can only handle 65,536 of those!

Therefore, in FAT formatting the sectors are gathered in *clusters* of 2, 4, 8, 16, 32, or 64 sectors:



The cluster concept is an administrative invention. They are necessary, to allow DOS to handle large disks.

They are also called *allocation units* . The number of sectors gathered in one cluster depends on the disk size:

Disk size (partition size)	Cluster size
< 255 MB	8 sectors (4 KB)
< 512 MB	16 sectors (8 KB)
<1024 MB	32 sectors (16 KB)
< 2048 MB	64 sectors (32 KB)

In Dos, the data area of the hard disk is divided into a specified number of clusters, which of necessity increase in size with the size of the disk. On modern hard disks, the clusters will usually be 16 or 32 KB, as illustrated above

---

Small clusters with FAT32

Top

The good news is that FAT32, found in the Windows 95 OSR2 and Windows 98, handles disk formatting much better than FAT16. With FAT32 it is possible to format hard disk *partitions* of more than 2 GB with small cluster sizes:

Partition	Cluster size

<8 GB	4 KB
8 GB - 16 GB	8 KB
16 GB - 32 GB	16 KB
> 32 GB	32 KB

Something else new in FAT32 is the moveable root directory, which can be of variable size. It involves active use of both FATs (I cannot explain how). Altogether, it should make it simpler and safer to change partition sizes. But the number of clusters per partition grows enormously in large partitions.

FAT32 can only be installed in a new PC, since the partition has to be formatted in a special manner. The file system is only available in the Windows 95 OSR2 (OEM Service Release 2) and in Windows 98.

- Next page
- Previous page

To learn more

Top

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6a.3

About file systems (continued)

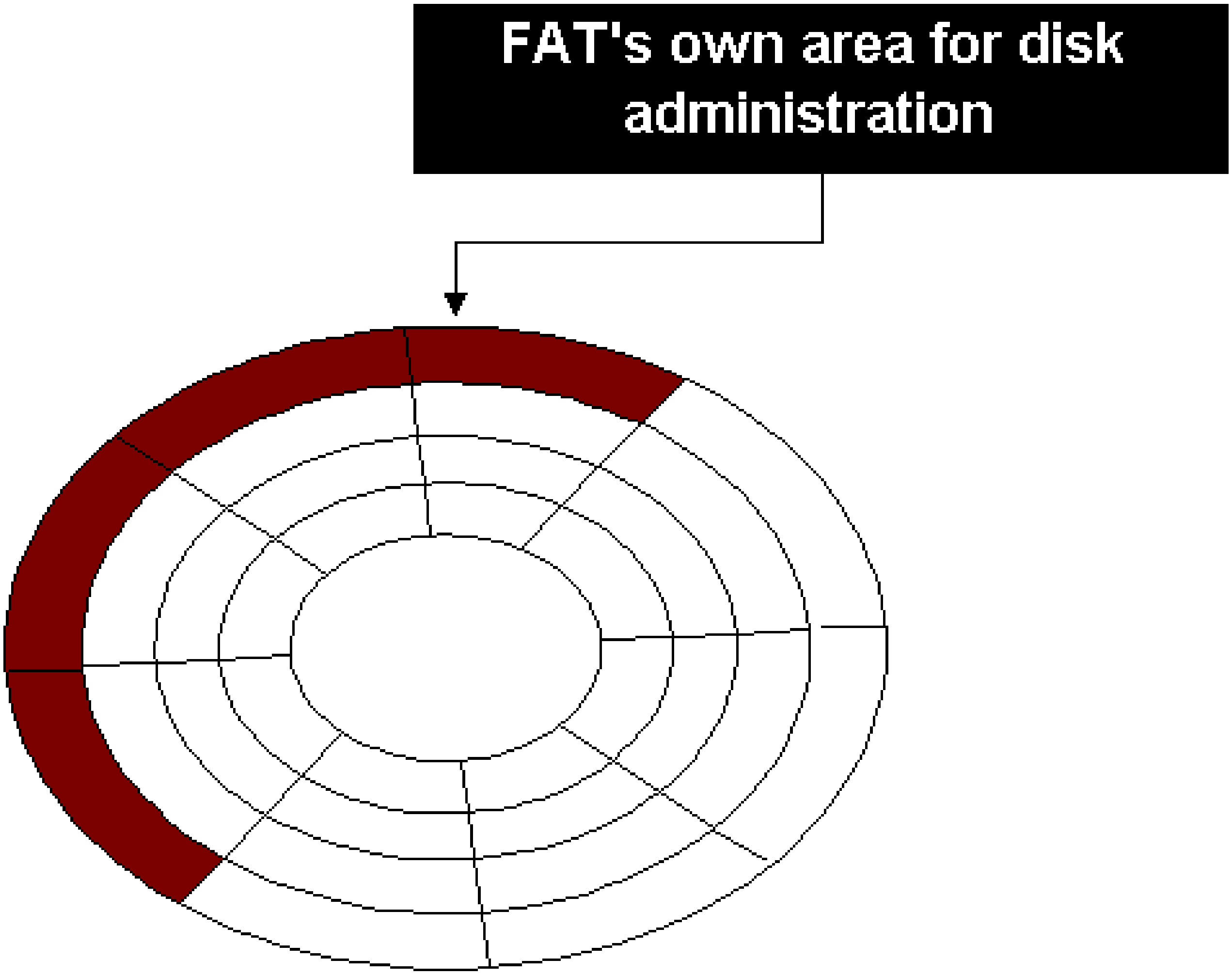
The contents:

- Working with the FAT formatted disk
- 4 areas different on the disk

- Next page
- Previous page

The FAT formatted disk

During formatting, all hard disk are divided into multiple sectors. The sectors must contain *both* user data and the file system administrative data. This is because in FAT, the administrative data are stored on the disk also:



Thus, the disk is divided in:

- Sectors, occupied by FAT administrative data.
- Sectors, which are user available for data storage (the data area).

The latter of the to parts obviously is the biggest.

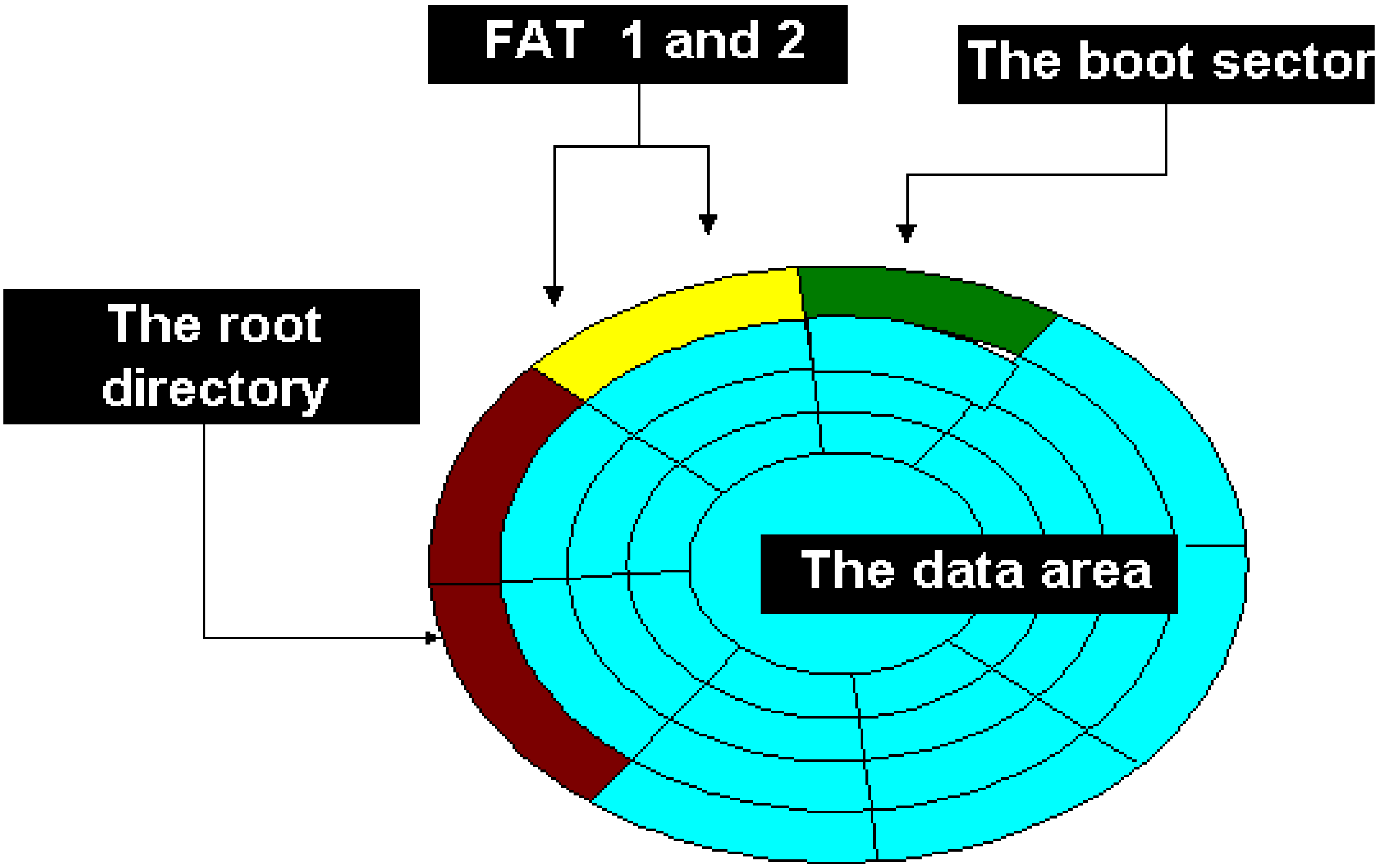
The four disk areas

Top

Each disk or disk partition contains four fundamental areas:

- The boot record, which is always in the *first sector* .
- FAT areas, of which there are usually *two* identical.
- The root directory.
- The data area.

In the data area all files and sub directories (beyond the root directory) are stored. The data area sectors are gathered in *clusters*, and this organization is illustrated here:



- Next page
- Previous page

To learn more

Top

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6a.4

About file systems (continued)

The contents:

- Boot record
- FAT areas
- Root directory and other directories
- The data area

- Next page
- Previous page

An example of the sectors in the four areas

Let us look at a FAT formatted hard disk with 160 MB: How are the sectors utilized?

Note: My calculations in this example are NOT 100% CORRECT, as a few bright readers have noticed. I am sorry. However, use them as an indication of how the disk is subdivided. One day I hope to give the correct picture...

The partition contains exactly 157.9 MB, if we calculate correctly. We are talking about 165,656,576 bytes. The total data storage area is divided in 323,549 sectors, 512 bytes each. If you multiply, that results in 165,656,576 bytes:

323,548 sectors x 512 bytes = 165,656,576 bytes.

This equals to 157.9 MB as you can see from this:

= 165,656,576 bytes divided with 1024 = 161,774 Kilobytes	161,774 Kilobytes divided with 1024 = 157.9 Megabytes

The file system now assumes control over these 323,548 sectors. The boot record occupies the *first sector* . In the following you find a brief description of the boot record and other administrative areas.

Boot record

Top

The first disk sector is always reserved for the *boot record.* It contains information about the disk or its partitions. A physical hard disk can be divided into different partitions. DOS, Windows 95 and NT treat each partition as a separate drive.

The boot record information enables the file system to handle the disk. At the same time, it includes a *small* program to use during system start-up. Here is a summary of that sector's contents (skip, if you do not understand):

8086 instruction (JUMP).
DOS name and version number.
Bytes per sector.
Sectors per cluster in the data area.
Number of reserved sectors.
Max. number of entries in the root directory.
Total number of sectors.
Media description (is this a hard disk?).
Number of sectors per FAT
Number of sectors per track.
Number of disk read heads.
Number of hidden sectors.
BOOT-strap program routine, which reads the hidden file (like IO.SYS), which starts the operating system.

The boot record is found on all disks, regardless of whether they are FAT or otherwise formatted. That sector contains the necessary description of the partition.

The FAT areas

Top

After the boot record, we get to the FAT areas. There are usually two identical FATs. FAT number 2 is simply a spare copy of number 1, since FAT is essential for the function of the disk.

The FAT file administration is actually a very simple system, but it is complicated to describe. Later, I will show some practical examples. Here is the first description. Even if you do not entirely understand the following, do not give up.

FAT consists of a table of whole numbers, which has 65,536 16-bit *entries* . Each of these entries contain information about a cluster.

The content of each FAT entry consists of a *whole number* . In the table below, they are written as four digit hexadecimal numbers, which show one of four options.

Possible FAT cluster entry	Value
The cluster is part of a file, the last in the file.	FFFF
The cluster is part of a file. You can read the number of the next cluster in the same file.	like A8F7
The cluster is empty, thus free.	0000

The cluster contains defective sectors.	FFF7

Example on reading a file

Top

When the file system has to read a file, it follows this routine. We imagine that the file occupies 4 clusters and it occupies cluster numbers 442, 443, 444, and 448. But how does the operating system read these addresses?

- Find the file directory entrance (from its file address).
- Read the first cluster, number 442 in the directory entrance.
- Look up in FAT under number 442. We find the number of the next cluster (443).
- Look up in FAT under number 443. We find the number of the next cluster (444).
- Look up in FAT under number 444. We find the number of the next cluster (448).
- Look up in FAT under number 448 to find FFFF (indicating the last cluster).

FAT always works in this way. Whenever a file has to be read, its location is read in the table. Every time a file has to be written to a disk, vacant clusters must be found for it, and the information is stored in FAT, to facilitate retrieval.

One of the great advantages of disk cache programs are, they always have a copy of FAT in RAM. In this way the disk cluster "map" can be read much faster than if the operating system had to read the FAT from the disk at each request.

The size of FAT

Top

Since each cluster has a FAT entry, the size of the FAT areas depends on the disk size. Each entry occupies 16 bits.

Let us return to the sector account in the example of a disk of 160 MB size:

The maximum FAT size is 128 KB, since 2^{16} files, 2 bytes each, equals 65,536 x 2 = 131,072 bytes or 128 KB. In our example, there turns out to be 40,400 *clusters* , since the disk partition is 160 MB.

We have two FAT's, at 40,400 X 2 bytes. That comes to a total of 161,600, and that will occupy 316 sectors.

The root directory and other directories

Top

The last administrative area on the disk is the root directory. Since there are always 512 file or directory *entrances* in the root directory, it is the same size on all hard disks. The root directory is unique in its fixed size and its location in the root. Other than that, it is a directory like any other.

Actually, a directory is a list of files and other directories. Thus, you can read the names of files and sub directories in the directory! The directory structure consists of a number of directory entries.

Let us look at these directory entries, each of which occupies 32 bytes. The directory entries are identical, whether they are in the root directory or a sub directory.

These entries, 32 bytes each, contain a lot of information like:

- The file name (in 8.3 format)
- File size in bytes
- Date and time of last revision

You can see the layout of the file entry on the illustration to the right. The 32 bytes are grouped in sections. This holds true for *all* entries, whether they point towards files or directories. This holds true for the root directory as well as all sub directories.

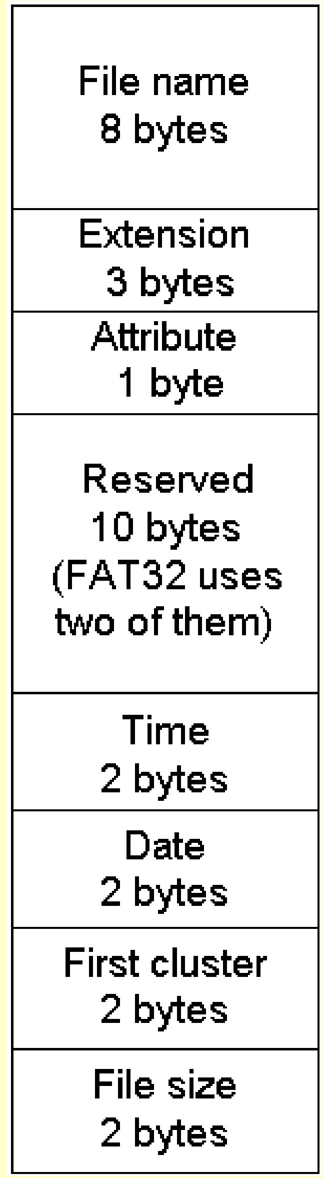
Note that we also find the number of the first cluster. This is important, because this is where the operating system starts to localize the file.

Remember the description of FAT above. You see that the start cluster number is read in the directory entry for the file.

Next FAT reads the numbers of cluster number two and so on, if the file is spread over additional clusters.

The location of any file is described in this manner: The first cluster is read in the directory entry (root or sub directory). The following cluster numbers are retrieved from FAT.

On FAT16 formatted hard disks, the root directory occupies 512 entries, which are 32 bytes each. Thus, it occupies 16 KB.



All sub directories have at least two entries. They are rather *special* , in that they refer to the directory itself and to its "parent" directory (in which it is a sub directory). The entries can be seen with the DOS command DIR.

The entry for the directory itself is seen as one dot. The entry for the parent directory is seen as two dots.

The data area

Top

The rest of the *disk* contains the most important part, the data area, where all files and sub directories are stored. The data area is by far the largest part of the disk.

The sectors in the data area are *allocated* in clusters. As mentioned before, the maximum number of clusters for data is 2^{16} = 65,536. Our hard disk is 160 MB. That results in 40,400 clusters, 8 sectors each.

All sub directory entries in the data area are organized in 32 byte files, which contain the same fields as the root directory entries.

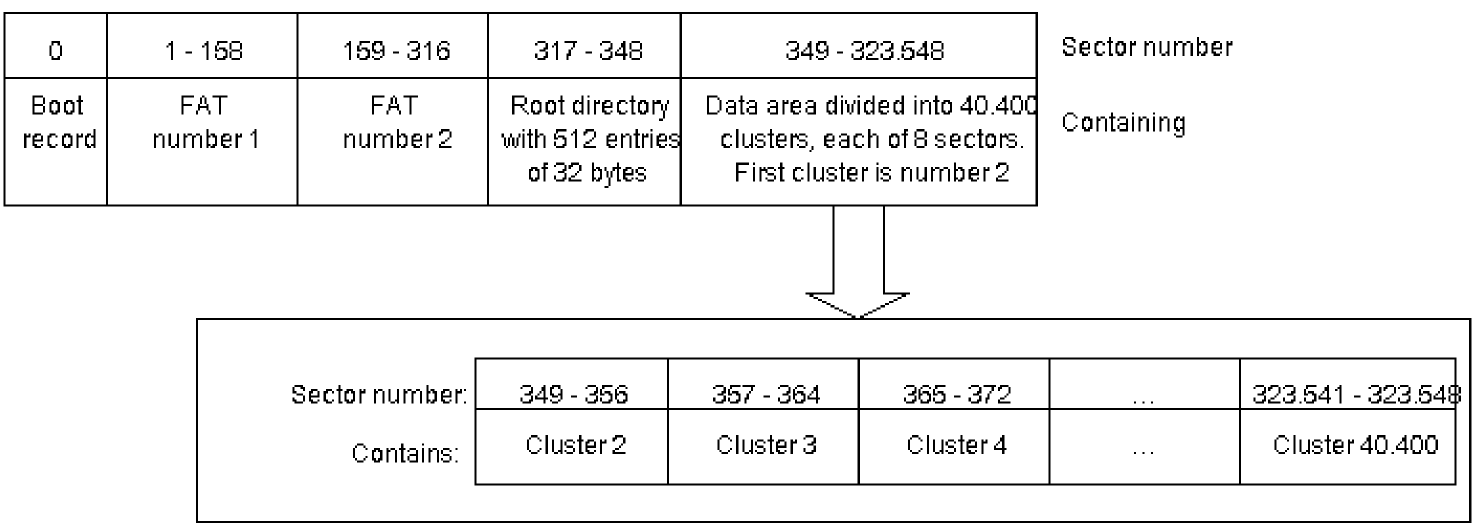
Completing the account

Top

The user has a 160 MB hard disk, but that is a somewhat theoretical view. Actually, the disk contains 323,548 sectors, 512 bytes each. They are distributed like this (or rather almost, there are minor errors in the calculations):

Area	Number of sectors	Sector number
Boot - record	1	0
FAT 1	158	1 - 158
FAT 2	158	159 - 316
Root directory	32	317- 348
Data area with 40,400 clusters of 4 KB	323,200	349 - 323,548

Here is a graphic illustration of the same distribution:



- Next page
- Previous page

To learn more

Top

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6a.5

About file systems (continued)

The contents:

- File fragmentation
- Defragmentation
- Windows 98 Defrag

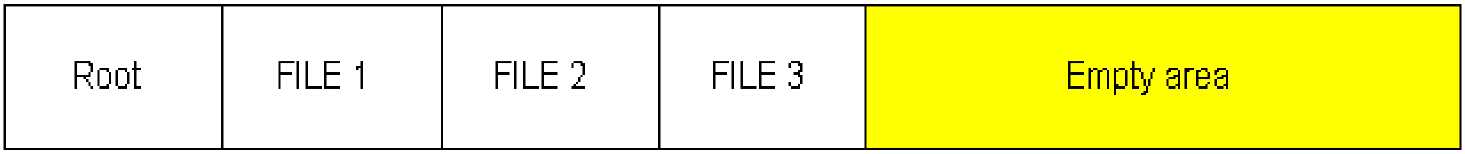
- Next page
- Previous page

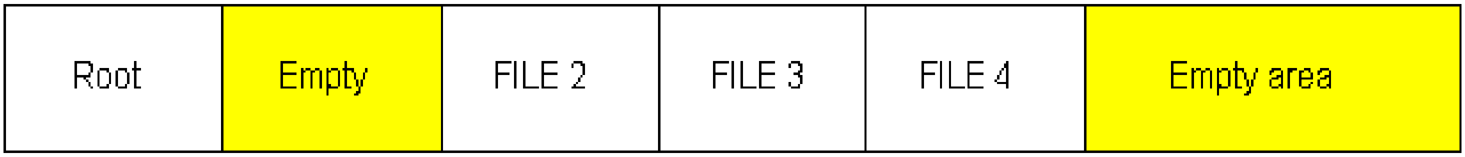
File fragmentation

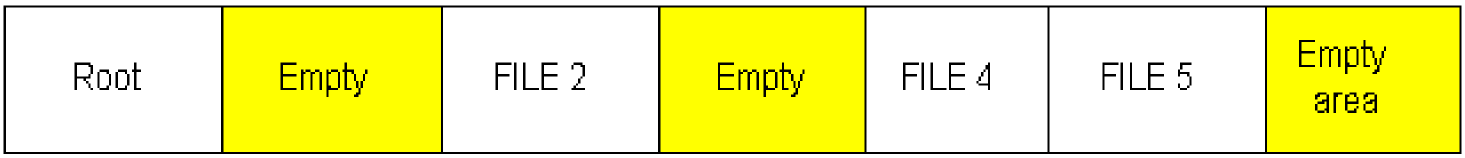
Top

When we work with FAT formatted disks, file *fragmentation* will occur all the time. One file can be several megabytes, thereby occupying more than one cluster. Maybe it requires 17 clusters. Ideally, the 17 clusters should be located next to each other. They can then be read at optimum speed, since that allows minimal movement of the read head. However, that is not the way it works.

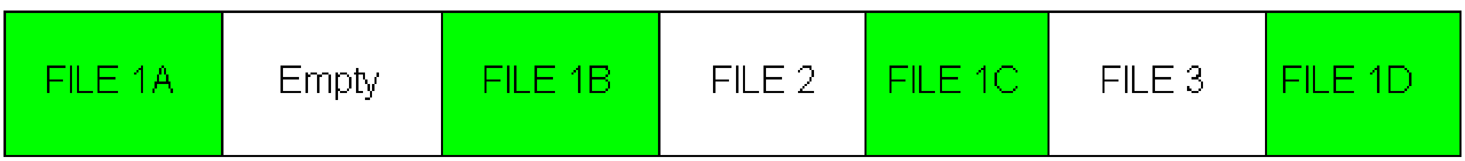
In actual operation, the individual files are broken up in multiple blocks, which are scattered across the disk. The problem increases with time. The more files you have on the hard disk, the more fragmentation you will experience. To begin with, vacant spaces appear between the files:







When you first write to a new hard disk, the file might occupy 17 clusters in sequence. The same will happen to file number 2, 3, etc., until there are no more vacant clusters on the disk. Then the file system must re-use clusters. That is done by finding empty clusters, where the contents have been erased. Thus, the file could be scattered in 17 clusters, *none* of which are in sequence. Here you see a file split in four disconnected clusters:



In the first DOS versions, when a new file had to be written, the file system always returned to the first vacant cluster to start a new file. That was done, to get optimum utilization of the disk. It also resulted in immediate and total file fragmentation. Since DOS version 3.0 the system was changed to fill the disk, before any vacant clusters were re-used. That delays fragmentation, but sooner or later it will occur anyway.

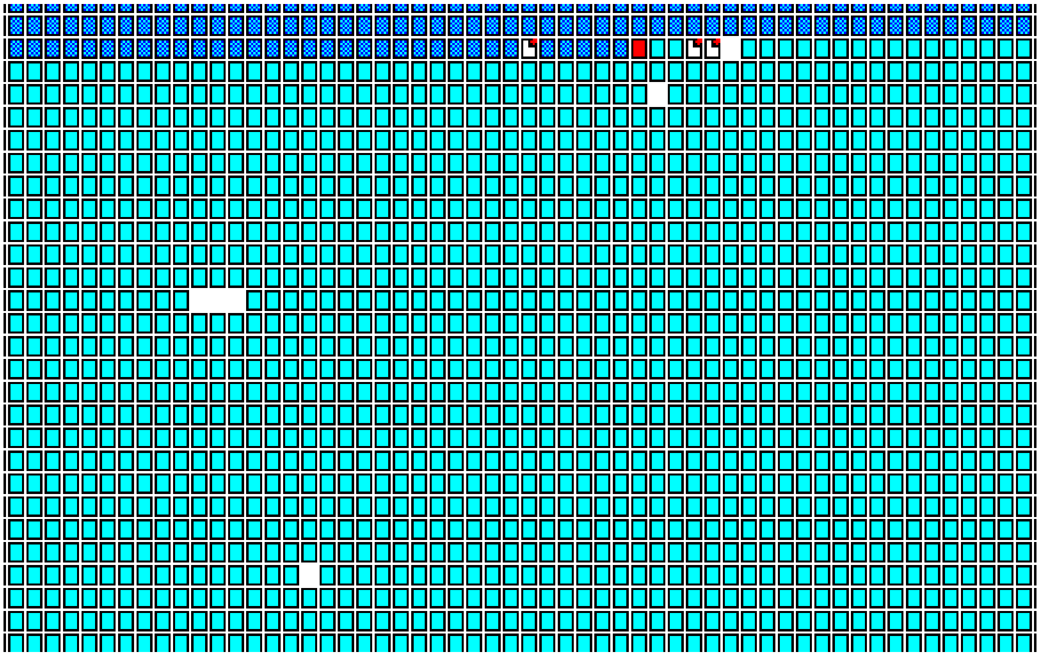
Defragmentation

Top

You can use the program DEFRAG to defragment the files on the disk. If you are a heavy PC user, it needs to be done often. I usually run SCANDISK first. That checks the file system for logical errors and repairs them. Scandisk will often find errors, so it does a good job.

Next defragment the disks with defrag /all. Both programs can be started with the command Start --> Run. Type in the command on the window: defrag /all

Here you see the defragmentation:



Run defrag weekly - that will keep your hard disks in good shape. Don't make the intervals too long. That can cause the disk to get messed up, especially if it is nearly full.

Windows 98 Defrag

Top

In Windows 98 the defragmentation was changed from the way it worked in Windows 95. Windows 98/Me monitors how programs are loaded. Opening Word, as an example, includes opening a large number of DLL and other program files.

With the defragmentation, all these files are placed in the right position to another on the disk, so they are loaded with optimal speed. It works very well, the programs are loaded 2-3 times as fast as before.

However, you have to defrag on a regular (weekly) basis and the process may take a long time.

- Next page
- Previous page

To learn more

Top

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6a.6

About file systems (continued)

The contents:

- Partitioning with FDISK
- More than one boot record
- An FDISK example
- FDISK /mbr
- The primary partition and booting
- Long file names with FAT32

- Next page
- Previous page

Partitioning with FDISK

Hard disks can be divided in more than one partition. That is done with the program FDISK, which is found in all PCs - regardless of which version of DOS, Windows , or OS/2 is the operating system. They all have FDISK.

FDISK can divide the hard disk in up to four partitions. In FAT16, the individual partition must not exceed 2 GB. Therefore it is often seen that the hard disk is not utilized 100%. Look at this picture of FDISK, which has partitioned a 2 GB hard disk. The illustration is in Danish, but you'll see the same in English:

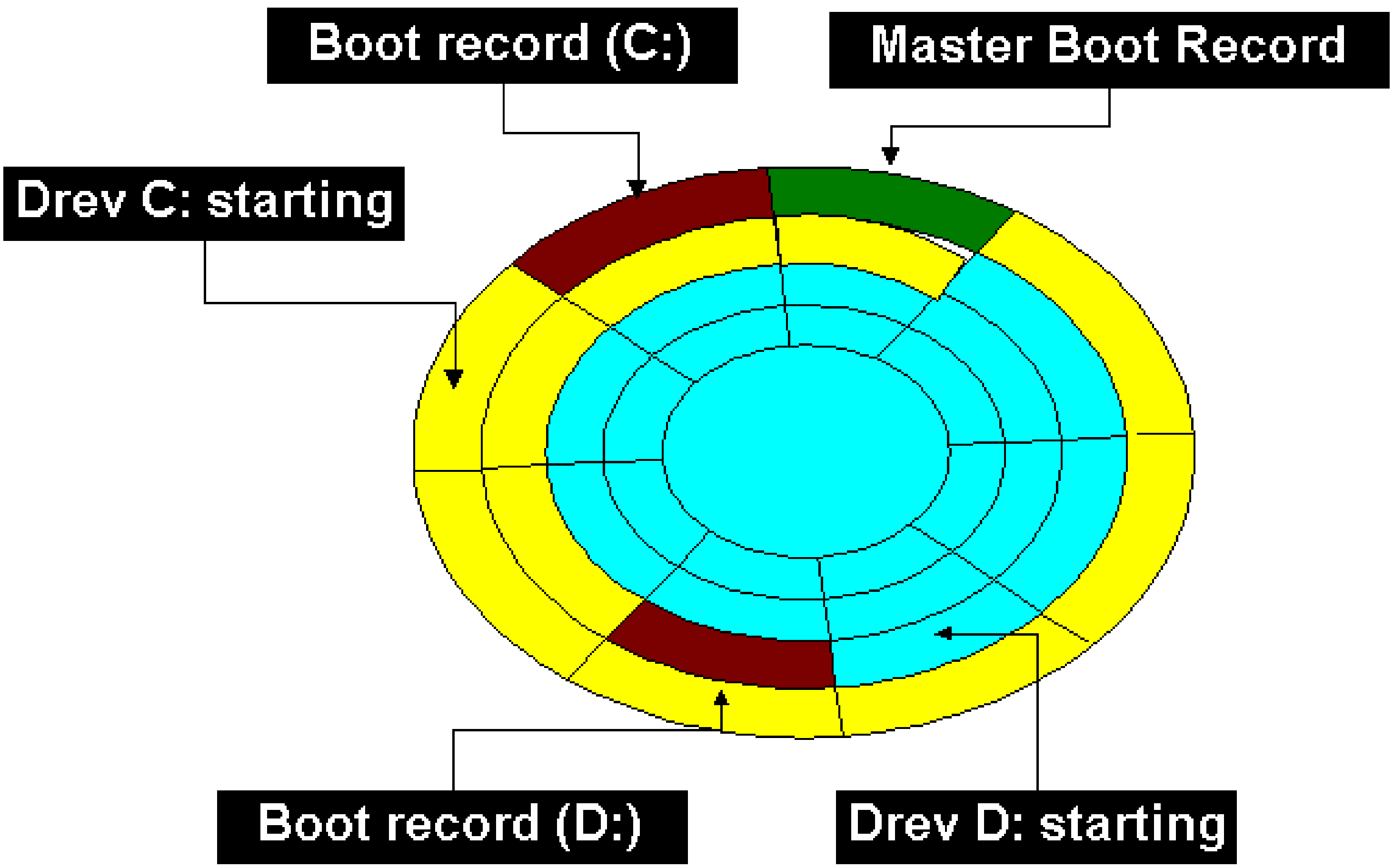
You can clearly see, that there are actually only two partitions. However, only the upper is assigned a drive letter (C:). The other partition consists of 43 MB unused hard disk, which FDISK identifies as Non DOS. It is not used, because you asked for a 2,020 MB partition. The remainder is left over. The 43 MB is not enough to bother to place in a new partition.

More than one boot record

Top

When FDISK has partitioned the hard disk, the file system must be able to recognize this partitioning. Information about the location of beginning and end of each partition is stored in the first sector (number 0), which is called Master Boot Record (MBR). Then, regular boot records are stored in the beginning of each partition on the disk.

Here is a disk divided in two *logical* drives, which will be named C: and D:. The Master Boot record is in the first sector of the physical disk. It contains information about the two partitions. In the beginning of each partition we find a new boot record, describing that partition.



An FDISK example

Top

You use FDISK to divide the hard disk in one or more partitions. FDISK writes a MBR in sector zero. That divides the rest of the disk in logical drives, each of which is regarded as a "real" drive by the operating system.

Let us look at the division of an old EIDE hard disk, as it was formatted using Windows 95. The harddisk was sold as 5.1 GB. Actually, it holds 4.8 GB. Through FDISK, this capacity is distributed in three partitions using FAT16. Here are the expressions, as used in Windows 95 version of FDISK:

First a *primary partition* is created. We choose to assign it maximum size. That is 2,047 MB, corresponding to 2,146,467,840 bytes. That becomes our C drive, which is activated, so we can boot from there.

We choose to establish an *extended DOS partition* for the rest of the disk.

The extended DOS partition must be divided in *logical DOS drives*. We choose to make the first logical DOS drive the maximum allowable size. The D drive will then be 2,047 MB, just like the primary partition is.

A smaller part of the hard disk still remains. We will make that into anaother logical DOS drive. That will have 813,561,344 bytes, or 775 MB. That becomes the E drive.

Now FDISK reports that the disk has three drives. C: is the primary partition, D: and E: are two logical DOS drives, which are in the extended partition.

The Physical Disk

If we look at the physical hard disk, we find that it has a total of 9,974,720 sectors, 512 bytes each. After the partitioning, these almost 10 million sectors are distributed as shown below:

Physical sector number	Contents
0	Master Boot Record, which describes the entire hard disk
1 - 4,192,866	Drive C:
4,192,867 - 8,385,732	Drive D:
8,385,732 - 9,974,719	Drive E:

Note, that each of three drives has its own disk administration divided in boot record, FAT, root directory, and data area. If we select the C drive from above, we can see here how the sectors are distributed in the C drive partition:

Physical sector number	Contents
1	Boot record
2 - 513	FAT 1 + 2
514 - 545	The root directory
546 - 4,192,866	Data area, which is divided in 32 KB clusters

FDISK /mbr

Top

Usually the Master Boot Record holds read-only information. It is written once by FDISK during the process of partitioning the drive, but after that it should be left unchanged. However, some programs do change the content of the MBR. This goes for:

- Virus, certain viruses are occupying the MBR giving them a safe position.
- System Commander and other multiple boot utilities.

The multiple boot utilities replace the MBR with code belonging to the utility. This way a utility like System Commander takes over the boot process and allows the user to install several operating systems at the same PC. It works fine, I can say; I have used it for a long time without any problems.

But how do I get rid of these MBRs, if I want to re-establish the original MBR. One way is to use the command fdisk /mbr. It simply re-writes the MBR and the other boot sectors.

I have used it against a virus a few times, having to boot from a floppy disk first. And I recently experienced a disk, where System Commander was installed and protected by a password! This disk was inaccessible even after FDISK'ing and formatting until we used fdisk/mbr. So please remember this command!

The primary partition and booting

Top

There will always be one primary partition on the hard disk. Booting must be from the primary partition and the operating system is read from here.

The hidden system files

The core of the operating system is stored in the two hidden system files, which are always found in a primary DOS partition. In traditional MS-DOS, the files are named IO.SYS and MSDOS.SYS. These files have the same names in Windows 95/98, but the contents are changed slightly compared to the traditional DOS. This review is from the old fashioned DOS, but tells something general about the boot process of an operating system.

The DOS system formatted disk contains two hidden system files. The first, IO.SYS, *must* be the first entry in the root directory. MSDOS.SYS *must* be on entry number two.

Start-up on disk

When the start-up program has finished POST (Power On Self Test) and the loading of BIOS routines, the boot process starts. It follows the following steps:

- MBR is read first. The sector number of the primary partition's boot record is found here.

- A small boot program is read from the primary partition's boot record. That activates the loading of the two hidden files.

- IO.SYS is saved to working memory. Besides creating an interface "downwards" to the BIOS programs, IO.SYS includes a small program called SYSINIT. This program starts the next steps of the boot process.

- Now MSDOS.SYS is read from the disk.

- Then SYSINIT looks in root directory for a file named CONFIG.SYS. All commands in CONFIG.SYS are executed, gradually configuring the PC to have a ready and usable DOS.

- Then SYSINIT looks for the command interpreter COMMAND.COM. If that is not found, we will get an error message about this. When it is found, AUTOEXEC.BAT, which contains the last information for personal configuration of the PC, is executed.

That was a little bit about the boot process.

OS/2 Boot Manager

With OS/2's FDISK edition, you can divide the hard disk into more primary partitions. That allows use of the special Boot Manager, which comes with OS/2. Even if you do not use OS/2, you can still use Boot Manager let us say to have DOS /Windows 3.11 on one primary partition and Windows 95/98 on another. They will both appear as C drives, but you can only see one at a time. This, you control with the Boot Manager.

I hope you understand the importance of FDISK. It is a good program to be fluent in. Altogether, it is important to understand the file system, the boot process, etc.

There are two excellent utilities - Partition Magic and System Commander, which give further facilities to change the partitions and the start-up sequences, etc.

Long file names with FAT32

Top

You can store long file names in Windows 95/98, which uses the VFAT file system. That is a 32 bit edition of FAT. VFAT was introduced with Windows 3.11, but the long file names did not become available until Windows 95.

The file systems in Unix, NT, and OS/2 have *always* been able to store long file names, but now Windows can do it too. Also VFAT is compatible with regular FAT, which is smart. You can exchange files with other PCs - regardless of whether they can use long file names or not.

Actually, the VFAT file system is much like regular FAT. But in a smart way Microsoft has been able to break the heavy 8.3 file name limitation, which limits regular FAT.

Physically, the file names are stored in a traditional 8.3 file name, which VFAT creates (without user control). The user can assign a long file name. As an example, a file is named "Ford Escort sales.doc". That will be translated to "FORDES~1.DOC", when the filename is registered by FAT.

The long file names may be up to 255 characters long, but they are translated to an "alias," which follows the traditional 8.3 FAT format. The trick is, that the long file name is written across *multiple* directory entries. Normally, one directory entry points towards one file, but in this case one file can occupy several root directories, each of which provides 32 bytes to the file name.

You should be happy about the long file names in Windows - it makes it much easier to identify saved files. The only "danger" is, that you *must not* defragment the hard disk with a DOS based application. Then the long file names are destroyed. The files still exist, but you can only find them under their 8.3 name and that is an annoying experience, especially if you have thousands of files.

However, you should not waste memory and disk space using filenames 50 characters long. Usually filenames of 15-20 characters work fine.

- Next page
- Previous page

To learn more

Top

Read Module 6b with a little about Windows 95/98.

Read Module 7a about the videosystem

Read about video cards in Module 7b .

Read about digital sound in Module 7c .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6c1.

About operating systems and driver programs

The contents:

- What is an operating system?
- The operating system recognizes hardware

- Next page
- Previous page

On the following pages:

- BIOS or driver programs
- Which operating systems?
- DOS control of hardware
- 32 bit drivers and installation

Click & Learn deals primarily with hardware. In these pages I will cover the operating system as it connects downward towards hardware. The operating system is closely associated with the ROM BIOS program routines, which are described in module 2a . The two program layers (operating system + BIOS) are called *system software* and it is very useful to understand their importance for the PC.

Let us start by studying what an operating system really is.

What is an operating system?

[top]

Traditionally the operating system consists of three parts:

Part	Function
Kernel	The low level core being loaded after the boot process. Has many functions such as control of the data flow between memory and I/O units.
Shell	The user interface
File system	A standard for disk formatting

The operating system can also be evaluated from these viewpoints:

- An operating system is a number of files, which are read from the hard disk at the end of the PC start-up routine.

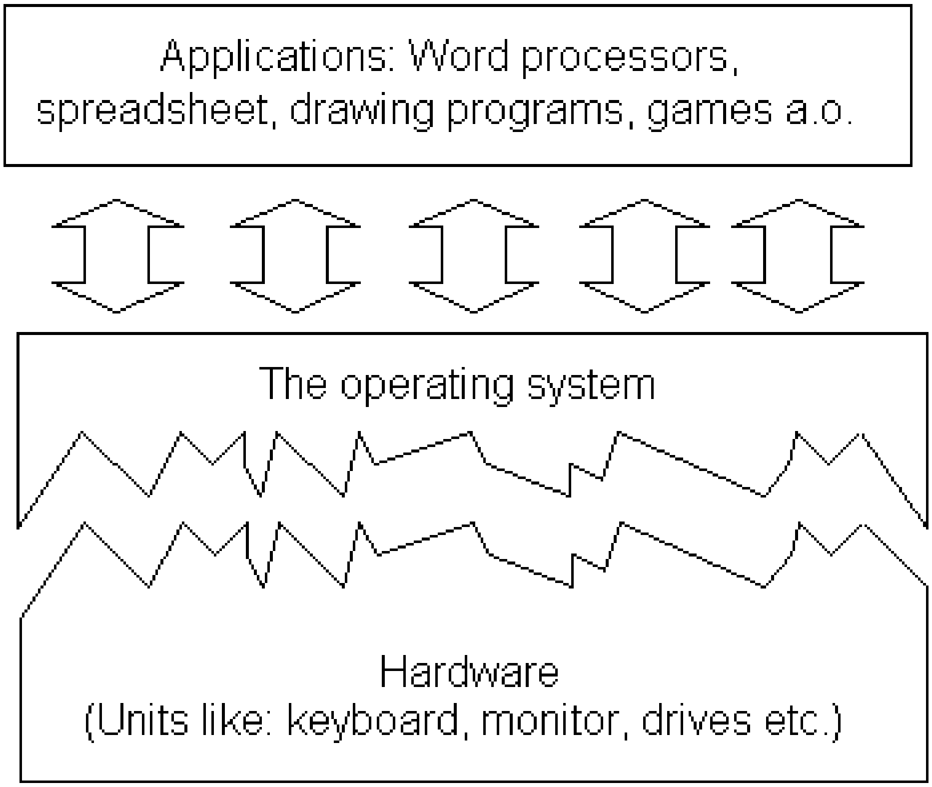
- An operating system is a program layer. It connects to the PC hardware, to facilitate optimal execution of the user programs.

The first definition does not say much. Let us start with the second: The operating systems links software and hardware together. It has to enable user programs, like Works, Office, etc., to function with all possible hardware configurations. You can imagine the relationship between hardware and user programs thus:

- Hardware is clumpsy and dissimilar. There are untold variations of PCs. They can have one or another type hard disk, CPU, video card, etc. All of these various PC configurations behave each in their own way.

- The user programs are 100% similar. They are off the shelf products, which expect the PC to respond in a certain manner.

How do we make these two layers work together? Can we eliminate, take out, the differences in the PC hardware, so a standard product like Works just functions? Yes we can. We read in an *operating system* - a system layer, which smoothes out and standardizes the hardware:



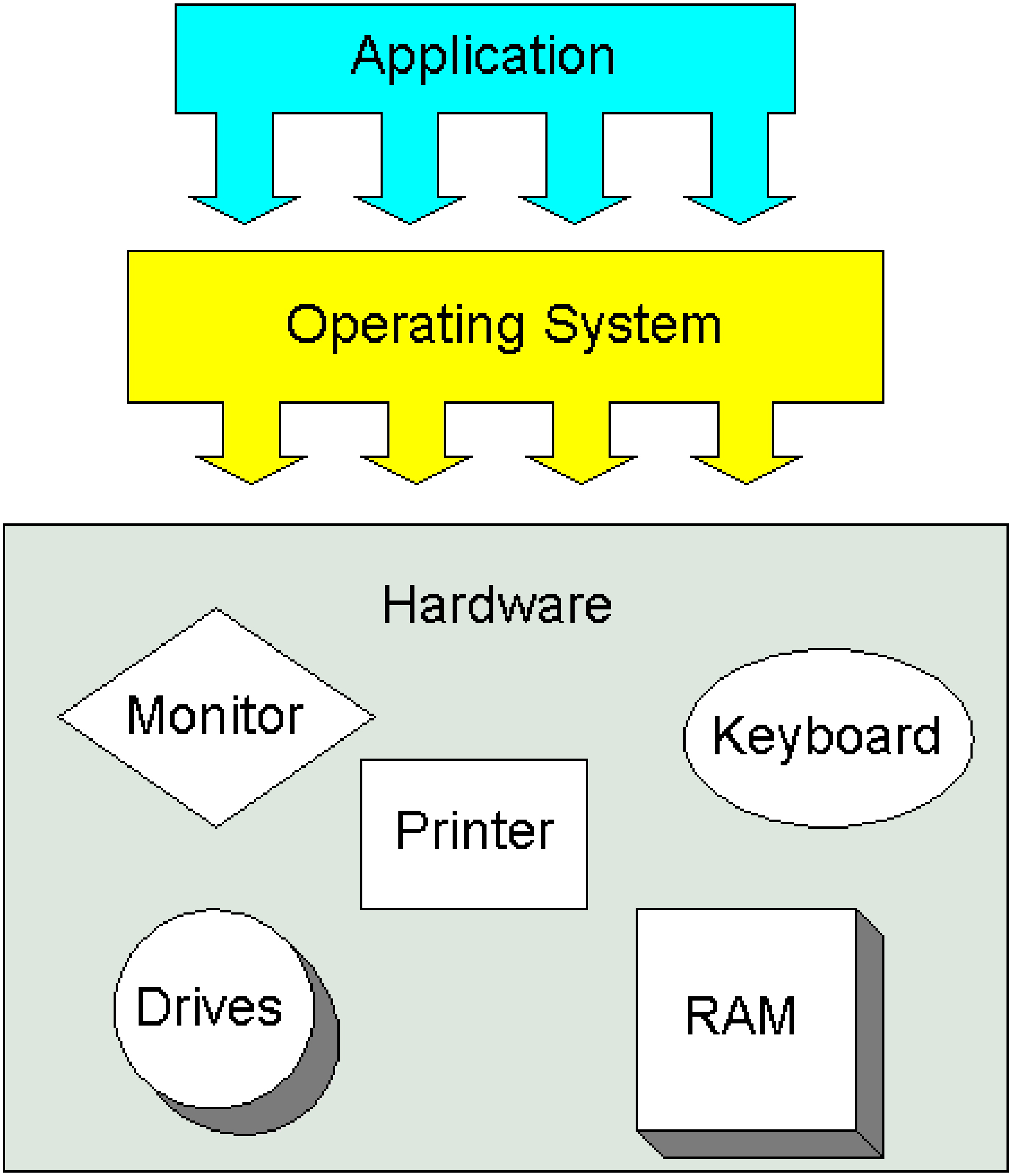
You should understand the operating system as a necessary layer, which smoothes out bumps and pot holes in your PC's hardware. This will give the user programs a stable, even work platform.

The operating system recognizes hardware

[top]

The PC's hardware represents *resources* relative to the user program.

Think of your word processing program: You want to print your text. The program issues a print order, expecting that the document will be printed as designed. The word processing program dispatches data according to your commands. How they are translated to signals understood by *your* printer - that is not the word processing program's problem. The printer is a resource relative to the word processing program. The connections to these resources is via the operating system. This holds true for all the resources, which are included in the PC hardware:



As you can see, the operating system has a very central function in the PC. So with that placement, it must be able to recognize all forms and types of hardware. There is no point in connecting a new mouse, if it does not work! Then what makes it work - the operating system. The system must recognize your mouse!

- Next page

- Previous page

Learn more **[top]**

Also see The Software Tips

Read of module 7a and module 7b about installation monitor and video card in Windows 95/98!

Read about chip sets on the motherboard in module 2d

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6c2.

About OSs and drivers - continued

The contents:

- System software
- BIOS or driver programs

- Next page
- Previous page

System software

[top]

Together, the operating system and the ROM BIOS program routines form the layer on which the user programs "rest." When the PC has to work, an operating system has to be read from a disk. There are many different operating systems to choose from. However, the BIOS is always placed firmly and centrally in the PC hardware.

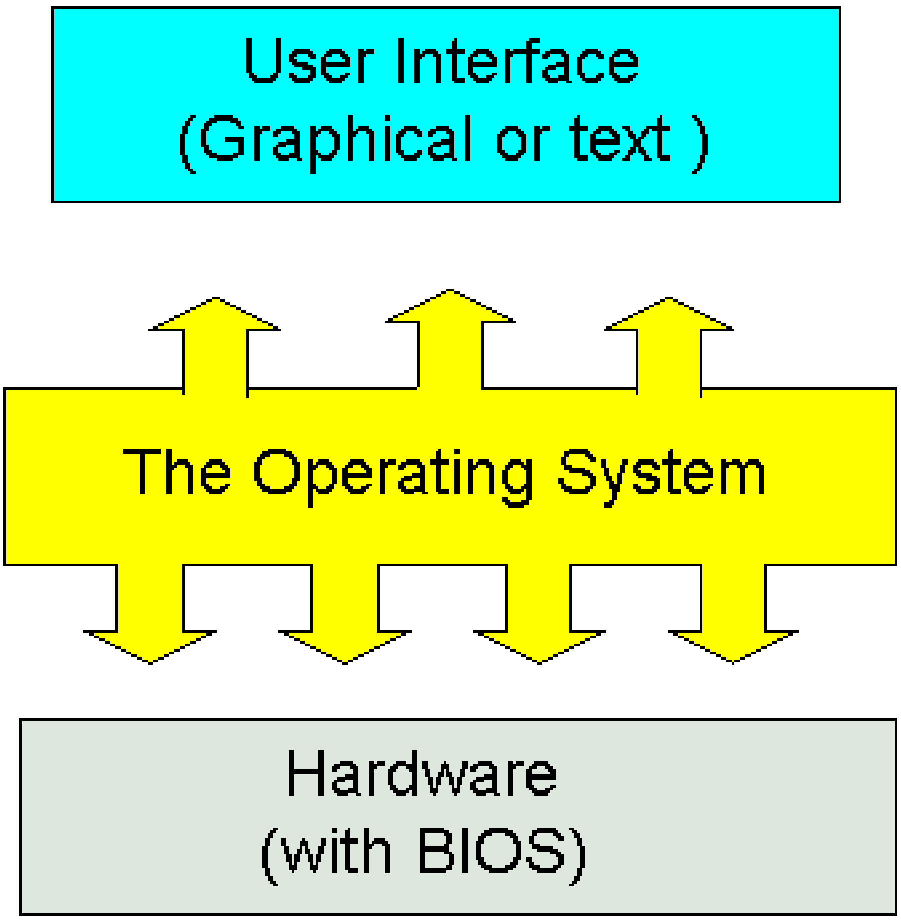
BIOS - firmware

One of the fundamental techniques in the PC design is the BIOS program layer. BIOS (Basic Input Output System) is a group of small programs, furnished by the PC manufacturer - also called *firmware* .

The BIOS routines are placed in the hardware - in a ROM chip - and are always available. Being stored in the hardware, they are functional regardless of which operating system they have to work with. So, in designing an operating system, one must pay close attention to the BIOS. The operating system must be able to work closely with the BIOS.

BIOS contains some very basic program routines, which handle data transfer between different hardware components. During PC start-up, the BIOS programs are the only accessible software. Later in the start-up process, the operating system is read. It will then take control of the PC. The operating system has to provide a *user interface* , on which the user programs can rest.

Thus, the operating system has two "faces": One pointing up towards the user and his/hers programs and one pointing down towards the system and hardware:



As computers have become more and more powerful, the user interface has become more graphic and user friendly. In a few years we will be able to address our commands directly to the operating system (you can do it already today with IBM's OS/2).

Thus, the "upwards" face of the operating system will change greatly - supported by technological development. The "downwards" face - the operating system's interface with hardware - will change less. At least, the fundamental principles are the same as in the childhood of the PC.

BIOS or drive programs

[top]

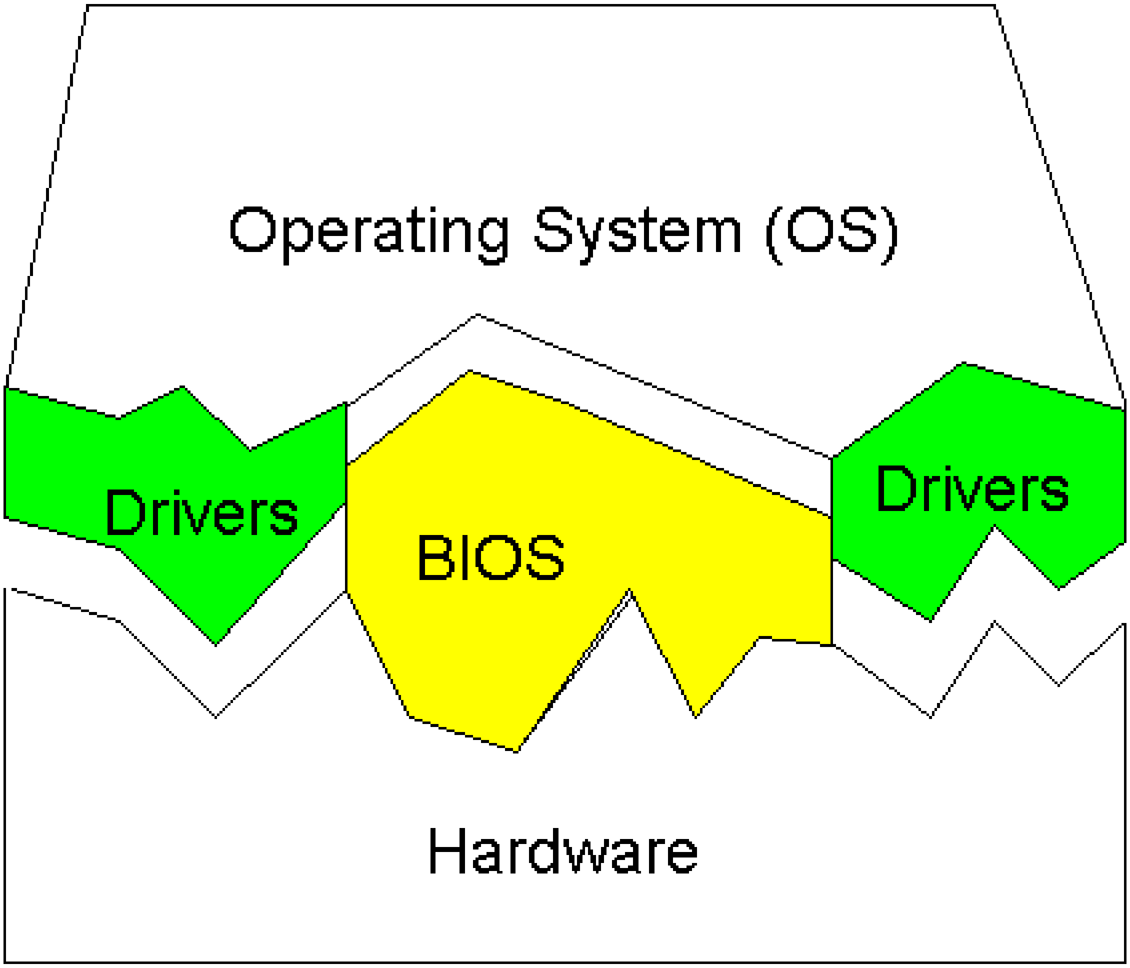
The operating system must be able to communicate with hardware. As we are going to see, this can be done in two ways:

- The operating system communicates directly with hardware through *drive programs.*
- The operating system utilizes the BIOS programs.

While BIOS is hardware specific program code, stored in hardware, the drive programs are small hardware specific program elements read from the disk together with the operating system.

Depending on which operating system is installed, both principles are used in various degrees. Since the BIOS programs consist of 16 bit code, it is typically DOS (a 16 bit operating system) which utilizes BIOS to a large degree. In the newer 32 bit operating systems, it is not efficient to use BIOS any more than necessary.

Here is a model, which shows the operating system with BIOS and drive programs (usually just called *drivers*):



As you can see, the driver/BIOS functions are closely associated with the operating system. So let us look at that on the following page.

- Next page
- Previous page

Learn more **[top]**

Also see The Software Tips

Read of module 7a and module 7b about installation monitor and video card in Windows 95/98!

Read about chip sets on the motherboard in module 2d

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6c3.

About OSs and drivers - continued

The contents:

- Which operating systems?
- DOS control of hardware

- Next page
- Previous page

Which operating systems?

[top]

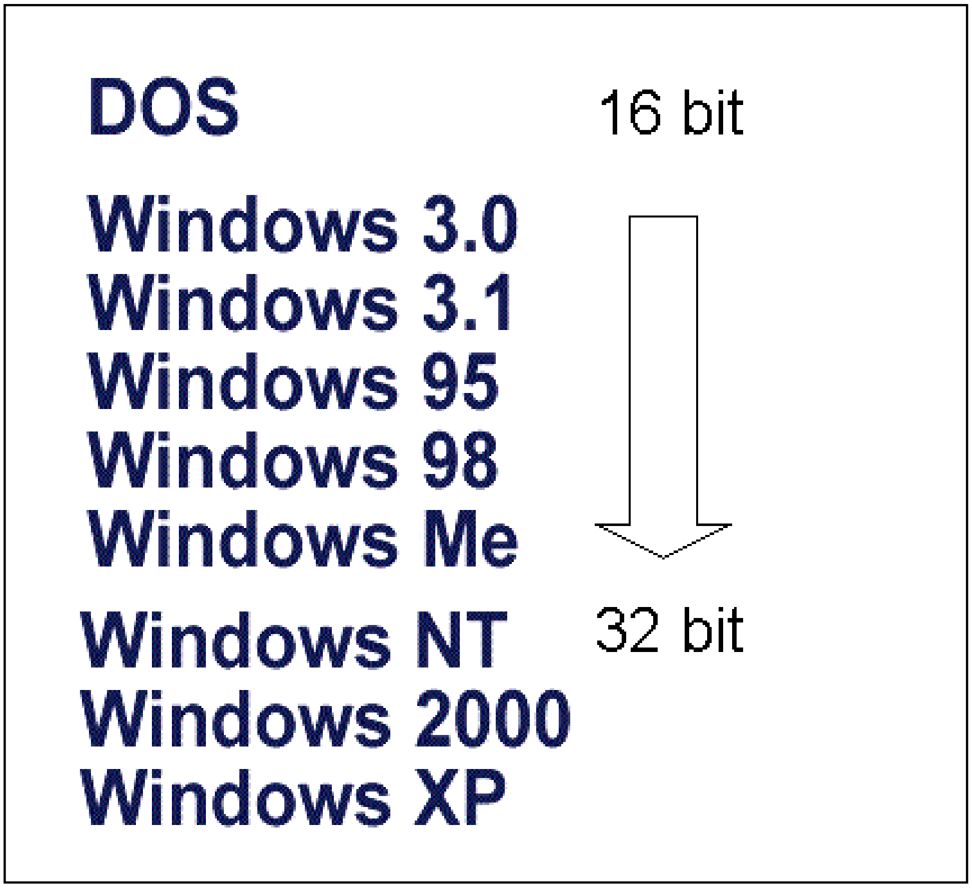
The operating systems have undergone a tremendous development since 1981. It all started with *DOS*, which was a 16 bit modification of a simple 8 bit operating system called CP/M.

DOS was further developed in the 1980s. Since around 1990 Windows came into the scene. Windows started as a GUI (*Graphic User Interface*) for DOS.

The PC booted with DOS as operating system. Then you could choose, if you wanted mouse and graphics on the screen with Windows . Windows was a supplement to DOS.

The Graphic User Interface (GUI) allows you to work with a mouse instead of writing long command lines like copy c:\texts*.* d:\textbak\ *.* /s/v/, which is the standard in text based operating systems (like DOS).

DOS was designed for 16 bit computers, which the first PCs were. With Intel's 80386 the 32 bit technology was knocking at the door. Modern PCs are designed for straight 32 bit program execution. So we have seen a gradual trend in the PC operating systems from 16 bit towards 32 bit and this affects hardware design.



OS/2 was a completely new-designed OS build on a 32 bit kernel (as UNIX and LINUX). It was originally designed by IBM and Microsoft together, but Microsoft abandoned the project in favour of their own Windows . In the early 1990s many people (including I) were very fond of OS/2. But it lost momentum as Windows 95 appeared.

Windows 95 was a radical development of Windows 3.11. Windows 95 was build on DOS, but the 32 bit components made up a big part of the OS. Together with Windows 95 came new 32 bit appliacations, which could not be executed within DOS.

Windows 98 and Me are further enhanced versions of Windows 95.

Windows 95 and Windows 98/Me are mostly 32 bit OS's, but with some 16 bit remnants.

Windows NT is a pure 32 bit OS from Microsoft. It was developed in a parallel run with OS/2.

Windows 2000 and XP are more popular versions of Windows NT. They hold the same user interface as Windows 98 end ME. They include all the DirectX technologies of Windows 98/Me which enables game and other multimedia applications to run.

Protected mode

The 32 bit programs we know from Windows 95/98/Me work in *protected RAM sectors*, with the CPU running in *protected mode* . This allows the PC to multitask - more than one program can run concurrently and independently. That is not possible in 16 bit operating systems, where the CPU works in *real mode* .

A brief comparison of 16 bit and 32 bit operating systems can look like this:

Operating system	DOS	32 bit operating system (NT, OS/2, UNIX)
Users	Single user	Multiple users
Program execution	16 bit single task in real mode	32 bit multitask in protected mode
Screen appearance	Mostly Text based (poor quality graphics)	Often GUI - graphic interface with high resolution graphics
Hardware handling	Primarily BIOS	Custom designed 32 bit drivers for each hardware component.

DOS control of hardware

[top]

DOS is quite simple to describe, since it principally consists of only 4 parts:

- A boot record, which activates the operating system.
- The file IO.SYS, which is interfaced to ROM BIOS with installation of *device drivers* .
- The file MSDOS.SYS. That is the core of DOS, handling the file system and program execution.
- The file COMMAND.COM, which provides the command line, the text based user interface.

When we talk about hardware control, it is done through IO.SYS. That is a program which reads the ROM BIOS code and converts it to DOS's own device drivers.

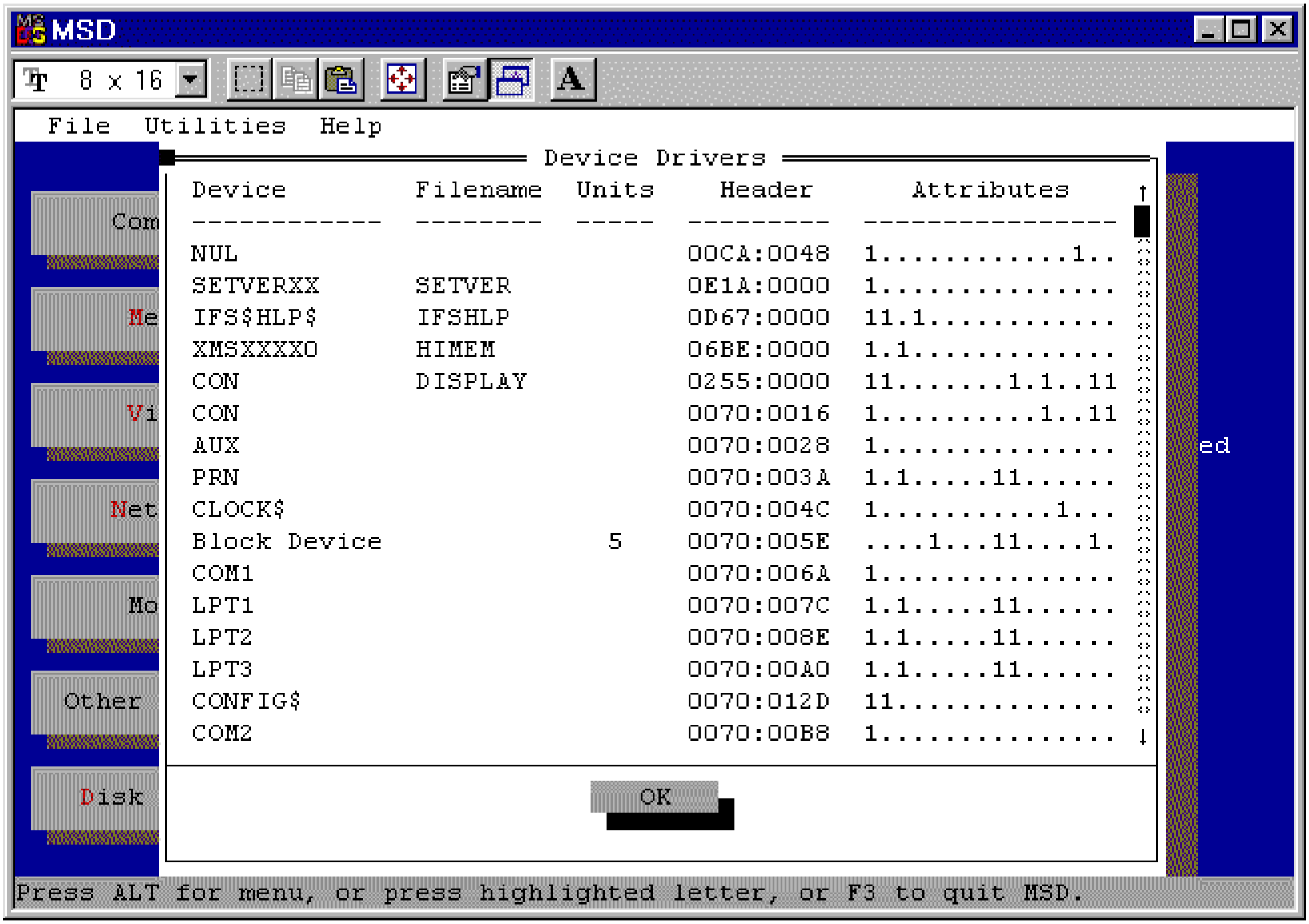
The smart thing about DOS is that the operating system can be *expanded* with external device drivers. IO.SYS reads them via the start-up file CONFIG.SYS. First device drivers are read from ROM BIOS. Then any possible additional drivers are read from disk. In that way DOS can handle hardware units which did not exist when the PC was originally configured.

A final option to handle hardware from DOS programs is to write special drivers for the individual *user program* . Many DOS games come with their own graphics drivers (they have to recognize all graphics standards on the market!). Another classic example is the word processing program WordPerfect, which in its prime (version 5.1) came with drivers to more than 500 different printers!

Unit	Example of DOS device drivers
Hard disk	BIOS
Video card	BIOS
Mouse	MOUSE.SYS

CD-ROM	ATAPI.SYS + MSCDEX.EXE
Printer	Internal drivers in the user program (like WordPerfect 5.1)

The device drivers can be seen with the program MSD. Here is a picture from Windows 95, where you can clearly see the names of the device drivers (CON, PRN, LPT1 etc.):



All these device drivers are in 16 bit program code.

- Next page
- Previous page

Learn more **[top]**

Also see The Software Tips

Read of module 7a and module 7b about installation monitor and video card in Windows 95/98!

Read about chip sets on the motherboard in module 2d

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 6c4.

About OSs and drivers - continued

The contents:

- 32 bit drivers
- Installation of new drivers

- Next page
- Previous page

---

32 bit drivers

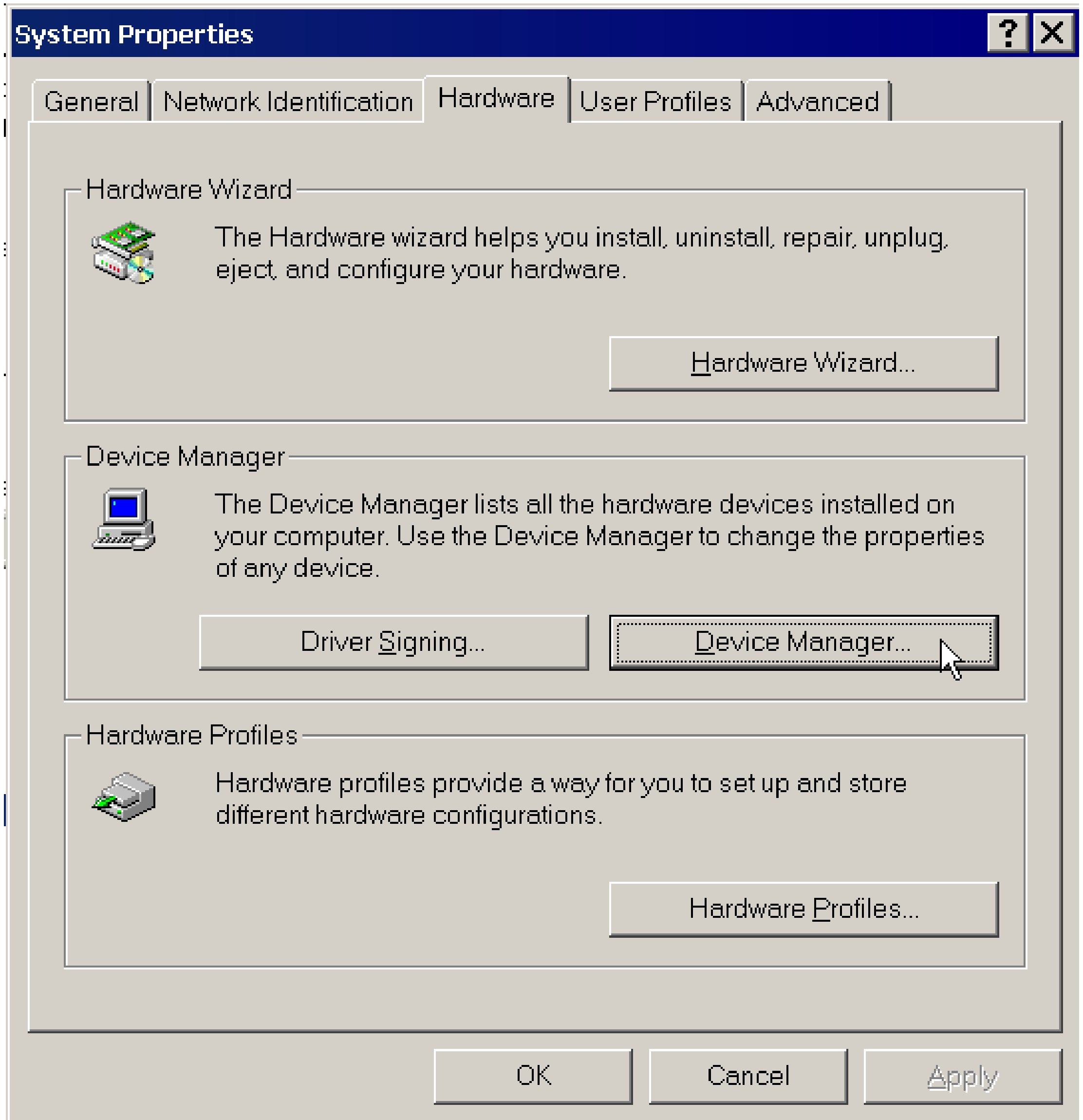
[top]

In 32 bit operating systems, you use 32 bit drivers instead of ROM BIOS. This means that software suppliers like Microsoft and IBM must be able to supply drivers to all conceivable hardware. The advantage is, that once the operating system has installed drivers, all user programs operate alike relative to hardware.

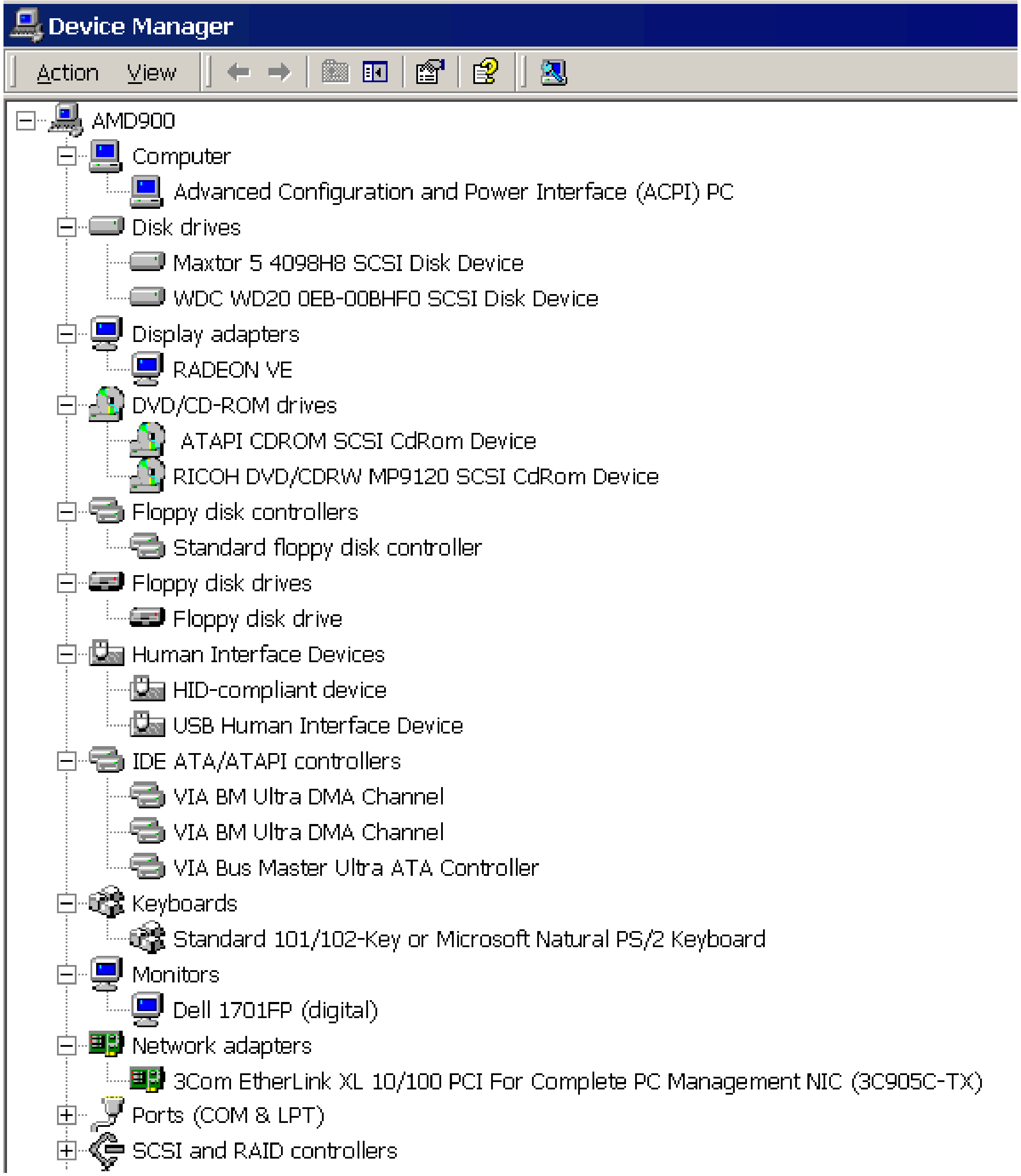
It is an enormous project to supply drivers. Especially OS/2 has suffered problems in getting the right drivers on the market. For many years, IBM for example did not supply OS/2 drivers for Canon printers. That was part of my reason to drop that operating system. Regarding driver installations, Windows 98 is unquestionably the best operating system.

Windows supports plug and play. The operating system searches the PC for hardware. Often all drivers (to CD-ROM, network controller, sound card, etc.) are installed automatically. The drivers can be seen under System in the control panel.

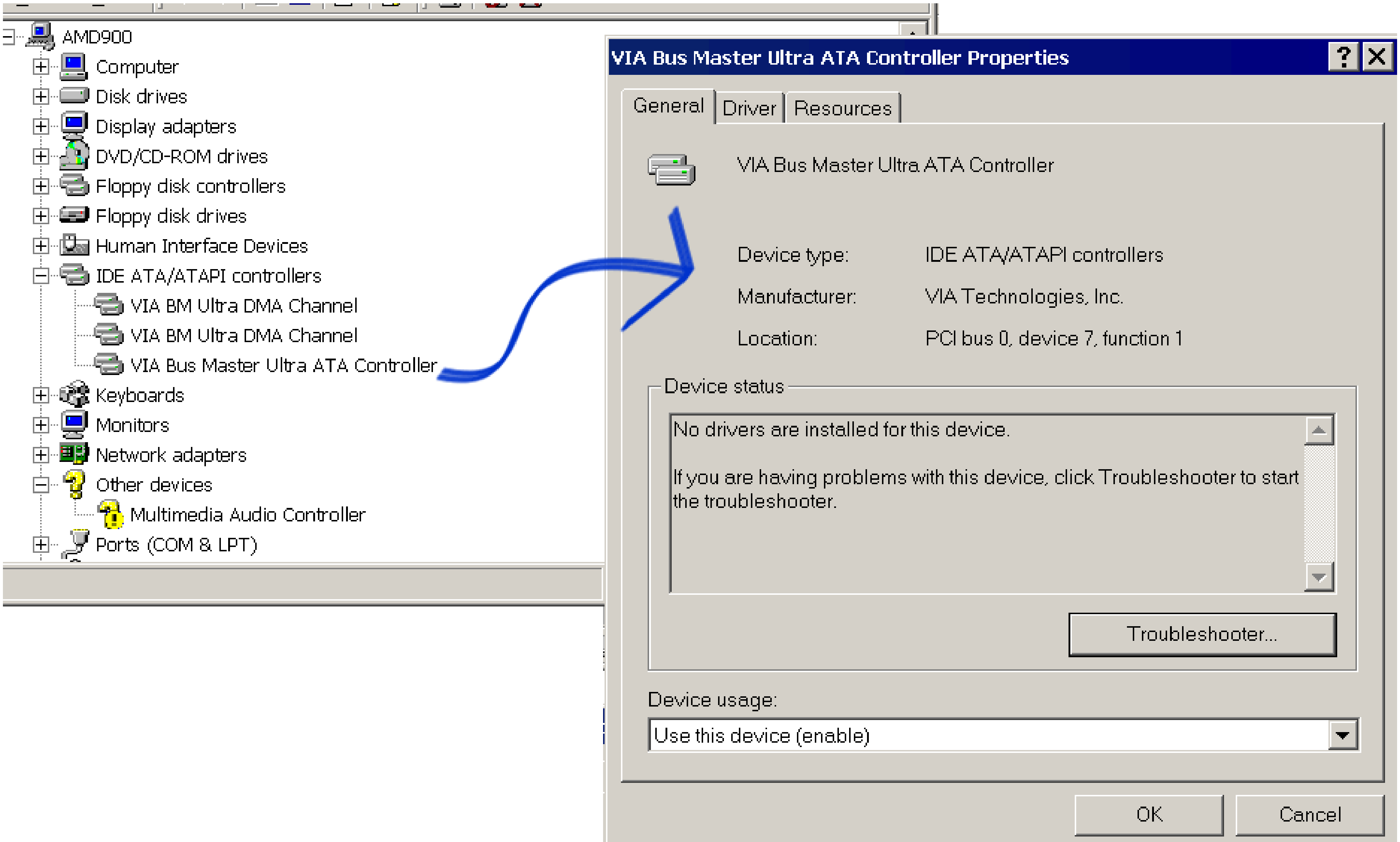
Let us look at a EIDE hard disk. The hard disk operation is regulated by an EIDE controller on the system board. Therefore, Windows must have a driver installed to this controller. We can find it easily. Go to: My computer -> Control panel -> System -> Computer. In Windows 2000 you should click on Device manager:



Then expand the entries to hardware units:



Actually, you can see a long list of drivers in the picture above. Windows has installed most of them during Windows installation. An VIA Bus Master IDE controller, which regulates the hard disk can be found:



I always have these drivers on the hard disk (in the folder C:\Disks\Drivers). That makes it easy to install them after an unforeseen but necessary re-installation of Windows .

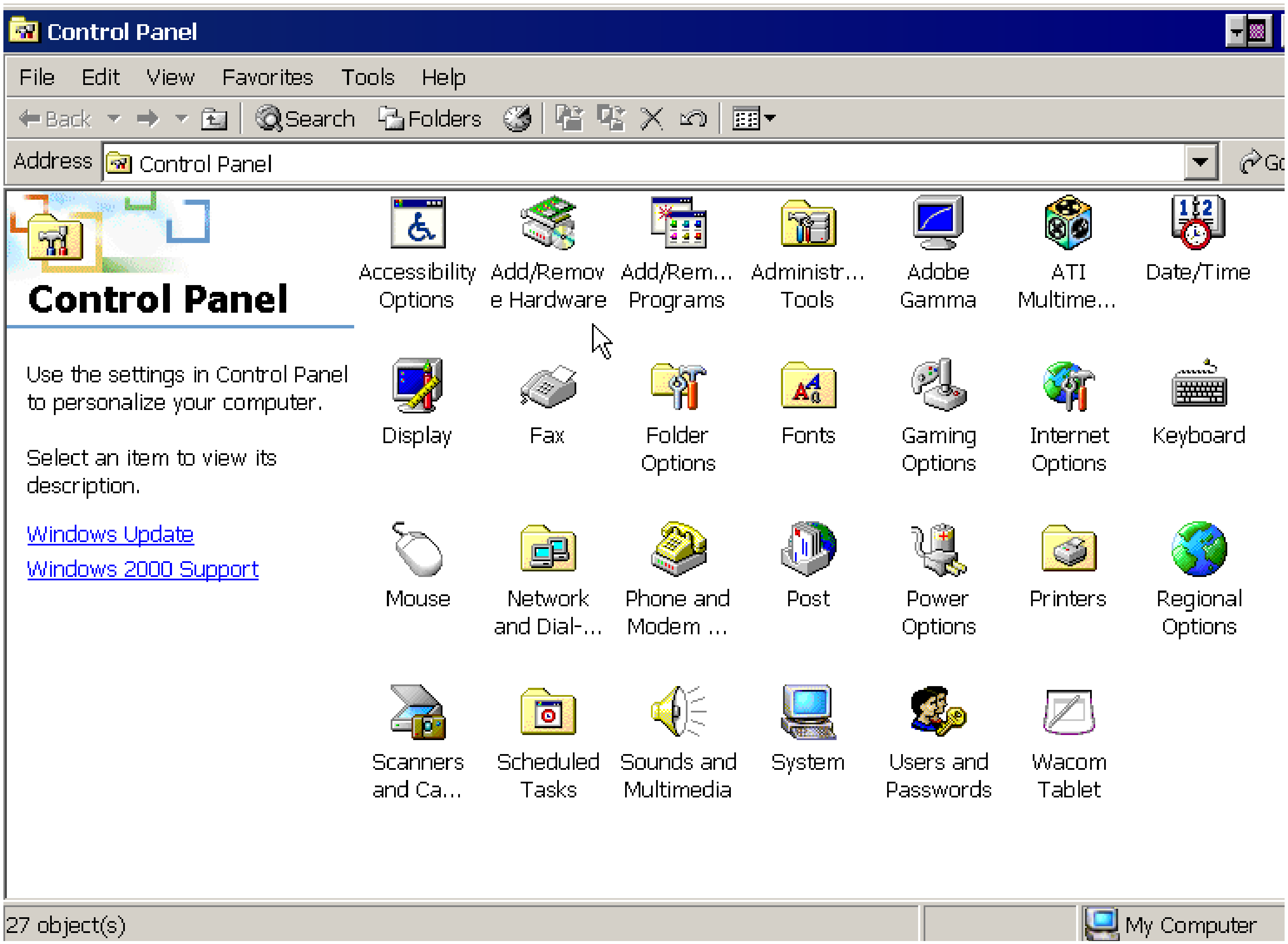
The quality of the drivers is very important. The drivers are extremely important for video cards. You often hear that a new driver has been developed for this video card and it improves performance by 40%. Then rush to download it (from the manufacturers Internet server) and install it.Don't forget to save it on disk for future use!

You also have to be cognizant about the system board chip set. Often Windows 98 installs a good standard driver, but new chipsets may contain facilities which require a new driver. That can be found on a disk, which comes with the system board or on the Internet.

Installation of new drivers

[top]

You install new drivers in Windows 98/Me/2000 with "add new hardware" found in My Computer -> Control panel:.



Don't let Windows search for hardware. Instead choose yourself. Then you have to select the particular hardware from the list and in the next screen click "Have diskette... " Learn this technique if you experiment with your PC and want maximal benefits from your hardware.

- Next page
- Previous page

Learn more [top]

Also see The Windows 98 page (module 6d)

Read of module 7a and module 7b about installation monitor and video card in Windows 95/98!

Read about chip sets on the motherboard in module 2d

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7

- About the video system
- About video cards
- About sound cards
- About digital sound and music.

The contents on this and following pages:

- Introduction
- Concepts and terminology
- Screen resolution, screen size etc.
- About colors, color depth, RGB etc.
- About refresh rates
- TCO standards
- LCD displays

- Next page
- Previous page

The video system (of which the monitor is a part) is one of the most important components in the PC. It affects directly your pleasure of working, and actually also your health. At the same time, the video system shows the biggest variation between different PCs. Read my coverage of this subject here in this module, which is subdivided into several pages.

Introduction

[top]

All computers are connected to some type of *display*, which usually is called the monitor. Monitors are available in many different types and sizes. The size generally goes from 12 to 21 inches diagonal.

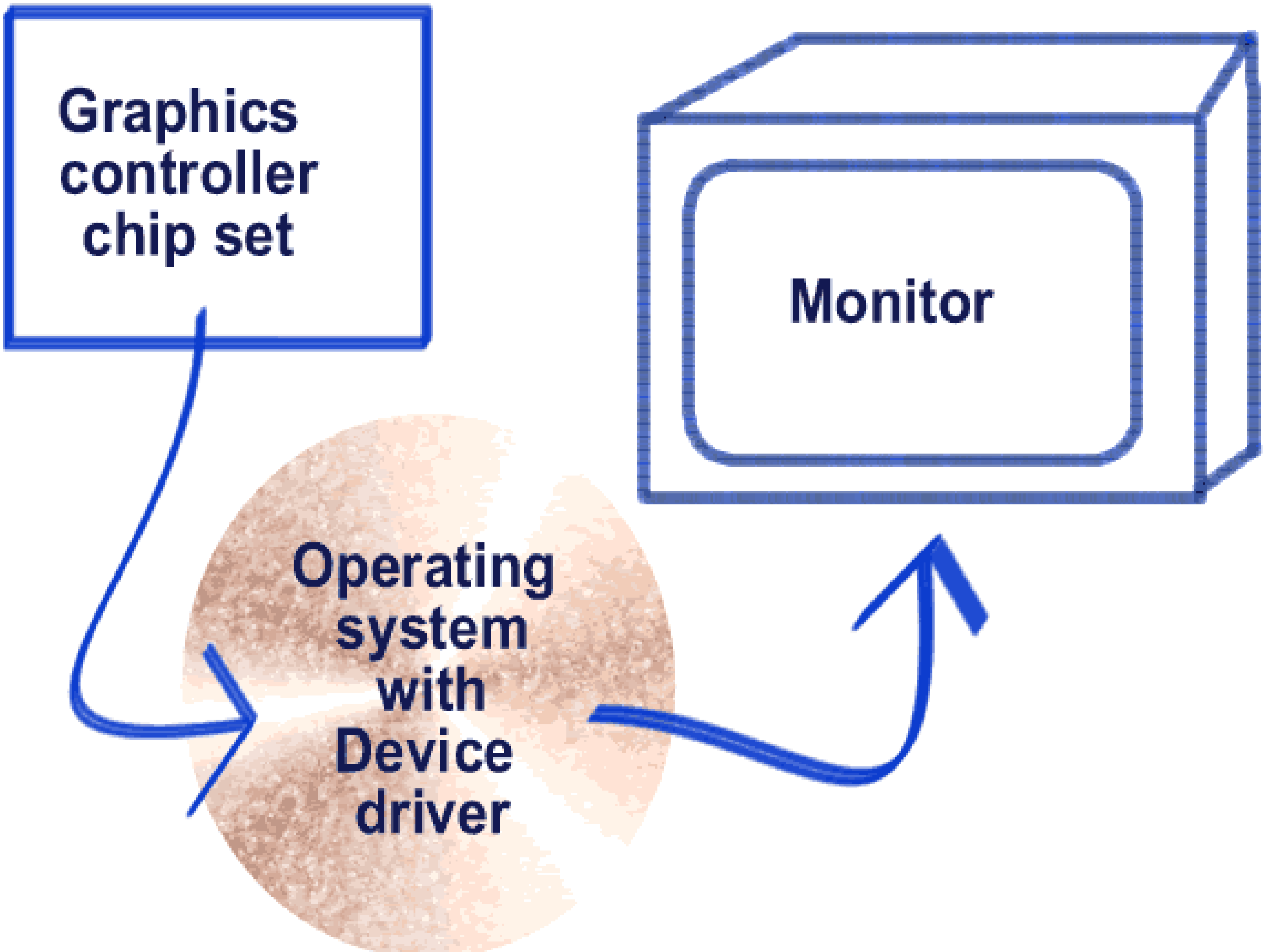
The monitor is a part of the computer video system. To understand how to obtain a good screen image, we need to look at the complete video system. It includes these three elements:

- The graphics card (also called the video card or video adapter). It is an expansion card, which generates electric signals to the monitor.

- The monitor itself, which is connected by a cable to the video card using some kind of *interface*.

- A *device driver* which Windows uses to control the video card, to make it send the correct signals to the monitor.

These three elements must be fitted and matched to achieve quality images. Even the finest and most expensive monitor will only render mediocre images if it is connected through a low quality video card. All video cards depend on the right driver and proper settings to function properly – otherwise the card will not perform well:



In these pages, I will review the complete video system. First you can read about the video image construction, pixels. resolution, and refresh rate. Those are very central subjects. Later, we will look at different monitor and video card types. Finally, we put it all together in Windows .

Fast development

The video system has developed as explosively as the rest of the PC since the 1980s. These improvements have occurred in different areas:

- **The monitors** – both the tubes and the electronics continue to improve, and the flat panel monitors has come along. The newer monitors render better images - sharper, with better resolution and better colors. Big plasma screens is a new and interesting technology.

- **The video cards** are getting faster. They can deliver better images, which the new monitors are capable of producing. The user gets more tuning options. New RAM types and buses will increase speed, and new features are added.

- **Video presentation, DVD, and 3D** games are other areas of development, which will change the video card standards.

The video system is a sub system in the PC, with its own technological development. At the same time, monitors and video cards are areas, where manufacturers and dealers often cut corners. As an ordinary user, you can improve your screen images significantly with careful planning. That holds true when you buy your PC - you must select your video system carefully.

It also holds true for existing video systems, where better drivers and software optimizing can help produce the optimal screen image. We will look at that in these pages.

- Next page
- Previous page

Learn more Top

Read about video cards in Module 7b .

Read about sound cards in Module 7c .

Read about digital sound and music in Module 7d .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7b.1

The video card

The contents:

- An introduction to the video card
- The video card supports the CPU
- About RAM on the video card
- RAMDAC or digital?
- Heavy data transport

- Next page
- Previous page

Three components in a videocard

The video card is just as important as the screen – and more often overlooked. During the years 1999-2001 the overall quality of video adapters have been improved. Earlier there was some very lousy products in the market. Follow my articles to know more of the video adapter!

A video card is typically an adapter, a removable expansion card in the PC. Thus, it can be replaced!

The video card can also be an integral part of the system board This is the case in certain brands of PCs and is always the case in lap tops. I have a clear preference for a replaceable video card in my stationary PC. However modern motherboard may include good integrated video chip sets. You just have to know which ones!

Regardless of whether it is replaceable or integrated, the video adapter consists of three components:

- A **video chip set** of some brand (ATI, Matrox, Nvidia, S3, Intel, to name some of the better known). The video chip creates the signals, which the screen must receive to form an image.

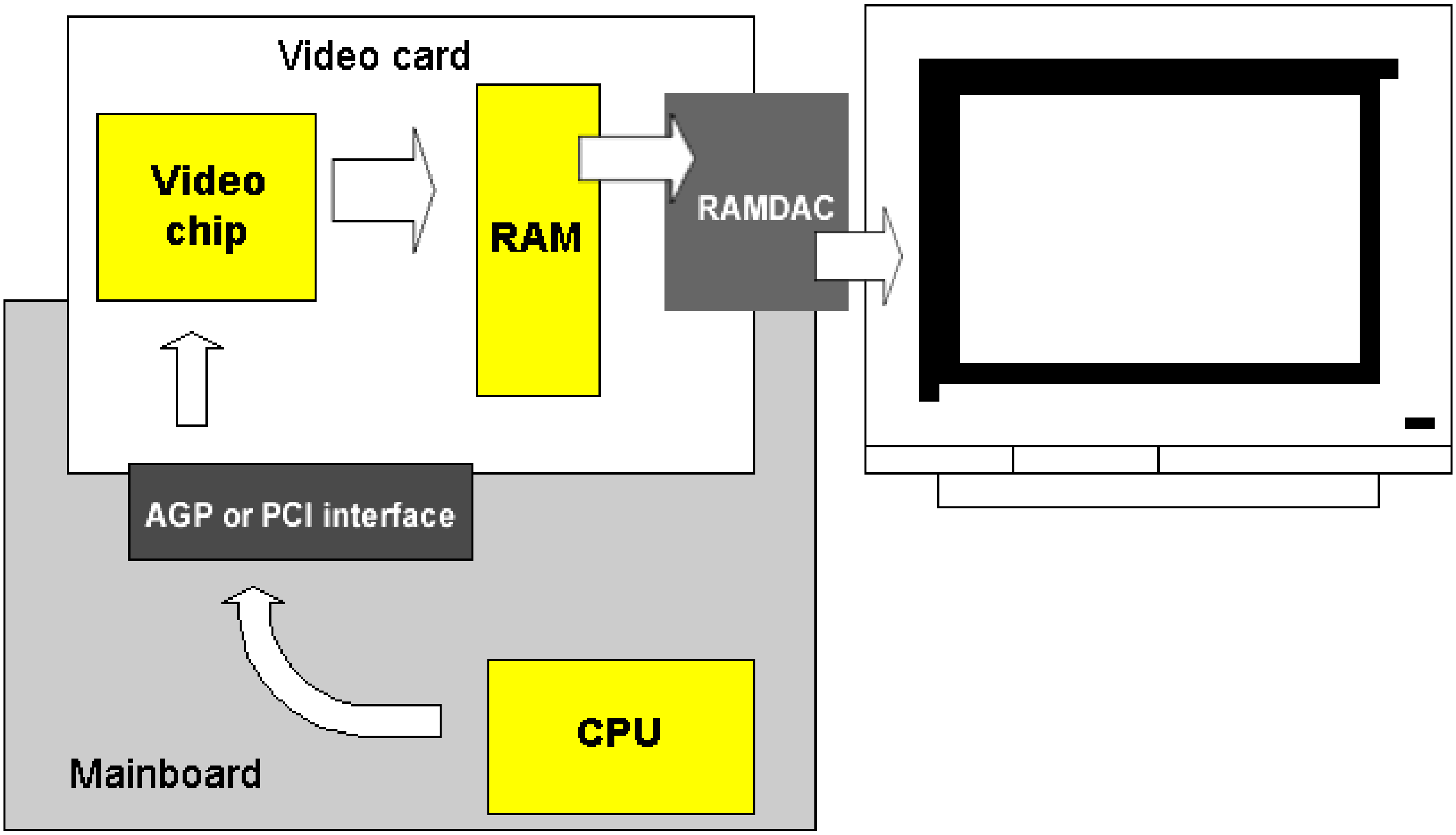
- Some kind of **RAM** (EDO, SGRAM, or VRAM, which are all variations of the regular RAM). Memory is necessary, since the video card must be able to remember a complete screen image at any time. Using AGP, the video card may use the main memory of the motherboard.

- A **RAMDAC** - a chip converting digital/analog signals. Using Flat panel monitors, you do not need a the function of a RAMDAC.

The video card supports the CPU

Top

The video card provides a support function for the CPU. It is a processor like the CPU. However it is especially designed to control screen images.



You could produce a PC without a video controlling chip and leave this work to the CPU. However, the CPU would be constantly occupied running the software that should generate screen images.

RAM on the video card

Top

Video cards always have a certain amount of RAM. This RAM is also called *the frame buffer*. Today video cards hold plenty of RAM, but before it was more important:

- How much RAM? That is significant for color depth at the highest resolutions.
- Which type RAM? This is significant for card *speed*.

Video card RAM is necessary to keep the entire screen image in memory. The CPU sends its data to the video card. The video processor forms a picture of the screen image and stores it in the frame buffer. This picture is a large bit map. It is used to continually update the screen image.

The amount of RAM

Older video cards were typically available with 1, 2, 4 or more MB RAM. How much is necessary? That depends primarily on how fine a resolution you want on your screen. For ordinary 2D use, 16 bit colors are "good enough." Let us look at RAM needs for different resolutions:

Resolution	Bit map size with 16 bit colors	Necessary RAM on the video card

640 x 480	614,400 bytes	1 MB
800 x 600	960,000 bytes	1.5 MB
1024 x 768	1,572,864 bytes	2 MB
1152 x 864	1,990,656 bytes	2.5 MB
1280 x 1024	2,621,440 bytes	3 MB
1600 x 1200	3,840,000 bytes	4 MB

Note that the video RAM is not utilized 100% for the bit map. Therefore, 1 MB is not enough to show a 800 x 600 picture with 16 bit colors, as the above calculation could lead you to believe.

Today video cards come with 4 MB, 8 MB or more RAM.

Using ordinary RAM, you saw speed improvements of the graphics card using 4 MB instead of 2 MB, if the resolution only was 800 x 600 or 1024 x 768. In this case data can be written to and read from the RAM simultaneously - using different RAM cells. With only 2 MB RAM, data sometime had to wait for a free cell.

3D - lots of RAM

Supporting the demand for high quality 3D performance many new cards come with a frame buffer of 16 or 32 MB RAM. And they use the AGP interface for better bandwidth and access to the main memory.

VRAM

Briefly, in principle all common RAM types can be used on the video card. Most cards use very fast editions of ordinary RAM (SDRAM or DDR).

Some high end cards (like Matrox Millennium II) earlier used speciel VRAM (Video RAM) chips. This was a RAM type, which only was used on video cards. In principle, a VRAM cell is made up of two ordinary RAM cells, which are "glued" together. Therefore. you use twice as much RAM than otherwise. VRAM also costs twice as much. The smart feature is, that the double cell allows the video processor to *simultaneously* read old and write new data on the same RAM address. Thus, VRAM has two gates which can be active at the same time. Therefore, it works significantly faster.

With VRAM you will not gain speed improvements increasing the amount of RAM on the graphics controller. VRAM is already capable of reading and writing simultaneously due to the dual port design.

UMA and DVMT

On some older motherboards the video controller was integrated. Using SMBA (*Shared Memory Buffer Architecture*) or UMA (*Unified Memory Architecture*) parts of the system RAM were allocated and used as frame buffer. But sharing the memory was very slow and the standards never became very popular.

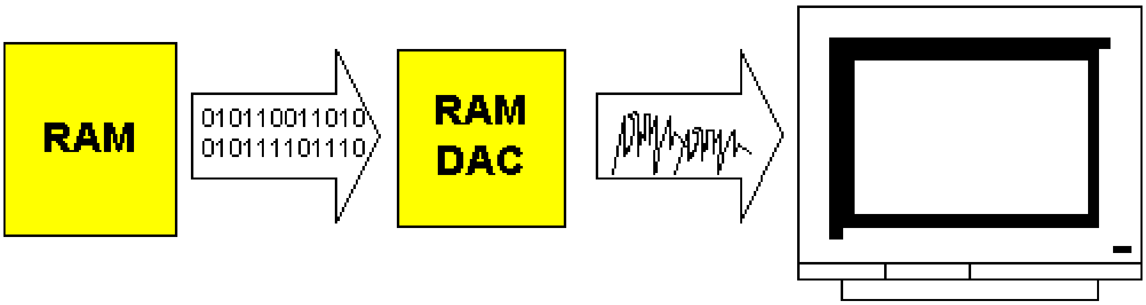
A newer version of this is found in Intel chip set 810 and the better 815, which also integrates the graphics controller and use parts of the system RAM as frame buffer. Here the system is called Dynamic Video Memory Technology (D.V.M.T.).

The RAMDAC

Top

All traditional graphics cards have a RAMDAC chip converting the signals from digital to analog form.

CRT monitors work on analog signals. The PC works with digitized data which are sent to the graphics adapter. Before these signals are sent to the monitor they have to be converted into analog output and this is processed in the RAMDAC:



The reccommandation on a good RAMDAC go like this:

- External chip, not integrated in the VGA chip
- Clock speed: 250 - 360 MHz.

Heavy data transport

The original VGA cards were said to be "flat." They were unintelligent. They received signals and data from the CPU and forwarded them to the screen, nothing else. The CPU had to make all necessary calculations to create the screen image.

As each screen image was a large bit map, the CPU had to move a lot of data from RAM to the video card for each new screen image.

The graphic interfaces, like Windows , gained popularity in the early nineties. That marked the end of the "flat" VGA cards. The PC became incredibly slow, when the CPU had to use all its energy to produce screen images. You can try to calculate the required amount of data.

A screen image in 1024 x 768 in 16 bit color is a 1.5 MB bit map. That is calculated as 1024 x 768 x 2 bytes. Each image change (with a refresh rate of 75 HZ there is 75 of them each second) requires the movement of 1.5 MB data. That zaps the PC energy, especially when we talk about games with continual image changes.

Furthermore, screen data have to be moved across the I/O bus. In the early nineties, we did not have the PCI and AGP buses, which can move large volumes of data. The transfer went through the ISA bus, which has a very limited width (read in module 2b about the buses). Additionally the CPUs were 386’s and early 486’s, which also had limited power.

- Next page
- Previous page

To learn more

Top

Read about monitors in Module 7a.

Read about sound cards in Module 7c .

Read about digital sound and music in Module 7d .

Read about FPU work in 3D graphics.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7b.2

The video card (continued)

The contents:

- About accelerator cards
- About video card and chips
- Card brands
- The video driver

- Next page
- Previous page

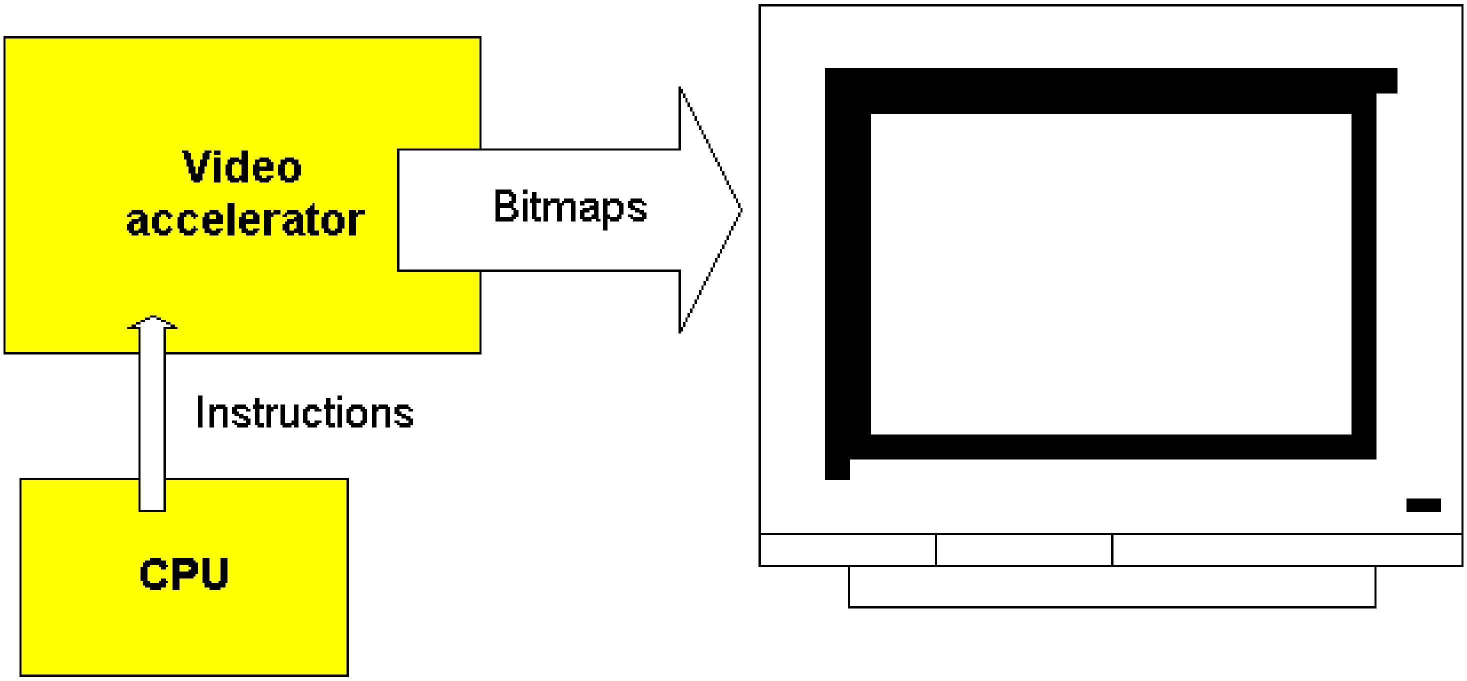
About the accelerator cards

Top

In the early nineties the *accelerator* video cards appeared. Today all cards are accelerated and they are connected to the CPU through high speed buses like PCI and AGP.

With accelerated video chips, Windows (and with that the CPU) need not calculate and design the entire bit map from image to image. The video card is programmed to draw lines, Windows , and other image elements.

The CPU can, in a brief code, transmit which image elements have changed since the last transmission. This saves the CPU a lot of work in creating screen images. The video chip set carries the heavy load:



All video cards are connected to the PCI or the AGP bus, this way providing maximum data transmission. The AGP bus is an expanded and improved version of the PCI bus - used for video cards only.

Modern video cards made for 3D gaming use expensive high-end RAM to secure a sufficient bandwidth. If you for example want to see a game in a resolution of 1280 x 1024 at 80 Hz, you may need to move 400 MB of data each second - that is quite a lot. The calculation goes like this:

1280 X 1024 pixels x 32 bit (color depth) x 80 = 419,430,400 bytes = 409,600 kilobytes = 400 megabytes.

Cards and chips

Top

There are many manufacturers of video cards and accelerator chips. Some produce both cards and chips, while others only make one or the other. I have tried a lot of different cards. The tables below illustrate my personal evaluation. Some may disagree with my evaluation.

First the best known video cards:

Vendor	Quality and price
ATI	High - Medium
Cirrus Logic	Low
Matrox	High - Medium
NVidia	High - Medium
S3	Medium

Here we see video card manufacturers:

Vendor	Quality and price
ATI	Medium - High
Diamond	Medium (High)
Matrox	High
Creative Labs	Medium (High)
Orchid	Medium (High)
STB	Medium - High
Britek/Viewtop	Low

You can use these tables, when you buy a PC and/or video card. Make sure to start with a quality video card!

The driver – almost the most important part

Top

The difference between good and mediocre cards is clearly visible in their software. The companies ATI, Matrox, and Creative Labs deliver excellent drivers with their cards. They allow their cards to provide optimum screen performance.

In contrast I can mention the ET 6000 accelerator chip, which was introduced in the mid 1990s. It had very fine specifications and scored very high in various tests. I bought a couple of cards with that chip, but I could never get them to work properly. The driver programs are poorly written, for example the refresh rate is not adjustable. Such cards are all right for low quality monitors, but not for monitors with high specifications. For these the refresh rate should be adjustable.

Here you see a section of a Matrox video card control box. The driver knows precisely which refresh rates the monitor will tolerate at different resolutions:

Resolution	Maximum Vertical Frequency
640x480	75 Hz
800x600	85 Hz
1024x768	85 Hz
1152x864	85 Hz
1280x1024	85 Hz
1600x1200	70 Hz

Another problem area is in the screen fonts, which come with the driver programs. Screen fonts are models for the letters seen on the screen. There are significant quality variations in this area. Again, ATI and Matrox are worth mentioning.

- Next page
- Previous page

To learn more

Top

Read about video basics in Module 7a.

Read about sound cards in Module 7c .

Read about digital sound and music in Module 7d .

Read about FPU work in 3D graphics.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7b.3

The video card (continued)

The contents:

- Video cards and chips, different brands
- Choosing the right video card is hard
- Spend your money well ..

- Next page
- Previous page

Video cards and chips, different brands

Let me comment on a couple of brands:

The Canadian firm ATI was among the first to produce accelerated video cards, when the graphic milieus came on board in the early nineties with Windows . One of the company's first chip was called *mach 8*. That had an 8 bit graphic processor, which was extremely fast relative to others at the time. But they were extremely expensive. I bought one!

Later, ATI presented *mach 32* and *mach 64*, which were 32 bit and 64 bit graphic processors, respectively. The next accelerator chip was called ATI Rage 128. It worked with a 128 bit bus connecting the onboard RAM and the processor and it holds a new type of cache as well. Later the RADEON chip was the top model. All the way, ATI has produced solid video cards with good quality drivers. Today they are available in many price ranges, including low cost editions. You will never go wrong with an ATI card (some gamers may not agree ...).

Matrox

The Matrox company is also Canadian, founded in 1976. They make excellent cards with their own accelerator chips. They only make a few models. Regardless of which Matrox card you buy, it is an excellent product. Matrox comes with good drivers. Obviously to be recommended. Matrox Mystique 220 and Millennium II as well as G400 were excellent cards I personally have been very satisfied with.

The Matrox Millennium G200 with its 128-bit DualBus delivers high end 2D, 3D and video performance. You find it in cards like Matrox Millennium and Marvel (including ports and software for Video editing). Great cards for office use!

The G400 (sixth generation of their graphics controllers) supports AGP 1X, 2X, and 4X. It has a dual 128 bit bus between the chip and the RAM allowing simultanously reading and writing. G400 has hardware-accelerated Environment Bump Mapping (!?!), which together with newest versions of DirectX should give a better depth perspective on screen. And G450 is even improved.

Tseng has made graphic chips for many years. In the good old DOS days, an ET 4000 card was one of the best on the market. It was equipped with Tseng's ET 4000 chip, which was excellent for DOS usage. Since then came the somewhat overlooked ET4000/W32 chip. I had good experiences with that on some low cost ViewTop cards. Tseng's latest chip was ET 6000, which mostly sold in the cheapest cards.

S3 was big name in graphic chips. They did not manufacture their own cards, but their chips are used in numerous cards. Companies like IBM, Diamond, Number Nine, and ViewTop/ Britek use S3’s different accelerator chips with widely varying results.

A small S3 Trio 64 chip was mounted in an old IBM’s PC 300. On paper, the controller is not very powerful. Yet, it produced a very fine image. Thus, the quality depends just as much on other video card design features as on the accelerator chip.

Choosing the right video card is hard

It is difficult to choose a video card, because there is such a multitude of different ones. And you can read *test reports* forever. Yet, they may not be particularly useful.

One of the best cards I ever laid my hands on was IBM's XGA graphics, which unfortunately was never available separately. Once they provided a working Windows driver, this card was in a class by itself. This was back in the early 1990s, where poor graphics was a big problem on PCs. Today most graphic controlling chip sets work fine.

One of the problems with video cards was that the same graphics chip may be used in both good and inferior video cards. This is especially true for S3 chips, which were used in fine cards from the vendor Number Nine and also was used in ViewTops discount cards. When a fundamentally good chip is found both in great and in mediocre cards, you cannot select the card based on what chip it uses!

Another problem was in the test methods, which the computer magazines use. They measure exclusively *speed* . Speeds are measured with special programs, which read how fast the screen image can be built, etc. That's fine – but it does not say much about image quality, as perceived by the eye. Is it sharp, bright, not flickering, comfortable? Those are more subjective and abstract qualities, which can never be evaluated by a test program.

You should choose a card based on its specifications. For example, can it deliver a 1024 x 768 image at 85 Hz? It should be able to do that, but not just in theory. It must also be able to do that in real life. Here is where the *driver* comes in.

Spend your money well

Top

When you read all these technical explanations, the *choosing* and *buying* remains. In my mind, the screen image is without doubt extremely important for your daily work. It is also an area with vast quality variations.

Often PC's have been advertised with a very cheap monitor and the very cheapest video card. Many would be happy with this equipment. The PC works fine and they may never have seen a high quality screen image.

I will strongly recommend that you invest a little more, to get a better video system. Specifically I would recommend a 15" or 17" LCD display or a 19" Trinitron monitor. It later could be a Mag or ViewSonic. Both make excellent Trinitron screens at reasonable prices. Combine the monitor with an ATI or maybe Matrox video card and you will have a good video system!

However, to make a good and lasting purchase, you have to understand your own video demands. Do you need:

- Office programs with 2D graphics?
- 3D games?
- DVD acceleration?
- DVI interface?

You should get a demonstration of the card and monitor you want to buy. Especially if it is in the low/medium price group, I will strongly recommend that you see it connected. Then evaluate the screen image. How sharp is it? Does it flicker? Ask about resolution, color depth and refresh rate. If the dealer cannot answer these questions, I would not trust him. Finally, find out which driver the card needs. Read on...

- Next page
- Previous page

To learn more

Top

Read about video basics in Module 7a

Read about sound cards in Module 7c

Read about digital sound and music in Module 7d

Read about FPU work in 3D graphics

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7b.4

The video card (continued)

The contents:

- 3D graphics
- 3D accelerated graphics cards
- Intel Combination 2D/3D graphic chip sets
- The fusion of graphics controller and motherboard chip sets?

- Next page
- Previous page

3D graphics

Top

3D images, where you can move around in space, is a technology, which is expanding to the PC world. Ordinary PC's today are so powerful, that they can actually work with 3D environments. Ordinarily, our screen images (such as in Windows 95/98) are two-dimensional. But we know 3D effects from movies and from some computer games.

Over the last few years, more and more 3D standards have arrived in the PC market. That includes:

VRML

A "language" which provides 3D space on the Internet.

DirectX 3D API

A Microsoft API which enables programming of 3D games for Windows 95/98. This API is especially used for flight simulators and racing games.

OpenGL

This is another API for 3D developers. OpenGL is needed for 3D-action games like Quake, Unreal etc.

For games, like Quake and others, it is extremely important that the accelerator card has drivers for both DirectX and OpenGL.

No office use

The 3D technologies are of no consequence for ordinary office programs. Actually, the ordinary video cards are optimized to show 2D images. 2D cards can construct 3D movements, but it will take time to bring the images to the screen. That is because of the very complex calculations needed. Therefore, hardware accelerators have been designed. They can give drastic improvements. Also, special functions are included in the video card chip, allowing it to calculate 3D movements lightning fast.

But many 2D video controllers, like Matrox G200, G400, and G450, Intel i752, S3 Savage3D and ATI Radeon, also have 3D accelerators built in - some of them very powerfull.

3D accelerated graphics cards

Top

All ordinary graphics cards can show 3D games. That is really no special trick. The problem is to present them smoothly and fast. If the PC's video card is made for 2D execution only, the CPU must do the entire workload of geometric transformations etc.! And that task can cause even the fastest CPU to walk with a limp.

When we talk about accelerating 3D programs, we are primarily talking about games. The three-dimensional games like Forsaken, Battlezone and Quake are very demanding for the PC to execute. The users want to see the games with good details in a large screen window and with as many FPS (*Frames Per Second*) as possible. See a description of the geometric transformations which occur in a 3D environment.

In recent years there has been an enormous development in 3D graphics cards. Let me briefly describe those here. Originally there were two types of graphics cards, which could be used for 3D acceleration:

- Combination 2D/3D cards. These are ordinary graphics cards, which have been equipped with extra 3D power.

- The pure 3D cards, which only work as accelerators.

The pure 3D cards required that there also is an ordinary (2D) graphics card in the PC. In beginning the pure 3D card yielded the best acceleration, but there soon came good combination cards into the market.

The real 3D card

The 3Dfx company has set the *standard* for 3D execution with their Voodoo accelerator chips. The first version was launched in 1997 and it set the standard for 3D acceleration.

The Voodoo^2 accelerator chip came in 1998 at also became an enormous success. This card cannot display 2D-images, so it needed to be installed in combination with an ordinary graphics card.

The Voodoo^2 cards are special in that they do not use AGP. Many think that this is a big flaw in the architecture. At least this means that you cannot show 32 bit color depth (which is 24 bit colors with 8 additional bits for transparency). However, the difference between 24 bit and 32 bit colors is not always visible - according to experts ... Bundling two parallel Voodoo^2 cards, each with 12 MB RAM, working in tandem (called SLI mode) in the same PC, yielded lots of power. That gave 3D games in a resolution of 1024X768 in 60 FPS.

The Voodoo^2 controller operated at 95 MHz and was produced using 0.35 micron technology. It was a chip set of three controllers (one pixel processor and two texture processors).

The Voodoo^3

The later Voodoo^3 controller operated at up to 183 MHz and was produced using 0.25 micron technology. It only supported AGP 2X, which gave a great disappointment within the press.

The Voodoo3 RAMDAC operates at up to 350 MHz.

Intel Combination 2D/3D graphic chip sets

Top

To relieve the CPU, you can add certain 3D accelerator characteristics to the chip on the graphics card. Many companies have done that in recent years; of them Intel are intersting.

Intel's i740

Back in 1997 Intel teamed up with a company called Real 3D to produce a new graphics chip. The i740 chip was constructed to give maximum performance within two demanding areas:

- 3D scenes
- Video playback

The processor allows parallel data processing and gives precise pixel interpolation. Using AGP, an i740 based controller will be able to process very large amounts of data at a high speed.

The i740 board works as a "normal" video card as well as 3D accelerator. However, it never became very popular and today it is outdated.

i752

The chip i752 (code named "Portola") is the new generation 2D/3D graphics controller from Intel. It should be 5 times better than the i740. It includes 2D graphics, 3D rendering and digital video acceleration. The i752 also features AGP 2X.

According to Intel the power of the 752 set is found in:

The 3D visual quality being enhanced using Intel’s new HyperPipelined 3D architecture. The Pixel Precise Engine includes new features as a 16 tap anisotropic filter, emboss bump mapping, texture compression, and texture compositing.

Enhanced digital video streams from a wide variety of input sources: VCR, camcorder, TV tuner, MPEG-2, and Web video streams. Software DVD is accelerated through a high-precision hardware-based motion compensation algorithm.

A 128-bit 2D engine and support for high-resolution flat panels.

You find the 752 graphics controller integrated in the Intel chip set 810.

i754

The chip i754 (code named "Portola") is the next generation of this Graphics chipset, featuring AGP4X.

A later high-end version is code named Capitola.

The fusion of graphics controller and motherboard chip sets?

Top

With the more intense focus on 3D graphics performance and development, it has come to many changes in the industry the recent years.

- In the summer 1999, chip manufacture S3 fusioned with card and multimedia manufacture Diamond.

- The chip manufacture 3dfx fusioned with the card manufacture STB Systems. Hence no other company will have the Voodoo chips for card production.

There seems to a trend in all this. The traditional business of graphics chips have to re-arrange to survive. Many vendors want to integrate the graphics controller with other processors on the motherboard.

Perhaps it all started, when Intel made their first own graphics controller, the ill-faited i740. Here Intel made an attempt to move into a new productional area, and this must have caused some worry among the producers of graphics controllers. On the other hand, integration of a graphics controllers with the CPU or the chip set is a logical thought. This disadvantege is the reduced flexibility; you cannot choose the graphics controller yourself since it is a part of the motherboard. The advantage is of course the reduction in price.

We have seen these actions supporting the trend:

- Intel's focus on the chip sets i810 and i815 which include graphics engines.

- In June 1999 Chip set producer SiS came with the 630 chip set for Celerons and the similar

SiS540 for AMD K6-III. These sets hold 2D/3D graphics engines as well as a lot of other features - all in just one chip.

- On April 11, 2000 the company S3 announced that they had agreed to transfer its PC graphics chip business to the successfull Taiwanese chip set producer VIA

- ATI's has also announced, that they are devoloping full chip sets integrated with their own graphics controllers. We have not seen them. ATI and Matrox seems to survive in the market by delivering high-end graphic chip sets without getting involved with the motherboard manufactures. NVidia lives them same way serving the market for game PCs.

- Next page
- Previous page

To learn more

Top

Read about video basics in Module 7a

Read about sound cards in Module 7c

Read about digital sound and music in Module 7d

Read about FPU work in 3D graphics

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7b.5

The video card (continued)

The contents:

- About the Windows driver
- About DDC
- About Quickres (a smart utility)

- Next page
- Previous page

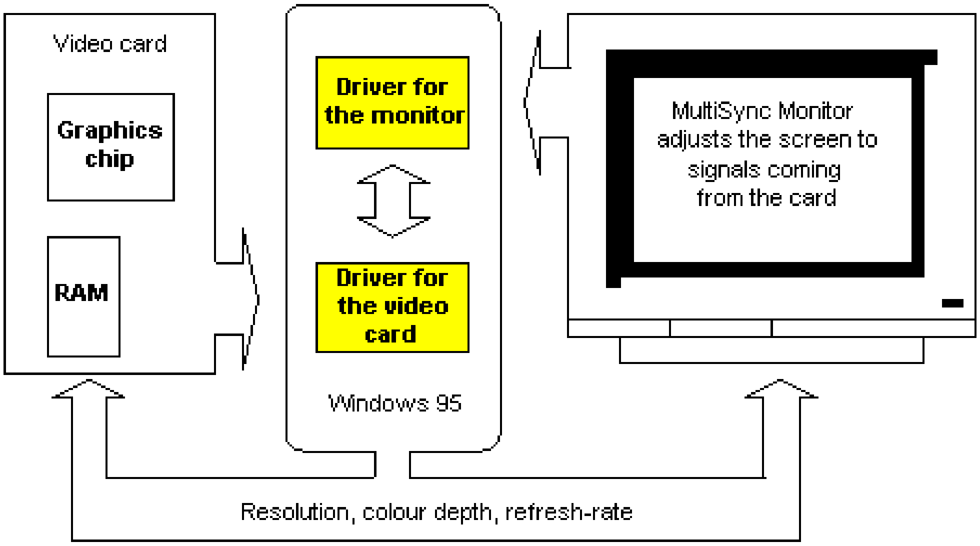
The screen image and Windows

Top

Once you buy *both* a good monitor *and* video card, you have to make them work together. That is done in Windows through driver programs. That part of the installation is extremely important, it requires attention.

If you leave it to Windows to install the necessary drivers, the result may be mediocre. Windows is so smart, so along the road it will find your hardware. And Windows will install drivers, when it encounters new software. Often some *standard drivers are installed* . They will make the software work, but no more.

The Windows drivers link video card and monitor together, and make them cooperate with each other.

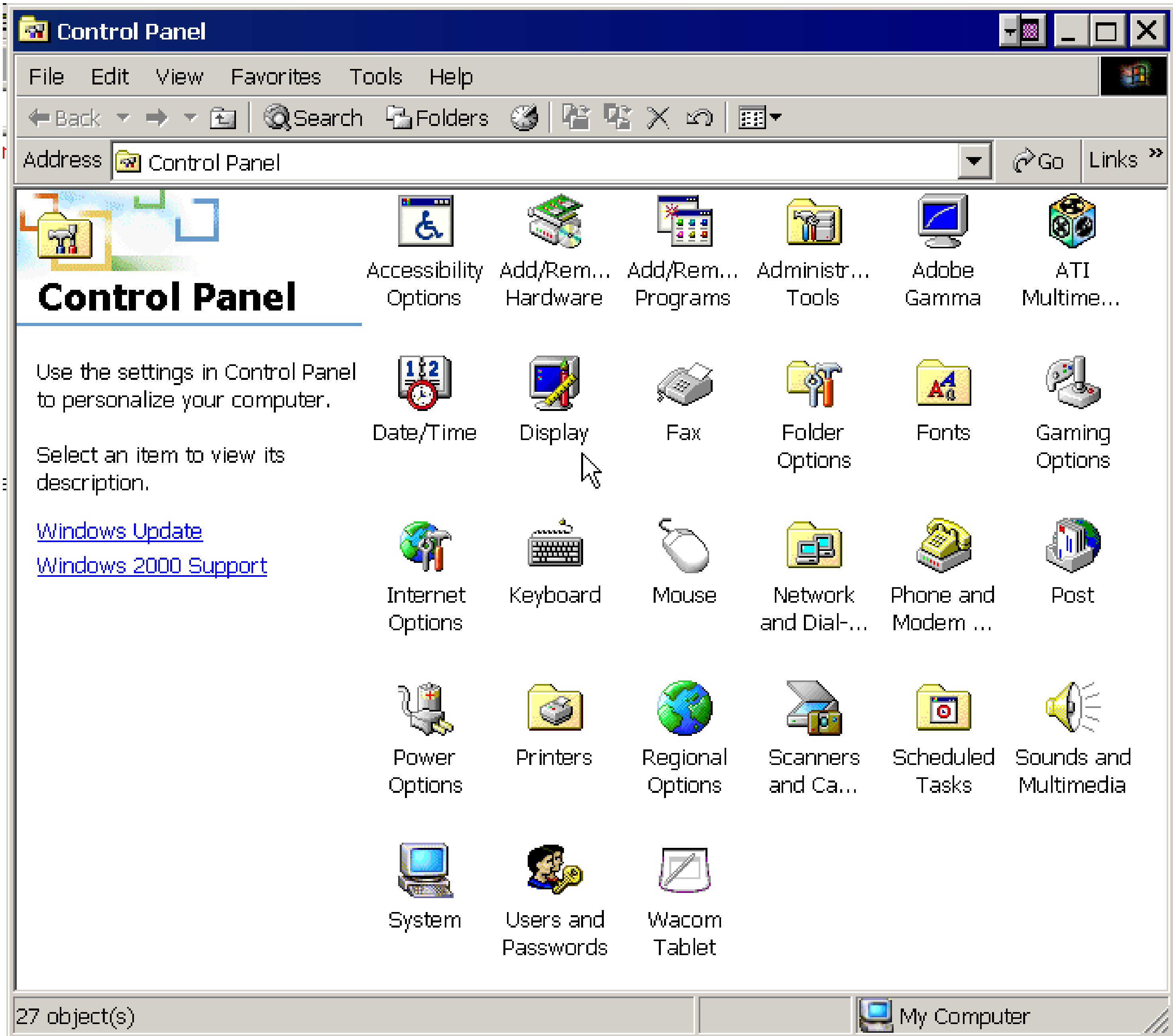


Checking the drivers

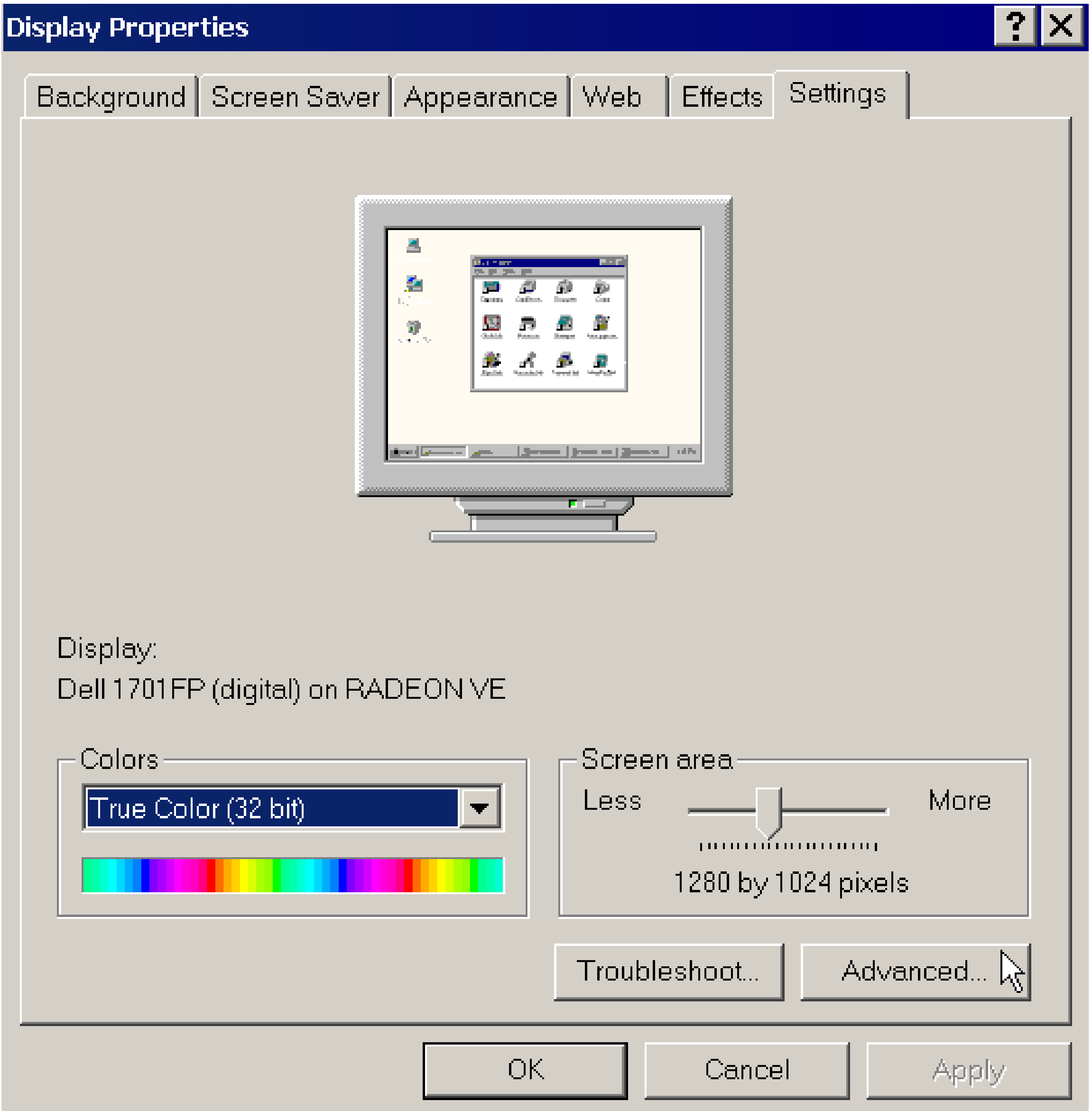
If you care about your screen image, you must make sure you have the correct drivers installed. We are talking about two drivers:

- Driver for the video card (the most important).
- Driver for the monitor (less important).

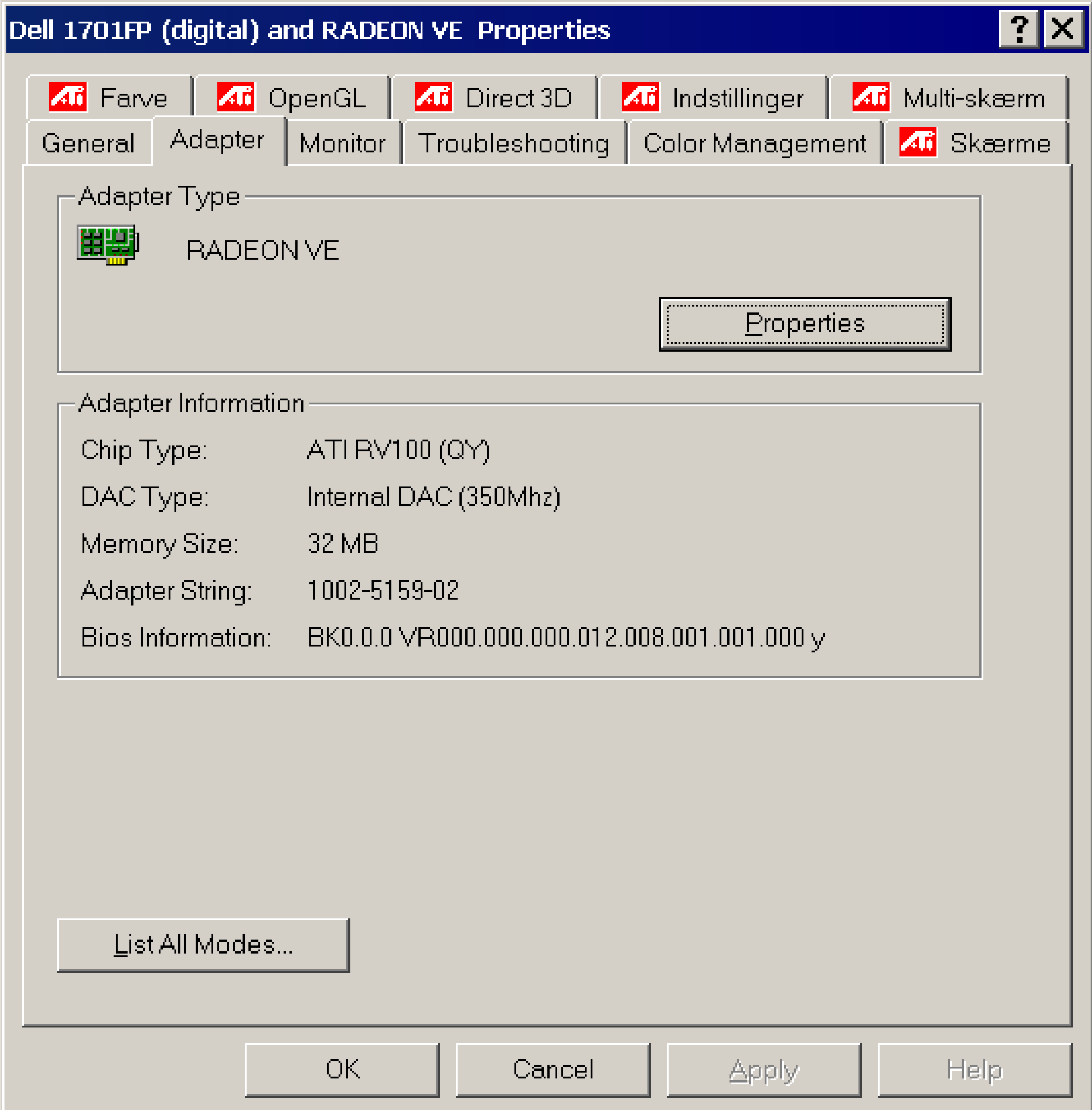
Both can be found in Windows, in My Computer -> Control Panel -> Display:



You should choose Settings and click on Advanced:



It opens a box, which the graphics driver may alter. Here you see my current settings. The ATI Radeon controller has istalled its own settings. The first five tabs are standard ones, but ATI has added six new tabs:

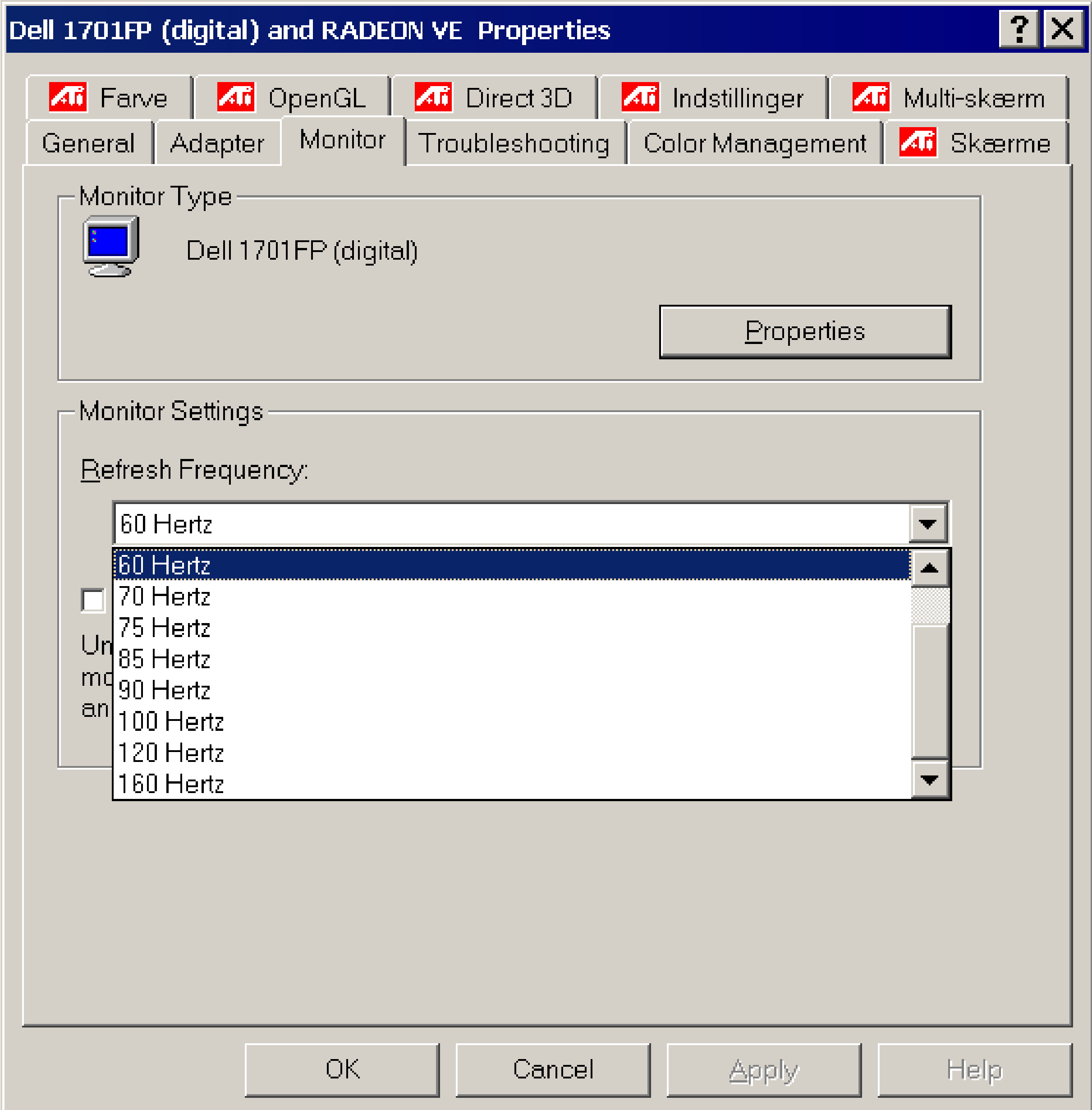


If you check the tabs Adapter and Monitor, you should find the names of your hardware and find the optimal drivers are installed. This allows Windows to get a full picture of the video system. Then the video card can deliver the optimal signals to the screen.

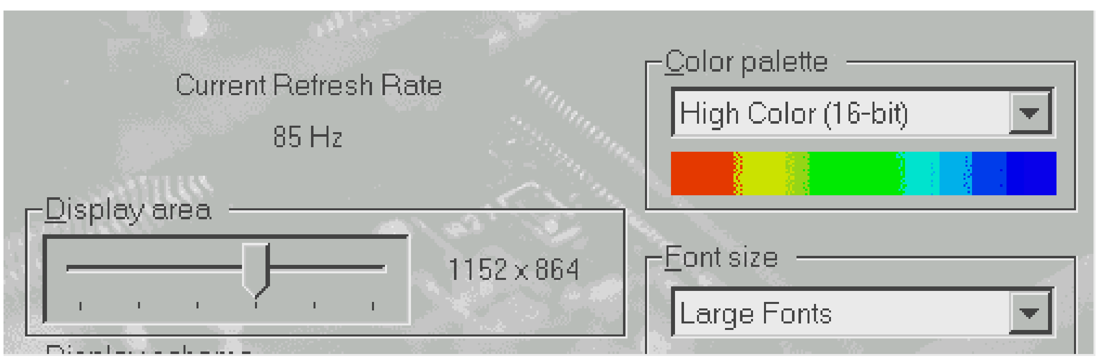
I once destroyed a 17" monitor by changing the video card. I adjusted the new card to deliver precisely the maximum ability of the monitor - according to the specifications. However, the monitor was a few years old. It had always run at a lower resolution and refresh rate, to which it must have adjusted itself. It did not work out - the electronics burned out!

Adjusting the refresh rate

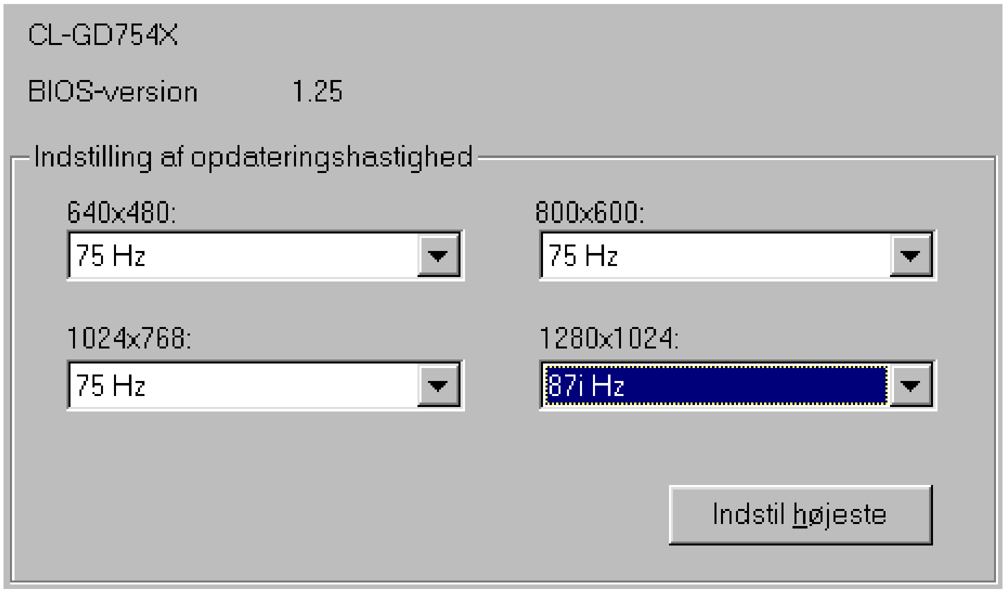
The standard driver in Windows often cannot adjust refresh rates. You can check it on the tab Monitor:



Often, a new driver has to be installed to exercise this option. Here is a Matrox Millennium II video card, with its own dialog boxes installed:



And here are the settings from my notebook, which has an adjustable Cirrus Logic video chip:



You *need* to install a driver program, which works specifically with your video card. Otherwise, you are guaranteed not to utilize your video card efficiently. Very few dealers seem to understand this concept. Nearly all PCs are sold with Windows standard driver installed and the video system will render absolute minimum performance!

DDC

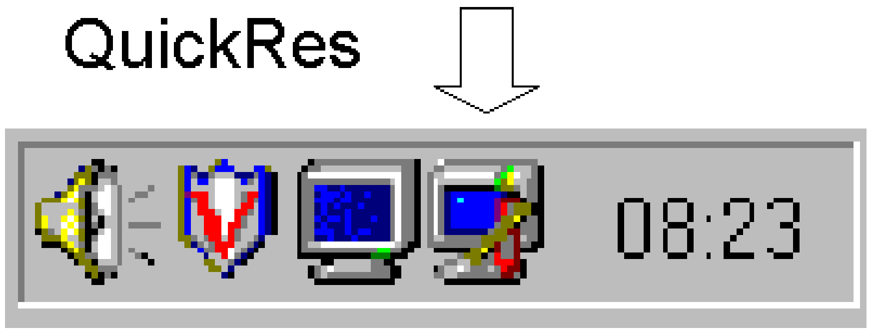
Top

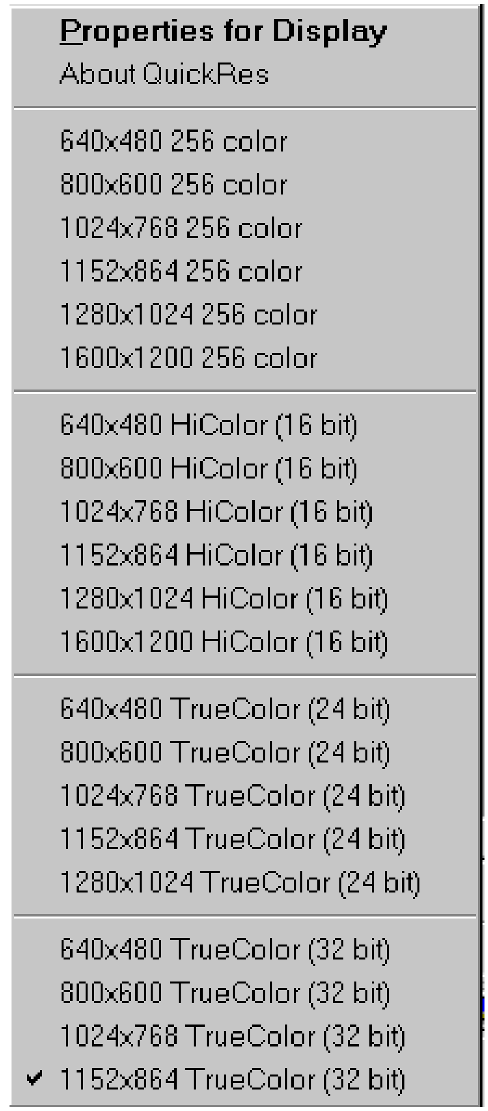
VESA DDC (Display Data Channel) are technologies, which should allow the video system to find the optimum adjustment, through communications with the video card. I do not think it is quite working yet.

QuickRes

If you want to experiment with different screen resolutions, you can install this program: QuickRes.exe. It is a small Windows utility application (only found in some versions of Windows), which you then have to run (double click on it).

Then, the program will appear as a small icon in the lower right corner of your screen (in the Systray):





When you right click on the icon, a menu will open showing all the resolutions and color depths you can choose from on your PC:

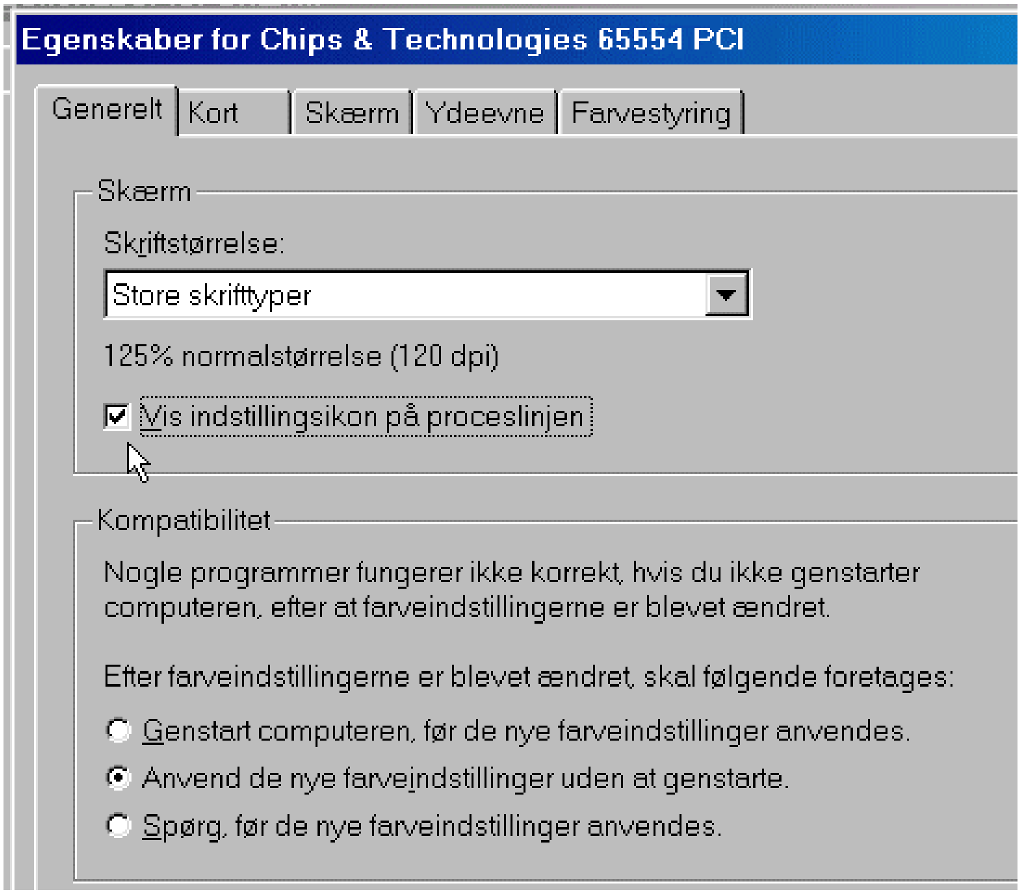
Here you can see, that the resolutions on my PC go from 640 x 480 up to 1600 x 1200. Color depth goes from 8 bit (256 colors) to 32 bit.

Note that the maximum color depth at 1600 x 1200 resolution is 16 bit and at 1280 x 1024 it is 24 bit.

Only the 1152 x 864 resolution can be seen in full 32 bit colors. This limitation is because this video card only has 4 MB RAM installed.

QuickRes is smart, because you can change resolution "on the fly". Normally Windows has to be re started, but here, the screen image just blinks a couple of seconds, then the new resolution is in place.

In some versions of Windows, you may install Quickres from the screen properties box, Settins, Advanced. Choose as here (Danish dump - I am sorry):



After clicking OK the little icon is in the Systray.

- Next page
- Previous page

To learn more

Top

Read about video basics in Module 7a.

Read about digital sound and music in Module 7d .

Read about FPU work in 3D graphics.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7c1.

PC sound

The contents:

- An introduction
- The synthesizer
- The A/D conversion
- Sampling the Wav files

- Next page
- Previous page

I do not claim to be an expert in sound cards. But I will try to describe what little I know about this technology. The sound capabilities of the PC are quite interesting. In the late 1990s, new and radical designs in sound technology appeared, and at the same time the MP3 wave swept the Internet society.

On the first pages we shall describe the traditional sound card concept (the Sound Blaster compatible sound card). Then follows something on the newer technologies.

Introduction

[top]

Sound cards have a minimum of four tasks. They function as:

- Synthesizer
- MIDI interface
- Analog-to-digital conversion during the recording (A/D).
- Digital-to-analog conversion during the playback (D/A). All elements are to be explained on these pages, so please read on.

The synthesizer

The synthesizer delivers the sound. That is, the sound card generates the sounds. Here we have three systems:

- FM synthesis, Frequency Modulation
- Wave table
- Physical modeling

FM synthesis

The cheapest sounds card use the FM technology to generate sounds simulating various instruments. Those are true synthesizers. The sounds are synthetic – it may sound like a piano, but it is not. FM synthesis is and sounds like the artificial sounds it consists of.

Wave tables - sampling

Wave table is the best and most expensive sound technology. This means that the sounds on the sound card are recorded from real instruments. You record, for example, from a real piano and make a small *sample* based on the recording. This sample is stored on the sound card.

When the music has to be played, you are actually listening to these samples. When they are of good quality, the sound card can produce very impressive sounds, where the "piano" sounds like a piano. Wave table is used in Sound Blaster's AWE card.

Physical modeling

Physical modeling synthesis is a third sound producing technology. It involves simulating sounds through programming. The process is supposed to be rather cumbersome, but it should yield a number of other advantages. The original Sound Blaster Gold card contains 14 instrument sounds, which are created from physical models.

Testing the sound

The basic quality of a sound card can be tested by playing a MIDI file. Then you can easily hear the difference. There is also a difference in how many notes (polyphony) can be played simultaneously. If you want to compose your own music on your PC, you use the sounds available on your sound card. The greater works you want to write, the more "voices" you will need. The SB AWE64 card has 64 voices, while SB16 only has 20 voices.

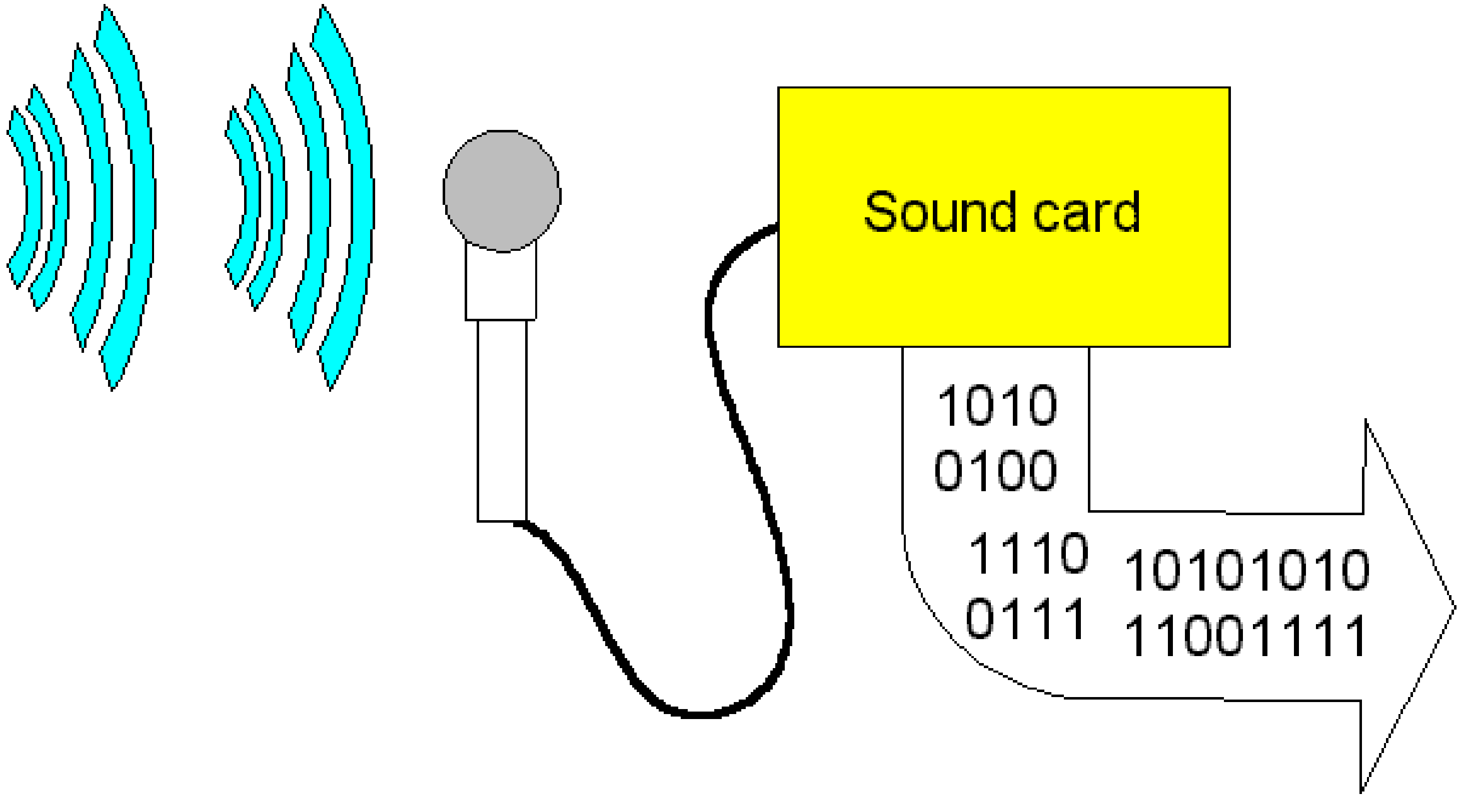
Some sound cards can import new sounds. They are simply downloaded to the sound card, which might have 512 KB (Sound Blaster AWE64) or 4 MB RAM (Sound Blaster AWE64 Gold) available for the user's own sounds.

The A/D conversion

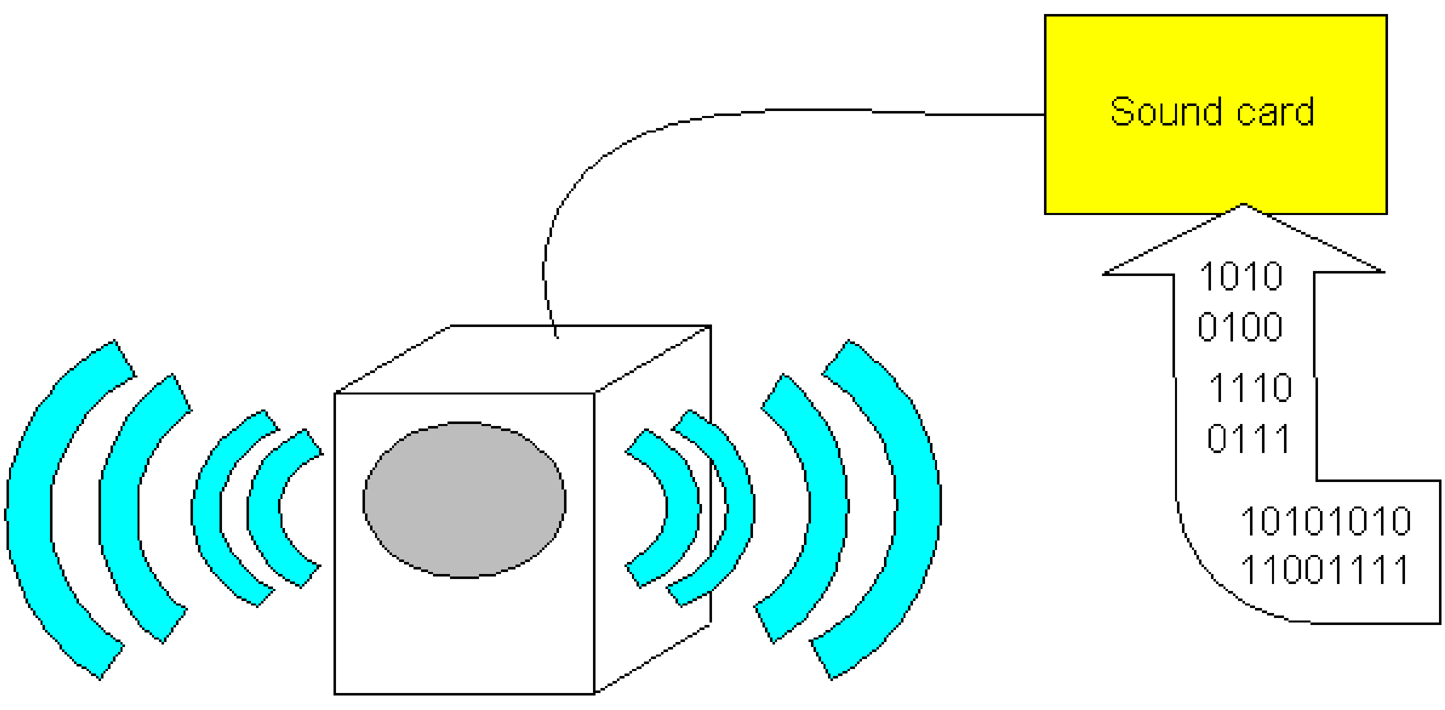
[top]

You need a A/D conversion, when analog sound signals are recorded, i.e. from a microphone. The other way around, the D/A-converter is used when the digital sounds have to be reproduced into a signal for the speakers amplifier.

The acoustic waves are collected by the microphone and lead to the sound card. Here it is converted into series of digital pulses, which eventually are saved in a file. This way a sampling is an analog-to-digital conversion:



During the playback, the bit stream from the sample file is converted to analog signals, which end in the loudspeaker.



When you connect a microphone to the sound card, you can easily record your own voice on the PC. The result is a small WAVE file which holds a digital recording of the sound, which reached the microphone.

The sound of your voice is analog, but the resulting the file is digital. The transformation from analog signals to digital data is done in the A/D converter of the sound card.

About sampling

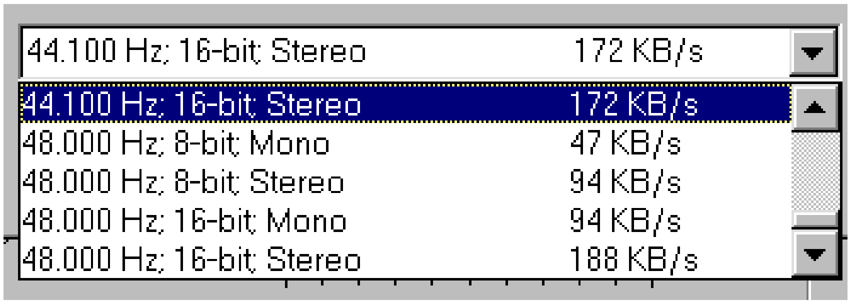
As mentioned is the basic concept of digital recording of sound is called sampling. You can record any sound you want into a sample (a Wav file) if you have a sound card and a microphone. The sampling can be done in various qualities:

- 8 bit or 16 bit sampling
- 11, 22 or 44 KHz (kilohertz)
- Stereo or mono

The number of kilohertz tells how many thousand times per second the sound will be recorded.

The quality of the sample

A sample is like a tape recording - it can be good or less good. The recording range from low end, as recorded on the cheapest cassette recorder, to hi-fi recordings in CD quality. Here is an image from my setup, where I can choose between different qualities for recording:



You record by sampling many times per second. The more frequently it is done, the better quality we get. The best would be infinitely sampling, which is not possible.

To record audio CDs the sampling is executed 44,100 times per second. This we call a 44.1 KHz sampling.

The quality is measured in kilohertz (KHz) and resolution (bit width) as you see above. The higher the KHz is, the better becomes the quality of the WAV file, but it also becomes bigger in file size. 8 or 16 bit sampling refers to how much data we spent on each sample. 16 bit gives a good quality.

File sizes

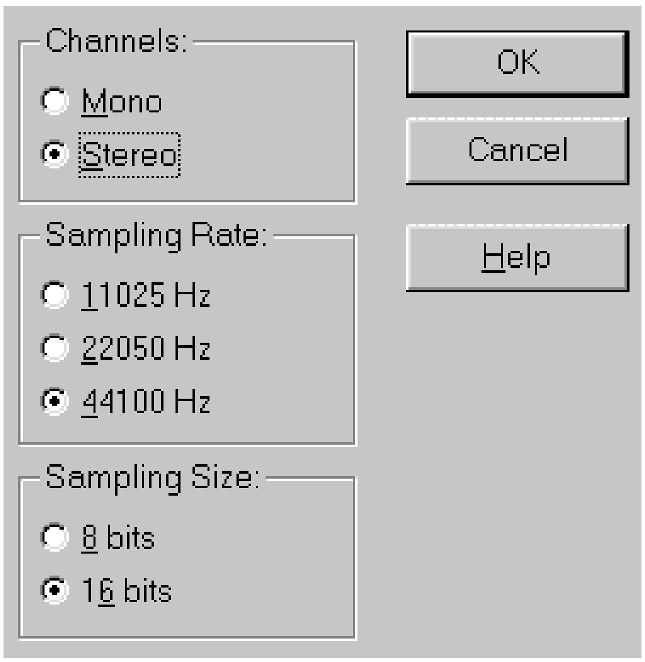
Using 2 channels stereo and 16 bit sampling at 44.1 KHz, the bit stream can be calculated as this:

2 channels X 16 bit X 44,100 samples per second = 176,400 bytes per second (since 8 bits make one byte).

This gives us the following file sizes of sampled stereo music in CD quality:

Replay	Number of bytes
One minute	10 MB
One hour	605 MB
74 Minutes	746 MB

Here you see the settings in a Wave program:

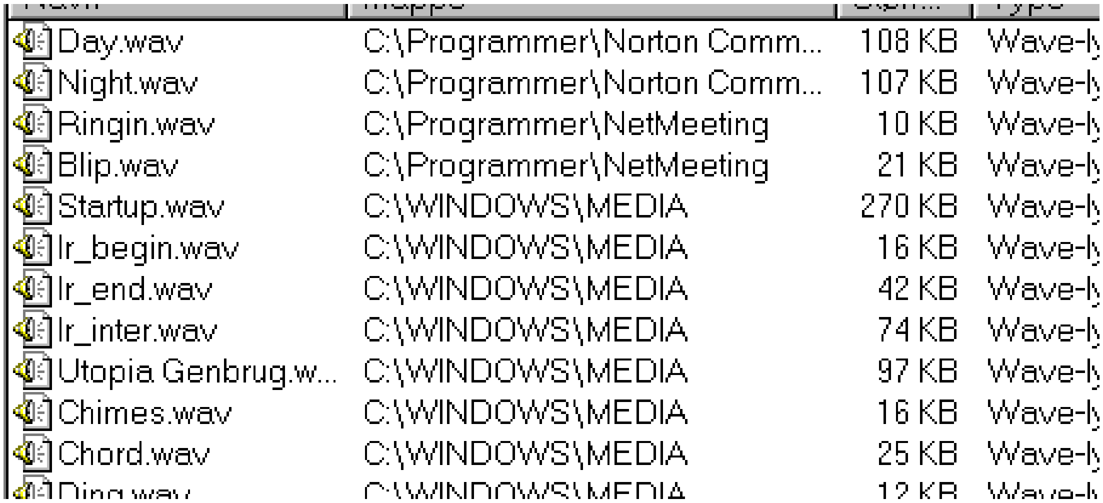


Stereo sampling at 16 bit and 44 KHz gives the best quality, but the Wave files will take up quite a bit more space.

The Wav files

[top]

If you look down in your PC, you will find plenty of Wav files. I did a little search, and this showed up:



Day.wav	C:\Programmer\Norton Comm...	108 KB	Wave-l
Night.wav	C:\Programmer\Norton Comm...	107 KB	Wave-l
Ringin.wav	C:\Programmer\NetMeeting	10 KB	Wave-l
Blip.wav	C:\Programmer\NetMeeting	21 KB	Wave-l
Startup.wav	C:\WINDOWS\MEDIA	270 KB	Wave-l
Ir_begin.wav	C:\WINDOWS\MEDIA	16 KB	Wave-l
Ir_end.wav	C:\WINDOWS\MEDIA	42 KB	Wave-l
Ir_inter.wav	C:\WINDOWS\MEDIA	74 KB	Wave-l
Utopia Genbrug.w...	C:\WINDOWS\MEDIA	97 KB	Wave-l
Chimes.wav	C:\WINDOWS\MEDIA	16 KB	Wave-l
Chord.wav	C:\WINDOWS\MEDIA	25 KB	Wave-l

All these Wav files contain sounds in a digital form - samples. They only contain very few seconds of sound, because of the file size, which must not grow too big.

A Wav file will sound the same no matter which sound card you may have, be it a sound card using FM Synthesis or Wave table. The sound is in the file and not in the sound card!

These samples above are used as sound effects within Windows. Similar samples are used as material on music CDs and in the MOD format of digital music.

AU is another file format for samples. MP3 files are highly compressed samples.

- Next page
- Previous page

Learn more **[top]**

Also see: Module 7d - about digital music: MP3s, MODs etc.

Read about video cards in Module 7b .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7c2.

PC sound - continued

The contents:

- The sound card - an adapter
- Newer sound cards

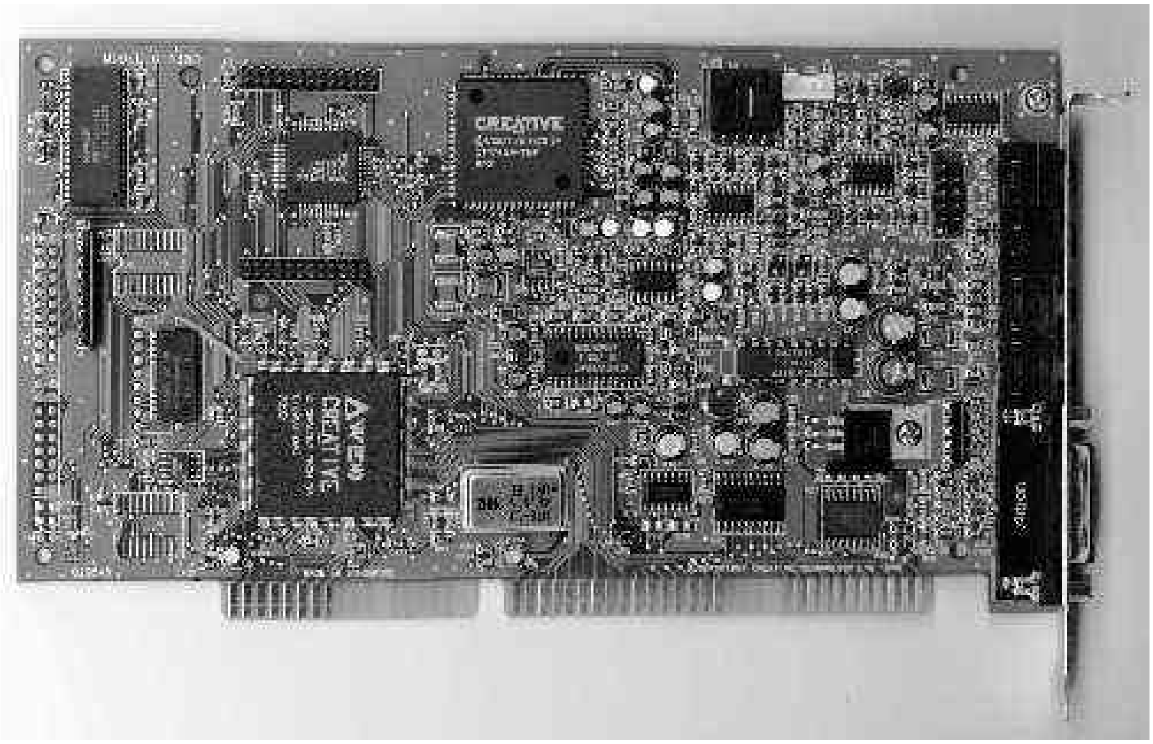
- Next page
- Previous page

The sound card - an adapter

[top]

In the previous module, we saw the principles of PC sound. Let us here look at the *sound card*, which is an adapter card.

The sound card used to be a ISA card (using the old ISA interface). Here you see the original Sound Blaster AWE64 Gold card:



The connectors may look different on different sound cards, but as an example: In the back of the AWE64 Gold card you find connectors to:

- Microphone input, a jack
- Line input, a jack
- Two phone jacks for active speakers
- A DB15 jack for MIDI or joystick.

Most sound cards typically have a 2 Watt amplifier built-in. It can push a set of earphones. An exception is the SB Gold card, where the amplifier is eliminated. It has no practical significance, since you probably want to attach it to a pair of active speakers.

The new sound cards

[top]

For many years PC sound has been totally dominated by the Sound Blaster card. All sound cards had to be compatible with Sound Blaster, or it would not sell. Obviously that is due to the numerous game programs, which require a SB compatible sound card.

The new sound cards break away from the Sound Blaster compatbility. This break involves many facets. Below I will describe some of the tendencies in the sound technology.

Sound over the PCI bus

New sound cards use the PCI bus. The SB compatibility used to require the old ISA bus, but this has been overcome. Creative Labs produce fine PCI-based SoundBlaster cards. With PCI you gain these advantages:

- The IRQ problems disappear.
- Signal/noise ratio can be improved with 5 dB.
- There is sufficient bandwidth (capacity for data transmission).

- The sound card workload for the CPU is less.
- We can drop the ISA bus, which takes up unnecessary space on the PC system board.

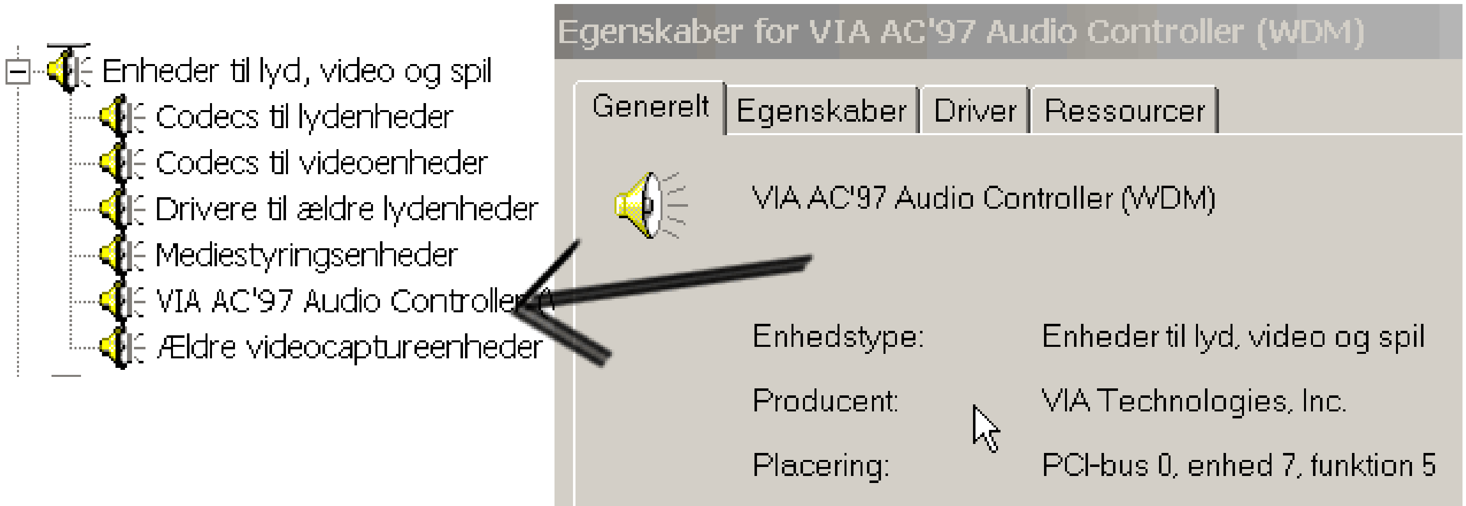
The problem in moving the sound to the PCI bus involved the existing software. First of all the old DOS games, which expected and demanded the Sound Blaster card with its well-known IRQ- and DMA numbers. The games did not work with the new cards, unless special solutions were implemented. However, the impact of this problem is gone. No more ISA-based sound cards are in production, and all games use the new standards for Windows sound.

Onboard sound chips

Many motherboard include sound card functions. This is i fine thing, if you only need sound for ordinary use. The quality is not as good as the sound from a $80-$100 sound card, but for many users it is fine!

On-board audio is found within some chip sets. For instance you find it in the much used VIA KT133 chip set for AMD processors. Here the VT82C686B south bridge I/O-controller holds built-in AC97 digital audio functions:

A Windows report on this:



Onboard sound chips is an in-expensive an simple via to incorporate sound facilities in your PC.

- Next page
- Previous page

Learn more **[top]**

Read Module 6a about file systems

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card

Also see: Module 7d - about digital music: MP3s, MODs etc.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7c3.

PC sound - continued

The contents:

- 3D sound

- Next page
- Previous page

3D sound

[top]

3D sound is a new hot area. You can create a very powerful illusion of 3D sound coming from just loudspeaker. This is done using new 3D processors on the sound card, which work with some very complex mathematical models. The sound comes from behind, from the front, from side to side - completely realistic.

The idea of 3D sound has to come from games, which especially are designed for it. The sound systems often include 4 or 6 loudspeakers. But they work fine with headphones too.

Sound Blaster Live is such a high-end, high fidelity 3D card coming from Creative Labs. It is a PCI-based card, and the best performance comes with the Four Point Surround sound system. Diamond MX300 is another.

SB Live!

The SoundBlaster top model sound card is called SoundBlaster Live!

It includes a lot of fine and powerful features:

- EMU10K1 accelerator chip

- Connections for four speakers.

- Digital DIN plug, which can be used for future high-end sound systems like Dolby ProLogic.

- SP/DIF digital phone plug, which can be connected to units such as DVD drives, DAT or MiniDisc for direct digital input.

- Plug for digital MPEG signal.

The EMU10K1 is as powerful as a Pentium 166 MHz CPU. It is an accelerator chip which relieves the PC's CPU when executing sound, such as DirectX activities that require a lot of processor power.

Creative Labs SoundBlaster Live is a phenomenal piece of hardware:

3D sound - better than stereo

In the 1950s stereo was invented. The music is recorded using two channels - a left and a right channel. Since then the aim has been to expand the sound into 3 dimensions.

This is possible. Only using two speakers you can create an illusion of "room". Many new sound cards are capable of giving 3D sound effects (i.e. Virtual Dolby). This way games can achieve even more realistic sound.

Diamond MX300 is a 3D sound card. It is constructed using accelerator chips from Aureal (Vortex 2 and A3D ver. 2.0). The first sound chip from Aureal was very revolutionary to 3D sound performance. It has been used by many vendors (such as Compaq). In 1999 the next generation chip was shipping.

The Diamond card was very well received. It should be just as good as the SoundBlaster Live! product, which has been in a class by itself since the introduction in 1998.

SoundBlaster Live works with an open standard for 3D sound called EAX (Environmental Audio Extensions). The MX300 card is compatible with this as well as with the Microsoft's standard for 3D sound.

3D environmental sound

The 3D sound card gives the listener an illusion of being in a landscape, where the sounds come from the front and the back. Sounds coming from up and down are difficult to reproduce.

The illusion is best when you use a four-set speaker system as SoundBlaster PC Works, which gives a very high quality at a modest price.

3D sound is also possible using only two speakers. The MX300 should be very good at this. The spacious sound is created using advanced mathematical manipulations, which need a good portion of CPU power. Hence the accelerator chip. This is called 3D Positional Sound. The best result should be achieved using headphones.

Another 3D effect is called Environmental Sound. Here the sound from a game is changed corresponding to the physical situation of the characters. If a person enters a tunnel, there may be an echo. In a big empty hall the sound is completely different. This way the games can send commands to the sound card, which adjust the feeling of the sound to the environment.

- Next page
- Previous page

Learn more

[top]

Read Module 6a about file systems

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card

Also see: Module 7d - about digital music: MP3s, MODs etc.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7c4.

PC sound - continued

The contents:

- USB sound

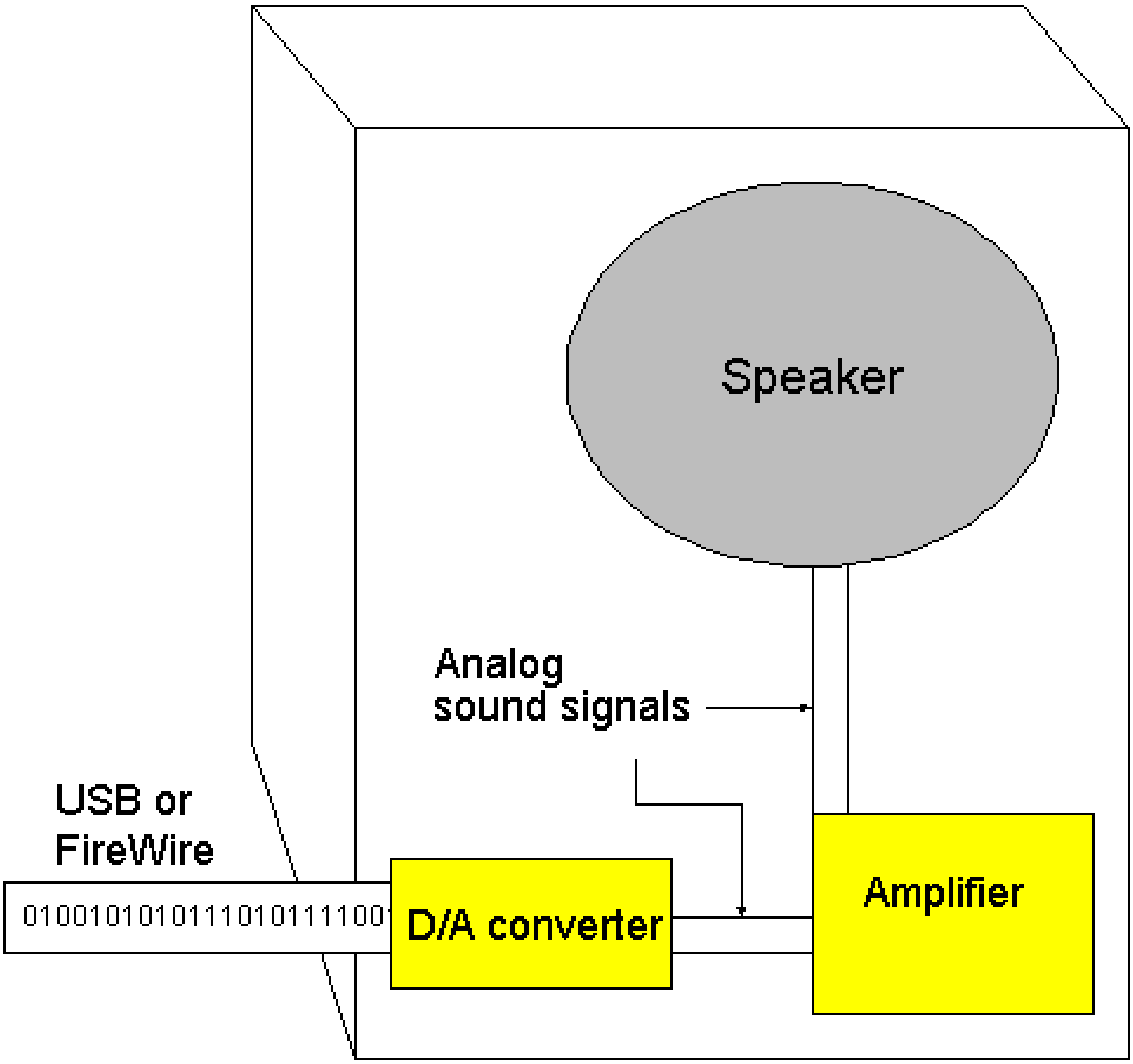
- Next page
- Previous page

Sound over the USB bus

You may experience very high quality sound systems using the USB bus.

The difference is that there is no sound card in the PC. You connect the speakers to a USB port instead of using the sound card.

Using this system, the sound signals are in digital form right coming from the harddisk or whereever, and they stay in this form when they are sent out on the USB channel and into the speakers:



Less noise

Inside the PC there is a lot of electric (static) interference from many sources. That can affect the integrity of the signals in the sound module.

With USB the noise sensitive digital/analog conversion will take place in the speaker, and this results in a superior quality. Both Philips and Altec Lansing produce USB speakers.

More CPU work

With USB sound you leave all the sound processing to the CPU. This "costs" some CPU power; however, modern CPUs are so powerful, that this is OK.

You find high fidelity loudspeakers with built-in amplifier and converter, which can receive pure digital signals (via USB). One could hope, that these speakers will be able to interpret data from hi-fi equipment, PC, TV/video and other sources. See the description of some of the finest speakers I ever have heard: A set ofPhilips USB speakers.

- Next page
- Previous page

Learn more [top]

Also see: Module 7d - about digital music: MP3s, MODs etc.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7c5.

PC sound - continued

The contents:

- DOS or DirectX

- Next page
- Previous page

DOS or DirectX

[top]

When so many games used to be DOS based, it primarily was because of the sound. Under DOS the programmer can modify and manipulate the sound card to a very high extent. It can be controlled very precisely, sounds can be mixed without interruption, and all kinds of effects can be designed. Here DOS proves very effective - the operating system permits direct control of the hardware.

The disadvantage with DOS sound is, that the hardware must be totally standardized. This gave the Sound Blaster card its great success.

Windows

In Windows all program instructions to hardware are executed through a programming layer

(API).

The first multimedia API would not allow mixing of sounds. Therefore the music in the Windows-based game had to be cut off, if there was a need for playing such a thing as the sound of an explosion. This put heavy restraints on programming creativity. Consequently DOS based game applications remained long into the Windows era. But it changed ...

DirectX

DirectX is a set of multimedia APIs (*application program interface*) developed for Windows . It is a collection of programs which enable much improved *low level control* over the hardware in games and other multimedia applications. DirectX has now reached version 6.1 and includes:

- DirectDraw
- DirectSound
- DiectSound3D
- DirectPlay
- DirectInput
- DirectSetup

These programs are designed to enable all possible image and sound effects.

The advantage of DirectX is that the applications can be written directly to Windows and simultaneously get maximum hardware control. Hence DirectX is very important to hardware manufactures. To make sure that the new products work together with all software, the drivers have support the latest version of DirectX.

With DirectX we should finally have eliminated the need for programs to rely on Sound Blaster compatibility.

DirectX comes in new versions every year.

In version 7.0 you find improved 3D acceleration of sound as well as picture with reduced CPU usage. The performance should be increased with 20% compared to version 6.1.

Windows 2000 was the first NT-based version of Windows to include DirectX.

- Next page
- Previous page

Learn more **[top]**

Also see: Module 7d - about digital music: MP3s, MODs etc.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7d1.

About digital music

The contents:

There are quite a few different kinds of music formats you can find on the Internet. A few of them are described on these pages.

- About the player

- Next page
- Previous page

About the player

[top]

When you have some of those files, you need a *player (a plugin)* to replay it on your PC. All versions of Windows have built-in players for the Wav and Midi files, so you do not have to think of that. Just double-click on the file, and the sound or tune is replayed:

However, the sound files MP3s and the MODs are much more interesting formats than Wav and Midi. But you need plugins, a little program to replay the tunes. These players are freely available on the Internet - I'll give you the links later.

Some players are only available as plugins to browsers. This goes for the Koan stuff. Others (MP3s and MODs) can be achieved as stand-alone players or plugins.

- Next page
- Previous page

Learn more	**[top]**

- Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page]	[Contact]	[Karbo's Dictionary]	[The Software Guides]

KarbosGuide.com. Module 7d2.

About MIDI and sequencing

The contents:

- MIDI

Next page
Previous page

Introduction

This page is about MIDI compositions, which are "real" pieces of music, written for playback with any sound card. MIDI is a standard in Windows, so any PC with a sound card can play these Midi files.

MIDI (Musical Instrument Digital Interface) is a specification, which was developed in the 1980s to communicate between synthesizers. Since then MIDI has also become a standard, which allows programs to play music through the PC sound card.

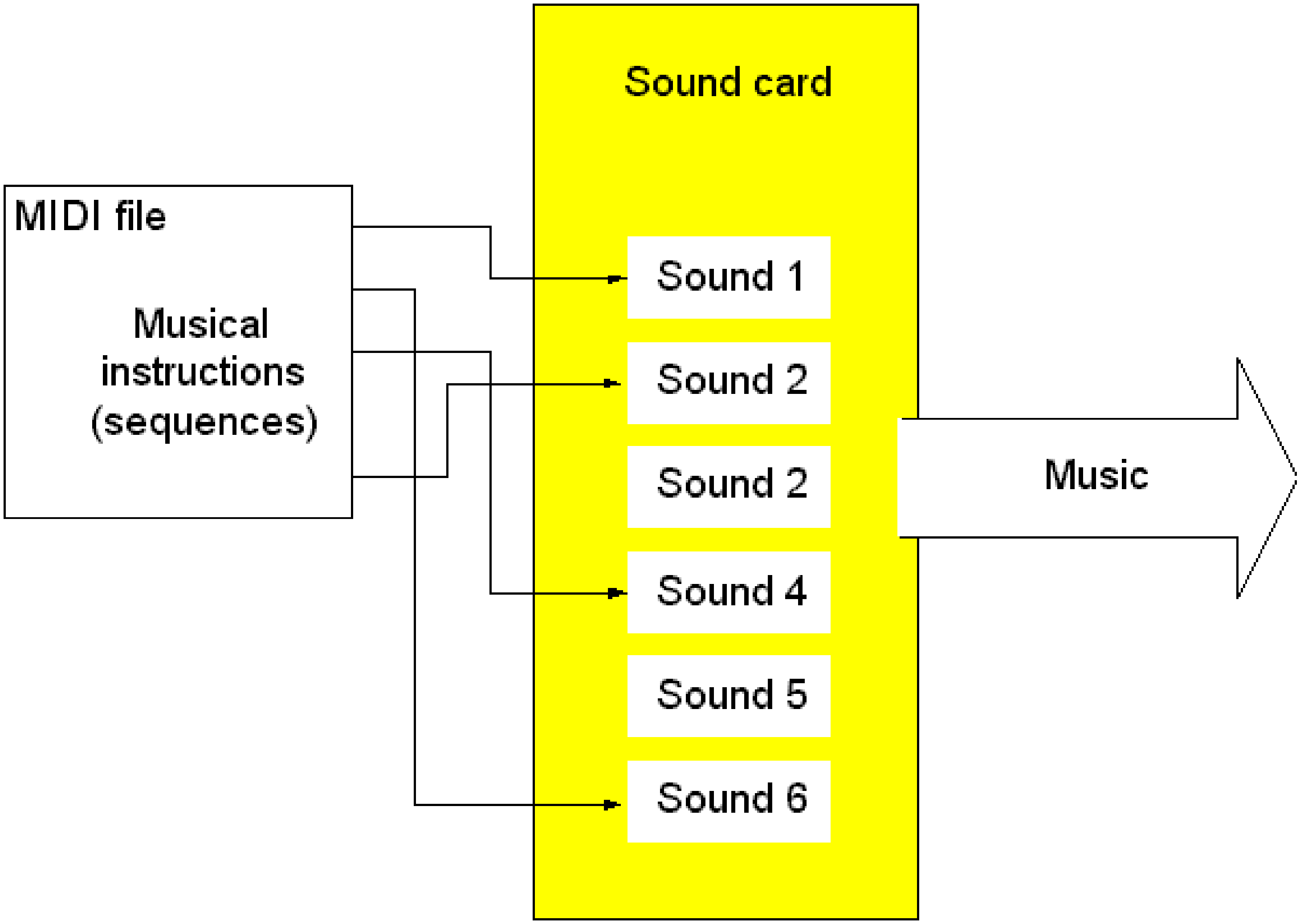
MIDI is a computer standard music format. You write compositions - musical events - in the MIDI format. The MIDI files do *not* contain the sounds but a description of how the music is to be played. The sounds are in your sound card. The MIDI file only contains *sequencing* information - *which instrument* it is played *how* and *when* .

For example a MIDI sequence can describe the hit on a piano key. The MIDI sequence describes:

- The instrument

- The note
- The strength of the key hit
- How long to maintain the note
- Etc.

The only thing which is not covered is the sound of the instrument - that is created in the sound card, and is totally dependent on the sound card quality:



Note level recordings

A MIDI recording is thus a recording of music on "note level," without sound. It is played by a module, such as a sound card, which can generate the sounds of the instrument. MIDI files do not occupy much space as compared with the pure sound (WAVE files). Therefore they are often used in PCs, on Internet etc.

You find a lot of MIDI music on the Internet. However, compared to MP3s the format is rather tame. There is rarely more than a few minutes of music in a MIDI file, and you soon get tired of the pieces, which all sound the same using the limited number of voices within your sound card.

The advantage of MIDI is that the file format is so standardized. If you have a sound card, no matter which, it will work. Depending on the quality of your sound card, a MIDI can sound good or lousy. Cheap sound cards have a chip on them which mimics the sounds of different instruments when you play a MIDI file. Newer sound cards use a Wave table chip which contains actual samplings of the instruments. The MIDI file is still limited to the around 120 instruments on the sound card.

MIDI interface for keyboards

[top]

A musical keyboard can be connected to the sound card with a connector. That is called a MIDI interface. You can buy special PC musical keyboards, or you can use one of the keyboards which are available in music stores. It will work as long as the MIDI connectors match.

You connect your DIN connector to the piano keyboard. In the other end of the cable is a DB15 connector to the sound card. Then you can play from the piano keyboard through the sound card. Of course it requires a program which can handle music, but it works.

I have tried it myself. The Sound Blaster AWE64 Gold comes with the program Cubasis. Once I connected an old and cheap piano keyboard (with built-in rhythm box) to the sound card, and everything worked through Cubasis. The keyboard acted as a "Local Synthesizer" in the program settings.

This keyboard is especially designed for the PC.

Links

[top]

Here is a link to Anselmo Salzani, who tries to create exciting music in the MIDI format. His page also includes a lot of other interesting music links: Brazilian MIDI music.

And a Dane: Anders Kornerups MIDI music

- Next page
- Previous page

Learn more [top]

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7d3.

MODs - digital music

The contents:

- About MODs
- Links to the music and software.

Next page
Previous page

The MODs

MODs another very interesting format. Originally made for Commodore Amiga computers and used for musical demos by a company that sold the music editor SoundTracker.

The source code of this program was cracked and illegal versions came. Today MOD is an interesting format offering more than four sound channels with both synthetic instruments and samples integrated.

Modules are digital music files, made up of a set of samples (the instruments) and sequencing information. The file tells the mod player when to play which sample on which track at what pitch, optionally performing an effect like vibrato, for example.

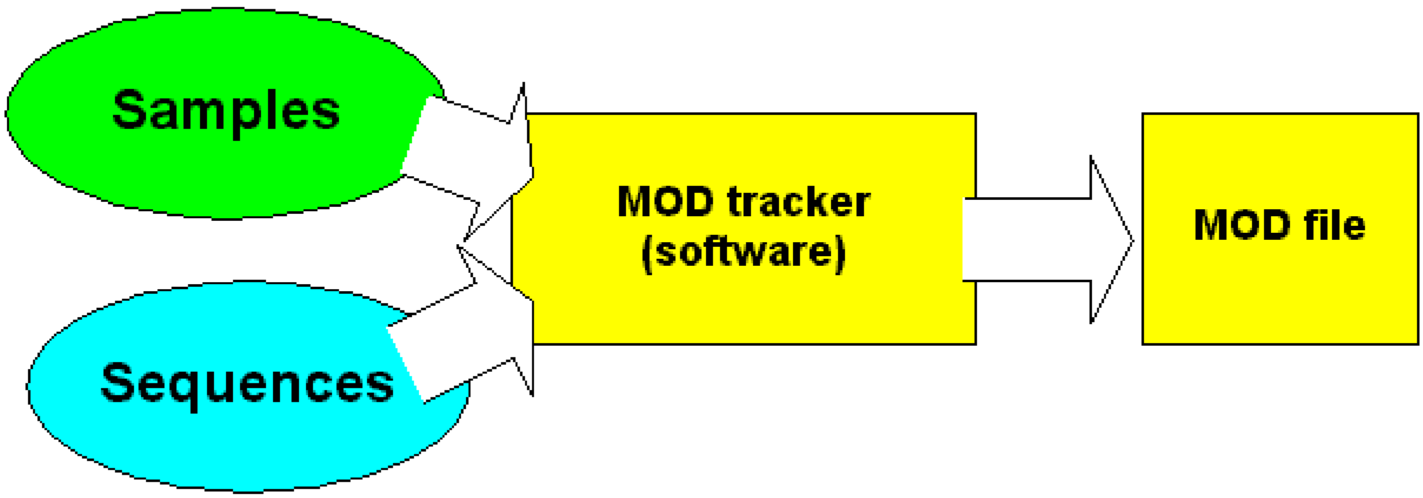
Thus MODs are different from pure sample files as WAV, which contain no sequencing information, and from MIDI files, which do not include any samples/instruments. MODs are extremely popular in the demo world and offer a way of making music of an acceptable level of quality rather cheaply.

With all the new high quality sound hardware new generations of musicians may produce a sound quality near that of the professionals.

The technique

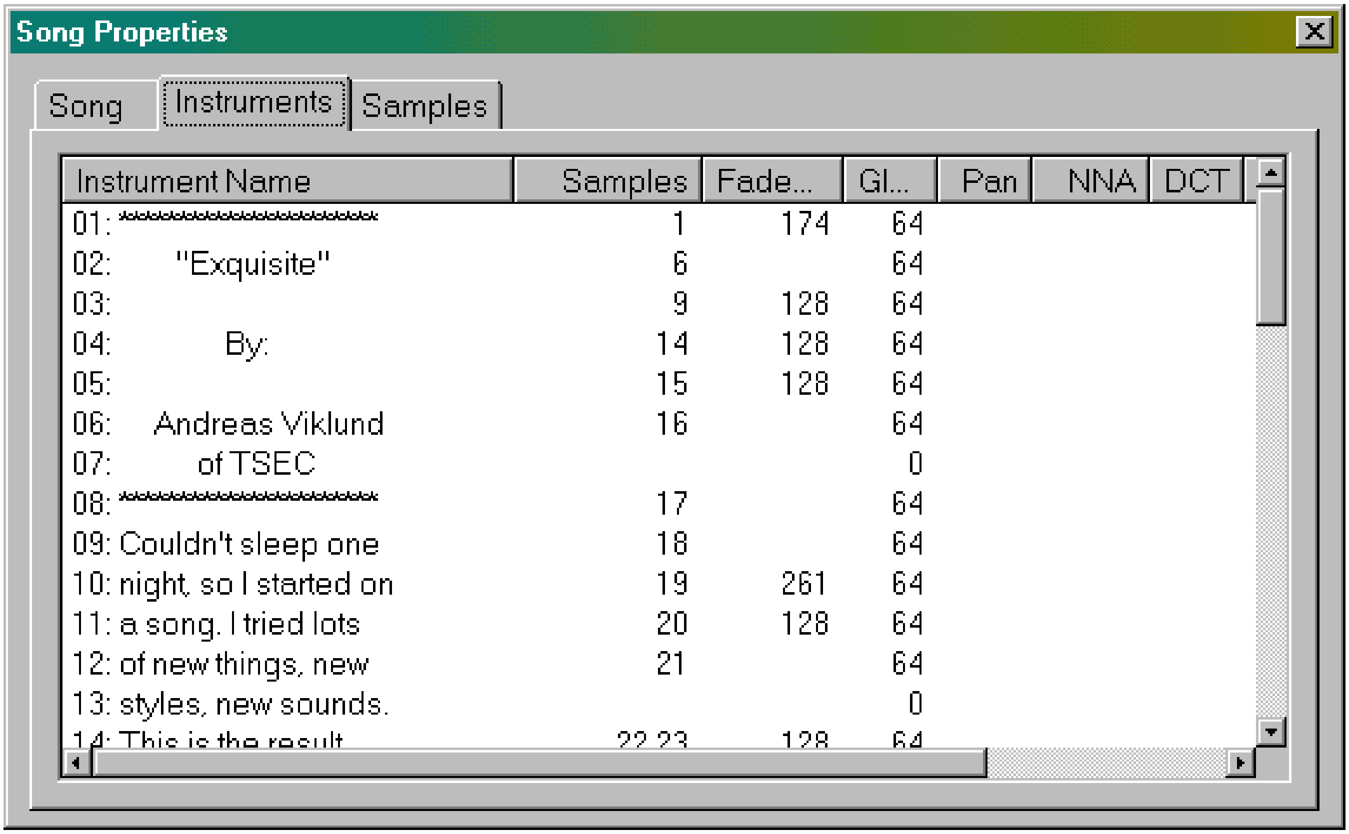
MODs' sequencing information is based on "patterns" and "tracks". A pattern is a group of tracks with a certain length, usually 64 "rows". The tracks are independent of each other. A four track MOD can play four voices or notes simultaneously. The patterns can be repeated in a play list reducing the file size.

The MOD files contain the instruments along with them in the form of samples. The samples are little WAV files of one note on an instrument, a beat on a drum kit, or perhaps a line of vocals. The MOD composer decides what samples he includes in the MOD file. He uses a *tracker* to make the tunes:



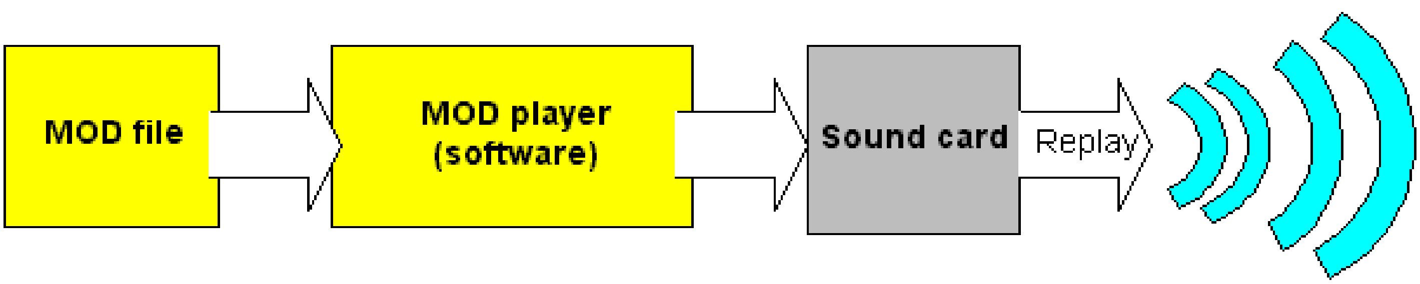
This way, the song will sound the same when played back on any computer, because the sounds as well as the sequences are included. Here is a MOD player:

And here I "look behind" one the tunes above:



You can download great music in the MOD format. The files are named .MOD or .XM. The most incredible is the file size. There obviously is a lot of compression in it, often you get more than a minute high quality replay out of a MOD file of just 100 KB.

All you need is a little software:



Links

[top]

Visit this MOD site where you find players and other MOD stuff.

Also check Great Swedish music in MOD format.

- Next page

- Previous page

Learn more [top]

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7d4.

The MP3s are dynamite

The contents:

- Introduction to MP3s
- Psychoacustic algorithms
- Ripping
- Links

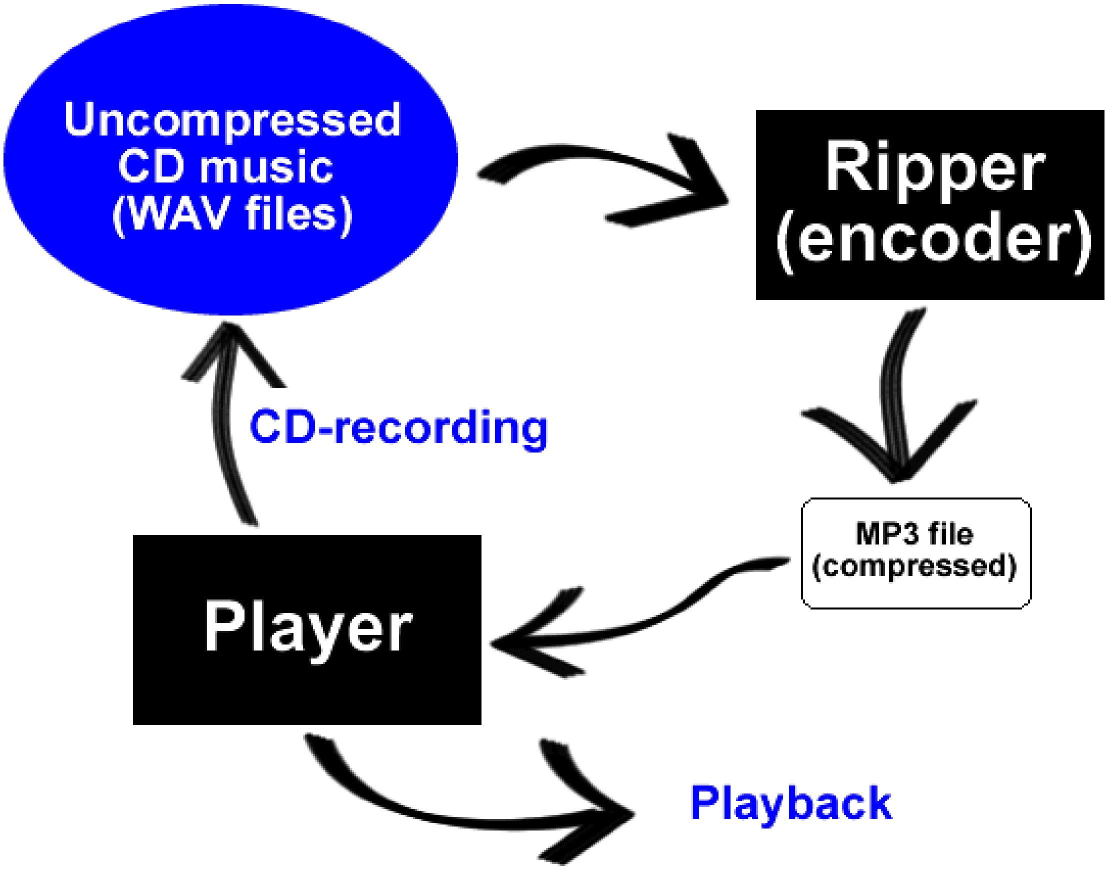
- Next page
- Previous page

Introduction

Since 1998 the MP3 standard has become more and more important, and an enormous success. The potential is even bigger - personally I believe, that MP3 end up being as popular as the Compact Cassette did in the 20th century.

MP3 is a system to give a huge compression of digital sound files. The compression is *lossy* (i.e. musical details are cut away). Yet MP3 delivers a sound quality (almost) as good as uncompressed CDs, due to the very intelligent psycho-acustic algorithm reducing the file size.

The MP3 format is very versatile; it can be hosted on any storage media and can be transferred on demand over the Internet. You use a *ripper* to encode MP3 files. These files can be played using a player like Winamp, MusicMatch or Windows Media Player. The MP3 files can also be decoded an used for CD-recording:



German research

MP3 means MPEG Audio Layer 3. It is an audio compression technology being a part of the MPEG-1 and MPEG-2 specifications. MP3 compresses CD quality sound by a factor of 8-12, while maintaining almost the same high-fidelity sound quality.

MP3 is developed by a German research institute called Frauenhofer. The company Thomson Multimedia has patented MP3 in USA and in Germany.

Effective compressions

Music on CDs have a bandwidth of 1.4 Megabit per second. It is calculated as 2 X 16 X 44100 bit/sec. This mean that one minute of music on a CD takes up 10 MB of data.

Using MP3 this bitstream is dramatically reduced (by factor 8 to 12). A typically MP3 file will need 128 per second. Hence one minute of music is reduced from 10 MB data to only 1 MB. Greater compression ratios are also possible for use on Internet etc. but here you will encounter a decrease in sound quality.

Standard MP3s hold approx. 1 minutes hi-fi music per megabyte.

This reduction is only possible using a set of compressions.

Lossy compression with psychoacustic algorithms

Overall we have to types of compression:

- Compression without loss
- Lossy compression

If we want compression without loss, we use systems like ZIP. This is very effective compression data files that hold plenty of redundant information. This could be Microsoft Word documents, they often zip very well. And when you unzip them, the document is identical to the original. You find similar compression within GIF and PNG graphics files, which compress many graphic images very well (but not photos).

However you do not find much redundant information in music files. A zip compression of raw music data (WAV files) may only yield 10% reduction in file size. Therefore we use a lossy encoding to reduce the music files sizes.

Lossy encoding mean that we take away music information (just as JPEG encoding take away image information from a photo). The goal is to remove music details you *would not hear* anyway!

Since MP3 offers *variable compression* you will find that the more you compress the music, more details are removed and lesser fidelity is the result.

Many ways to MP3

The MP3 standard tells what design a MP3 file should have. It *does not* tell how to produce the file. This indicates that we may experience quite different quality from different encoders.

The most important principle in MP3 compression is the *psychoacustic* selection of sound signals to cut away. Those signals, we are unable to hear are removed. These include weaker sounds that are present but are not heard because they are drowned out (masked) by louder instruments/sounds.

Many encoders use the fact that the human ear is most sensitive to midrange sound frequencies (1 to 4 KHz). Hence sound data within this range is left unchanged.

An other compression used is to reduce the stereo signal into mono, when the sound waves are so deep, that the human ear cannot register the direction. Also the contents of *common information* in the two stereo channels is compressed.

The Huffman algorithm reduces the file size by optimizing the data code for the most often used signals. This is a lossless compression working within the MP3 system.

Pirating or legal

All over the Internet you find pirate copies of commercial music. This is not very good since it is illegal and may stop the development of the technology. At www.mp3.com you only find legal music, but there is lots of it!

Napster

A great online music community was created around downloading and sharing MP3 files. This was Napster, and it was illegal. You cannot give away copies of your MP3s to anyone, unfortunately.

Napster had to close down several times in 2000 and 2001 due to law suits from the music industry.

SDMI

Secure Digital Music Initiative. This standard was developed by Sony, EMI, and three other big companies.

It is a security certification which can be used on MP3 files and other formats. It should help to prevent illegal copies of music. With SDMI a MP3 file can, as an example, be designed so it only can be copied three times.

SDMI is to built into MP3-players as Rio and MP3-man. Here it verifies the SDMI-sigature on MP3 files. However, the system allows replay of "illegal" MP3s as well.

The RIO player

In October 1998 the American organization RIAA (Recording Industry Association of America) tried to stop Diamond Multimedia from selling this great little thing. It is a MP3Man, just like a portable CD player:

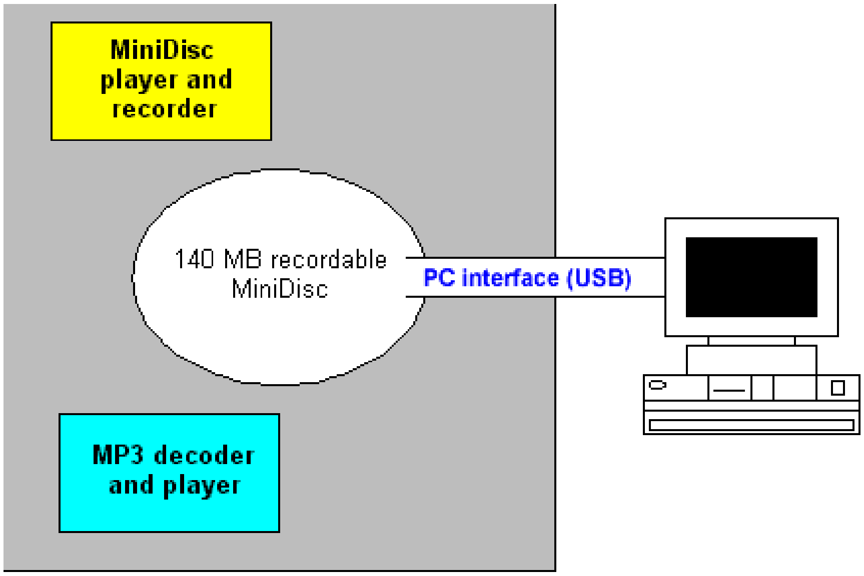
Holding the music in 32 MB of flash memory the player has no moving parts. Without moving parts, it could play for about 15 hours on a single alkaline AA battery.

The RIO was a revolutionary new device. Later MP3 decoders have come in many (better) versions, including mobile phones and digital cameras as well...

Karbo's Player

My own favorite device would be a Sony MiniDisc recorder holding MP3 playback software and an interface to the PC. With the 140 MB MiniDisc you will have a great medium for musical storage. Of course Sony has to protect their music division, but the MiniDisc could be so good in this setup.

The Minidisc uses it's own compression algorithms much similar to MP3, but in my setup you would be able to copy the already-encoded MP3 files directly to the Minidisc.



The device should also connect to my HIFI stereo set as well as to the cars sound system.

Winamp

I use the little program called Winamp and replay from harddisk through my PC's loudspeakers:

Microsoft's Windows Media Player also plays MP3s. But It is not as smart as Winamp is.

Microsoft tried to "kill" the MP3 format introducing their own Windows Media Audio (WMA) format, which is similar to MP3 but not compatible. This attempt to incorporate yet another "digital area" in Windows has failed - not many people prefer WMA to MP3, and this pleases me. Microsoft produces great software, but they should not monopolize everything.

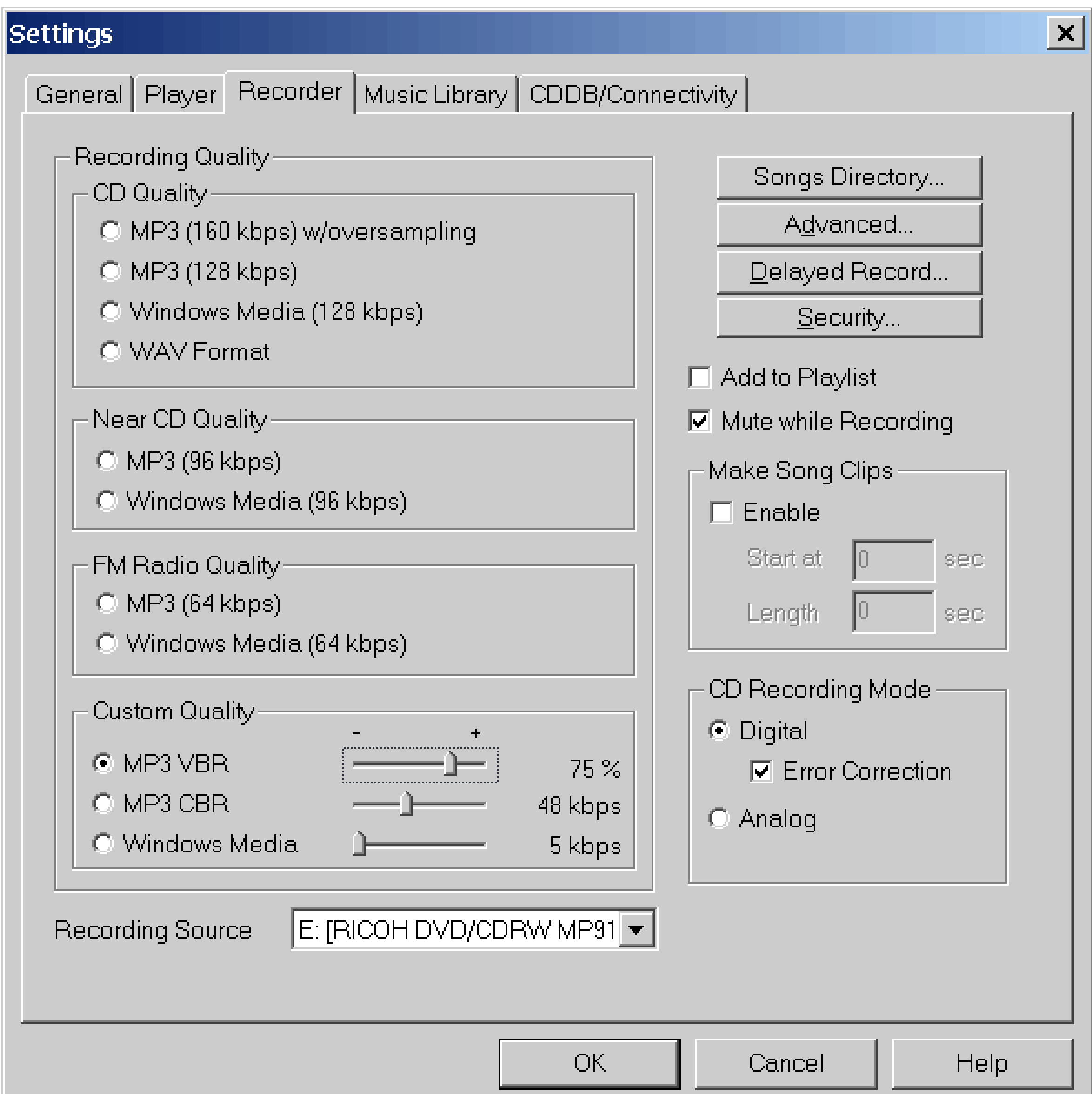
A good player should be able to produce *playlists.* A playlist is a little text file, which lists a sequence of songs that are to be played continously. The playlist is a file with the extension M3U. It can be edited using Notepad etc.

Ripping

To produce MP3s you use a ripper. You load a music CD into the CDROM drive. The software finds a CDDB database on the Internet and finds the artist name and the title of the disk and each song.

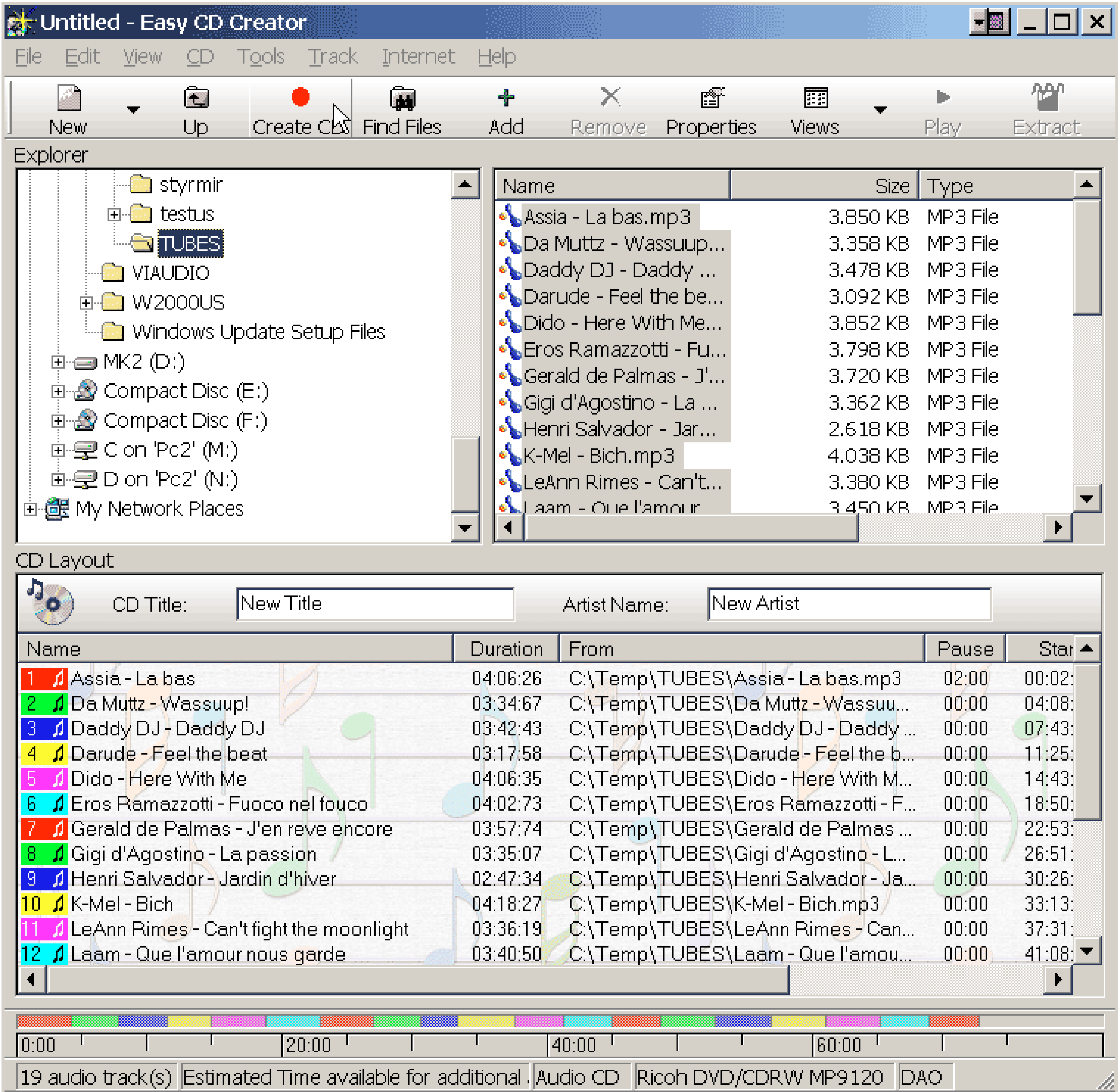
You just tell the ripper which tracks to rip, and the recording starts. I use MusicMatch:

I used to rip at a constant bit rate (CBR) 128 kbs, which worked fine. Experts tell me that I should use a variabel bit rate (VBR) setting of 75%. It should produce the best sound quality, using many bits when the music is complex and fewer when it is simple. Here you see the MusicMatch settings for ripping:



The MP3 format is extremely easy to use. Considering the explosive developments within Internet and electronics in general, MP3 must hold a revolutionary potential capable of transforming the music industry quite a lot.

We use MP3 for backup storage of our music. When we need a copy of a CD, we "burn" it from MP3's. Here we use Adaptec Easy CD Creator, which works fine:



MP3pro

In 2001 a new and updated version of the MP3 standard was introduced. Using better compression, it should deliver same sound quality from files half the size.

Here you see my DVD-player which plays CD-ROMs filled with MP3 files:

Links

[top]

Get a MP3 player

You will have to find the MP3s yourself - start with www.mp3.com

Get a ripper from MusicMatch

Microsoft's new player " Windows Media Player " also plays MP3s.

- Next page
- Previous page

Learn more [top]

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

KarbosGuide.com. Module 7d5.

About digital music - Koan

The contents:

- Koan music.
- Links to the music and software.

Next page
Previous page

Koan music

This something entirely different from MP3 etc. Koan is genuine computer music, with much more potential than the "flat" MIDI files earlier mentioned.

Koan is an electronic music standard. And this represents a fascinating technology developed by the British company SSEYO. Koan requires the addition of a *plug-in* to your browser to enable playing the files. Koan music is written to designated sound cards. The Sound Blaster AWE is the best as far as I know. I just have an ordinary Sound Blaster 16, and there is also a lot of good Koan music for that.

Koan is "live" music - it changes every time you play it. You can compare it with an aeolian harp, where the wind and thus the tone is different each time it is used.

> "I too think it's possible that our grandchildren will look at us in wonder and say:
> You mean you used to listen to exactly the same thing over and over again?
> Brian Eno 1996

The Koan music consist of small files, which start a process in the PC where they work. There may be 8 hours of music in a 12 KB file! So it is not the music itself which is contained in the file.

Rather the files contain some structures, frames if you wish, about a composition. These frames are activated in your PC's math processor. Then the music is generated within your PC, differently each time you play it.

Soundcard compatibility

Koan music is written specifically for a certain sound card. So you must have either Sound Blaster 16, 32 or 64. or a few other makes. Here again is a good argument to stay with the SB sound cards. They will give the fewest problems. The music is Internet suitable, since the files are small. I have found music in the category ambient, that is long electronic music sequences. They can be very quiet and meditative, but they can also be more rhythmic. If you:

- Like electronic music á la Tangerine Dream and Brian Eno
- Have a SB sound card and speakers in your PC system

Then you ought to try some Koan software. It is really simple to install and requires only a little space.

How do I do?

I write this on three premises:

- You have a Sound Blaster sound card and speakers.
- You are on the Internet and use either Netscape or MS Explorer.
- You know how to download and extract (unzip) files.

My installation example is based on software for Sound Blaster 16 and Netscape in 32 bit Windows 95 edition (Netscape Gold, version 3 or 4). It may sound complicated as I describe it. However it is really quite simple:

You want to install SSEYO software, so you can play the small SKP files with exciting music. First get the following: 32 bit Koan Software for Windows 95, and SB16 . You have to find the file on SSEYO home page. New versions arrive all the time.

It is a self extracting Exe file about 300 KB big, which you place in some temporary folder. That file will be deleted after installation. Run the file (it is called knp1032.exe), which will install the necessary plug-in in Netscape.

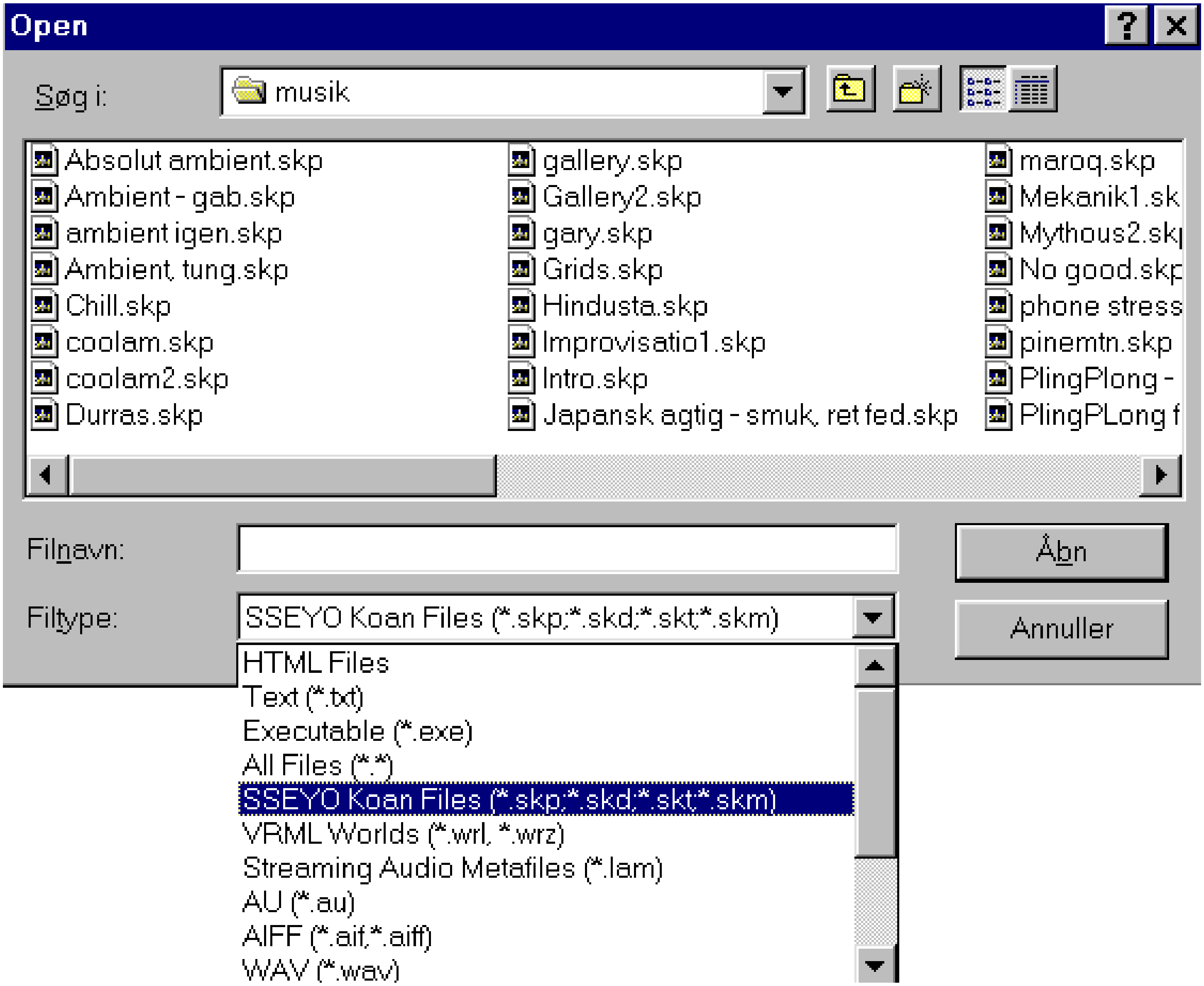
Now you can go on the net and for example retrieve the starter package on the same server, which includes some SKP files. Each of those represents hours of electronic music.

The best result is obtained with a AWE64 sound card and a pair of good speakers. I purchased a set of Altec Lansing ASC45. They are two tiny satellites with a heavy sub woofer, and giving fantastic sound - that is hi-fi!

Play back of Koan music

Once you have installed Koan plug-In, Netscape can play the Koan files! It sounds backwards. You would think that the filter should work in Windows 95, but no - the music has to be played through an Internet browser. I use navigator for Koan (it works there) while I have chanced into Internet Explorer for surfing....

You save your Koan files in a folder (in my computer: \web\music). Then, in Netscape, you press Control+o (for *open*). Now you have to modify the file type, to activate the filter:



Now you just select the melody, and Netscape will play it. You need not be on the Internet, you just use the browser to play the music. The music can run in the background all day, while you do something else.

If the SKP files are associated with Netscape, you can play them directly by double clicking on them.

Links

[top]

At the SSEYO Koan home page you can find plugins, tunes and information, including software to let you write your own Koan music.

- Next page

- Previous page

Learn more [top]

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 8

The contents:

- Editing photos with Photoshop

- Next page
- Previous page

UNDER CONSTRUCTION.

- Next page
- Previous page

Learn more Top

Read about video cards in Module 7b .

Read about sound cards in Module 7c .

Read about digital sound and music in Module 7d .

Karbosguide.com. Software Tip

Fixing photos using Photoshop

- Next page
- Previous page

When you look at your digital images, you'll probably find that a few of them could benefit from some editing.

The fish eye

One of the problems you may experience consist of unwanted *geometrical distortions*.

The lens of the digital cameras is rather "short", it covers a wide angle. This often gives geometrical distortions, especially photographing big objects at a short distance. If a square object fills all the image, you will see that the lines of the figure no more are parallel.

You may download the photo by right clicking on it, then you can practise the operation on your own PC, if you have Photoshop installed. We use this photo of St. Victor in Marseille:

It is not the photographer, who has had to much pastis. The photo has been taken on a very short distance. That has resulted in a significant distortion in the left part of the image.

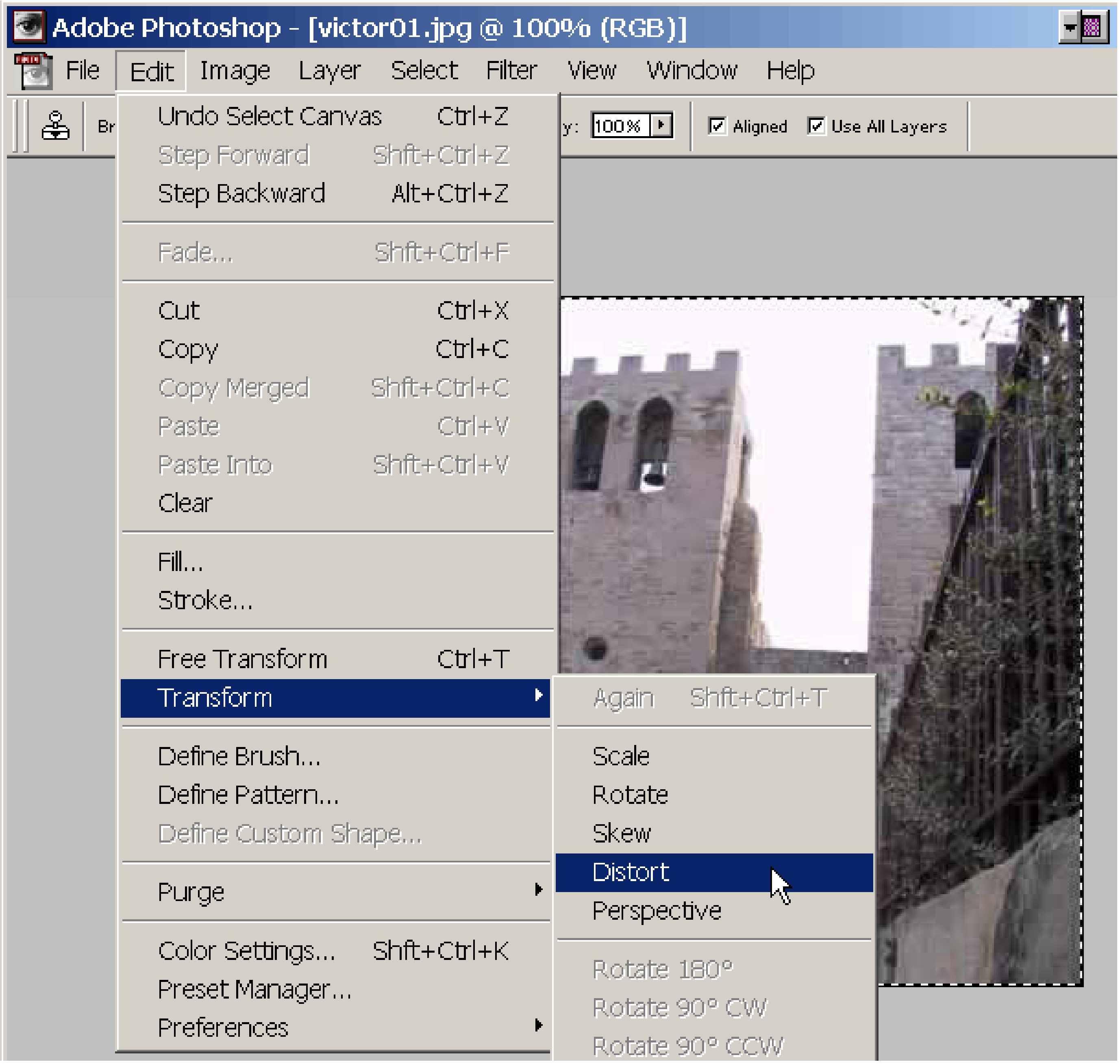
Too bad; otherwise the photo is OK, but now it is useless. Or what?

Use Photoshop

The photo is not to be wasted, it just has to be corrected. The great imaging editor Photoshop has the tools.

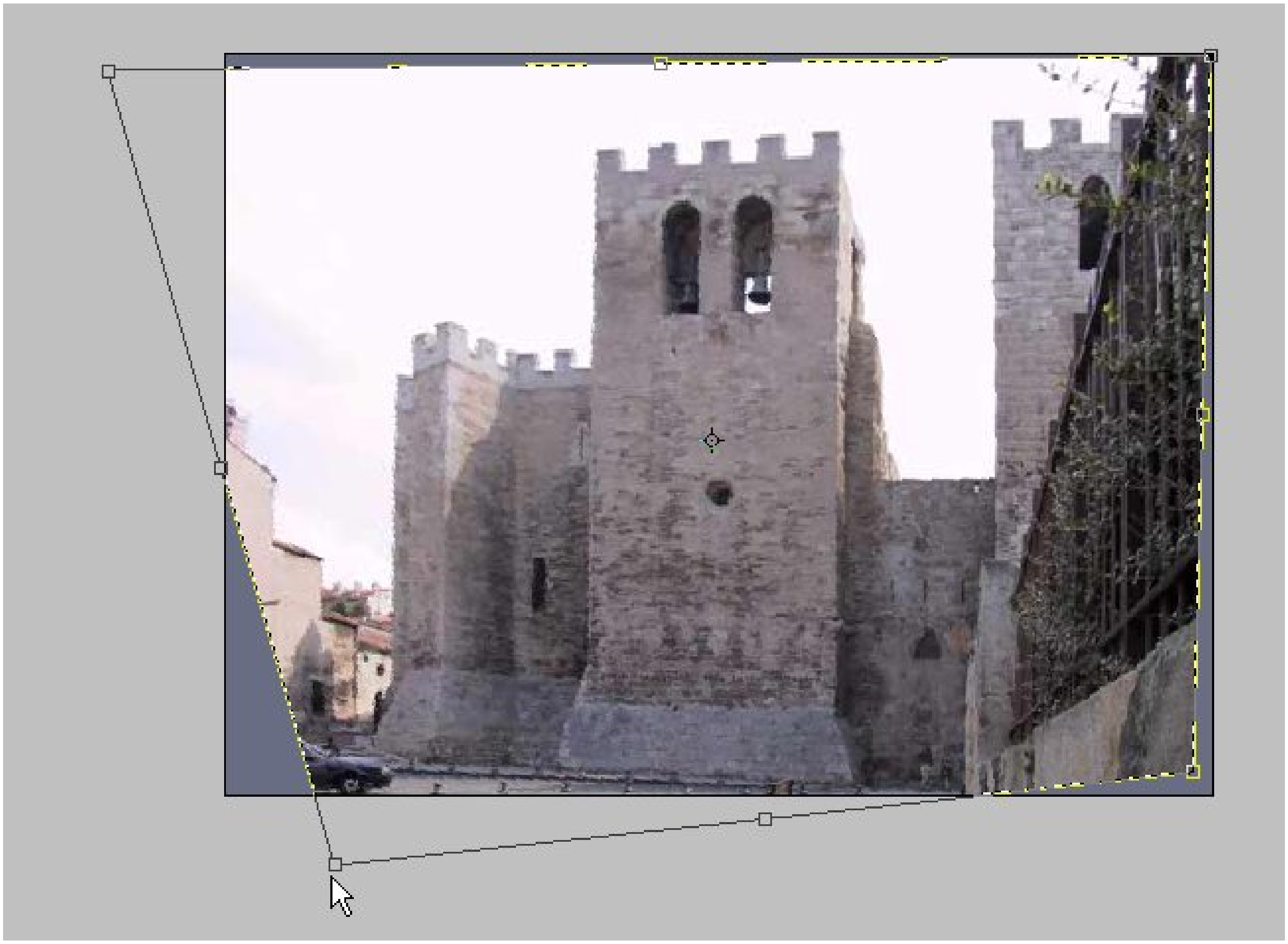
You open the photo in Photoshop and choose Select all (Control+a).

Next choose Edit --> Transformer --> Distort:



Now you see eight handles in the corners of the image. You may drag one of the handles, and the image is geometrically distorted. This is how this tool works.

You should apply a distortion which corrects the photo. Something like this:



You always apply a transformation by hitting [Enter]. The the image is OK from a geometrically point of view:

You may need to do a little further editing. In the left corner there is a grey are which is an unwanted effect of the transformation:

It has to be corrected. Either you crop the photo, so this part disappears, or you may use the clone brush to edit the areas. All this is described in my Photoshop book), which is available in many European languages (not English, unfortunately).

Finally the image should be saved in a high quality JPEG file. Delete the old version - you should not keep to versions of the same photo.

The Photoshop function Transformation is great and very versatile. You may use for many very different task. Here is the resultat of our little exercise, which can be performed within a minutes time:

The original photo:

After the treatment:

PS: All images on this page are heavily compressed. So the photo quality is not very good.

If you want to see a beautiful photo from same location, please click here.

- Next page
- Previous page

Photos taken with Canon S20 digital camera

- Next page
- Previous page

I am very happy with this photo. Here you see a reduced version of it.

Originally the size is 2048 x 1536 pixels. Here it only covers 850 x 638 pixels to make it more web-friendly. The original image file is of 900 KB; here it is reduced to a mere 61 KB.

The phot is taken using our little Canon S20 digital camera. It is from the inside of the church St. Victor in Marseille, Provence.

The exposure time was 0.7 second og the aperture was f/2.90. But these settings were applied automatically by the camera:

I find the photo wery nice with great color tones.

I am often surprised by the quality of the images coming from this little camera. It almost always produces good photos, no matter what I demand from it. Here is another example, taken in Saint-Raphaël by night:

The image data are:

- Shutter Speed: 1/3.33
- Aperture: f/3.50

I can recommend having a digital camera to all photo interested persons. It is so great!

- Next page
- Previous page

Karbosguide.com. Software Tip 18

- Next page
- Previous page

Use MSConfig to alter the Windows start-up.

Windows 98 has a new powerful tool called MSConfig, which a lot of people not are aware of.

In tip number 3 I showed how to clean up all the temporary Internet files including the subdirectories, cookies, and other temporary files. I have got several comments on this tip, which many people find very useful.

Temporary disabling commands

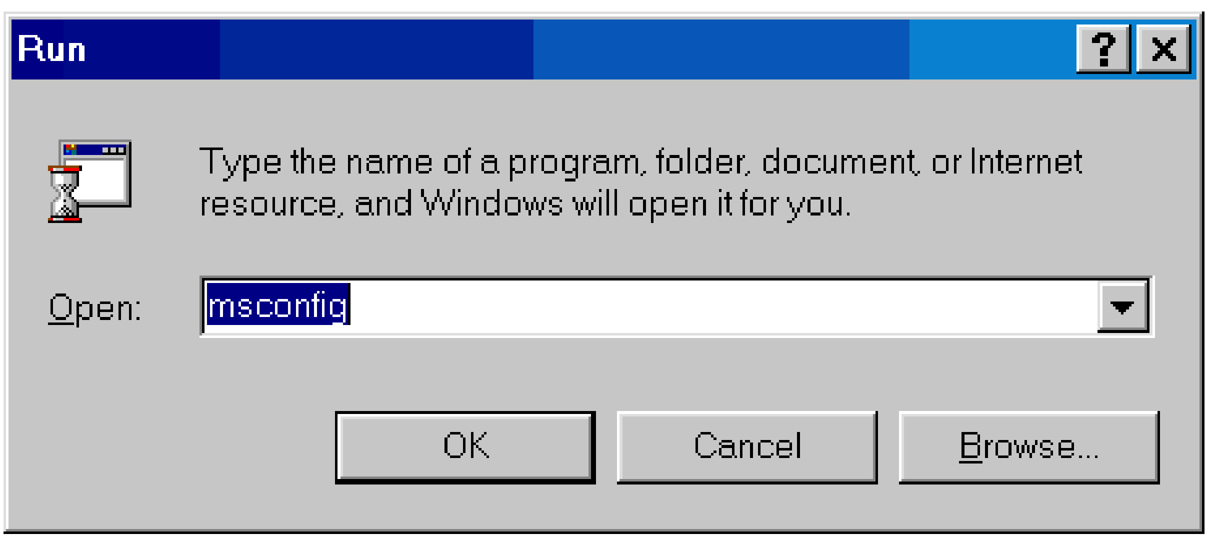
Sometimes it is nice temporary to disable these DOS commands, as Mr. Kokusai of Japan wrote me and asked how to it.

Many programs rely on the folder C:\Windows\temp during the install process. If they also include a re-boot, it may cause problems if your Autoexec deletes all the files in C:\Windows \temp.

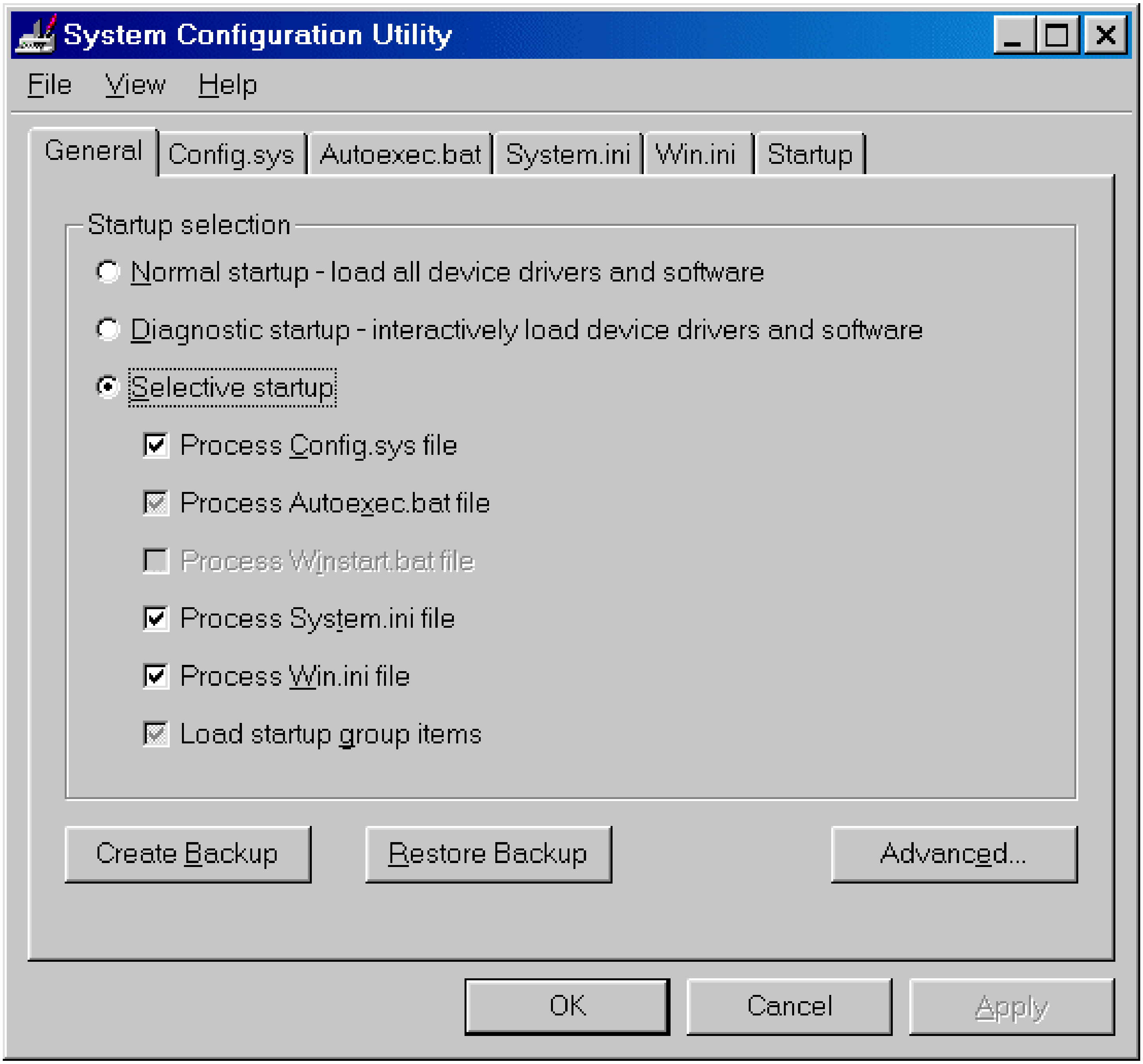
The solution to this is MSConfig.

Find MSConfig

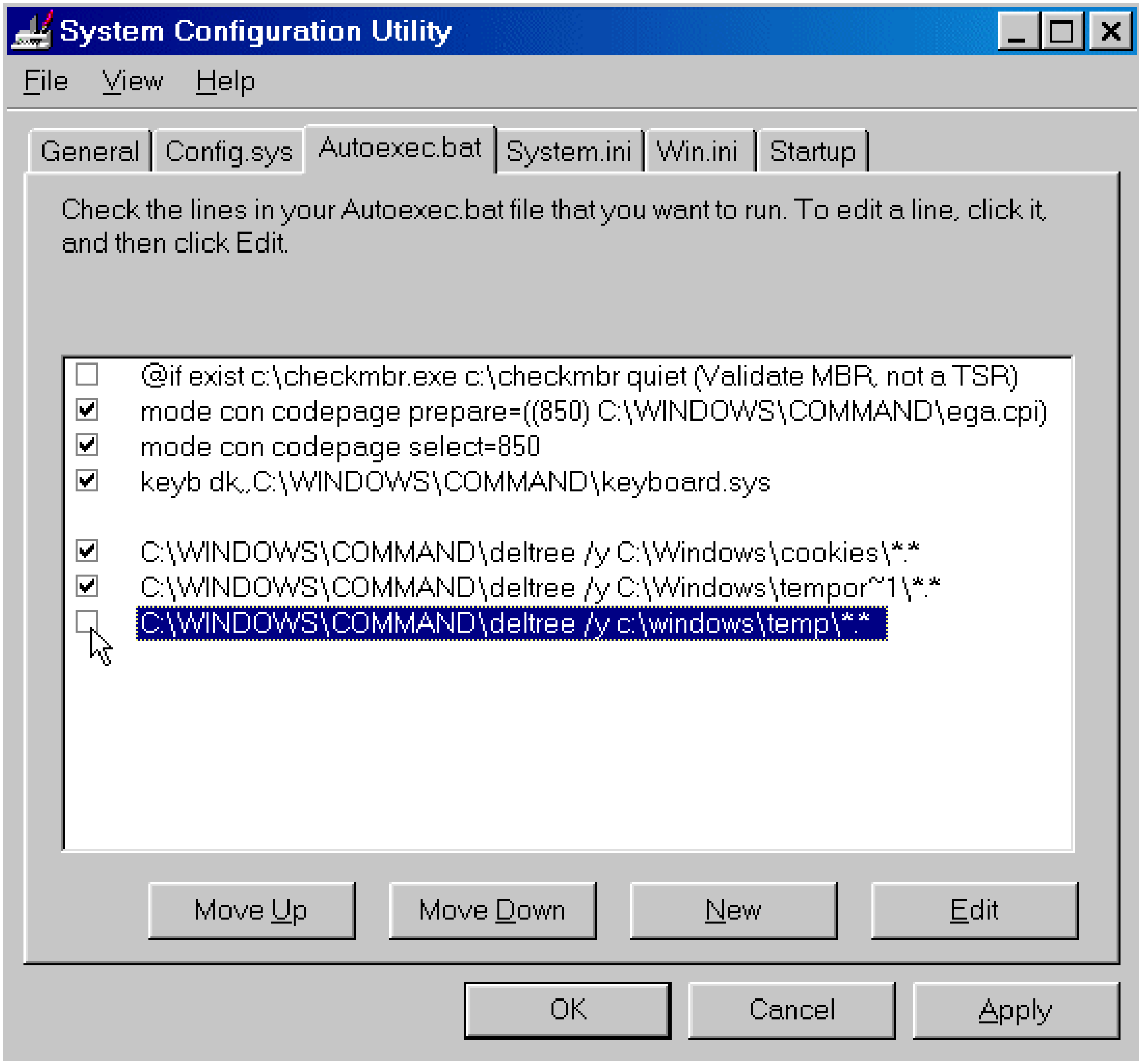
You have to use Start --> Run and type the command msconfig like this:



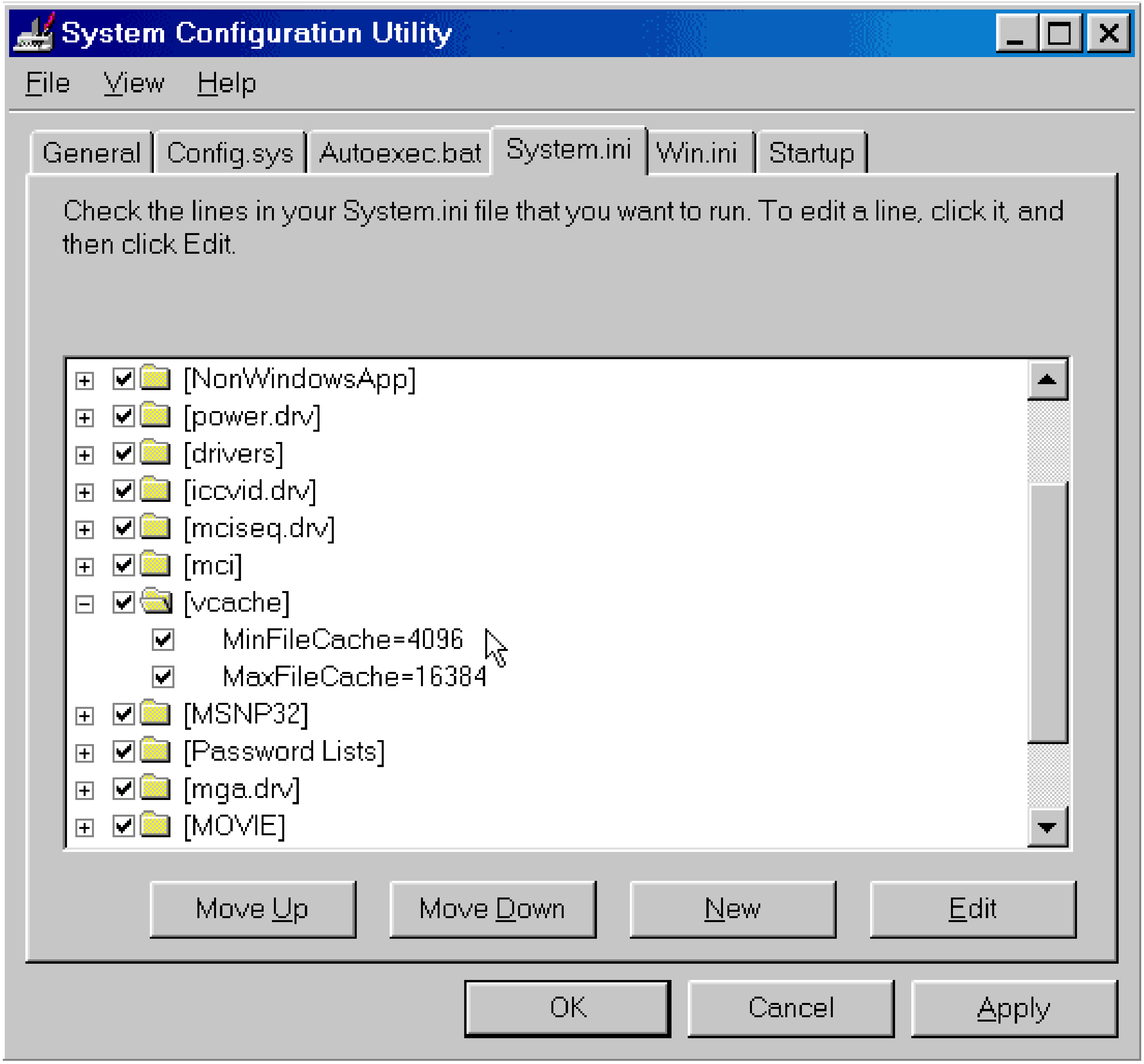
The program contains of six tabs. On the first, you choose Selective startup:



The second tab lets you enable/disable driver or other calls placed in Config.sys. The same is the case with tab number three Autoexec.bat. Here I see the three lines discussed earlier. Here I disable one of them:

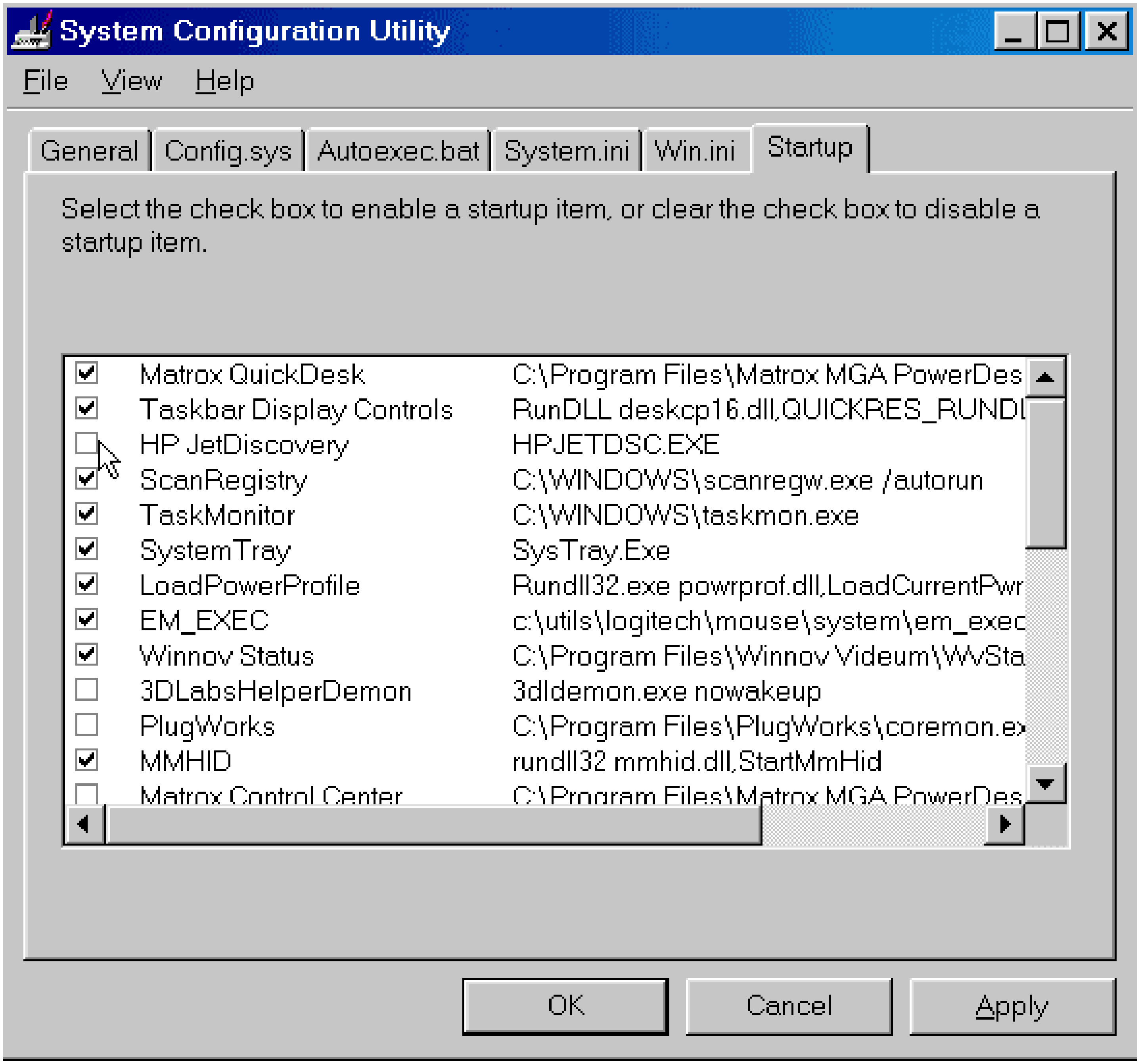


You can also add/edit/disable commands to System.ini. here you see my favorite setting for disk cache:



The commands in Startup

A very important tool is the last tab. Here you can disable programs, that normally are loaded during the Windows startup. This is a very handy tool for clearing some of those background program, since many of them are redundant. They are installed by varies types of software, but often they only delay the startup and represent a waste of memory and other ressources:



- Next page
- Previous page

Karbosguide.com. Software Tip 17

- Next page
- Previous page

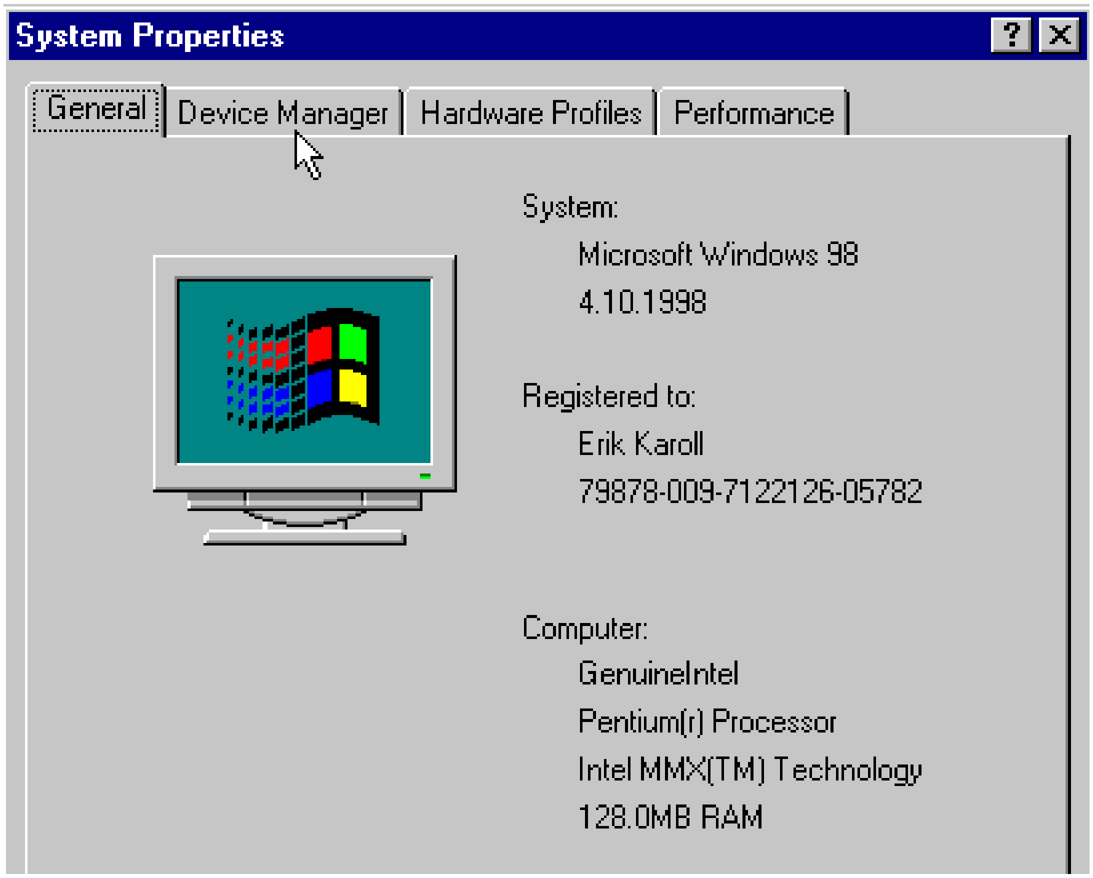
Enable DMA on your harddisk

It is extremely important to activate the DMA function in Windows!

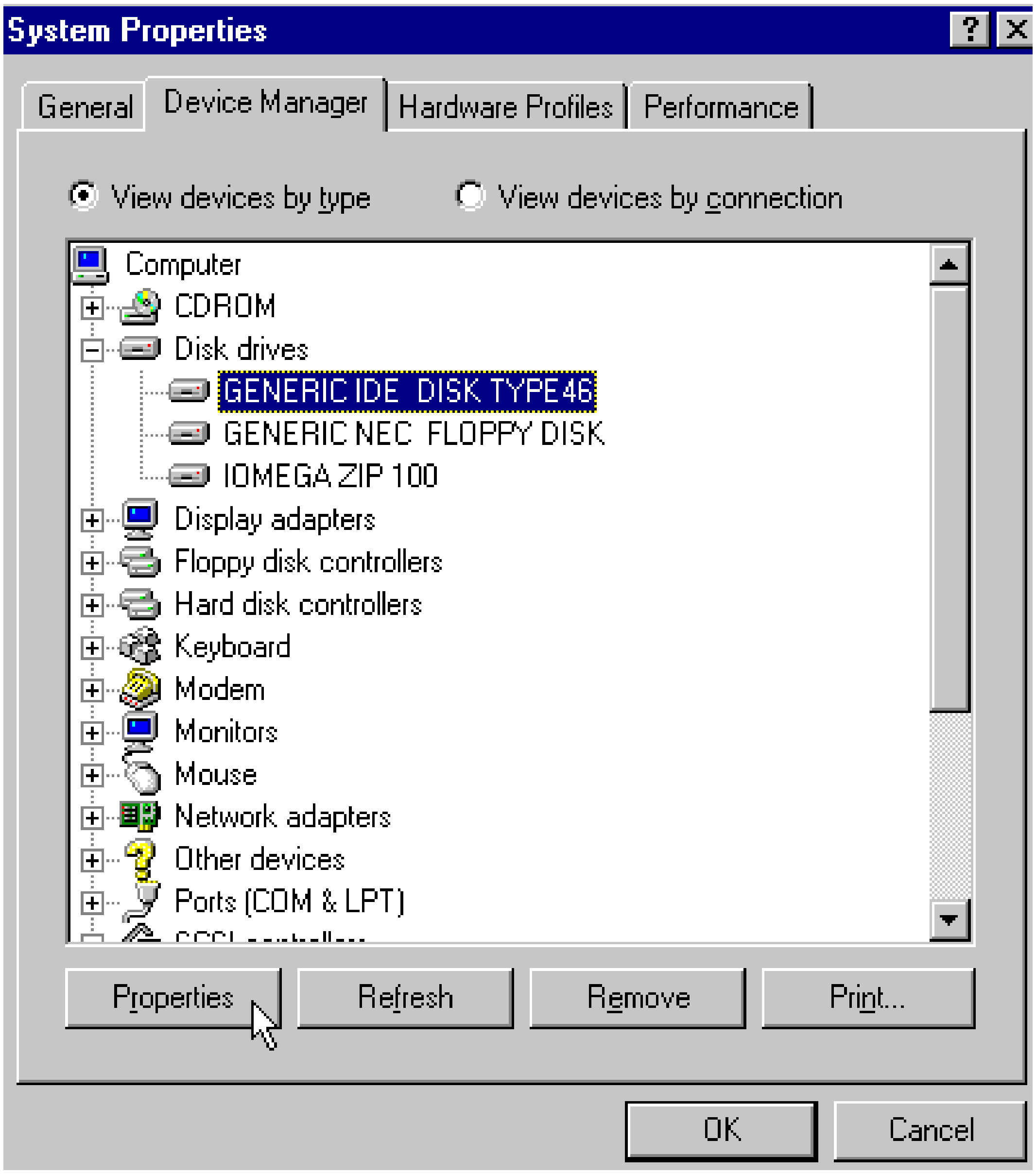
A few readers have reported problems with harddisk after enabling DMA; however I still recommend it since it gives such better performance. If you experience unstability, which will happen in some cases, please disable DMA.

Some of the combinations of motherboards and HDDs are unstable in this mode, and seems to be critical to the temperature of the HDD. Even 5 degrees celcius can have an impact on stability, and if the fan of your computer is dusty or underdimensioned, you might get this effect. Replacing a 40-wires HDD cable by 80-wires UDMA-66 cable will decrease the electrical noises in the IDE channel and make your computer more stable in DMA mode.

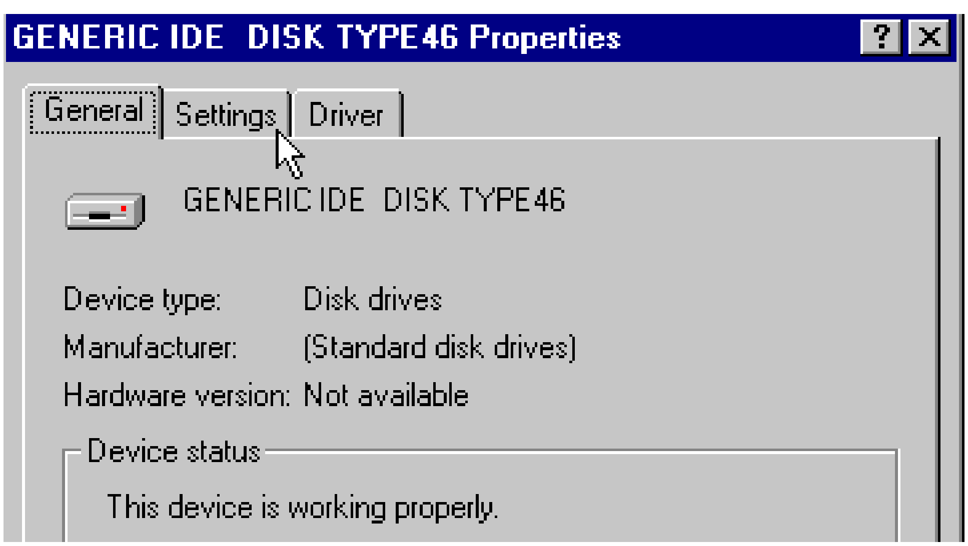
For example press Windows+Break to open the System Properties dialog box. Select the Device Manager tab:



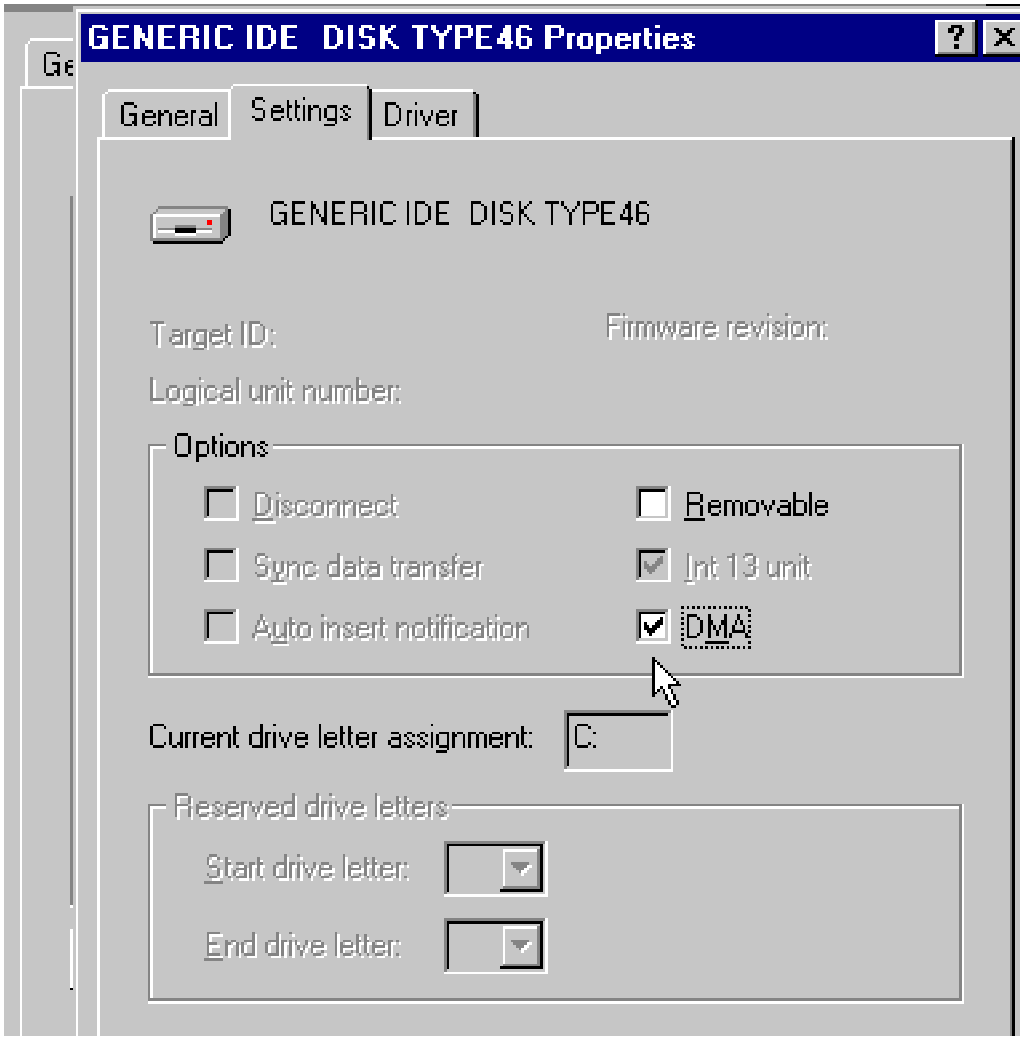
Double click on Disk drives, then select your hard disk and click on Properties:



Now select the Settings tab:



Verify that you have a check mark in the DMA box:



DMA transfer between hard disk and RAM is much more effective than when the DMA channel is not in use. With DMA a write or read operation can be executed in two to four clock ticks. Without DMA it will cost the CPU a minimum of 16 clock ticks per operation! That makes a big difference.

Notice that this enabling of DMA for hard disks and CD-ROM is far from automatic. In some motherboards it is activated automatically when Windows is installed, but that is probably the exception. So check for yourself that it is activated!

OBS: Some people report that enabling DMA to hard drives causes problems. On newer PCs with good cooling there should be no problems.

A tip from a reader

We never experienced problems ourselves from enabling DMA on any of our PCs. However here is a comment, that might prove a useful addendum:

You forgot one very important part that is required to enable DMA in Windows 98. After you check the DMA box in the device mamager- disk drive window, do the following:

To make sure that DMA doesn't cause any internal Windows conflicts (assuming the drive supports DMA), you are going to need to add these lines to the Mshdc.inf file at the bottom of the [ESDI_AddReg] tag (if they aren't already there):

HKR,,IDEDMADrive0,3,01
HKR,,IDEDMADrive1,3,01
HKR,,IDEDMADrive2,3,01
HKR,,IDEDMADrive3,3,01

Only add the bottom two lines (the bolded ones) if you have more than two IDE drives (HD/CD devices) connected to your system.

Within the DISKDRV.INF file, add the same lines under the [DiskReg] entry. If the lines are already there, you don't need to continue on, because you are done. Otherwise, copy the two modified files to another directory.

I recommend creating a new directory named inf2 in your windows system directory. Then remove all of the items under the Hard Drive Controllers and all of the Hard Drives (only hard drives, not removable media) from under the Disk Drives heading.

Then reboot your system.

When the computer's Hardware Wizard pops up to reinstall the drives, set it so you can choose from the compatible drivers list and check the date on the drivers. They should be for the same day you modified them. If they are, great, use them, otherwise you will need to point them at the copies (the inf folder is invisible at this point, which is why you need the second copies) and use those. Thanks guys!

David Gillespie

--- We hope this will be useful, and thanks to David!

- Next page
- Previous page

Karbosguide.com. Software Tip 16

- Next page
- Previous page

Running out of space on my hard disk...

People often think they lack space on their hard disk. But just as often the *real* problem is that their data and programs are terribly disorganized. The most common errors are:

Wrong partitioning of the hard disk.

In Windows 98 you should only have one partition, a C drive, where all data are located. The subdivision in D- and E- drives will in most cases add to wasted space and mess.

Duplicates

If program packages like Netscape Communicator, Internet Explorer, not to speak of Microsoft Office are present in different versions - no wonder you run out of space. I see that in many places.

General mess

Lack of structure and disk cleaning. People install all kinds of programs retrieved from the Internet and borrow from each other - of course that is fun. But if you do not have a firm division of the hard disk for different kinds of data, it might be impossible to clean up when you need to.

A model for hard disk space utilization

Your hard disk should be one big partition (a C drive), which in Windows 98 and 2000 often is formatted with FAT32. Then *all* installation and other data storage is done according to these guide lines:

Program folders

The folder C:\Windows contains what the name implies.

- C:\Program Files contains all program packages.

- C:\Utils contains *all* small programs (fax, cookie killers etc.), which do not install in C:\Program Files.

- C:\Temp contains all the temporary stuff, whatever is in for review and testing etc.

- C:\My Documents\Download\Program1 is temporary and contains the downloaded and *extracted* files for Program1.

When programs are "accepted" as being of lasting value, they are moved to C:\Utils\Program1 etc. And the temporary folders are deleted.

Program files from CD-ROM

The folder C:\Disks contains installation diskettes/CD-ROM's for important programs. Though you should keep the CD-ROMs coming with the programs, it is much more convenient to have a copy on the hardisk.

Your documents

A folder like C:\Texts contain all user documents set up in a *dynamic* structure of multiple sub folders. You continually need to create new sub folders and move and delete among the older ones, so your document structure fits your work needs.

I prefer *not* to use C:\My documents for storage of documents. This folder is involved in many operations from within Windows and other programs; hence it often gets cluttered with a lot of files which are of temporary character. Use a folder like C:\Texts for all your documents, images etc. having a permanent character.

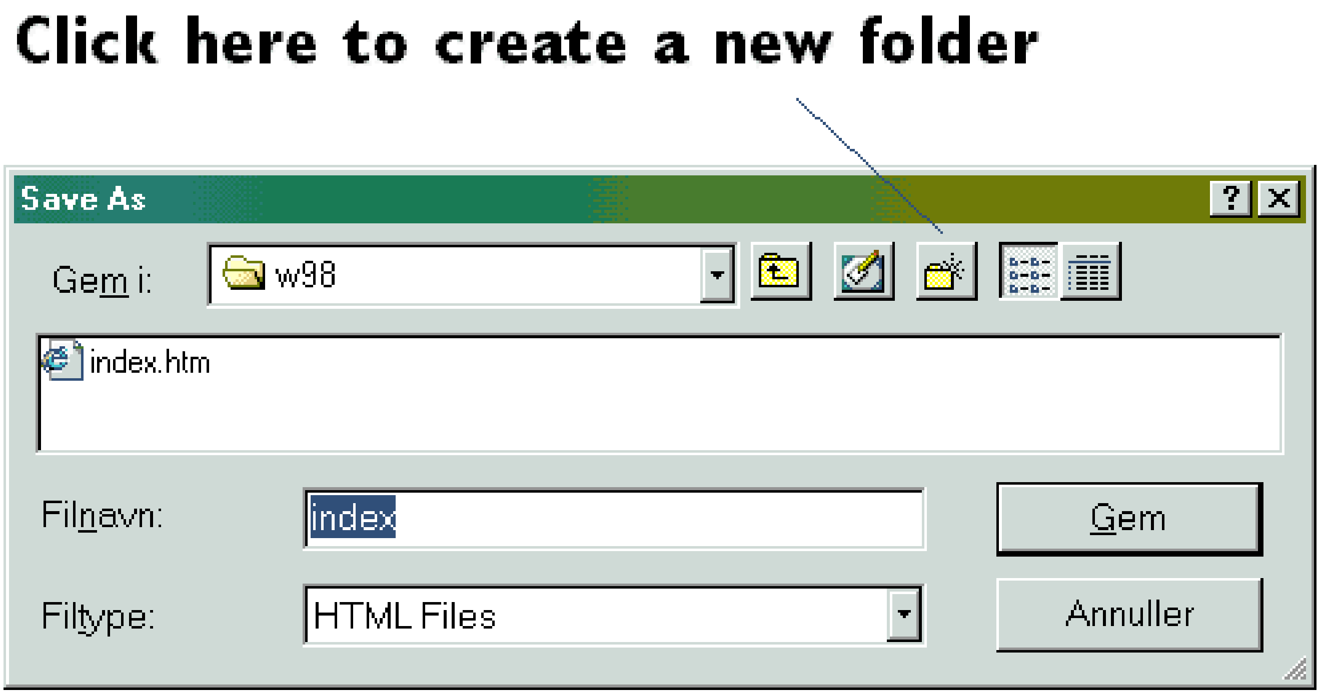
- D:\Backup contains *all* backups.

You should backup your documents daily/weekly on another harddisk than where the originals are kept.

The point is to maintain the system one hundred and ten percent!!! And that is really not hard - it just requires some presence of mind and careful thought during download/installation.

About creating folders

Always remember to create new folders - you can do that directly from the download dialog box as well as from the Save As dialog box. Use this button:



Is it worth the effort?

All this boring trouble - like having to plan ahead every time. It pays off ten fold in stability and surplus in daily work. The alternative is repeated reinstallation of the whole shebang, data loss and wasted time. That is just the way it is . I speak from more than 12 years experience.

- Next page
- Previous page

Karbosguide.com. Software Tip 15

MacroExpress 98

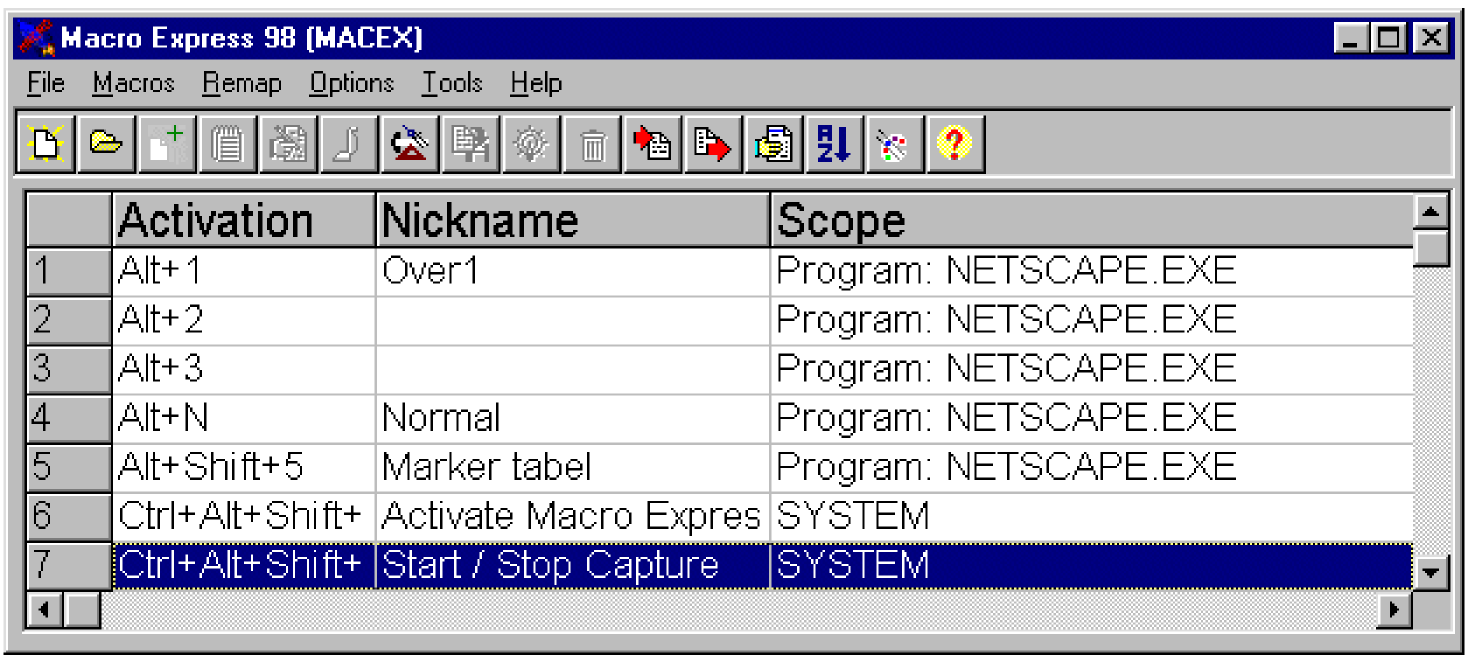
- Next page
- Previous page

This program is quite smart. It is about developing macros, which can handle the repetitive keyboard/mouse operations.

A small, very simple example: If I write my home pages in an editor like Composer; I have to forever format headings with the *heading* 1, 2, and 3 formats. The only way I can do that is through the menu or tool bar. To grab the mouse every time is not practical while I write, and menu choices require some complicated keyboard entries. All I wish is to make:

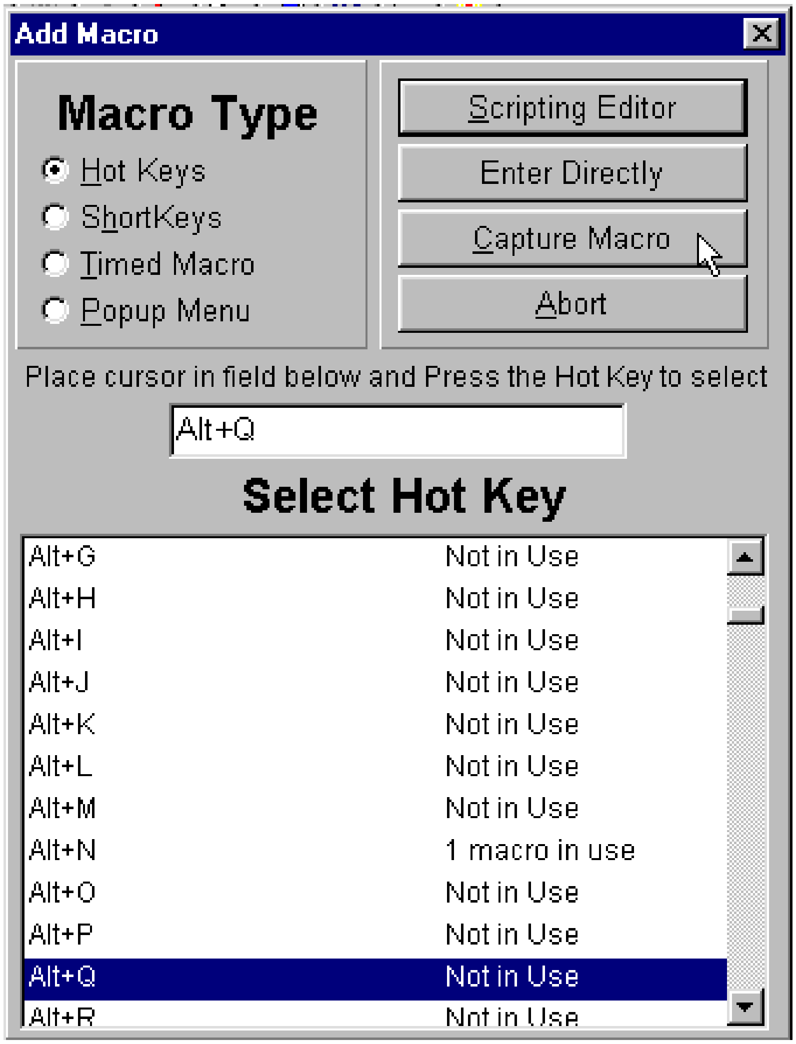
- [Alt]+1 give heading 1
- [Alt]+2 give heading 2
- [Alt]+3 give heading 3

That's how the *styles* work in Word. But with MacroExpress 98 it is very easy to add this functionality. You see the screen image here:

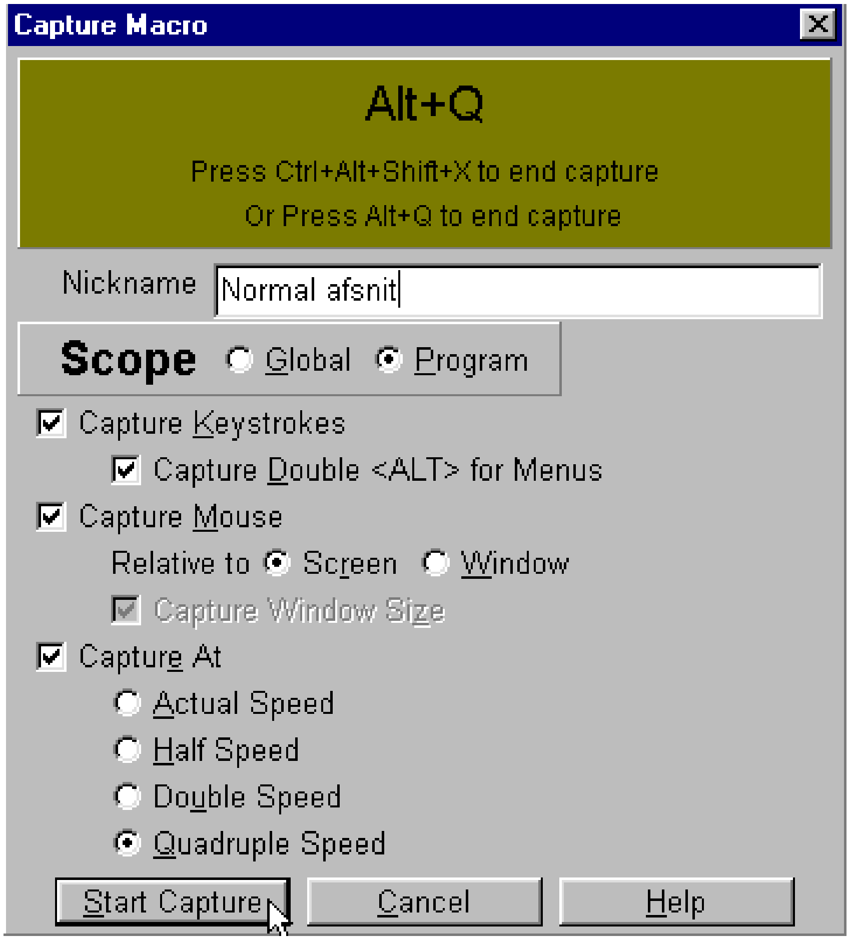


You can select whether the macros will work in *all* programs, or whether they only work in a designated program. Probably the last option is the smartest. I need to add that the program places some demands on the user. It takes a little while to understand how it works. You see MacroExpress 98 *records* an image of for example your mouse movements. So if you for example change the menus or program set-up, you might risk that the results are quite different when you play the macro. Here is a brief description:

You start the recording in this screen image. Here I will add a macro to the keyboard command [Alt]+q:



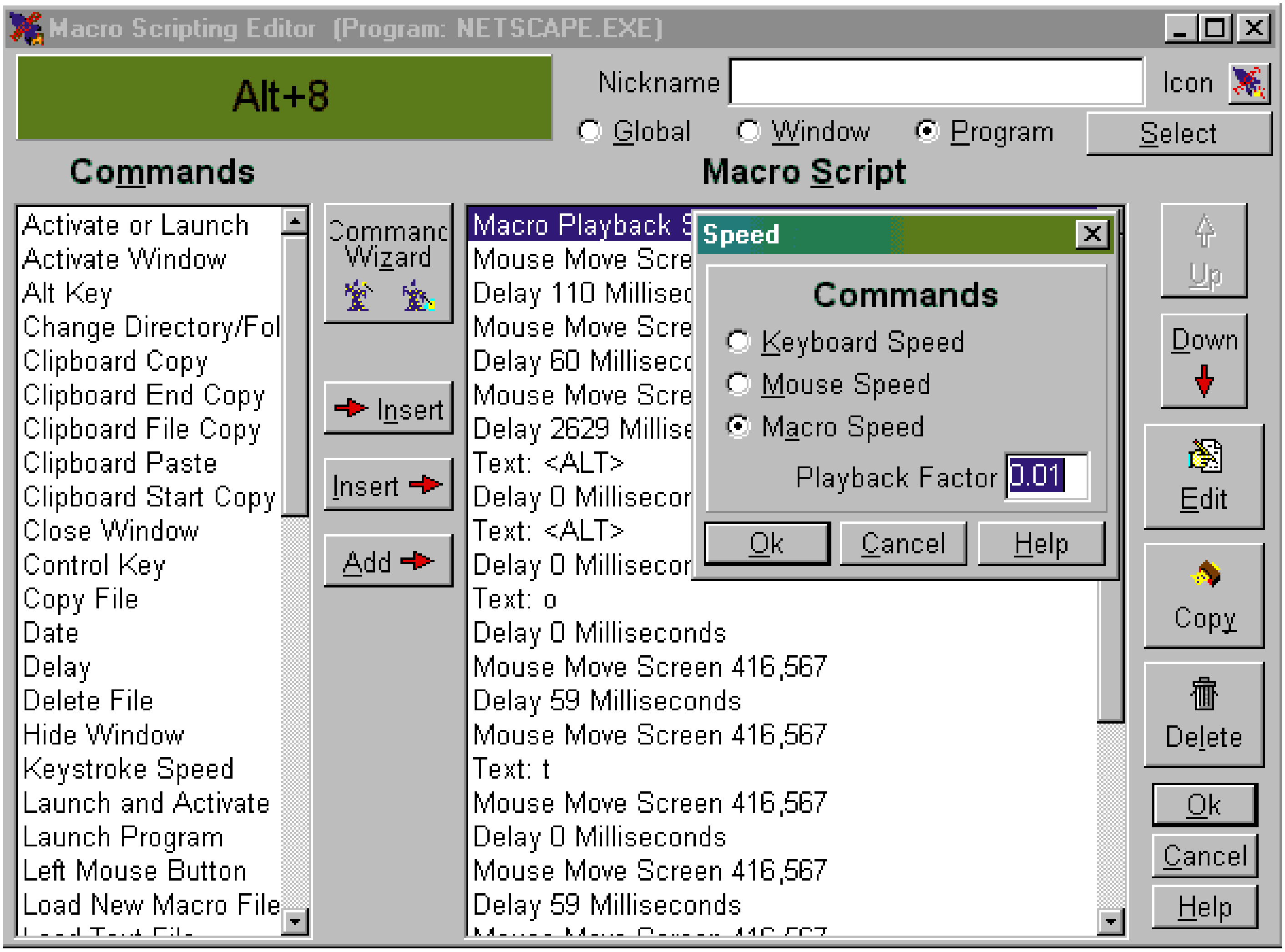
The next image looks like this:



Then the macro is recorded through the keyboard/mouse commands for that program. When you are finished, you close by pressing [Alt]+q (in this case).

Then the macro works. When you are in the given program you press [Alt]+q, and the macro plays. If you used the mouse during recording, you will see (surely to your great surprise) the mouse work on its own. It takes a little while, but the macro works just like you recorded it.

Now it just remains to speed it up. I do that by double clicking on the macro. The first command regards the speed, and I set the value to 0.01, to make the macro work at the speed of lightning. I can also delete all *Delays*:



I can recommend the two programs here - they are invaluable if you work a lot with your PC.

It is even easier, when you can type in the keys directly:

- Next page
- Previous page

Karbosguide.com. Software Tip 14

- Next page
- Previous page

Windows Autotexts

It is terribly impractical to have to repeat the same operations over and over again. Some programs (like Word) gives lots of options for customizing their execution with *macros, autotexts,* and *keyboard shortcuts*. You can read about that in "Word 97 short course," where I have given much emphasis on these options.

But other - many other- Windows programs do *not* offer these macro etc. options. That is the case with a program like Composer, which I use to write my home pages. There is no question that the program is excellent for writing home pages. But its functionality is quite primitive. There are hardly any keyboard shortcuts, and no macros. Many Windows programs are designed like this, and it can really be quite trying to fight your way through a larger assignment.

You can get help. I am talking about two programs, which can spruce up any Windows program. I can equip Composer with just the same autotexts and keyboard shortcuts, which I enjoy so much in Word:

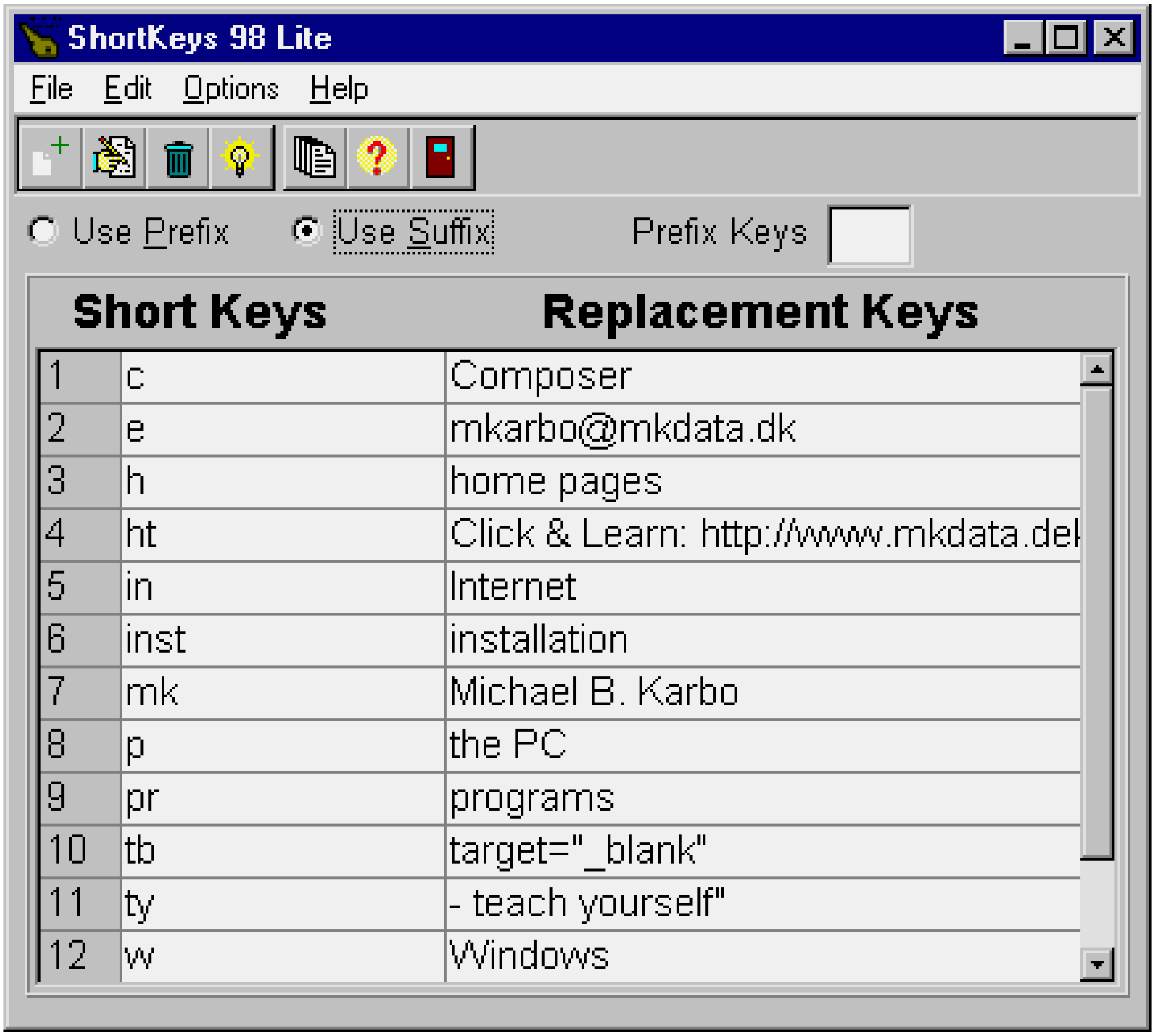
- ShortKeys 98
- MacroExpress 98

Both programs require some getting used to, but then they really *work*. Unfortunately only for 30 days, unless you pay for them. Let me show you some examples of how I use them. First the simplest of the programs:

ShortKeys 98

[top]

ShortKeys is quite similar to the Autotext function in the word processing program Word. The idea is that you make an abbreviation of those words or expressions you use most frequently. The smart thing about ShortKeys is that it works in *all* Windows programs. Look at this illustration:



Every time I need to write the word Windows, I just press w. When I hit the space bar after the w, the abbreviation is replaced immediately and automatically with "Windows". Likewise:

- h gives the text home pages,
- ht give gives the text Click & Learn: http://www.mkdata.dk/english
- etc.

You just have to get used to the program, since it sometimes may surprise you. But that is soon corrected, and it really saves you time. That is certainly worth its $20 - register it right away!

- Next page
- Previous page

Karbosguide.com. Software Tip 13

The FAX program - what happened in Windows 98?

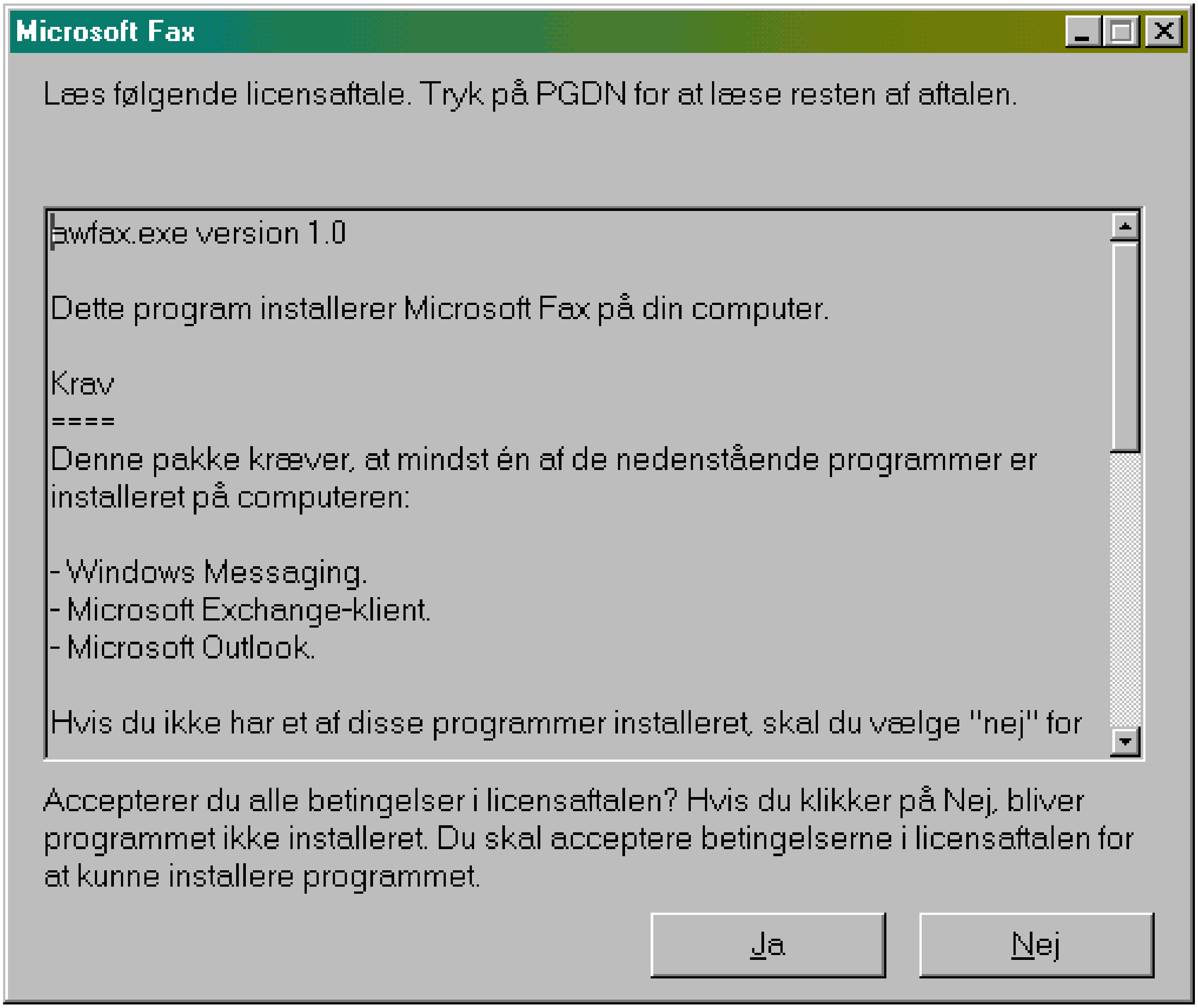
- Next page
- Previous page

Windows FAX is the program which enables you to *send* fax using your modem. That is a smart item, since fax by modem can appear much more appealing to the receiver than fax from a machine. Machine fax has to be scanned first, and that results in a loss of quality.

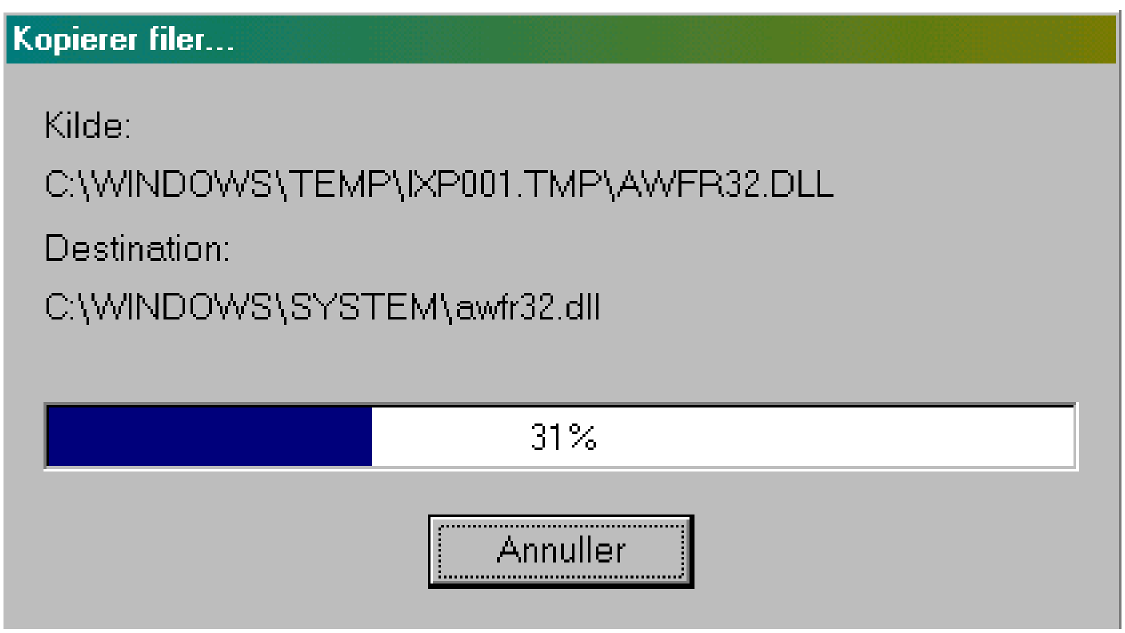
The FAX function was quite simple to install in Windows 95, but for some reason that program has been omitted in Windows 98. Well it is not really gone, but you have to look hard for it. Then you have to install it yourself. Peculiar, but here is how:

On the Windows 98 CD-ROM you need to enter the folder \WIN98\TOOLS\OLDWIN95\MESSAGE. There you find the Awfax program, and you need to click on that:

After you have accepted the license agreement, you can now install Microsoft Fax:



Now follows the installation:



After you restart your PC, Microsoft Fax will be ready to use. It works by installing an extra *printer*:

When you have written a document in your word processor, you can print that to the fax instead of your regular printer. I hope this works (my own ISDN modem can not work with fax).

One question remains: Why all these capers to hide the program so far out of the way in Windows 98? God's ways are unfathomable!

- Next page
- Previous page

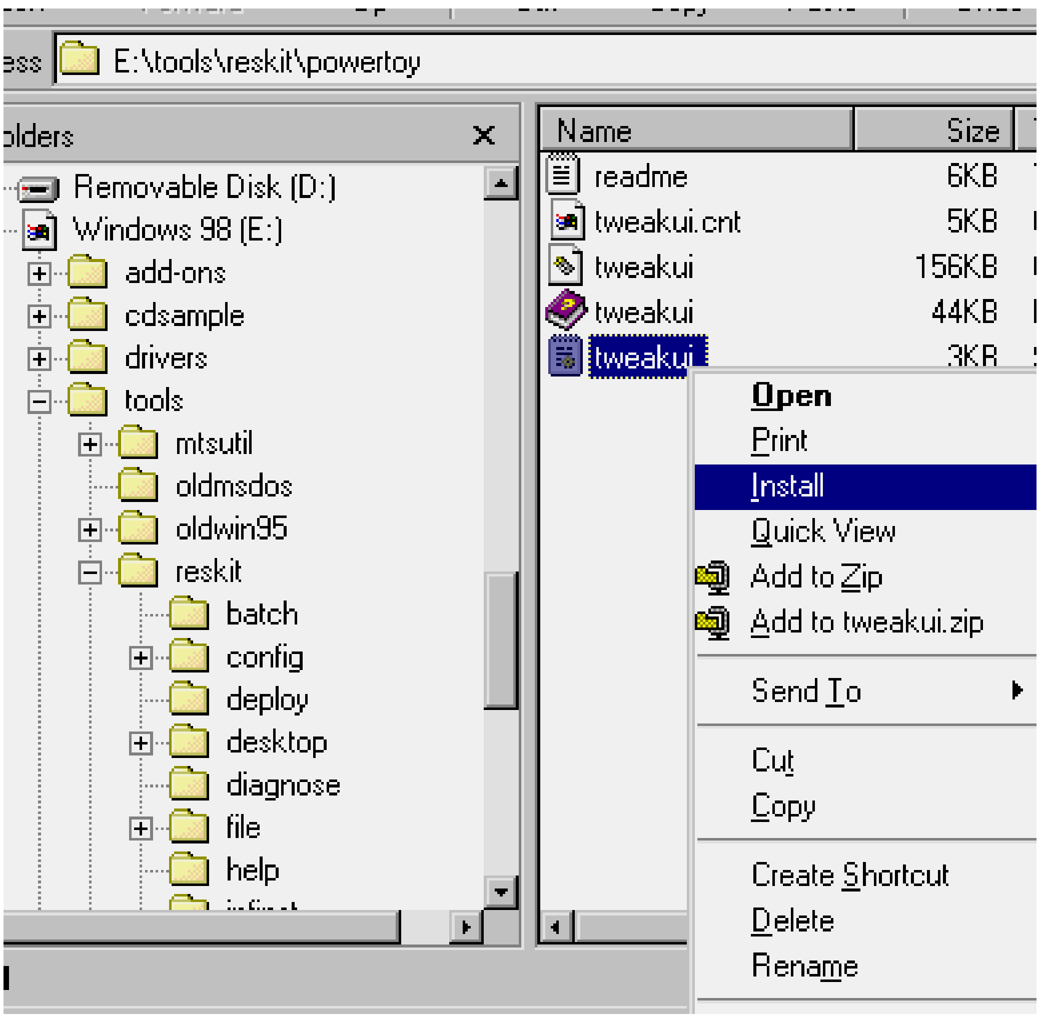
Karbosguide.com. Software Tip 12

Tweak UI

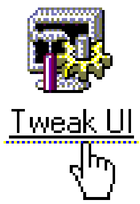
- Next page
- Previous page

TWEAK UI, which is found on the 98 CD, is a program to customize the user interface. It is included in the collection of small programs, which is known as "Power toys." In Windows 98 these programs have been included in the CD-ROM, but they are still well hidden. First look at the installation here.

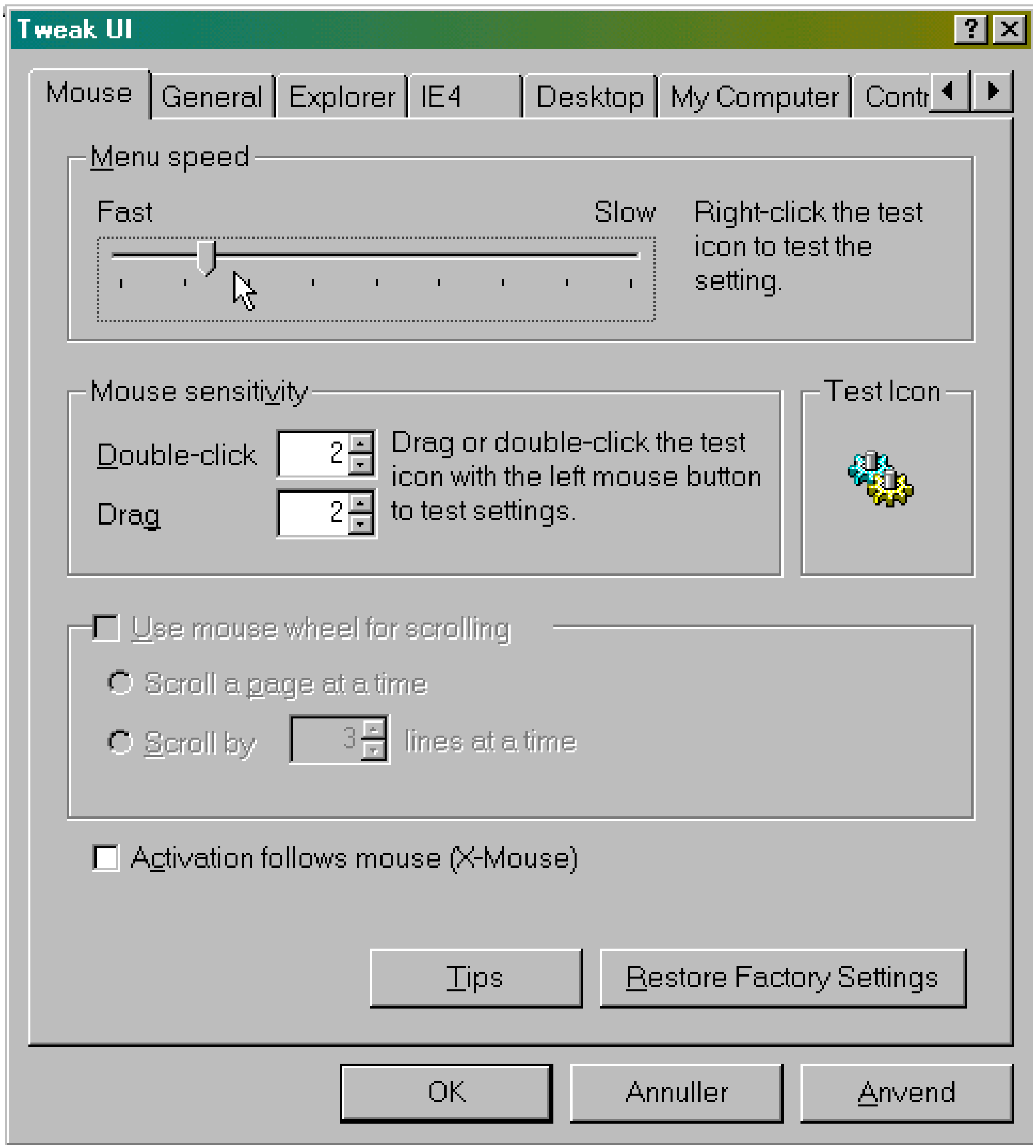
You need to find an installation file named tweakui.inf. Right click on that and select Install:



Then the program will be installed quickly, and now you can find it in the Control Panel:



Here you can change a lot of parameters concerning the behavior of the user interface:



Enjoy yourself!

- Next page
- Previous page

Karbosguide.com. Software Tip 11

Single click in Explorer

- Next page
- Previous page

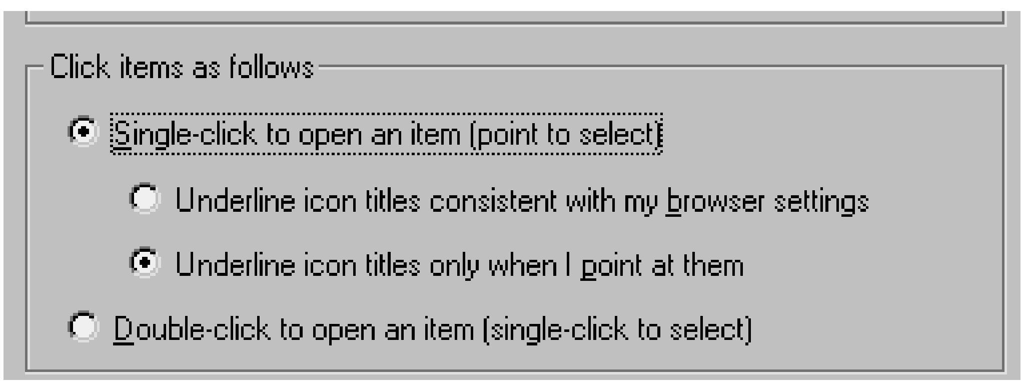
I use the single click system in Explorer, as you can see in some of the screen dumps.

This causes the folder names to be underlined, which means that they work as links on a home page. One click opens them. Try that.

To activate it you need to select View -> Folder Options in Explorer.

On the General tab select:

Then click Settings..., and select:



Click OK and it works. You just have to get used to the new interface. You just select objects by letting the mouse cursor rest on them - without clicking. Give it a chance!

Moreover you can adjust the "reaction time," that is the time between touching an object and selecting it. You need to find the program TWEAK UI, which is on the 98 CD. There you can adjust "menu speed" - increase it. It causes the selection to appear faster. Now there is no excuse for the impatient souls, who might think that the pointing method is too slow compared to the traditional click. Read later about TWEAK UI.

- Next page
- Previous page

---

y

Karbosguide.com. Software Tip 10

- Next page
- Previous page

Folders with names in capital letters

A (small) novelty in Windows 98 is that you can use capital letters in folder names. Notice the second to last folder in the picture below:

Windows 95 did not allow that, but now it is acceptable.

Personally I should prefer, that only smallcaps were allowed both in files and folders names. It would make things a lot easier.

You should also remember that on many web-servers (Unix/Linux-based) there is a difference between lower cases and capitals. The file Index.htm and the file index.htm are not the same. Therefore it is good only to use lower cases, both in file and folder naming. If you mix the cases it can be difficult to remember whichyou have used.

- Next page
- Previous page

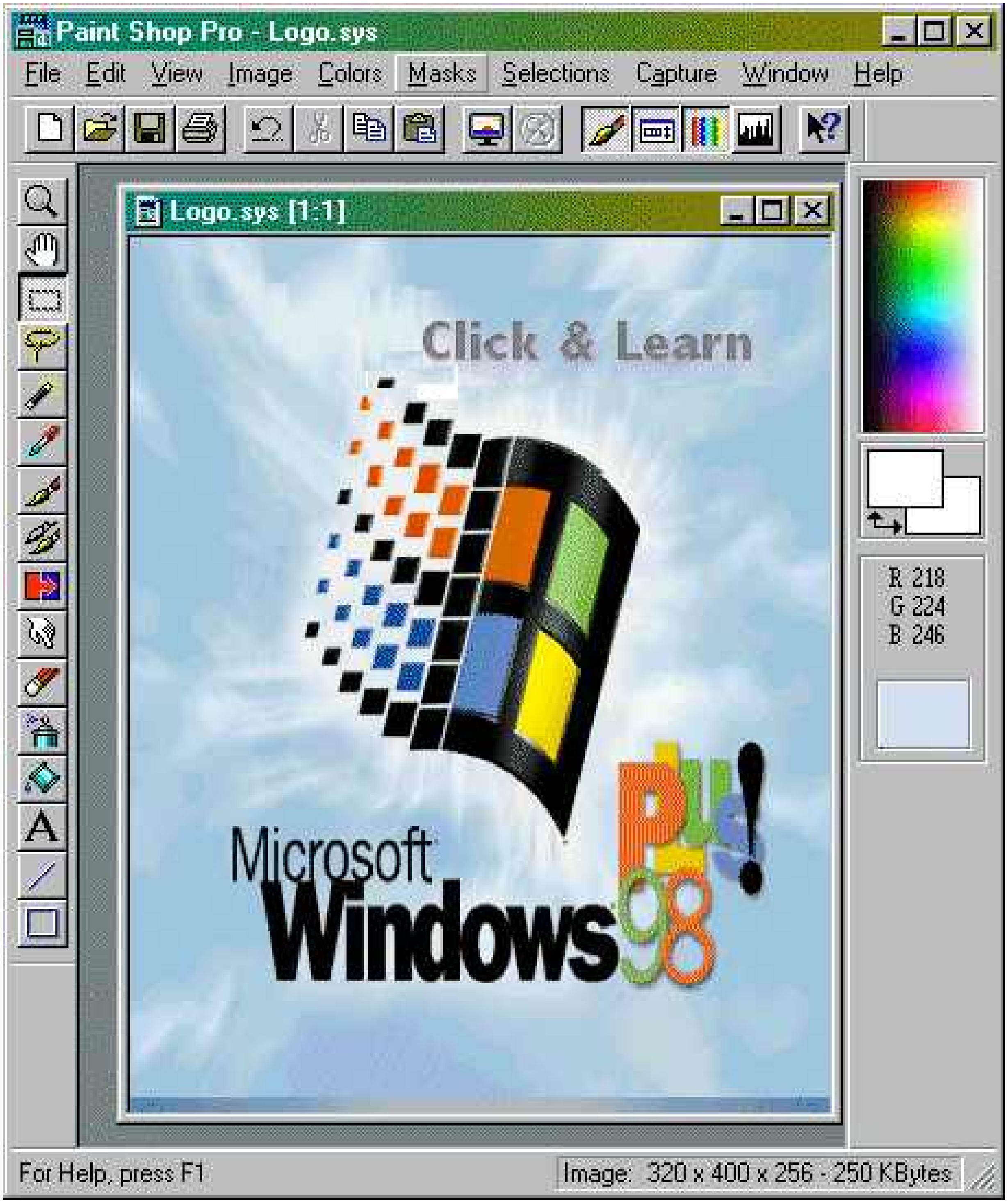
Karbosguide.com. Software Tip 9

Change the Windows opening screen

When you turn your PC on you see a screen like this:

Since it a simple piece of bit map graphics, you can alter this image. The file is named C:\Logo.sys, which means it is in the root directory. Open it in a graphics program, and alter it.

Just beware that this picture is in a special format, that is 320 X 400 pixels. When shown, it is scaled up to double width like 640 X 400. You must consider that when you modify the picture - it will be distorted when shown. Here you see the picture during editing:



If you use Windows 98, you have to create the file Logo.sys yourself and place it in the root directory. It will work then as well.

In the same manner you can change the closing picture. It is named C:\Windows\Logow.sys.

You can clearly see the peculiar tall format:

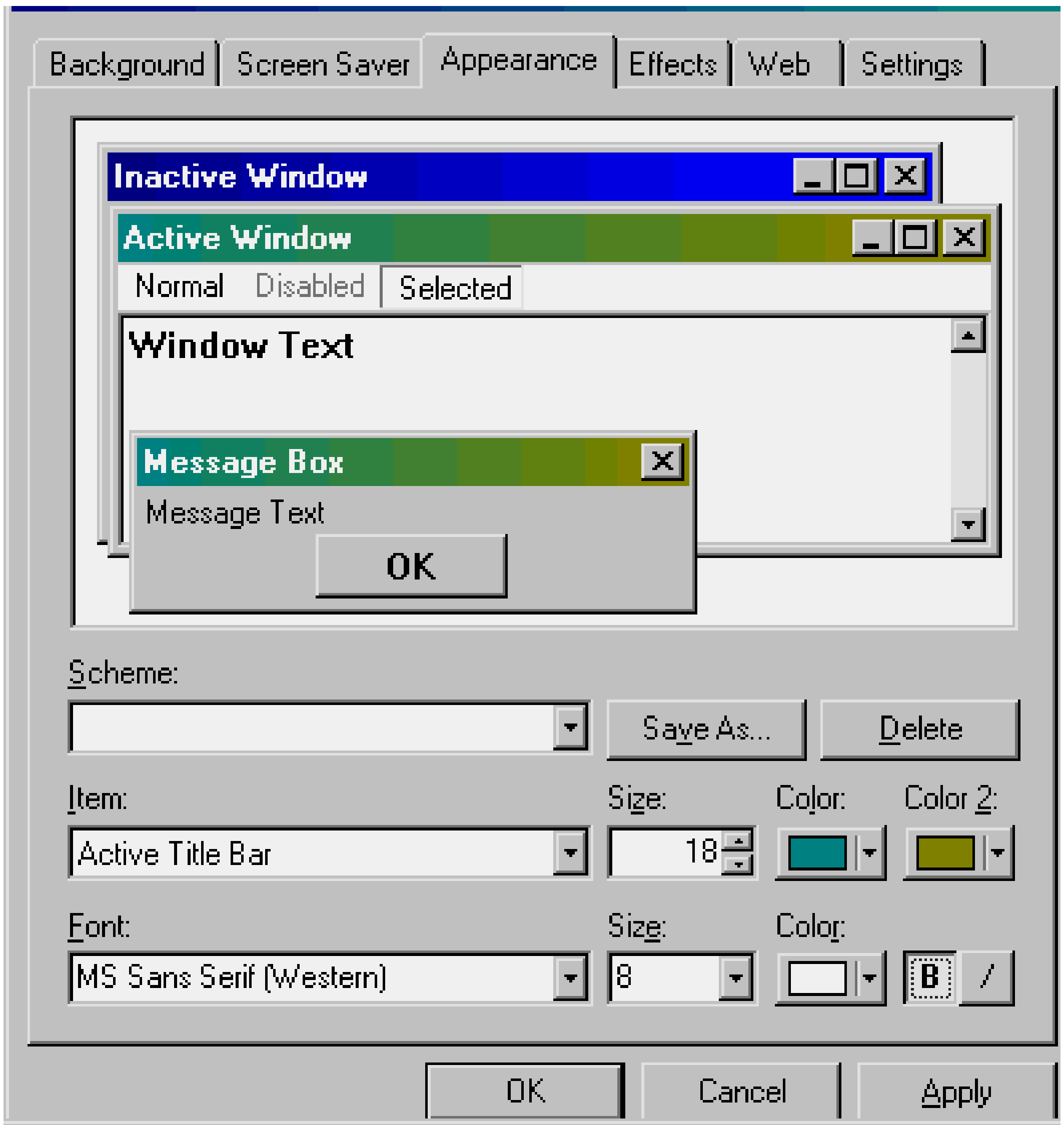
- Next page
- Previous page

Karbosguide.com. Software Tip 8

- Next page
- Previous page

Color changes in the title bar

A new detail in the user interface is that you can select different colors for different screen window portions. Select Display Properties -> Appearance. Select Menu and color 1, then select Active Window and color 2. as shown below:



That gives quite a neat effect.

- Next page
- Previous page

Karbosguide.com. Software Tip 7

Choose your own start page

- Next page
- Previous page

When you open your browser, some kind of home page is automatically written in. That is true regardless of which browser you use. It is not quite unimportant which home page you choose, and actually you have four options:

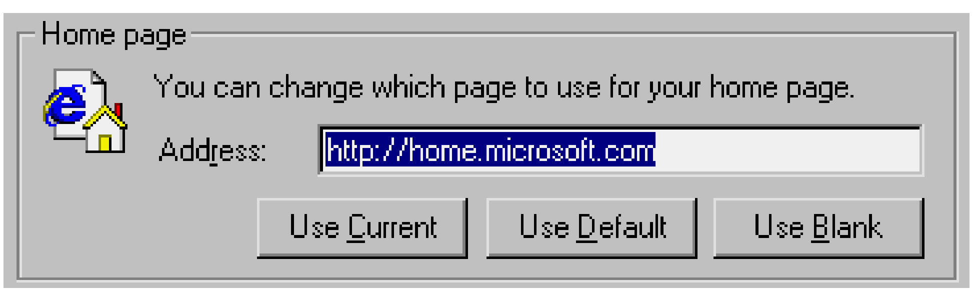
- A blank page
- A home page saved in your PC
- Some fixed page from the net
- The browsers default page.

I will highly recommend that you experiment with these options. The **blank page** is the easiest way out. The browser starts with a blank page and is ready immediately. A **local home page** is a good option. You make your own home page with the links you use most frequently. Then the browser will open that as your starting page. I have such a local home page myself, but it is on my desktop as is described elsewhere here.

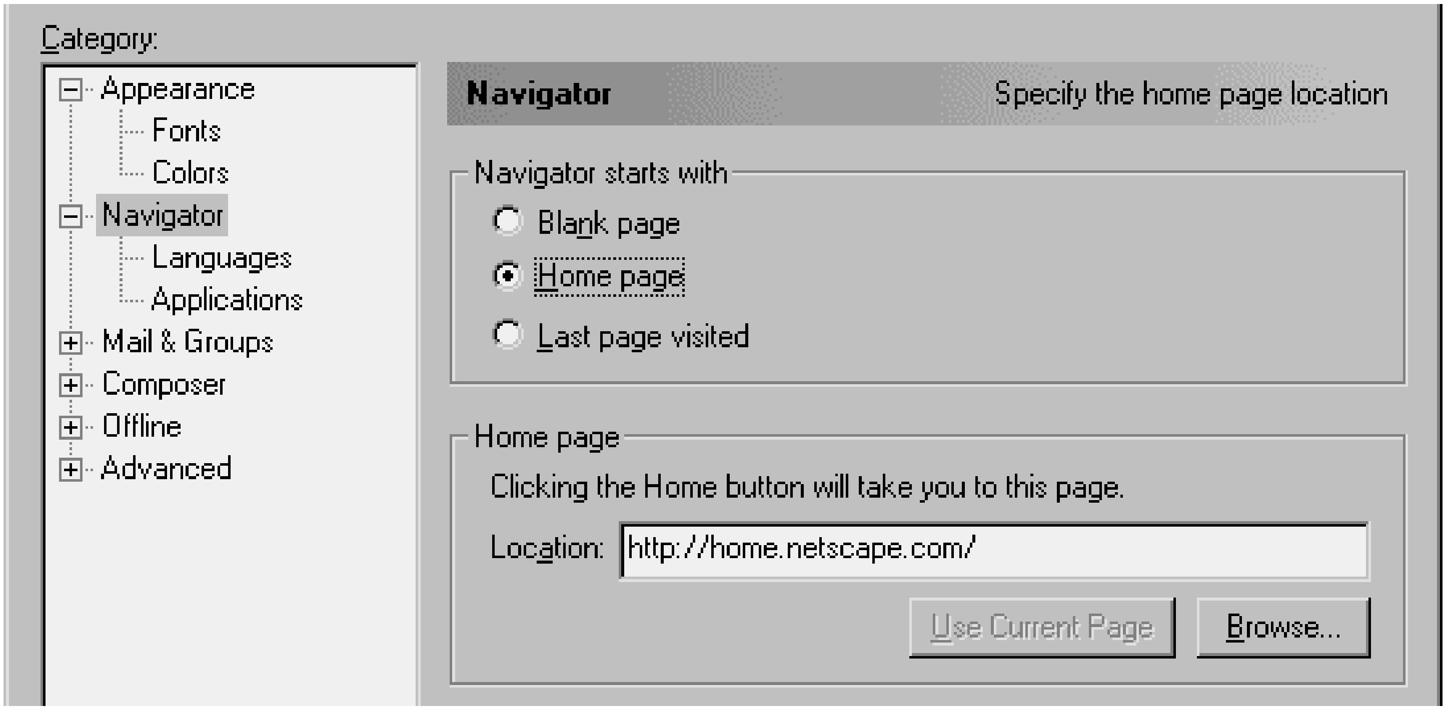
The **fixed start page from the net** can be an exciting enterprise. There are really many who like for you to use *their* web site as start page. That gives many visits to their page, which in turn signals success and might result in earnings through advertisements.

Finally you can choose to let the browser select a start page - but that is least desirable solution. Neither Microsoft's nor Netscape's start pages are particularly exciting. Make a better choice yourself!

How do you do this? Well, in Internet Explorer you need to select View -> Internet Options. The rest is easy:



Above the easiest is to browse to the page you want as your start page and then click on the button Use Current. In Netscape Navigator you similarly select Edit, Preferences...:



Here it is similarly easy to browse to the start page and then click on the button Use Current Page.

- Next page
- Previous page

Karbosguide.com. Software Tip 6

Fixed folder for downloading in Internet Explorer

- Next page
- Previous page

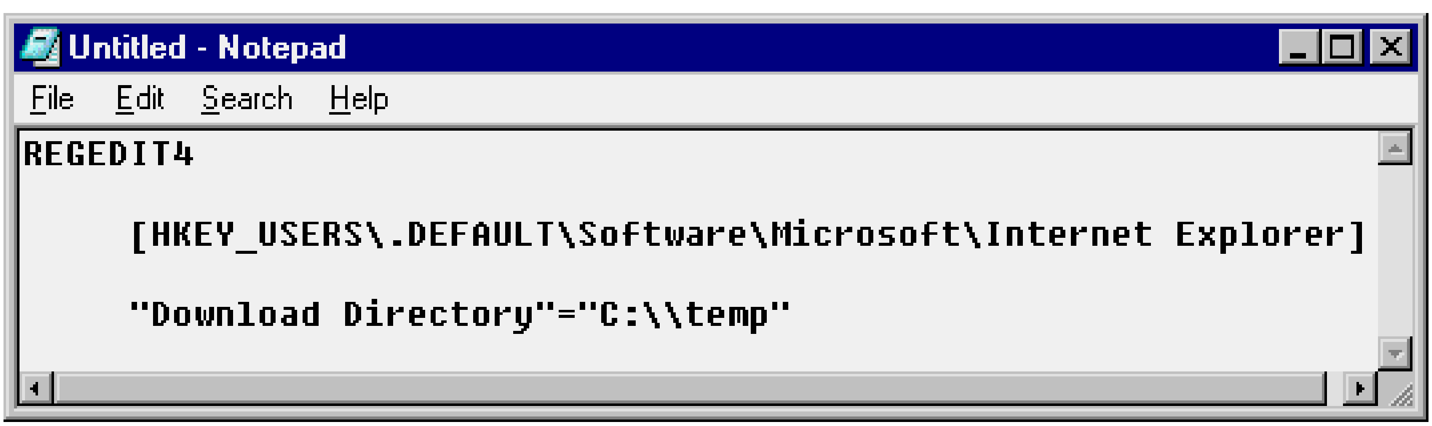
When you are in the Internet with Internet Explorer and look at a file, which is not a home page, you have the option to download the file. But in which folder does it end up? Maybe you have noticed that it appears to be random. Or maybe you have discovered that it always saves in the *last* folder you downloaded to. You see Windows saves that information.

I want Internet Explorer to always suggest C:\temp for saving downloads, since I always use that folder for this purpose. So how do I make the Internet Explorer program accept that? Unfortunately I need to go to the registry database. That is just a small thing, and here I will show you how:

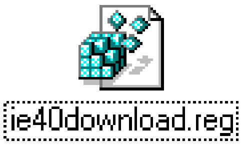
Open a new text document in Notepad, or whatever program you use for that kind of thing. Then write these lines, precisely as you see them here (It is easier to copy them from here, then you will get no errors):

```
REGEDIT4

[HKEY_USERS\.DEFAULT\Software\Microsoft\Internet Explorer]

"Download Directory"="C:\\temp"
```

Here you see it:



Now save the file, for example on the desktop. It MUST have the suffix .REG. I call it ie40download.reg. The moment you rename the suffix, the file will be included in the registry database, as you can see from the icon:



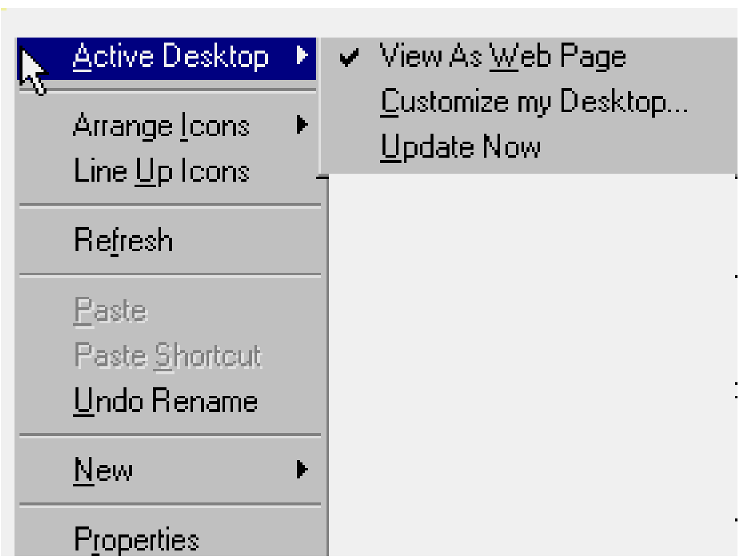
Now you just have to double click on the file to run it. Confirm with OK twice, then everything is in its place. In the future the download folder will be C:\temp (or whatever you wrote in the file). This remains true until you save in another folder. Then you have to run the file again to re-establish C:\temp as the default folder.

- Next page
- Previous page

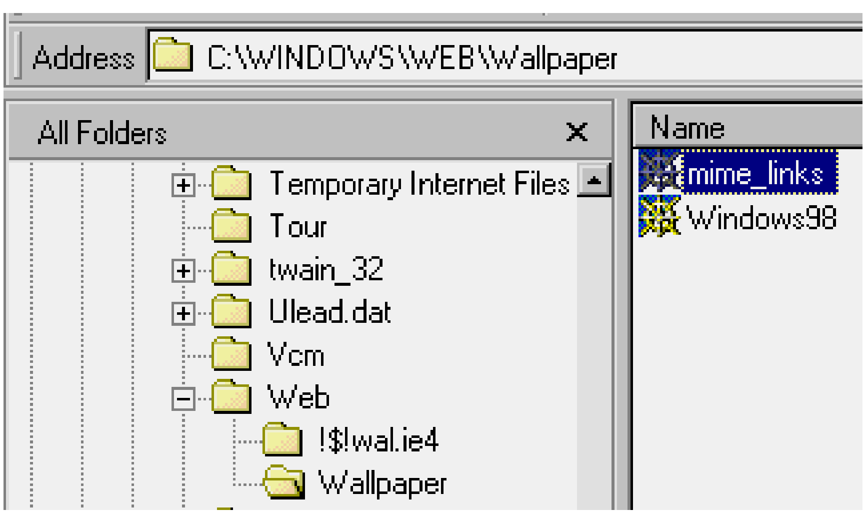
Karbosguide.com. Software Tip 5

Use the desktop

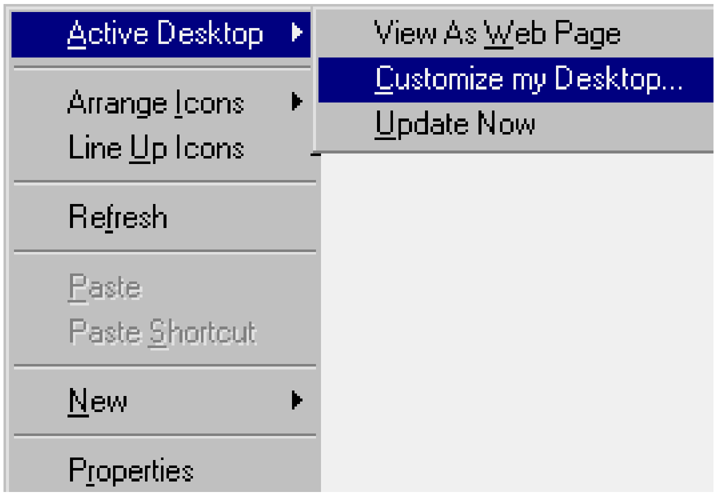
If you collect all your favorite Internet addresses in an HTML document, you can select that as background. Start by turning Active Desktop on:



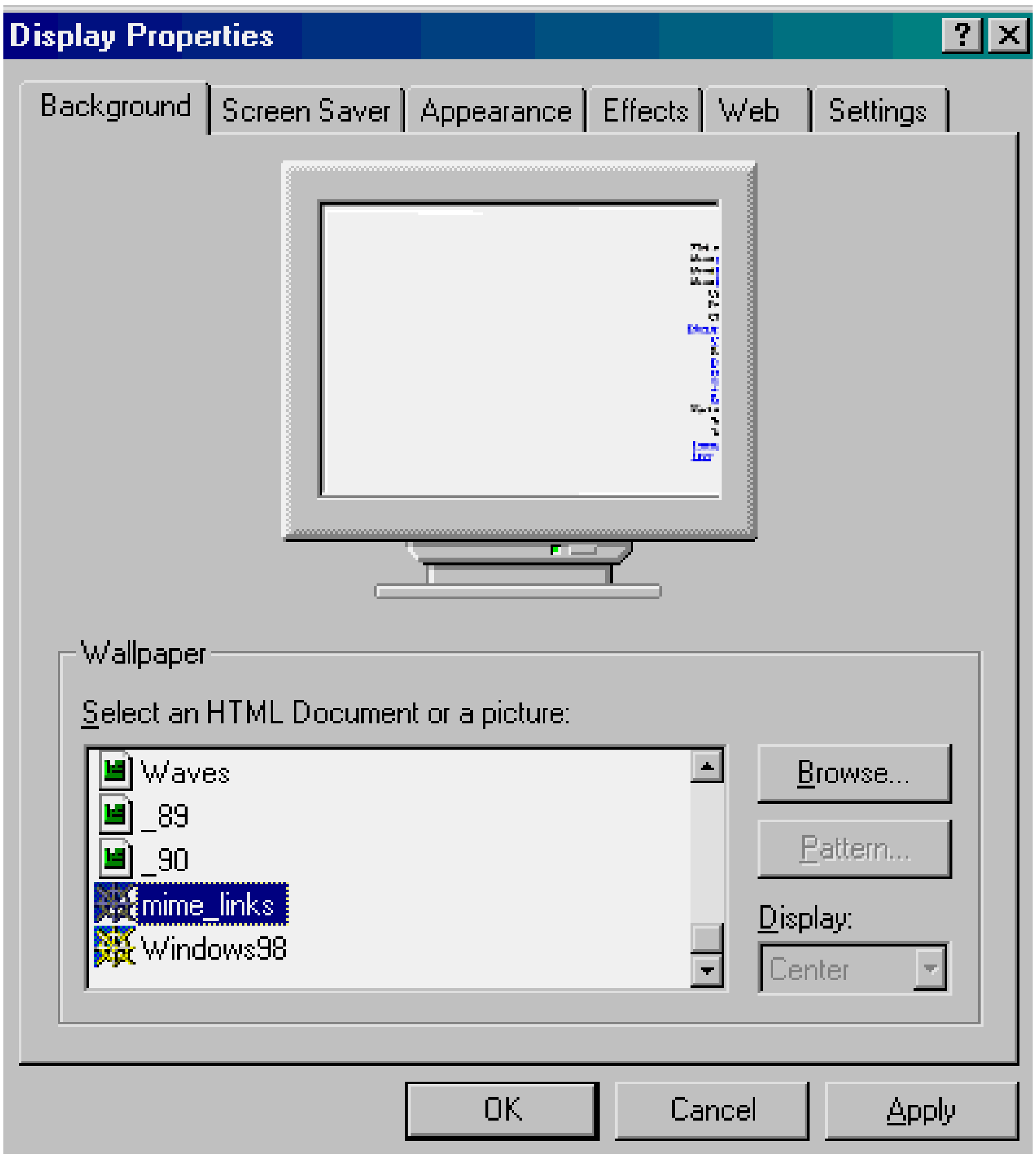
Now create your HTML document. Save it in the folder C:\Windows\Web\Wallpaper:



Then select:



Here you need to click on the Background tab and identify your HTML document:



When you click OK you will be off and running! I had to do some manipulations to get the links out in the right side. I inserted a table with two cells (columns). Then I placed a graphics item (the same color as the desktop) in the left cell. Then the links came in the right cell.

Now it works. Here is how the desktop looks now. You see the 6 links on the right:

- Next page
- Previous page

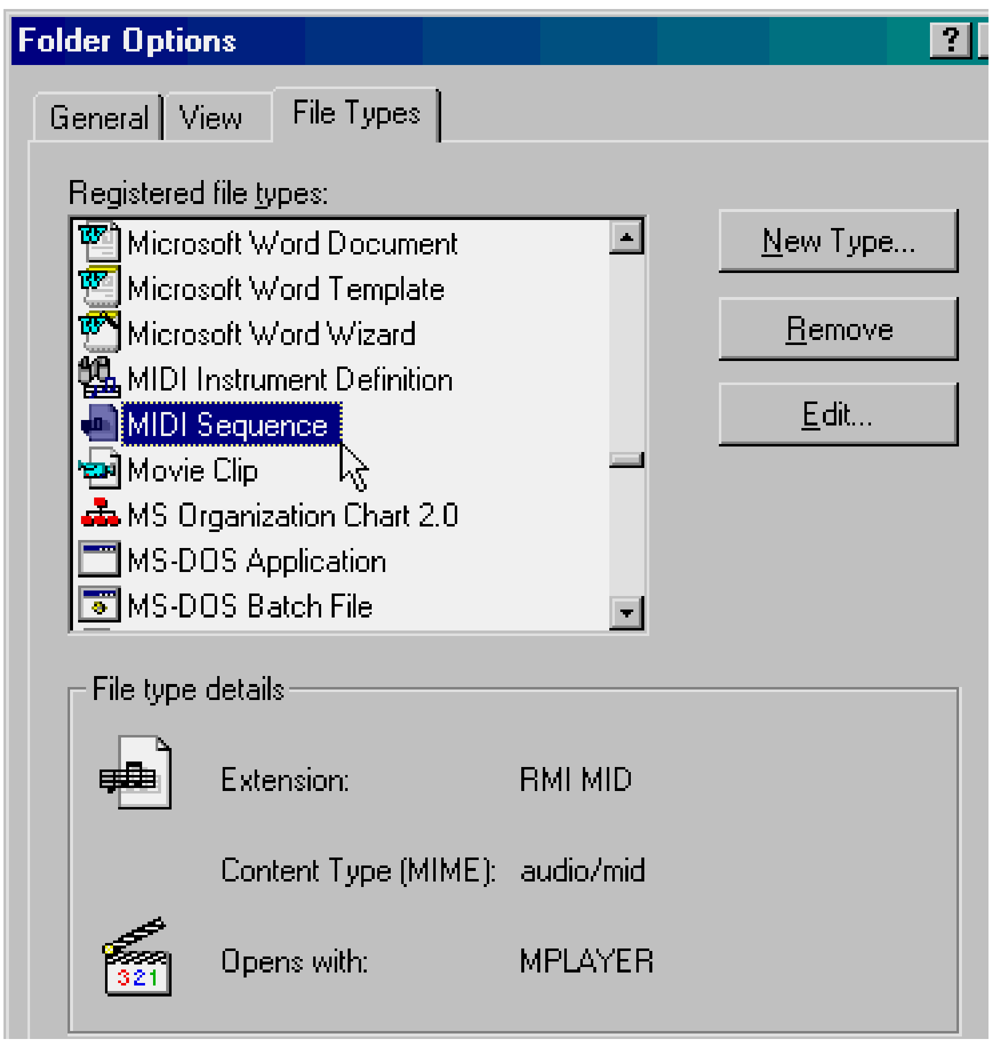
Karbosguide.com. Software Tip 4

About file types
- showing only some of them!

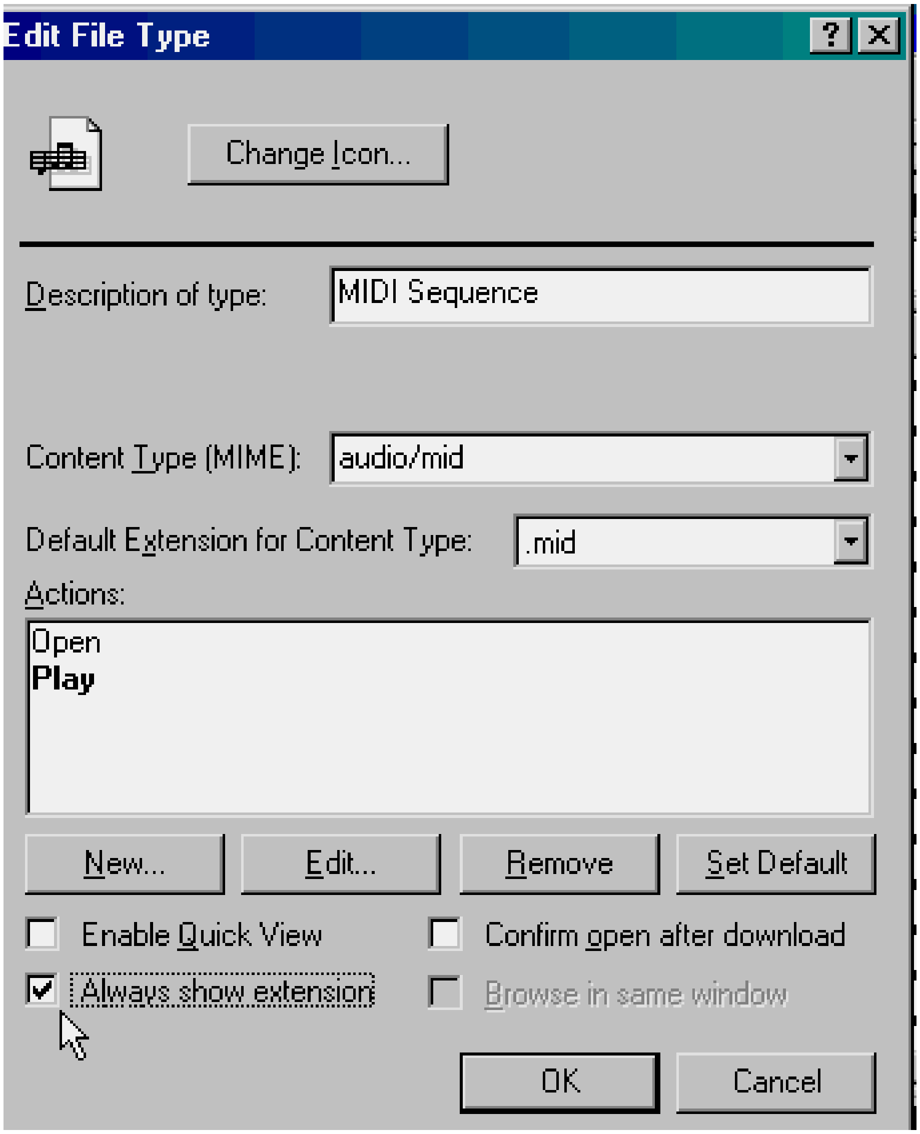
- Next page
- Previous page

By default Explorer does not show the suffix to the name for registered file types. However you can add that you always want to see the suffix for certain types of files:

In Explorer select View → Folder Options... and select the tab File Types. Then find the file type in the long list. Here I find the MiDi files, where I always want to see their suffix:



Then click on the Edit... button and check as seen below:



Then click twice on OK. From now on this file type will always be identified including the suffix in Explorer.

- Next page
- Previous page

Karbosguide.com. Software Tip 3b

About temporary files - get rid of them (II)

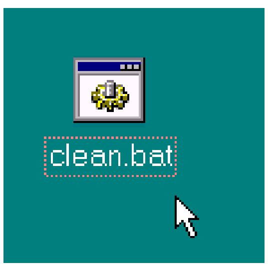
- Next page
- Previous page

Here is another tip to cleaning up Windows' temporary files. Thank you to Mr. Frank Fallon of Australia, who has a site with similar Windows tips. This works with Windows 98 and Me:

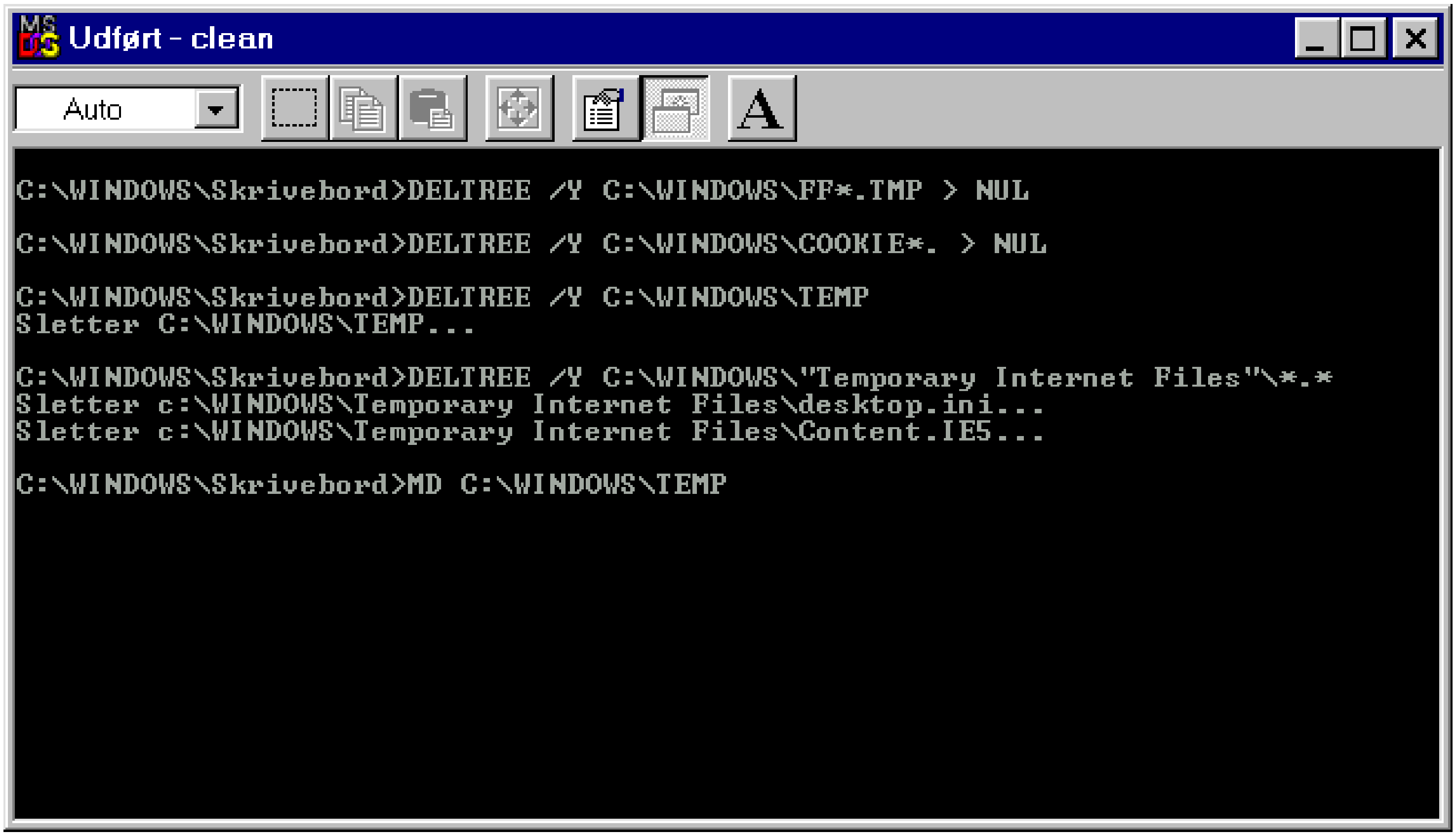
To clean up temporary files, you use these command lines:

```
DELTREE /Y C:\WINDOWS\FF*.TMP > NUL
DELTREE /Y C:\WINDOWS\COOKIE*. > NUL
DELTREE /Y C:\WINDOWS\TEMP
DELTREE /Y C:\WINDOWS\"Temporary Internet Files"\*.*
MD C:\WINDOWS\TEMP
```

I prefer to place them in a batch file. I have the file clean.bat placed on my desktop on my Windows 98-based PC. It holds the mentioned lines, and is easy to execute, being placed at the desktop:



When I run the batchfile, the commands are executed one by by one:



It really cleans up - very fast and very efficiently.

Mr. Fallon also recommended a handy temp cleaner. It is a little program which may take care of these problems. Download it here, if you want to test it.

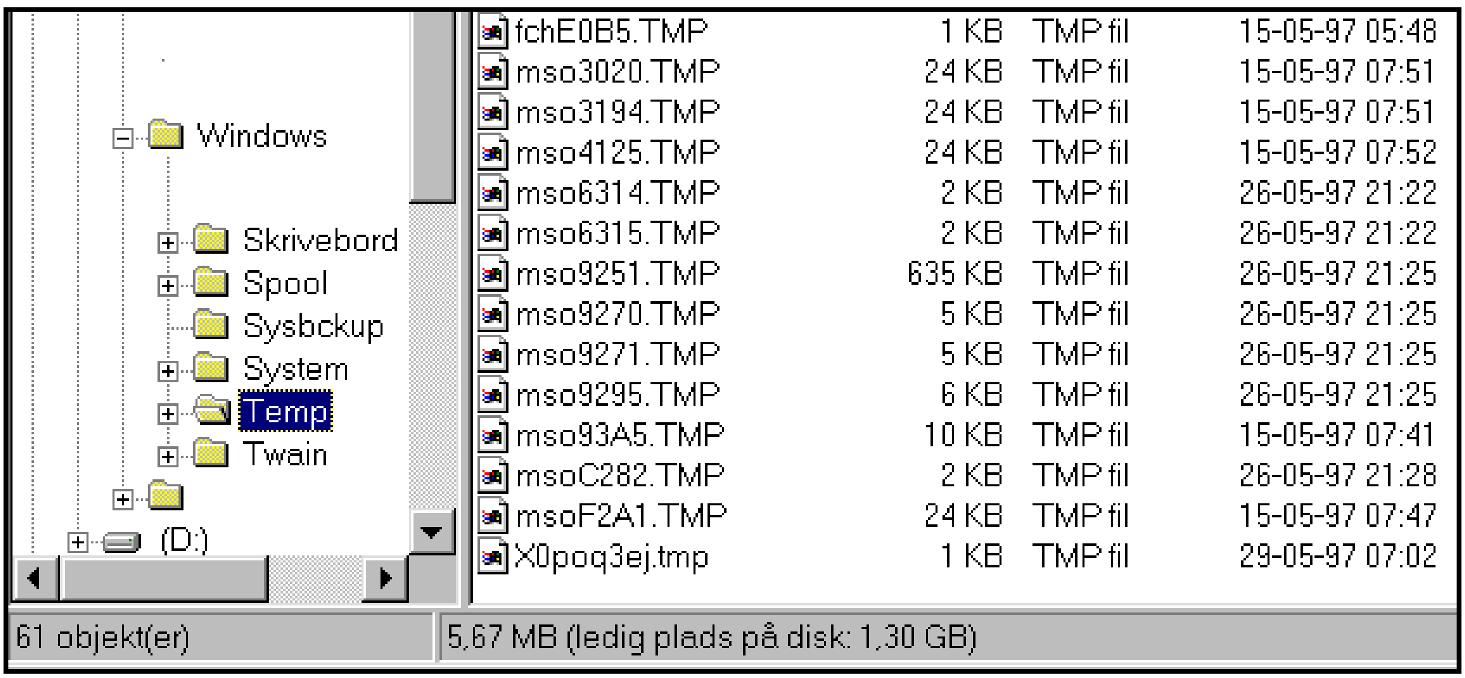
- Next page
- Previous page

Karbosguide.com. Software Tip 3a

About temporary files - get rid of them (I)

- Next page
- Previous page

Windows continuously creates temporary files. They are temporary files, which really need to be deleted. However, they are never deleted automatically - certainly not when Windows crashes, as it sometimes does. They are in the folder C:\Windows\Temp. You ought to check it routinely:



It is to your advantage to delete these temporary files. They just take up space and there can be hundreds of them.

However, the problem is that you may not be able to delete all temporary files while Windows is running - some of them may be active. Therefore I recommend this simple method: put a line in your Autoexec.bat! You can find the file Autoexec.Bat (Autoexec) through Start -> Find. Right click on it and choose edit. Then type the line shown below and save the file:

```
echo Y | del c:\windows\temp\*.*
```

This will cause all temporary files to be deleted at any start-up. The echo command adds a "Y" into the del command, so you do not need to confirm with a "Y" to execute the delete.

You may also want to delete the subdirectories using this command:

```
deltree /y c:\windows\temp\*.*
```

Use the Autoexec.bat

Windows also has tools for cleaning the harddisk. However I still recommend the following addition to the Autoexec.bat. This only works in Windows 98, since Windows Me has no Autoexec.bat.

Autoexec.bat is a text file, and it has to be to be edited. If you worked with your PC ten years ago, you would not ask what Autoexec.bat is. We used it all the time to tweak more free memory out of the start-up.

Autoexec.bat is the central start-up file in any DOS-based computer. It holds a number of "lines" (written in simple text).

Each line in Autoexec.bat is to be executed during the start-up - one by one. Hence the file type is BAT, which stands for batch. A batch file holds one or several of lines of commands. And here we add three lines into the Autoexec.bat.

To open Autoexec.bat, you find Windows Explorer and highlight C:\ (the root directory) in the left frame. In the right frame you find Autoexec.bat. Highlight it and make a rightclick on it. Then choose edit, and it is opened for you in Notepad.

Make your changes and save the file, which then works after re-boot.

Using Autoexec.bat is smart since it cleans up every time you boot the pc. The first line to add:

```
deltree /y C:\Windows\temp\*.*
```

This line deletes all the temporary Internet files including the subdirectories, which all the time is created in C:\WINDOWS\Temporary Internet Files:

```
deltree /y C:\Windows\TEMPOR~1\*.*
```

Finally add

```
C:\WINDOWS\COMMAND\deltree /y C:\Windows\cookies\*.*
```

This line clears out the cookies, which my browser collects all the time. The PC has to reboot for the Autoexec.bat to work. You can monitor the deletion on screen during the start up (hit [pause] key to freeze the screen).

Three lines

You may use these three lines in my Autoexec.bat:

```
C:\WINDOWS\COMMAND\deltree /y C:\Windows\tempor~1\*.*
C:\WINDOWS\COMMAND\deltree /y c:\windows\temp\*.*
C:\WINDOWS\COMMAND\deltree /y C:\Windows\cookies\*.*
```

The first line empties the browsers cache - a good thing. The second line cleans up after Windows in general. The third line clears out the cookies, which the browser collects all the time.

All use is on your own responsability - you may experience some problems using this tip, so please test it.

The next tip shows another option. Also read more about deleting these files in Tip 18.

- Next page
- Previous page

Software Tip 2b

About the Disk Cache

- Next page
- Previous page

Especially in Windows versions 95 and 98 it is important to understand the of the relationship between:

- Size of and control of Disk cache
- The free memory
- The size of the swap file

What is disk cache?

The cache is a portion of RAM, reserved for cache (buffer) for the hard disk. The disk cache is necessary, since it speeds up the hard disk a lot. However, it should not be bigger than 8 or 16 MB.

The problem is that the disk cache really gobbles up RAM. In Windows 98 it can easily eat up 20-25% of your RAM. An that is a total waste of RAM.

In Windows 98 you can limit the size of your cache. This is done by editing the file System.Ini, which is found in C:\Windows. Double click on it and scroll down until you reach the text [vcache].

Then type in the two lines you see below and save the file. Do it soon. This is important!

```
[Non Windows App]

[vcache]
MinFileCache=8096
MaxFileCache=8096

[display]
```

The change takes effect when you restart Windows. I am convinced that 8 MB disk cache is sufficient - at least when you use the FAT32 file system.

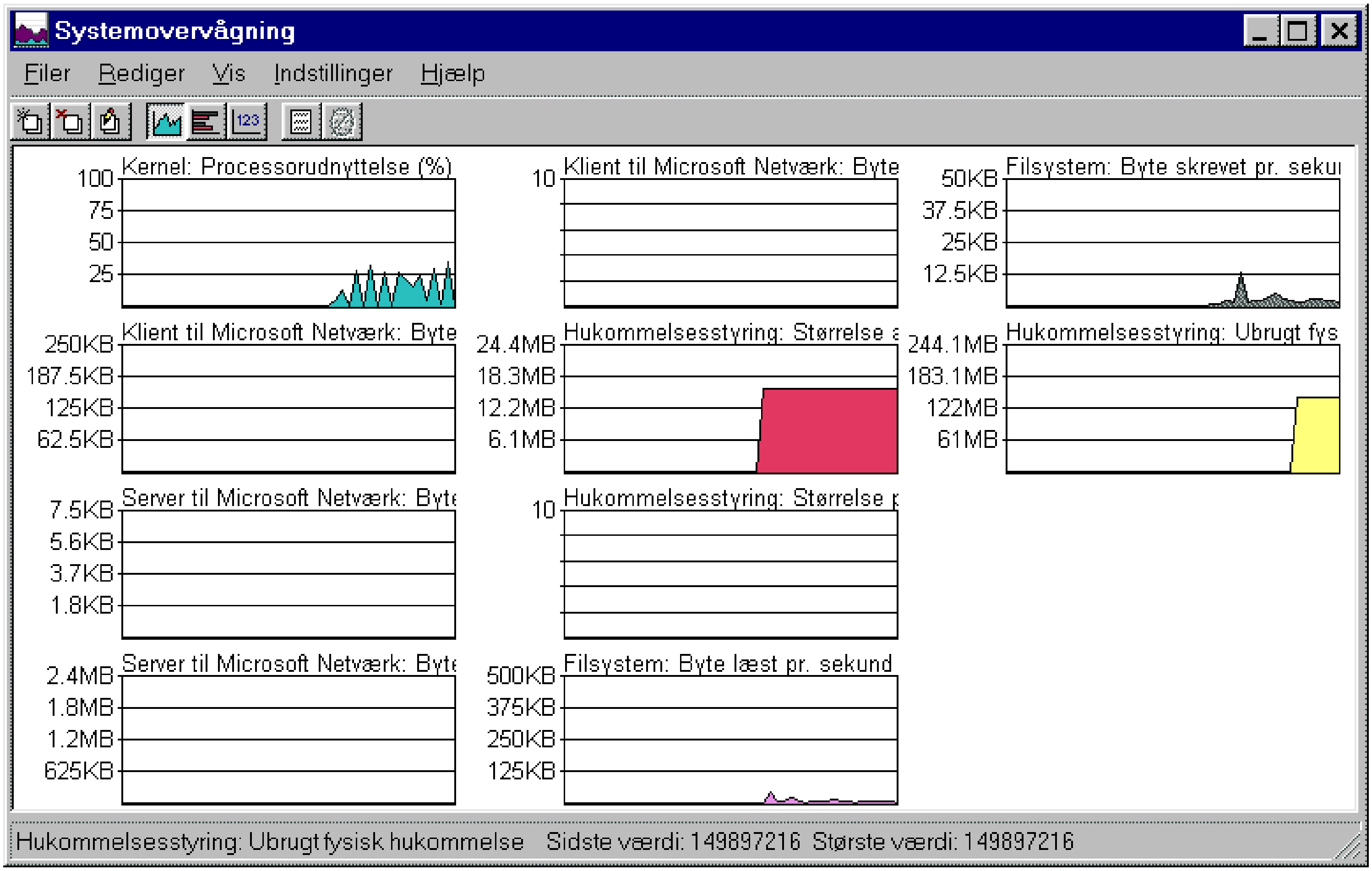
If you use Windows 95 with the old FAT16 file system, you should probably maintain just 1 or 2 MB of disk cache.

In Fat 32 the vcache holds a permanent copy of the whole FAT table, which occupies full 2 MB under the FAT32 file system. With this size FAT the Vcache has to be 4 MB big, if there shall be room for other things besides the FAT.

Read about file systems and about the cache copy of FAT in same module.

The System Monitor

You can watch your use of memory with the excellent tool *resource meter* . You find it by going to: Start -> Programs -> Accessories -> System Tools -> System Monitor. You can add elements in the edit menu. That will allow you to see available memory and the swap file, as illustrated below (you may print this page, although the figure is in Danish):



You should check available *memory* and the size of the *swap file* over a period of time. Do this daily for a while and see how big the swap file gets.

It is also a good idea to check the disk cache, so that it does not occupy more than 16 MB (or less). If the disk cache only occupies 8 or 16 MB, you can easily calculate your actual RAM usage by keeping track of *available memory* and the *size of the swap file.*

- Next page
- Previous page

Karbosguide.com

Sitemap
Guestbooks
Support
Photo Gallery

Hardware Guides:

Digital Data
Motherboard logic
CPU and RAM
Disks and drives
PC Interface
System Software
Monitors & Sound
Digital Imaging

Others:

Software Guides
Search
Dictionary
About us
My Super PC

Karbosguide.com

Welcome to Michael Karbo's Online Service. Here you will find a modern online-magazine with more than 500 illustrated articles for the critical reader!

Use our menu to your left or the sitemap. You may also follow any of the links listed below. We hope that you appreciate our work!

- Start studying the design of a PC motherboard.
- Learn about harddisks and other drives.
- Learn about the PC I/O system.
- Learn about the PC video system.

News ... Sitemap

See our guestbook and add your comment.

Corrections

Misspellings, typos or other corrections. Please report!.

Karbo's Newsletter:

Sign up!

Pentium 4 and AthlonXP.

More than 50 photos of old cars free to download!

The MP3 article is re-written.

Editing photos with Photoshop

Cleaning Windows Me for temporary files ...

All modules 7 re-written

KarbosGuide.com. Module 2e1.

About RAM

Here you can read about:

- What is RAM?
- About RAM types

On the following pages:

- About SIMMs
- DIMMs
- PC100 RAM and further
- Rambus
- DDR

- Next page
- Previous page

What is RAM?

[top]

This page should be read together with modules 2a, 2b, 2c, and 2d, which deal with system board, system bus, I/O bus and chip sets. When we talk about motherboard and chip sets, we cannot ignore RAM. Warning: RAM and RAM chips is a very complicated, technical subject area. I can in no way give a complete, comprehensive description of this subject.

RAM is our working memory storage. All the data, which the PC uses and works with during operation, are stored here. Data are stored on drives, typically the hard drive. However, for the CPU to work with those data, they must be read into the working memory storage, which is made up of RAM chips. To examine RAM, we need to look at the following:

- RAM types (FPM, EDO, ECC, and SD RAM)
- RAM modules (SIMM and DIMM) in different versions
- RAM and the system bus

First, let us look back in time. Not too many years ago, Bill Gates said, that with 1 MB RAM, we had a memory capacity, which would never be fully utilized. That turned out to be untrue.

Historical review

Back in the 80s, PCs were equipped with RAM in quantities of 64 KB, 256 KB, 512 KB and finally 1 MB. Think of a home computer like Commodore 64. It had 64 KB RAM, and it worked fine.

Around 1990, advanced operating systems, like Windows , appeared on the market, That started the RAM race. The PC needed more and more RAM. That worked fine with the 386 processor, which could address larger amount of RAM. The first Windows operated PCs could address 2 MB RAM, but 4 MB soon became the standard. The race has continued through the 90s, as RAM prices have dropped dramatically.

Today. it would be foolish to consider less than 32 MB RAM in a PC. Many have much more. 128 MB is in no way too much for a "power user" with Windows 95/98, it is important with plenty of RAM. Click here to read about the swap file and RAM considerations. Windows 98 is a little better at handling memory, but still a lot af RAM is a good thing.

RAM types

[top]

The traditional RAM type is DRAM (*dynamic* RAM). The other type is SRAM (*static* RAM). SRAM continues to remember its content, while DRAM must be refreshed every few milli seconds. DRAM consists of micro capacitors, while SRAM consists of off/on switches. Therefore, SRAM can respond much faster than DRAM. SRAM can be made with a *rise time* as short as 4 ns. It is used in different versions in L2 cache RAM (for example *pipe line BURST Cache SRAM*).

DRAM is by far the cheapest to build. Newer and faster DRAM types are developed continuously. Currently, there are at least four types:

- FPM (*Fast Page Mode*)
- ECC (*Error Correcting Code*)
- EDO (*Extended Data Output*)
- SDRAM (*Synchronous Dynamic RAM*)

A brief explanation of DRAM types

[top]

FPM was the traditional RAM for PCs, before the EDO was introduced. It is mounted in SIMM modules of 2, 4, 8, 16, or 32 MB. Typically, it is found in 60 ns or 70 ns versions. 60 ns is the fastest and the one to use. You cannot mix different speeds on the same Pentium motherboard.

EDO (Extended Data Out) RAM is an improvement of FPM RAM. Data are read faster. EDO extends the time that output data is valid, which betters timing issues between the CPU and RAM and this way improves the performance.

By switching from FPM to EDO, one could expect a performance improvement of 2 to 5 percent. EDO RAM was usually sold in 60 ns versions. A 50 ns version was available at higher cost.

EDO has now been replaced by the even faster SDRAM.

ECC RAM is a special error correcting RAM type. It is especially used in servers.

SDRAM (synchronous DRAM)): The replacement for DRAM, FPM, and EDO RAM types. SDRAM "locks" (synchronizes) the memory access to the CPU clock. This way we get faster data transfer. While one portion of data is tranported to the CPU another can be being prepared for transfer.

SDRAM comes only in 64 bit modules (long 168 pin DIMMs). SDRAM has a access time of only 6-12 ns. The performance improvement over EDO RAM was a mere 5 percent running at 66 MHz. At 100 and 133 MHz it proves better.

DDR RAM is clock doubled version of SDRAM, which is replacing SDRAM during 2001-2002.

RAMBUS (RDRAM) is a more futuristic RAM type. Intel and others had great expectations from this type, but it flopped in 2000-2001.

---

8 or 9 bits per byte?

Normally you figure 8 bits to one byte. For many years, a ninth bit has been added as *parity* bit in the RAM blocks to verify correct transmission. That way you have to transmit 9 bits, to store 8 bits in the old 30 pin RAM chips. And it takes 36 bits to store 32 bits in the larger 72 pin chips, which increases the cost of the RAM chip by about 12%.

If your motherboard requires 36 bit modules, you must respect that. Fortunately, most system boards accepts 32 bit modules, so this creates no problems.

RAM and motherboard

[top]

You cannot freely install your desired RAM type. RAM is controlled by the chip set on the motherboard, so you must install a type, which matches your motherboard. Furthermore, RAM chips come in different sizes, which must match the system board.

On modern system boards, RAM is installed on SIMM or DIMM modules. Before, small individual DRAMs were used. There was usually room for 36 small chips on the system board. That made it cumbersome to install new RAM. Then, someone figured out to install RAM chips on cards, which are easily installed. First came the SIPP modules. They had multiple pins, which fit in the motherboard. Since then came the SIMM modules. They are mounted on a card, which has an edge connector. They fit in sockets on the motherboard, and anyone can install them.

RAM speeds

[top]

RAM speed is measured in ns (*nano seconds*). The fewer ns, the faster is the RAM. Years ago, RAM came in 120, 100 and 80 ns. Today, we are talking about 60 ns and faster.

It becomes complicated to describe the relationship between RAM speed and the ability of the system bus to utilize fast RAM. I will gloss over that. But here is a table which illustrates RAM speed, relative to clock speed:

Clock speed	Time per clock tick
20 MHz	50 ns
25 MHz	40 ns
33 MHz	30 ns
50 MHz	20 ns

66 MHz	15 ns
100 MHz	10 ns
133 MHz	6 ns

Peak Bandwidth

[top]

Here you see the maximal peak bandwidth of the three well known RAM types. The figures illustrates the absolutely maximal transfer from RAM to the L2-cache - in peaks, not as continuously transferred.

RAM type	Max. peak bandwidth
FPM	176 MB/sec
EDO	264 MB/sec
SD	528 MB/sec

- Next page
- Previous page

Learn more [top]

Also see module 3. An illustrated Guide to CPUs from 8086 to the Pentium-III.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2e2

About SIMM RAM

The contents:

- About SIMMs
- Number of chips per module
- Buswidth 32 bit

- Next page
- Previous page

SIMMs

[top]

SIMM (Single Inline Memory Modules) were first made in 8 bit editions. They were small cards with 1, 2 or 4 MB RAM. They were connected to the motherboard with a 30 pin edge connector. The modules were 8 bit wide. This meant that 16 bit processors (286 and 386SX) needed 2 SIMMs in a pair. Thus, there was room for two modules in what is called a *bank*.

32 bit processors (386DX and 486) need 4 of the small 8 bit SIMMs in a bank, since their banks are 32 bit wide. So, on a typical 1st generation 486 motherboard, you could install 4 X 1 MB, 4 X 2 MB, or 4 X 4 MB in each bank. If you only had one bank (with room for 4 modules), it was expensive to increase the RAM, because you had to discard the old modules.

---

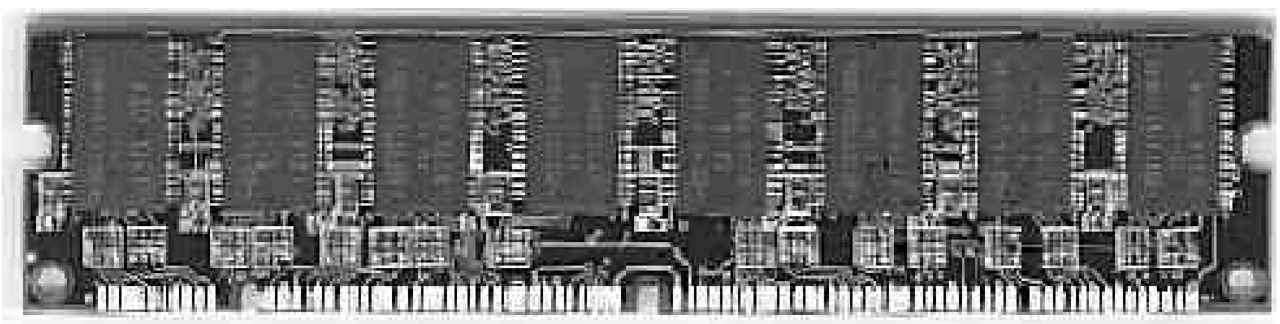
32 bit modules

With the advent of the 486 processor, demand increased for more RAM. Then the larger 32 bit modules came into use. A 486 motherboard could still have 4 SIMM sockets, but when the modules were 32 bit wide, they could be installed one at a time. This was quite ingenious.

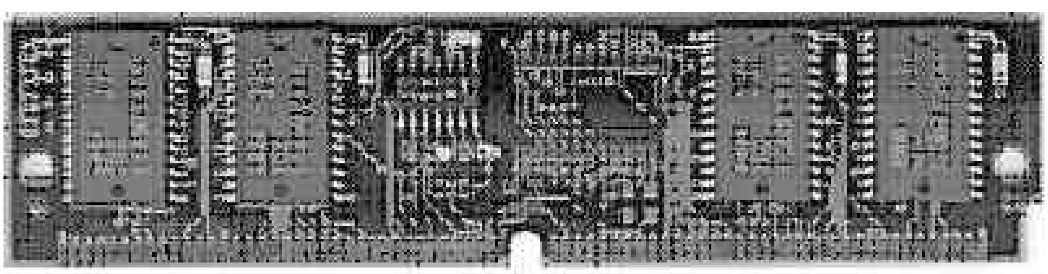
You could add different types of modules and still use the old ones. Also, since the 486 motherboard ran at only 33 MHz on the system bus, the RAM module quality was not so critical. You could mix 60 ns and 70 ns modules of different brands without problems.

Here you see a couple of SIMM modules. On top is a 64 bit module (168 pins - don't try to count them). Next is a 32 bit module with a 72 pin connector. Below is an 8 bit module with a 30 pin connector:

64 bit SDRAM:

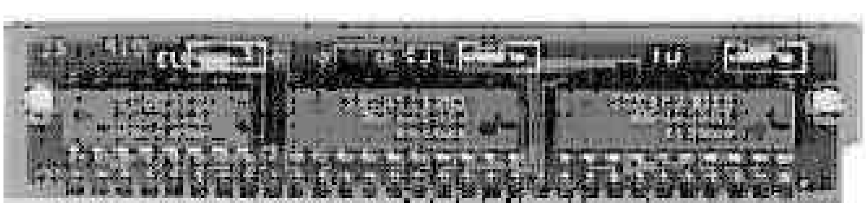


32 bit DRAM:



and

16 bit DRAM:



Number of chips per module

[top]

Some SIMMs have more chips on the module than others. Looking at just the 32 bit modules, we find modules with 2, 4, 8 or chips on each side. SIMMs with 2 MB, 8 MB and 32 MB are double sided. There are chips on both sides of the module, and all these chips are 16 Mbit ones.

The newest DIMM-modules holds 64 Mbit RAM chips. This way a 32 MB module is made of only 4 chips since 4 X 64 / 8 = 32.

Pentium motherboard with SIMMs

On the Pentium motherboard, the system bus is 64 bit wide. Therefore, the 32 bit SIMMs are installed in pairs. Since the standard motherboard only has two banks with a total of four SIMM sockets, RAM expansion possibilities are limited. NOTE: never use different speed RAM modules on the Pentium motherboard. All modules *must* have the same speed. Here you see a few configurations on an old Pentium motherboard with four SIMM sockets:

Bank 1	Bank 2	Total RAM
16 MB + 16 MB	-	32 MB
16 MB + 16 MB	32 MB + 32 MB	96 MB
32 MB + 32 MB	32 Mb + 32 MB	128 MB

Certain motherboards (like TYAN) have 6 or 8 SIMM sockets. That provides more RAM expansion flexibility.

- Next page
- Previous page

Learn more [top]

An illustrated Guide to CPUs from 8086 to the Pentium-III: Module 3.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2e3.

About new fast RAM

The contents:

- DIMMs
- PC100 RAM
- PC133 and VC133
- Intel and PC133

On the following pages:

- Rambus
- DDR

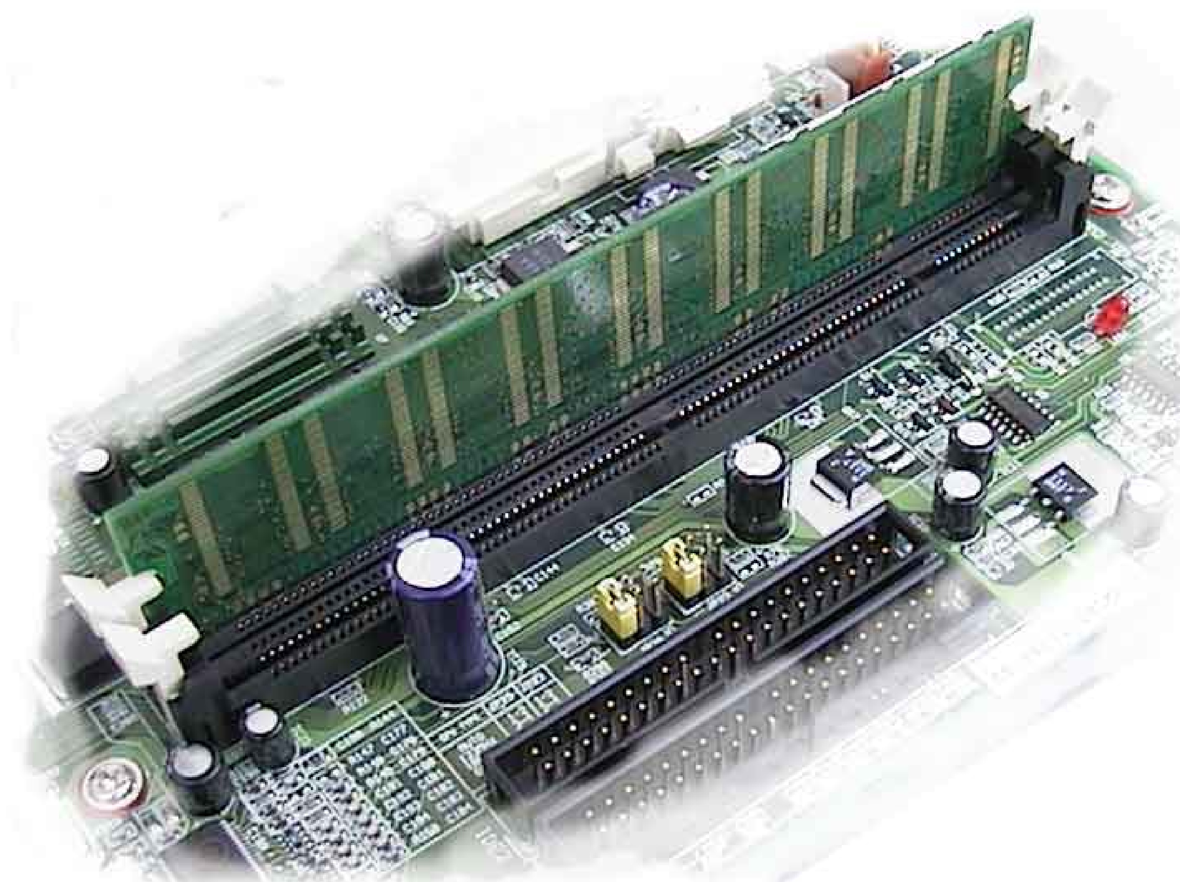
- Next page
- Previous page

DIMMs

[top]

The most used modern RAM type, SDRAM is made in 64 bit wide modules called DIMMs (Dual Inline Memory Module).

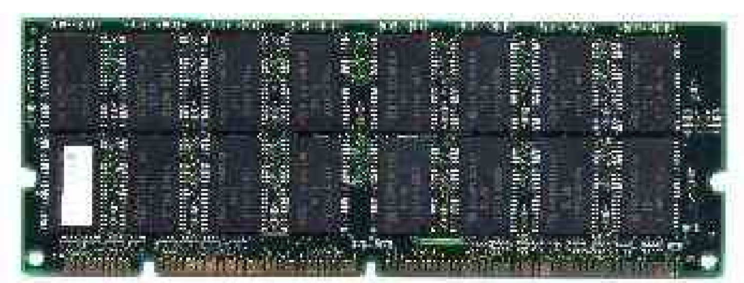
They have a 168 pin edge connector. Here you see one module:



Since the DIMM modules are 64 bits wide, you can install one module at a time. They are available in 8, 16, 32, 64, 128, 256 MB, and 512MB with 6, 8, 10, and 12 ns speed. There are usually 2 -4 DIMM sockets on a motherboard.

The advantage of SDRAM is increased speed. That allows you to increase system bus speed. With 60 ns EDO-RAM, you can run at a maximum of 75 MHz on the system bus, while SDRAM speed can increase to 133 MHz and above. Also the SDRAM work synchronous with the system bus for a better performance.

Most chip sets are made for SDRAM. Some motherboards had both SIMM and DIMM sockets. The idea was that you could reuse old EDO RAM in the SIMM sockets, or choose to install SDRAM in the DIMM sockets. They were not designed to mix RAM types although it works at some boards.



Above: a 64 MB DIMM-module holding 32 chips each of 16 Mbit (32 X 16 Mbit / 8 bit = 64 MB).

It is better to use DIMMs made of the new 64 Mbit chips. A 64 MB module is this way made of only 8 chips (8 X 64 Mbit / 8 bit = 64 MB).

Fast RAM

[top]

Intel have managed to speed up the processors power by factor 200 times the last ten years. That is a lot, but it is a problem that RAM memory technology only has improved by factor 20 in the same period.

Today we hope and dream of new fast RAM types, that will help us to get the full potential from our powerful PCs.

PC100 RAM

The first attempt to improving RAM speed was the PC100 standard. With chip sets like BX the system bus speed has come up to 100 MHz. Hence Intel has made a new standard called PC100. Only 8 ns SD-RAM modules that are constructed according to these standards are guaranteed to work at 100 MHz. In some articles this RAM is described at 125 SD-RAM.

SPD

The new DIMM-modules include a EPROM-chip holding information about the module. This little 8-pin chip works as a SPD (*Serial Presence Detect*) - a unit storing information about the RAM type. The idea is that BIOS can read this information and this way tune the system bus and the timings for a perfect CPU-RAM performance.

You can find a program, that tests the contents of the SPD at this c't homepage. It works with the Intel chip sets holding a 82371 south bridge like BX and GX.

Another program is called DIMM_ID.

PC133

The PC133 RAM running at 133 MHz is the latest version of SDRAM. Specifications are made by VIA, Micron, NEC, Samsung, SIS, Acer Labs and other vendors. The first production (from Corsair, June 1999) used 7.5 ns RAM modules from Micron.

VIA supports the PC133 RAM with their Apollo Pro Plus chip set (693A). Later they launched support for PC266 DDR RAM!

Also AMD's K7 Athlon may use PC133 RAM with the VIA KX133Pro chipset.

VC133

Virtual Channel 133 is another flavour of the PC133 standard. The modules holds a small cache of superfast SRAM. According to tests, these modules perform very well, but due to unknown reasons, it never became popular.

Intel and PC133

Originally Intel planned to by-pass PC133 RAM in their roadmaps. They intended to migrate from PC100-based chip sets (like BX) to Rambus-based chip sets (like i820).

For a period of 12 months in 1999-2000, Intel experienced several disastrous incidents from their attempt to implement

Rambus in chip sets and motherboards. During this period they were forced (by taiwanese motherboard manufactures) to adapt the PC133 standard.

The chip set i815 was the result of this revision of strategies.

Intel's problem is that they have "sold their soul" to Rambus Inc. According to their agreement, until 2003 Intel can only implement other RAM types than RDRAM if the bandwidth is less than 1 GB/sec. This agreement does not include server chipsets, from what we understand.

- Next page
- Previous page

Learn more [top]

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2e4

About Rambus RAM

The contents:

- Rambus
- High clock rates

- Next page
- Previous page

Rambus RDRAM

[top]

While the CPUs has become around 200 times faster in ten years, the RAM performance has only gone up 20 times. So we need new RAM types. But which?

Many vendors decided to go for DDR RAM as described in . Where DDR RAM is a development of the existing SDRAM technology, Intel choose RDRAM, which represents a much more revolutionary change in RAM design.

Intel and RDRAM without succes

Intel is comitted to the Rambus RAM, which also is called RDRAM (Rambus Direct RAM), nDRAM, or RIMM (Rambus Inline Memory Modules).

RDRAM is an advanced technology patented by a company, who sells the technology to other chip manufactures for a 2% in license. In 1997 Intel signed a contract that apparently commits them to support RDRAM only in all new chipset up to 2002.

Originally AMD also expected to support the Rambus RAM for its Athlon processor. But having seen all Intel's problems with the technology, AMD is not so keen on the Rambus anymore. However, RDRAM is already used in Sony PlayStation 2 and in Nintendo64 machines. In the Sony PlayStation 2 you find 32 MB of RDRAM delivering a bandwidth of 3.2 GB/sec.

During 1999 and 2000, Rambus was not very successful. In fact, Intel has suffered a serious set-back due to their commitment to the Rambus design. The chip set i820 "Camino" became a little disaster.

Intel failed to produce a reliable way to interface SDRAM to the 820 chipset. The MTH (Memory Translator Hub - which translated RDRAM bus to SDRAM modules) had some timing or noise issues that caused unreliable operation. Intel replaced CC820 boards with VC820 boards (with 128MB RDRAM included) as the CC820 use the MTH and SDRAM while the VC820 used RDRAM.

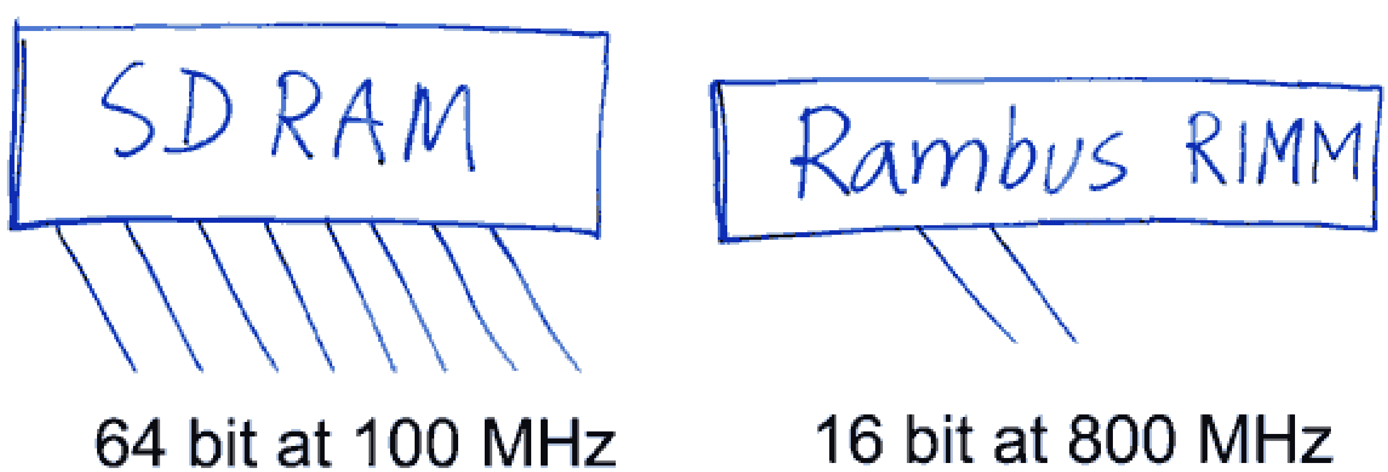
But, on the paper, Rambus sounds great:

Intelligent Rambus design

RDRAM is developed from the traditional DRAM, but the architecture is completely new. It has been streamed and optimized to yield new performance.

The RAMBUS-design gives a more intelligent access to the RAM, meaning that units can "prefetch" data and this way free the CPU some work. The idea is that data is read in small packets at a very high clock speed.

The RIMM modules are only 16 bit wide compared to the traditional 64 bit SDRAM DIMMs, but they work at a much higher clock frequency:



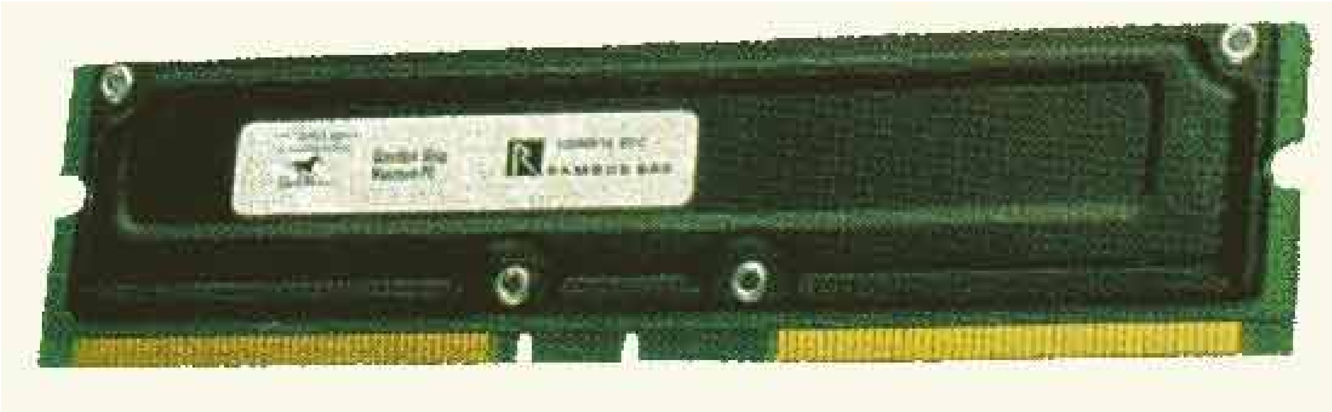
The Rambus modules work on 2.5 volts, which internally is reduced down to 0.5 volt when possible. This helps to reduce heat and radio signals.

The RIMMs hold controlling chips that turns off the power to sections not in use. They can also reduce the memory speed if thermal sensors report of overheating.

CRIMMs

All RAM slots have to be full; this is new, with RAMBUS we have to fill in blank modules in slots which are not in use. The blank modules are called CRIMMs (with a 'C' for *continuity*).

The RIMM modules hold 184 pins.



The RDRAM chips have to be placed very close to the CPU to reduce radio noise. This indicates, that RIMM technology is rather sensitive; Intel seems to have made that discovery as well.

High clock rates

As mentioned, the modules are only 16 bit wide, but they work at 600, 700 and 800 MHz. Actually a PC800 RIMM runs on a 400 MHz clock using both rising and falling edges, being clockdoubled just as DDR RAM.

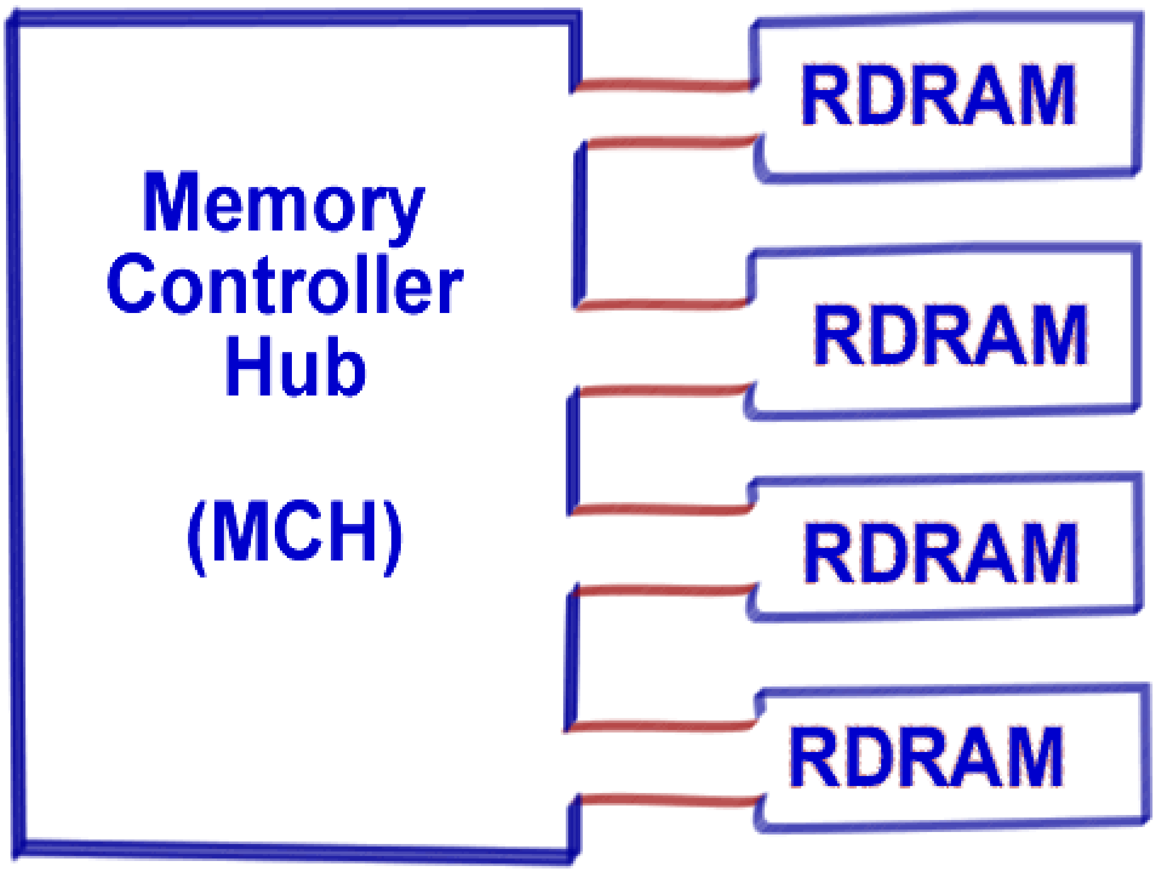
More confusing the PC600 RIMM actually runs on a 266/532 MHz clock, and the PC700 works at 366/712 MHz.

PC800	800/400 MHz
PC700	712/366 MHz
PC600	532/266 MHz

This gives the bandwidth of up to 1.6 GB per second - compared to the 500-800 MB/sec of PC100 SDRAM - of a single Rambus channel. You may find a chart comparing the bandwiths of different RAM types in the next page.

Multichannel memory design

You may bundle four channels to a 64 bit wide RAM bus giving 6.4 GB/sec:



This is not possible using the existing RAMBUS-based chip sets like i820. They only operate with one RAMBUS channel onboard. The high-end chip set i840 operates with dual RDRAM channels, as the up-coming i850 will.

RIMM in the future, says Intel

GigaHertz versions of Rambus RAM will probably follow, so the technology has potential for much higher bandwidths.

In 1999 it seemed that Intel was having big problems with the Rambus technology in the ill-faited i820 chip set (the so-called "Caminogate" tragedy). Hence they were forced to support PC133 RAM as seen in the i815 chipset.

Poor performance so far

Unfortunately it was soon obvious that the i815 chip set with its PC133 RAM was performing slightly better than the i820 chip set with its still very expensive RDRAM. You have to use dual Rambus channels (as in upcomming Intel chip set i850 "Tehema") to benefit from a higher bandwidth. But this doubling is also possible from using DDR RAM.

A test between a i840-based dual Rambus PC and a Micron DDR-based PC gave the same result; all benchmarks were better on the DDR system.

So far Rambus RAM is of no big interest. It is too expensive, and there is nothing to gain from it. However the Rambus technology stil is quite promising, but the prices has to come down, and it better be soon. DDR RAM is closing in.

Intel claims that DDR is to slow for the new Pentium 4 processor. It would require dual channel DDR RAM to get the required bandwidth. And dual channel DDR RAM meens a 128 bit wide bus, which is no good solution. The north bridge and the motherboard would be loaded with hundreds of signal lines.

In 2001 RDRAM is being used with great success on the GB850 Pentium 4 board and RDRAM prices are tumbling steadily.

RDRAM 2.0

Rambus plans to speed up the bandwith a factor two using a Quad Rambus Signaling Level. This should happen without any increase in clock frequency.

- Next page
- Previous page

Learn more [top]

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2e5

About DDR RAM

The contents:

- DDR RAM
- PC2100 RAM
- Intel not allowed
- Comparing bandwidth

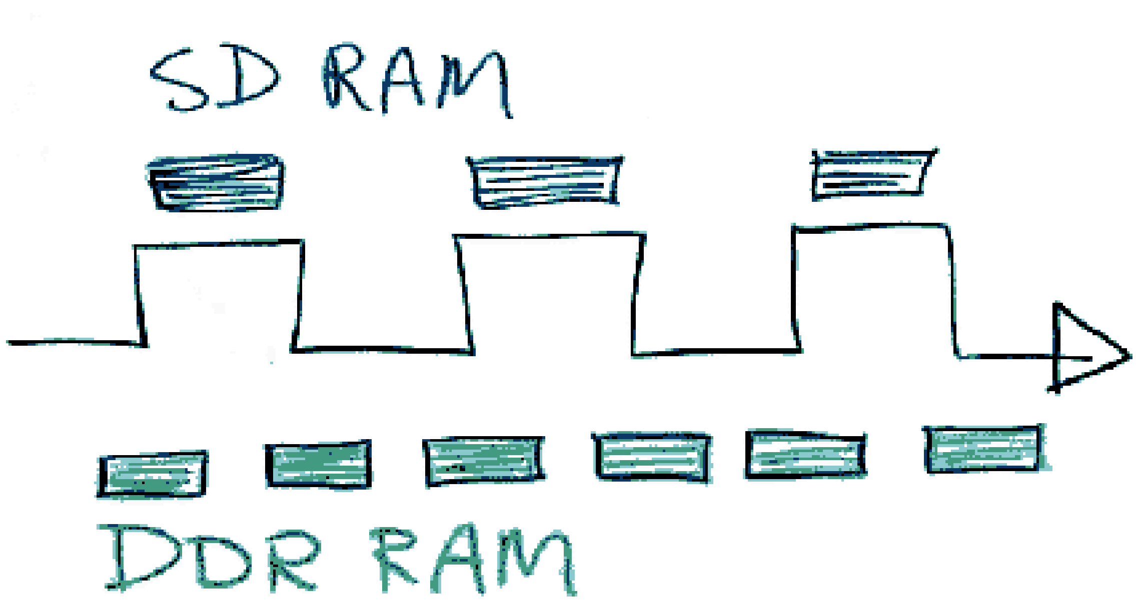
- Next page
- Previous page

DDR RAM

A very interesting RAM type is the DDR RAM, which is expected to hit the market in 2001.

DDR stands here for *Double Data Rate*. It is a technology that transmits data on both sides of a tact signal.

This way the performance has been doubled; a 133 MHz SDRAM chip can very easy become altered to a 266 MHz DDR chip:



It should be pretty easy for the market to change for DDR RAM. The modules look like and operate quite similar to existing SDRAMs. We just need new chipsets to start the migration.

However, the modules hold 16 pins more than SDRAM do, so they do not fit into the same sockets.

PC2100

The Taiwanese company VIA, which produces chip sets and CPUs and who are fighting Intel, is fully supporting the DDR RAM strategy. Soon we shall see 266 MHz moduls (being "overclocked"133 MHz SDRAM modules).The 266 MHz modules reaches a 2.1 GB/sec bandwidth. Hence they are to be sold as PC2100 RAM.

Other terms used are:

- DDR200 (200 MHz)
- DDR266 (266 MHz)
- DDR333 (333 MHz)

VIA expects DDR to be used in all segments of the pc market. Intel, who is behind the Rambus technology, only expects to use DDR in large servers, where you find several

Gigabytes of RAM installed, and where RAM price really matters.

No Intel here

Intel is dedicated to the Rambus technology. In the summer 2000 it was revealed that Intel has comitted itself to the RAMBUS technology so they *cannot* implement DDR! This goes for all future desktop PCs until 2003, according to their agreement with Rambus Inc. Only the 64 bit server Itanium processor and it succesors Foster and McKinley are using DDR RAM.

We hope that Intel will change their strategy. We expect DDR-SDRAM to be cheaper than Rambus RAM for quite some time; yet it should give the same performance. Rambus represents a sophisticated technology, but with prices 5 times higher it is not a low-end product. Intel produces great chipsets for desktop PCs like i815E, and it would be sad if they abandoned this market. We want Intel and PC2100!

Reports in the summer 2000 told that Intel has licensed VIA to develop DDR-enabled chip sets for Pentium 4.

Evolutionary changes of design

Where RDRAM requires completely new production plants, DDR represents an evolutionary progress. The chip manufactures may re-use their SDRAM fabs for the production without many problems.

Hence it seems quite natural and in tune with the previous changes in RAM technology that we use the DDR standard for a couple of years. Before Rambus (or something even better) enters the market.

Comparing bandwidth

Below you see the theoretical bandwidts of different RAM types. However, SDRAM does not perform as good as the figures show. This is due to latencies; the CPU and other units cannot read the data at these speeds; they have to wait some clock circles in between each reading before the data transfers start. The same goes for DDR RAM.

RAM type	Theoretical max. bandwidth
SDRAM 100 MHz	100 MHz X 64 bit= 800 MB/sec
SDRAM 133 MHz	133 MHz X 64 bit= 1064 MB/sec
DDRAM 200 MHz (PC1600)	2 X 100 MHz X 64 bit= 1600 MB/sec
DDRAM 266 MHz (PC2100)	2 X 133 MHz X 64 bit= 2128 MB/sec

DDRAM 366 MHz (PC2600)	2 X 166 MHz X 64 bit= 2656 MB/sec
RDRAM 600 MHz	600 MHz X 16 bit= 1200 MB/sec
RDRAM 700 MHz	700 MHz X 16 bit= 1400 MB/sec
RDRAM 800 MHz	800 MHz X 16 bit= 1600 MB/sec

DDR-II

A new version of DDR RAM is scheduled for 2003. Using another technique, it should be possible to double the performance!

- Next page
- Previous page

Learn more [top]

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3a1.

About CPUs

To understand the data processing methodology, an understanding of the design and function of the CPU is essential. The following subjects will be covered on these pages.

The contents:

- What is a CPU?
- Intro to CPUs from 1st to 7th generation

- Next page
- Previous page

The module is divided in several sub modules, which all together ought to be read as a unit.

What is a CPU?

[top]

The CPU is certainly the most important PC component. CPU stands for *Central Processing*

Unit. Let us briefly study that name:

- It is a processor, because it processes (moves and calculates) data.
- It is central, because it is the center of PC data processing.
- It is a unit, because it is a chip, which contains millions of transistors.

Speed, speed, speed

Without the CPU, there would be no PC. Like all other hardware components, the CPUs are continually undergoing further development. You can see the explosive technological development in data processing most clearly in the development of newer and faster CPUs. The CPUs have for years doubled their performance about every 18 months (Moore's Law), and there are no indications that this trend will stop.

When we now look at all the CPUs from a broader perspective, we can see that:

- The CPU history is closely tied to the companies IBM and especially *Intel*.
- The CPUs have their roots back to Intel's chip *4004* from 1971.
- You can identify seven or eight CPU generations up till today.
- The *compatibility* concept has been important throughout the development.

CPUs - brief review

[top]

CPU history starts in 1971, when a small unknown company, Intel, for the first time combined multiple transistors to form a *central processing unit* - a chip called Intel 4004. However, it was 8 years before the first PC was constructed.

PCs are designed around different CPU generations. Intel is not the only company manufacturing CPUs, but by far the leading one. The following table shows the different CPU *generations*. They are predominantly Intel chips, but in the 5th generation we see alternatives:

PC	CPUs	Year	Number of transistors
1st. Generation	8086 and 8088	1978-81	29,000
2nd. Generation	80286	1984	134,000
3rd. Generation	80386DX and 80386SX	1987-88	275,000
4th. Generation	80486SX, 80486DX, 80486DX2 and 80486DX4	1990-92	1,200,000

5th. Generation	Pentium Cyrix 6X86 AMD K5 IDT WinChip C6	1993-95 1996 1996 1997	3,100,000 -- -- 3,500,000
Improved 5th. Generation	Pentium MMX IBM/Cyrix 6x86MX IDT WinChip2 3D	1997 1997 1998	4,500,000 6,000,000 6,000,000
6th. Generation	Pentium Pro AMD K6 Pentium II AMD K6-2	1995 1997 1997 1998	5,500,000 8,800,000 7,500,000 9,300,000
Improved 6th. Generation	Mobile Pentium II Mobile Celeron Pentium III AMD K6-3 Pentium III CuMine	1999	27,400,000 18,900,000 9,300,000 ? 28,000,000
7th. Generation	AMD original Athlon AMD Athlon Thunderbird Pentium 4	1999 2000 2001	22,000,000 37,000,000 42,000,000

Please notice that the mobile CPUs as well as Pentium III CuMine include very large on-die L2-caches. These caches consist of millions of transistors.

We will now see what the CPU really does.

- Next page
- Previous page

Learn more [top]

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3a2.

About CPUs - continued

The contents:

- How does a CPU work?
- 8086 compatbility
- CISC, RISC and VLIW instructions

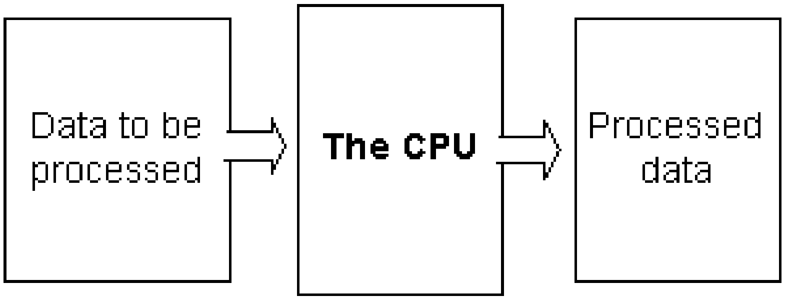
- Next page
- Previous page

How does a CPU work?

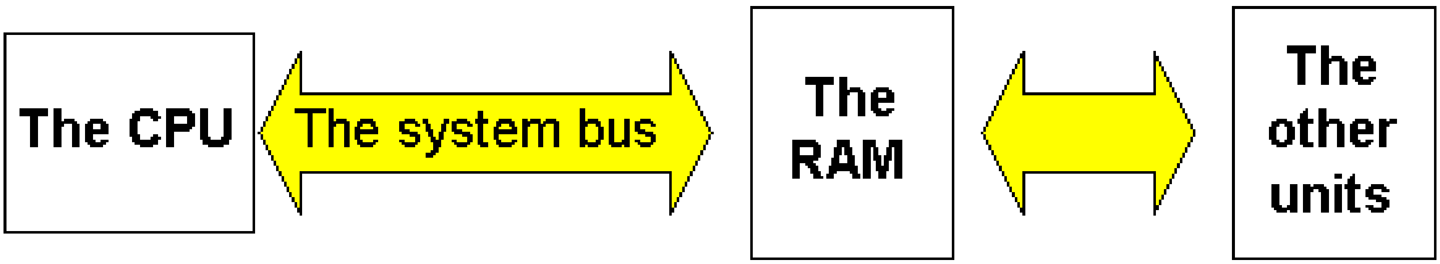
[top]

The CPU is centrally located on the motherboard. Since the CPU carries out a large share of the work in the computer, data pass continually through it. The data come from the RAM and the units (keyboard, drives etc.). After processing, the data is send back to RAM and the units.

The CPU continually receives *instructions* to be executed. Each instruction is a data processing order. The work itself consists mostly of *calculations* and *data transport:*



Data have a path to the CPU. It is kind of a data expressway called the *system bus.* You can read more about the system bus in module 2b.



Two types of data

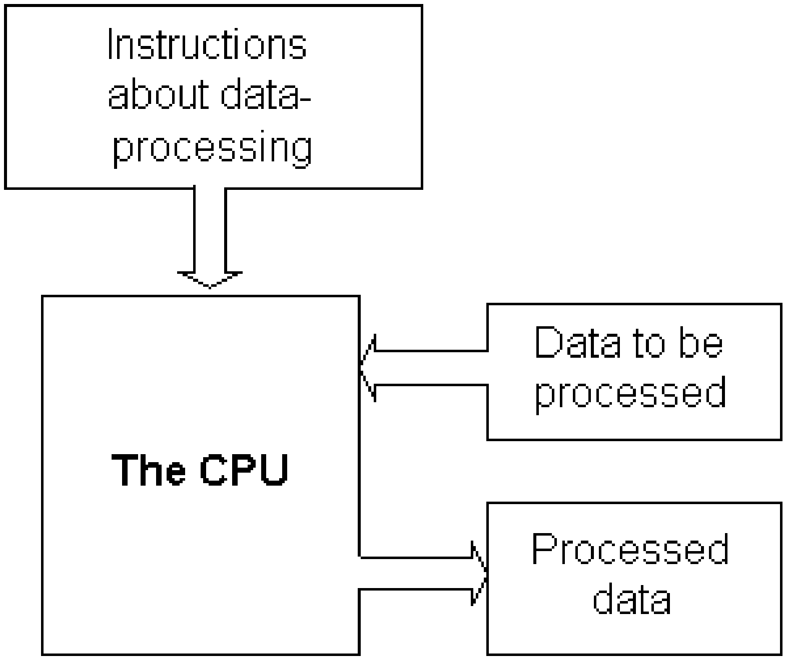
[top]

The CPU is fed long streams of data via the system bus. The CPU receives at least two *types* of data:

- Instructions on how to handle the other data.
- Data, which must be handled according to the instructions.

What we call instructions is *program code*. That includes those messages, which you continuously send to the PC from the mouse and keyboard. Messages to print, save, open, etc.

Data are typically user data. Think about the letter, which you are writing to Aunt Karen. The *contents*, letters, images, etc., are *user data*. But if you click "print," you are then sending program code (instructions):



8086 compatible instructions

[top]

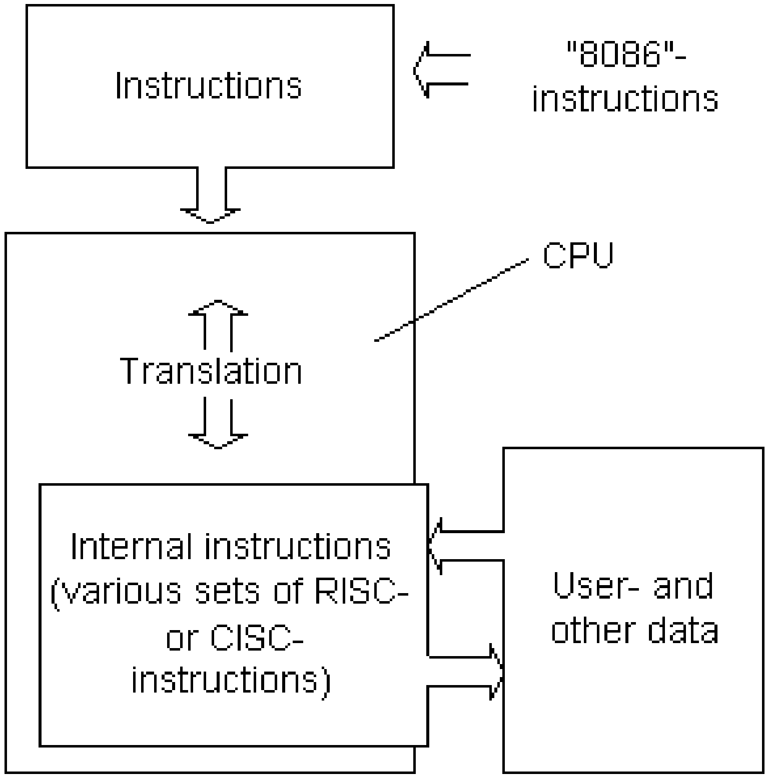
The biggest job for the CPU consists of *decoding* the instructions and *localizing* data. The calculations themselves are not heavy work.

The decoding consists of understanding the instructions, which the user program sends to the CPU. All PC CPUs, are "8086 compatible." This means that the programs communicate with the CPU in a specific family of instructions.

These instructions, originally written for the Intel 8086 processor, became the blueprint for the "IBM compatible PC" concept. The 8086 from 1978 received its instructions in a certain format.

Since there was a desire that subsequent CPU generation should be able to handle the same instructions which the 8086 could, it was necessary to make the instruction sets compatible. The new CPUs should understand the same instructions. This *backwards compatibility* has been an industry standard ever since. All new processors, regardless of how advanced, must be able to handle the 8086 instruction format.

Thus, the new CPUs must use much effort to translate the 8086 instruction format to internal instruction codes:



CISC, RISC, and VLIW instructions and their handling

[top]

The first CPUs had a so called *Complex Instruction Set Computer* (CISC). This means that the computer can understand many and complex instructions. The X86 instruction set, with its varying length from 8 to 120 bit, was originally developed for the 8086 with its mere 29000 transistors.

More instructions have been added within new generations of CPUs. The 80386 had 26 new instructions, the 486 added 6 and the Pentium another 8 new instructions. This meant, that programs had to be rewritten to use these new instructions. This happened for example with new versions of Windows . Hence, some programs require a 386 or a Pentium processor to function.

You should also see module 3e09 on MMX, 3DNow! and other extensions to the set of instructions.

Reduced Instruction Set Computer (RISC)

The RISC instructions are brief and the same length (for example 32 bit long, as in Pentium Pro), and they process much faster than CISC instructions. Therefore, RISC is used in all newer CPUs. However, the problem is that the instructions arrive to the CPU in 8086 format. Thus, they must be decoded.

For every new CPU generation, the instruction set has been expanded. The 386 came with 26 new instructions, the 486 with 6 new instructions, and Pentium with 8 new instructions. These changes mean that some programs require at least a 386 or a Pentium processor to work.

VLIW

A Very Long Instruction Word processor uses instruction that are long. The idea is to put many instructions together in one. Then the processor can fetch several instructions in one operation and be more effecient. Normal non-VLIW processors only receive one instruction per *word* . A word is an amount of data transmitted to the processor, and the VLIW processor receives several instructions in each word.

To re-order the instructions you use a software compiler. This principle works fine in more special processors such as DSPs. These chip perform the same operations over and over again.

A CPU is a general-purpose processor, and the VLIW design becomes extremely complex in this case. Hence, Intel has had many problems with their 64 bit Itanium processor, which comes in VLIW design. Another company to use VLIW is TransMeta with their portable Crusoe processor.

- Next page
- Previous page

Learn more	**[top]**

Click for Module 3b about CPU improvements

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3a3.

About modern CPUs

The contents:

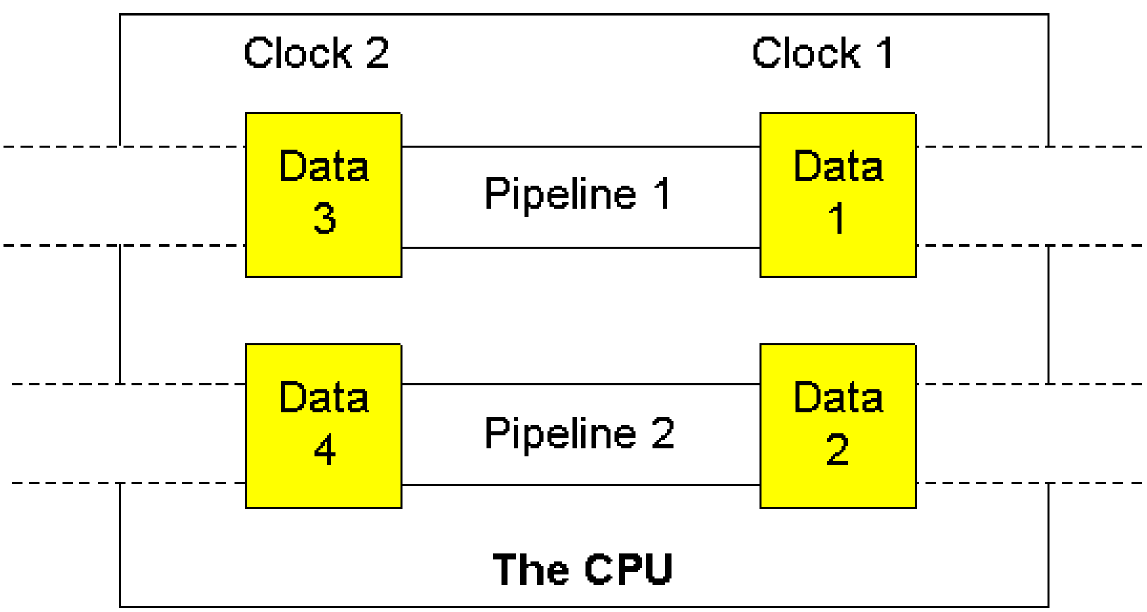
- Dual pipeline
- Floating point unit - FPU
- Graphic overview of the processors

- Next page
- Previous page

Dual pipeline: More work per clock stroke

There is also a continuous optimizing of the instruction handling process. One is that the clock frequency increases, as we will see later - the faster, the better. But what can the CPU do in one clock tick. That is critical to its performance. For example, a 386 needed 6 clock ticks to add a number to a sub total. A job which the 486 manages in only two clock ticks, because of more effective instruction decoding.

5th and 6th generation CPUs can execute more than one of those operations in one clock tick, since they contain more processing lines (pipelines), which work parallel:



Please also read the section about MMX, about 3DNow!, and Katmai instructions.

Floating point unit - FPU

[top]

The first CPUs could only work with whole numbers. Therefore, it was necessary to add a mathematical co-processor (FPU), when better math power was needed. Later, this FPU was built into the CPU:

CPU	FPU
8086	8087
80286	80287
80386	80387
80486DX	Built in
80486SX	None
Pentium and thereafter	Built in

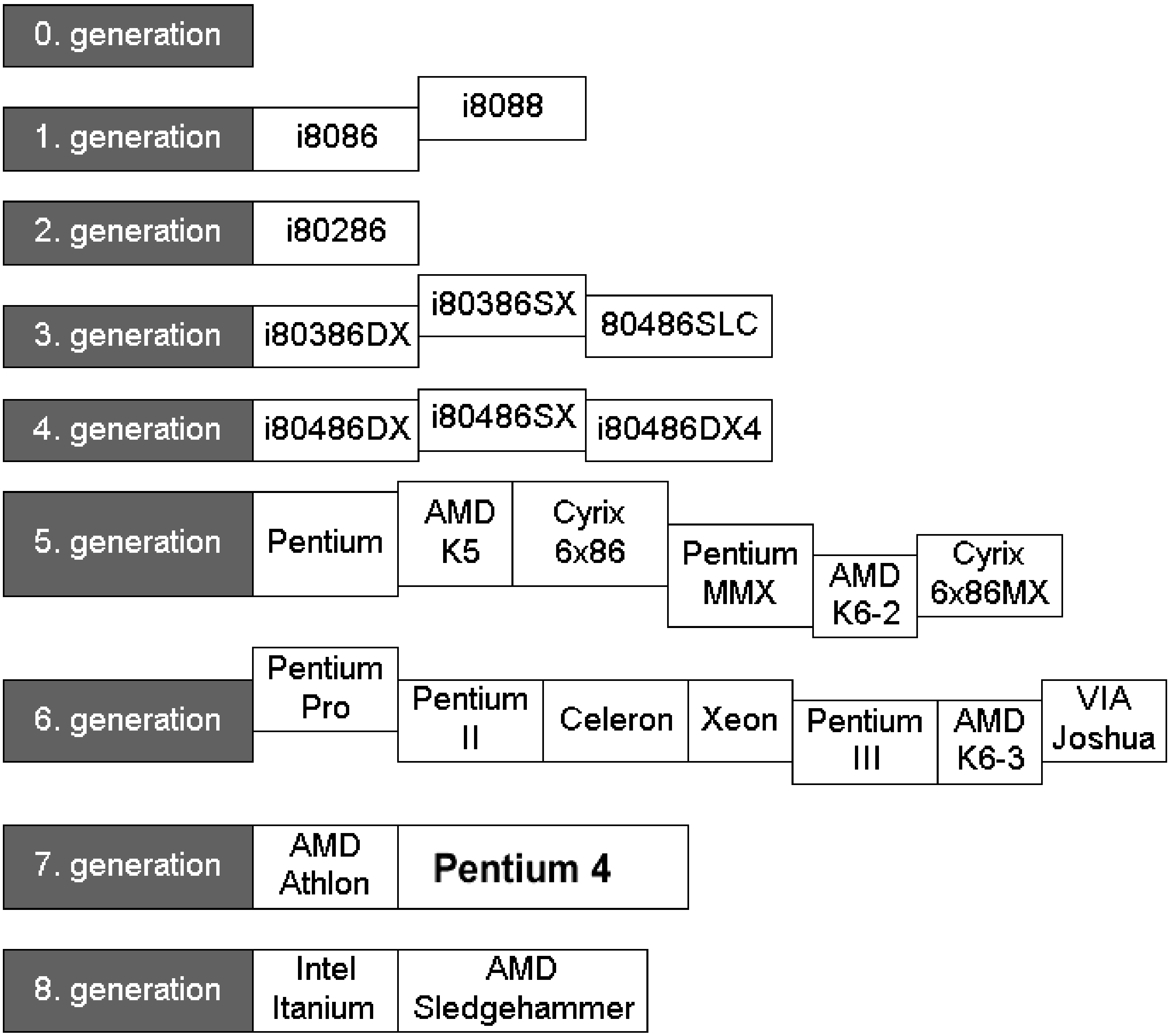
It is said that Intel's CPUs have by far the best FPU units. Processors from AMD and Cyrix definitely have a reputation for providing sub standard performance in this area. But, you may not utilize the FPU. That depends on the applications (user programs) you are using. Common office programs do not use the floating point operations, which the FPU can handle. However, 3D graphics programs like AutoCad do. And all 3D-games like Quake rely heavily on FPU perfomance! Read more of this subject here.

Therefore, if you use your PC in advanced design applications, the FPU performance becomes significant. For some users, it is only of limited importance.

Graphic overview of the processors

[top]

There are CPUs of many brand names (IBM, Texas, Cyrix, AMD), and often they make models which overlap two generations. This can make it difficult to keep of track of CPUs. Here is an attempt to identify the various CPUs according to generation:



- Next page
- Previous page

Learn more **[top]**

Click for Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3b1.

The CPU – developments and improvements

The contents:

- Clock frequency and -doubling

- Next page
- Previous page

Intro

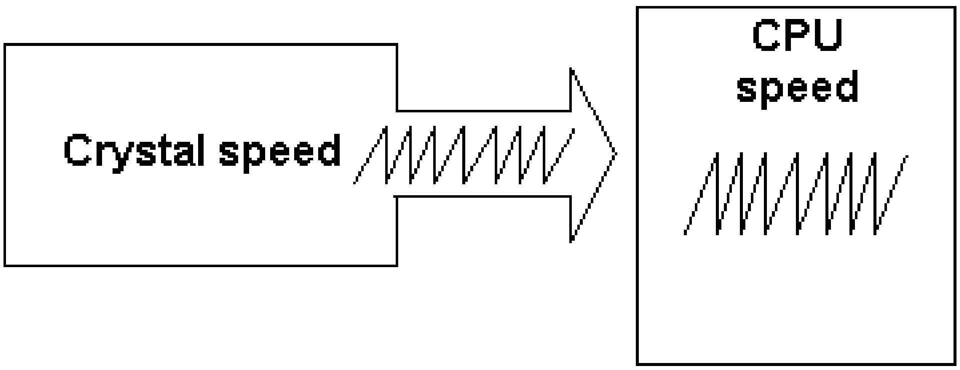
If you have to improve a CPU – and that happens all the time – it is not only a matter of technical development. There are many bottlenecks in and around the CPU, which are continually being bettered.

To understand these technological improvements, one must remember that the CPU is a data processing gadget, mounted on a printed circuit board (the motherboard). Much of the data processing takes place inside the CPU. However, all data must be transported to and from the CPU via the system bus. But what determines the speed of the CPU?

Clock frequency

[top]

We know this from the ads: "A Celeron 466 MHz." The 466 MHz is the *clock frequency.* Actually, there is a small crystal on the motherboard. which continually ticks to the CPU at a steady number of *clock ticks* per second. At each clock tick *something happens* in the CPU. Thus, the more ticks per second – the more data are processed per second.



The first CPUs worked at a frequency of 4.77 MHz. Subsequently then, clock frequencies rates rose to 16, 25, 50, 66, 90, 133 and 200 MHz to the best today, which operate at almost 2000 MHz. Clock frequencies are still being increased. In a few years we will have CPUs operating at 3 GHz and more.

To reach these very high clock frequencies, one has to employ a technique called clock doubling.

Clock doubling in the CPU

[top]

The problem with the high clock frequencies is to ensure that other electronic components keep up with the pace. It is rather simple to make data move very fast inside a chip where the print tracks are microscopic. But when we move outside the chip, other problems appear. The other components must be able to keep up with the pace. When the frequency gets too high, the circuit board print tracks start acting as antennae and various forms of "radio noise" appears. Briefly, it becomes expensive to make the rest of the hardware to keep up with these high frequencies.

The solution to this problem was to split the clock frequency in two:

- A high internal clock frequency, which governs the pace of the CPU.
- A lower external clock frequency, which governs the pace on the system bus.

Intel's 80486DX2 25/50 MHz was the first chip with clock doubling. It was introduced in 1992 with great potential. For a lower price you could acquire a chip, which provided 90% of the 486DX50 performance. The DX50 ran at 50 MHz both internally and externally. The DX2 ran at just 25 MHz on the system bus. This enabled lower cost motherboards. Also RAM speed demands were lower.

Clock doubling occurs inside the CPU. If the motherboard crystal works at 25 MHz, the CPU will receive a signal every 40 nanosecond (ns). Internally in the CPU, this frequency is doubled to 50 MHz. Now the clock ticks every 20 ns inside the CPU. This frequency governs all internal transactions, including *integer unit*, *floating point unit*, and all *memory management unit* operations as well as others. The only area still working at 25 MHz are external data transfers. That is transfers to RAM, BIOS and the I/O ports.

RAM speeds

The speed of the CPU is also connected to the RAM. The ordinary FPM RAM and EDO RAM can functioned at a maximum of 66 MHz (possibly 75 MHz). Therefore, Pentium and similar CPUs were "clocked up" with factors from 2 to 5 internally.

In 1998 the PC100 RAM was introduced together with new motherboards and chip set. This RAM works at 100 MHz, and using the clock factors 3.5, 4 and 4.5 we had CPUs running at 350, 400 and 450 MHz. The Intel CPUs Pentium II, Celeron, and Pentium III can operate with clock factors of up to 8.

With chip set designs like i815 the internal clock frequency operates independently of the FSB (front side bus) connecting the CPU to the north bridge of the chip set. Hence we do not need to talk about clock doubling anymore, and the clock frequencies of the CPU reaches 1700 MHz and above.

- Next page
- Previous page

Learn more **[top]**

Also see Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

KarbosGuide.com. Module 3b2.

The CPU – developments and improvements

The contents:

- Cache RAM
- Cache overview

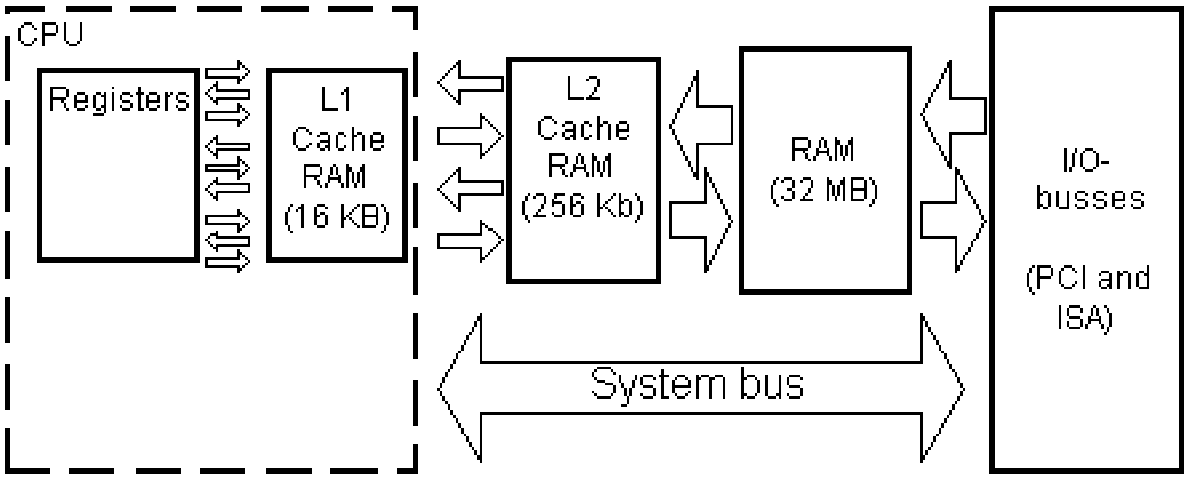
- Next page
- Previous page

About CPU cache RAM

[top]

The CPU must deliver its data at a very high speed. The regular RAM cannot keep up with that speed. Therefore, a special RAM type called *cache* is used as a buffer - temporary storage. To get top performance from the CPU, the number of outgoing transactions must be minimized. The more data transmissions, which can be contained inside the CPU, the better the performance. Therefore, the Intel 80486 was equipped with a built in mathematical co-processor, *floating point unit* and 8 KB L1-cache RAM. These two features help minimize the data flow in and out of the CPU.

Cache RAM becomes especially important in clock doubled CPUs, where internal clock frequency is much higher than external. Then the cache RAM enhances the "horsepower" of the CPU, by allowing faster receipt or delivery of data. Beginning with 486 processors, *two* layers of cache are employed. The fastest cache RAM is inside the CPU. It is called L1 cache. The next layer is the L2 cache, which are small SRAM chips on the motherboard. See the illustration below of a traditional Pentium PC:



How much RAM

The L2 cache can cache a certain amount of RAM. How much is determined by the chip set and the so-called TAG-RAM, the circuit controlling the cache.

One of the most popular chip sets for the original Pentium was Intel´s 82430TX. it worked very well - except for detail. it could not cache more than 64 MB RAM. If you added more RAM to the PC, it was not cached by the L2 cache. Hence, using more than 64 MB of RAM on a TX-based motherboard decreased the performance.

This situation has caused a lot of rumors about Windows not being able to use more than 64 MB RAM. However: Windows 98 can use up to 2 GB RAM! The only problems with the amount of RAM has come from poorly designed chip sets as the TX.

Cache overview

[top]

L1-cache first appeared in Intel's 80486DX chip. Today, bigger and better CPU cache is a natural step in the development of new CPUs. Here we only see the internal caches, i.e.

cache integrated to the CPU and working at the full clock speed.

CPU	Cache size in the CPU
80486DX and DX2	8 KB L1
80486DX4	16 KB L1
Pentium	16 KB L1
Pentium Pro	16 KB L1 + 256 KB L2 (some 512 KB L2)
Pentium MMX	32 KB L1
AMD K6 and K6-2	64 KB L1
Pentium II and III	32 KB L1
Celeron	32 KB L1 + 128 KB L2
Pentium III Cumine	32 KB L1 + 256 KB L2
AMD K6-3	64 KB L1 + 256 KB L2
AMD K7 Athlon	128 KB L1
AMD Duron	128 KB L1 + 64 KB L2
AMD Athlon Thunderbird	128 KB L1 + 256 KB L2

- Next page
- Previous page

Learn more **[top]**

Also see Module 3c about the 5th generations CPUs (Pentiums etc.)

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3b3.

The CPU – developments and improvements

The contents:

- Areas of development
- The CPU – speed measurement
- CPU changes - historical review
- 80486DX4

- Next page
- Previous page

Areas of development

[top]

In the following table, you see some of the technologies, which can be improved in the CPU design. Note that internal means inside the CPU. External speed, etc. refers to features immediately outside the CPU – on the motherboard.

Development area	Significance	Example
Internal clock frequency	speed of data processing *inside* the CPU.	800 MHz
External clock frequency	Speed of data transfer to and from the CPU via the system bus (or Front Side Bus).	133 MHz
Clock doubling	That the CPU works x times faster internally than externally.	6.0 times (like above)
Internal data width	How many data bits can the CPU process simultaneously.	32 bits
External data width	How many data bits can the CPU receive simultaneously for processing	64 bits
Internal cache (Level 1 cache)	Large and better L1 cache, which is a small fast RAM. It works as a buffer between CPU and regular RAM.	64 KB
External cache (Level 2 cache)	Larger and better implemented L2 cache, place on-die in same chip as CPU.	256 or 512 KB
Instruction set	Can the instruction set be *simplified*, to speed up program processing? Or can it be improved?	RISC code More pipelines MMX instructions 3DNow! or SSE

The CPU – speed measurement

[top]

When we look at a CPU, its speed is the most significant feature. All newer CPUs can do the same. You can run Office 2000 in Windows 98 on a 486 CPU. It would be quite slow, but it is possible.

Speed is the primary difference between newer CPUs. Speed improvement is a product of the above mentioned technologies (such as clock frequency and bus width).

The old Speed Index

There are many, many ways to measure CPU speed. The subject is boundless. For years, Norton's Speed Index was used. That is a test, which can be run on any PC with the Norton Utilities Sysinfo program.

In the table below, you see a number of the most common older CPUs. You can see how they are designed regarding clock speed and bus width. The last column shows their *Norton Speed Index* (SI). That is a relative number, which can be used to compare different CPUs. It is not used for modern CPUs.

CPU	CPU speed	Clock doubling	System bus speed	Data width	SI
8086	4.77 MHz	1	4.77 MHz	16 bit	1
80286	12 MHz	1	12 MHz	16 bit	8
80386DX	25 MHz	1	25 MHz	32 bit	40
486 DX2-66	66 MHz	2	33 MHz	32 bit	142
5x86-133	133 MHz	4	33 MHz	32 bit	288
Pentium 75	75 MHz	1.5	50 MHz	64 bit	235
Pentium 90	90 MHz	1.5	60 MHz	64 bit	278
Pentium 100	100 MHz	1.5	66 MHz	64 bit	305
Pentium 133	133 MHz	2	66 MHz	64 bit	420
Pentium 166	166 MHz	2.5	66 MHz	64 bit	527
Pentium 200	200 MHz	3	66 MHz	64 bit	629

Newer CPUs are compared by their clock frequency or by more more sophisticated ratings.

CPU changes - historical review

[top]

This describes briefly the changes throughout the early CPU generations:

8088 and 8086

The 8086 from 1978 was the first 16 bit CPU from Intel using a 16 bit system bus. However 16 bit hardware such as motherboards were too expensive and even non existing at this time, where the 8 bit microcomputers were the standard.

In 1979 Intel reengineered the CPU so it fit with existing 8 bit hardware. The first PC (in 1981) had this 8088 CPU. The 8088 is a 16 bit CPU, but only internally. The external data bus width is only 8 bit giving compatibility with existing hardware.

Actually the 8088 is a 16/8 bit CPU. Logically it could have been named 8086SX. The 8086 was the first total 16 bit CPU in this family.

80286

The 286 from 1982 was also a 16 bit processor. It gave a big advance relative to the first generation chips. The clock frequency was increased, but the major improvement was in optimizing instruction handling. The 286 produced much more per clock tick than 8088/8086 did.

At the introductory speed (6 MHz) it performed four times better than the 8086 at 4.77 MHz. Later it was introduced with 8, 10 and 12 MHz clock speed being used in the IBM PC-AT from 1984.

Another innovation was the ability to run in *protected mode* - a new work mode with a "24 bit virtual address mode", which pointed towards the later shift from DOS to Windows and multitasking. However you could not change from protected back to real mode without rebooting the PC, and the only operating system to use this was OS/2.

80386

The change to the 386s came October the 17th 1985. The 80386 was the first 32 bit CPU. From the traditional DOS PC's point of view, this was not a revolution. A good 286 ran as fast as the first 386SXs - despite the implementation of 32 bit mode.

It could address up to 4 GB of memory and had a better addressing (in bigger chunks) than the 286. The 386 ran at clock speeds of 16, 20 and 33 MHz. Later Cyrix and AMD made clones working at 40 MHz.

The 386 introduced a new working mode besides the *real* and the *protected* modes of the 286. The new mode called *virtual 8086* opened for multitasking since the CPU could generate several virtual 8086s running in each their own memory space.

The 80386 was the first CPU to perform well with the early versions of Windows .

80386SX

This chip was a very popular discount edition of 386DX. It has only 16 bit external data bus contrary to the DX 32 bit. Also, the SX has only 24 address lines, Therefore, it can only address a maximum of 16 Mb RAM. It is not really a true 386, but the cheaper motherboard layout made it very popular.

80486

The 486 was released April the 10th 1989. Generally speaking, the 486 runs twice as fast as its predecessor - all things being equal. That is because of better implementation of the x86 instructions. They are handled faster, more in RISC mode. At the same time bus speed is increased, but both 386DX and 486DX are 32 bit chips. A novelty in the 486 is the built in math co-processor. Before, that had to be installed as a separate 387 chip. The 486 also held 8 KB of L1 cache.

80486SX

This was a new discount chip. The math co-processor was simply omitted.

Cyrix 486SLC: Cyrix and Texas Instruments have made a series of 486SLC chips. They used the same set of instructions as did the 486DX, and they run at 32 bit internally, like the DX. However, externally they run at only 16 bit (like a 386SX). Therefore, they can only handle 16 MB RAM. Furthermore, they only have 1 KB internal cache and no mathematical co-processor. Actually they are just improved 286/386SXs. They are not cloned chips. There are substantial differences in their architecture compared to the Intel chips.

IBM 486SLC2: IBM had their own 486 chip production. The series was named SLC2 and SLC3. The latter was also known as *Blue Lightning*. These chips could be compared to Intel's 486SX, since they did not have a built-in mathematical co-processor. However, they had 16 KB internal cache (compared to Intel's 8). What reduced their performance was the bus interface, which was from the 386 chip. SLC2 runs at 25/50 MHz externally and internally, while the SLC3 chip runs at 25/75 and 33/100 MHz. IBM manufactured these chips for their own PCs in their own facilities, licensing the logic from Intel. The chips were not sold separately.

DX4: Further 486 developments

[top]

Intel's DX4 processors represented an improvement on the 80486 series. The clock speed was tripled from 25 to 75 MHz and from 33 to 100 MHz. Another DX4 chip was speeded up from 25 to 83 MHz.

Contrary to what you might think, the DX4 were not named for a quadrupling. They were named this way because of the registry of Intel's 80486 and 80586 names. The DX4 name is separated from that context, so it could be patented. If DX3 referred to a tripling, this would not work. The same type of problem caused the next generation chip to be named Pentium, rather than 80586.

The DX4 has 16 KB internal cache and operates on 3.3 volt (they will tolerate 5 volt, to accommodate existing system boards). DX and DX2 have only 8 KB cache and require 5 volt with inherent heat problems.

5X86: AMD has made a series of so called 5X86 CPUs. Those are improved 486s, which approach the 5th generation chips, hence their name. Their 120 MHz model is noteworthy. It could easily be tuned to run at 160 MHz.

- Next page
- Previous page

Learn more [top]

Click for Module 3d about the clock frequencies

Click for Module 3e about 6th generations CPUs (Pentium IIs etc.)

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. .

Module 3c. About the 5th generations CPUs

With Intel's Pentium from 1993, a new era began in the continued CPU development. In these pages, we will look at different variations and further development of 5th. generation CPUs.

- The original Pentium

- Next page
- Previous page

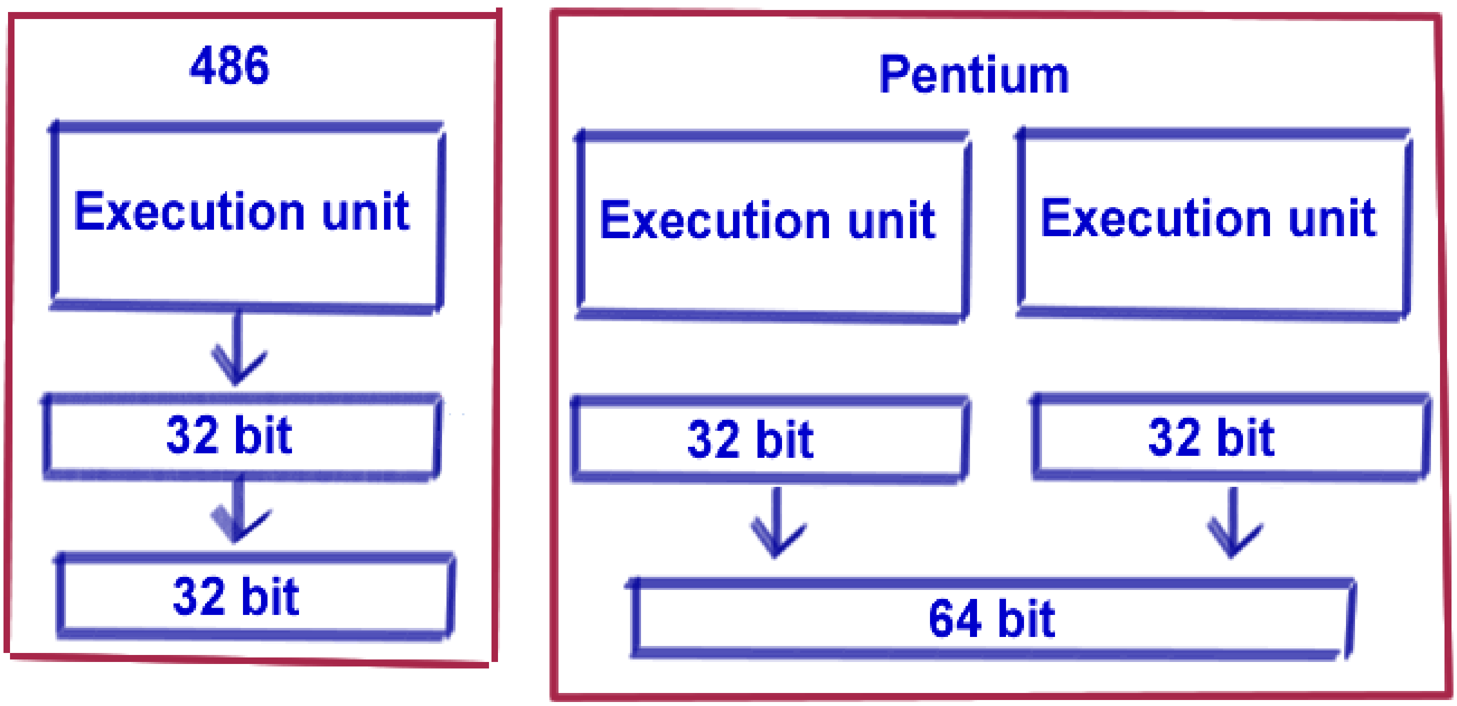
Pentium Classic (P54C)

[top]

This chip was developed by Intel in Haifa, Israel and was released on March the 22th 1993.

The Pentium processor is super scalar, meaning that it can execute more than one instruction per clock tick. Typically, it handles two instructions per tick. In this respect, we can compare it to a double 486.

At the same time, there have been big changes in the system busses. The width has doubled to 64 bit, and the speed has increased to 60 or 66 MHz.



This has resulted in a substantial improvement from the 486 technology.

Two versions to start with

Originally, Pentium came in two versions: a 60 MHz and a 66 MHz. Both operated on 5 Volt. This produced a lot of heat (it was said that you could fry an egg on them!).

The next Pentium (P54C) generation worked with an internal clock doubling of 1.5 times. These chips ran at 3½

Volt. This took care of the heat problem. However, heat coming from the CPU has been a problem ever since.

With these the first P5 processors, Intel carried two Pentium lines; some running at 60 MHz on the system bus (The P90, P120, P150, and P180) and others with 66 MHz system bus (the P100, P133, P166 and P200).

- Next page
- Previous page

Learn more **[top]**

Or continue with the 6th generation CPUs. Click for Module 3e.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. .

Module 3c2. About the 5th generations CPUs - continued

- The Cyrix 6x86
- The AMD K5

- Next page
- Previous page

Cyrix 6x86

[top]

This is a low cost alternative to Pentium. The chip from the Cyrix company, which was introduced February 5, 1996, is a cheap Pentium copy.

The chip was Pentium compatible, since it fitted into a Socket 7. When Cyrix suggested a 6th generation in their naming, it was because the 6X86 employed advanced techniques, which were not found in Intel's Pentium. Thereby Cyrix got improved performance from their chip with the same clock speed.

The Cyrix 6X86 was marketed using a comparison to Intel's clock frequency.

The 6x86 chips had lower internal speed than model name stated. Below, you can see the data for the different models:

Cyrix model	CPU speed	Clock doubling	System bus speed
P120+	100 MHz	2	50 MHz
P133+	110 MHz	2	55 MHz

P166+	133 MHz	2	66 MHz
P200+	150 MHz	2	75 MHz

An interesting detail was, that the 6X86 P200+ was the first CPU to run a system bus speed above 66 MHz. However it was difficult to find motherboards with chip sets capable of this, so the chip never achieved an important position in the market.

Cyrix 6X86s were known for poor performance regarding floating point operations. There also were problems with Cyrix and NT 4.0.

In my experience, the 6x86 did quite a good job with common office programs in Windows 95. I was very satisfied with the P166+ I had. Of course I would prefer a genuine Pentium 166, but I was not willing to pay three times the price at that time.

The 6X86 was later improved with Dual Voltage (like Pentium P55C). This reduced power consumption and heat generation.

Also see the article on Cyrix M3. The company Cyrix was in 1999 taken over by Taiwanese chip producer VIA.

AMD

[top]

AMD is another CPU brand, which has become very important. Their Pentium like chips offered Intel tight competition. AMD used their own technologies, and hence they are not clones. They had these series:

- K5, corresponding to the classic Pentiums (with 16 KB L1 cache and no MMX)
- K6, K6-2, and K6-3 which compete with Pentium MMX and Pentium II
- K7 Athlon, from August 1999, which is not Socket 7 compatible.

K5

K5 was Pentium copy. The old K5 was for example sold as PR133. This means, that the chip should perform like a Pentium P133. However, it only runs 100 MHz internally. It still has to be installed in the motherboard like a P133.

AMD's K5 also existed as PR166. As the name suggests, it was intended to compete with Intel's P166. It operated at just 116.6 MHz internally (1.75 X 66 MHz). According to the highly respected German magazine *c't,* issue *3.97 page 20,* it actually ran at least as fast as the P166.

This was due to an optimized cache and other new developments. The only feature on which it could not match the P166 was in floating point operations. These are typically necessary in 3D calculations in AutoCAD and similar applications.

PR133 and PR166 cost far less than the similar Pentium models, and they were very popular in low budget machines.

- Next page
- Previous page

Learn more **[top]**

Or continue with the 6th generation CPUs. Click for Module 3e.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. .

Module 3c3. The 5th generations CPUs - continued

- The P55C - MMX
- IDT WinChip
- Voltages - dual voltage

- Next page
- Previous page

Pentium MMX (P55C)

[top]

The P55C Pentiums were introduced January 8, 1997. MMX is a new set of instructions (57 new integer instructions, four new data types, and eight 64 bit registers), which expand the capabilities of the CPU. It is an addition to the original Pentium set of instructions.

The MMX instructions were designed for multimedia programs. The programmers can utilize these instructions in their programs. These allow the Pentium to provide improved program execution.

Both Cyrix and AMD use MMX in their 6th generation CPUs (K6 and M2). Programs, which are written with MMX instructions, can still be run on, for example, a Pentium without MMX. However, execution is slower with the traditional instructions.

Please, READ MORE ON MMX HERE.

More L1 cache and higher clock frequency

Compared to the Pentium Classic, the Pentium MMX were further improved with 32 KB L1 cache (the old one had 16 KB). There were also other improvements in the CPU. These improvements together meant 10-20% better performance at similar clock speeds. The clock frequency of the new processors were 166, 200 and 233 MHz.

Dual voltage

The P55C required a new motherboard. Not because of MMX - that is pure software, but because of changes in the power supply. The P55C operated with *dual voltage* technology. To reduce heat generation, this chip requires two different voltages: 2.8 Volt to *the nucleus* and 3.3 Volt to the I/O section. The old motherboards for the P54Cs have only one voltage to the CPU. Thus, the new CPU requires a new motherboard.

Tillamook

For use in laptops Intel has a special power-saving version of the Pentium MMX. The so-called Tillamook processor is manufactured using 0.25 micron technology, and you find it in 266 and 300 MHz versions.

IDT Winchip

[top]

IDT was another smaller company to produce low-priced Pentium MMX-like CPUs. Their first WinChip C6 was introduced in May 1997. The company wanted to deliver 200 MHz Pentium MMXs for $50.

About the IDT WinChip C6 CPU

- 5.4 million transistor
- 0.35 micron, 4-layer metal CMOS technology
- Socket 7 compatible
- 200 MHz

About the WinChip 2 3D, released May 19, 1998

- Socket 7 processors running at 266 MHz and 300 MHz
- 0.25-micron process technology
- 2.5-volt IBM Blue Logic technology
- 6 million transistors
- Die size only 88mm^2 making it the smallest x86 processor in the world
- Superscalar MMX and 3DNow!
- Fully pipelined floating point unit
- 100 MHz bus support IDT expected to continue the development of their WinChips. They wanted to double up the L1 cache for better performance and to introduce a *superpipeline*

technology, which also will speed up the whole thing.

We never saw many IDT chips in my country.

In 1999 the company was taken over by VIA who integrates the IDT technology in their Cyrix processor line.

Voltages - dual voltage

[top]

One of the most important CPU technologies is the continually thinner wires inside the chip. With thinner wires, the CPU can operate at lower voltage. That results in a smaller CPU generating less heat and with the ability to operate at higher speeds. A step in this development is the design of dual voltage chips:

- The interface to the I/O bus, which always requires 3.3 volt.

- In internal CPU parts, it is advantageous to reduce the voltage as much as possible. This can be done because of the extremely thin wires in the CPU. The Socket 7 motherboards have a two part voltage regulator to match the needs of the CPU. Here are some selected CPUs and their voltage requirements:

CPU	Internal voltage	I/O voltage
Pentium MMX	2.8 Volt	3.3 Volt
AMD K6	2.8/2.9 Volt	3.3 Volt
Cyrix 6X86MX	2.8 Volt	3.3 Volt
Pentium II "Klamath"	2.8 Volt	3.3 Volt
AMD K6-2	2.2 Volt	3.3 Volt
Pentium II and III	2.0 Volt	3.3 Volt
Pentium III "CuMine"	1.6 Volt	3.3 Volt

- Next page
- Previous page

Learn more [top]

Or continue with the 6th generation CPUs. Click for Module 3e.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

---

www.karbosguide.com

Module 3c4. The 5th generations CPUs - continued

- Chip production
- Moore's Law
- Various notes about CPUs

- Next page
- Previous page

Chip production

[top]

It takes a long time to manufacture a CPU. 5 to 50 million transistors must be placed on a tiny silicon wafer. Actually, it required 90 workdays 24 hours round-the-clock to produce a Pentium CPU.

CPUs are manufactured in large wafers containing maybe 140 to 150 CPUs. Usually 110 to 120 of these perform perfectly. The rest are discarded. The wafers are burned, etched, and treated in hour long processes - layer by layer. In the CPU there are up to 20 layers of silicon wafers with millions of micro transistors.

Process technology

The CPUs are processed using CMOS technology with smaller and smaller "wires". The result is smaller "dies" (the little area inside the chip holding all the transistors) with more and more transistors. The power consumption goes down, and the clock frequency goes up.

CPU	Process technology	Number of transistors	die size
486	1.0 micron	1,200,000	79 mm^2
Intel Pentium	0.5 micron	3,100,000	161 mm^2
Cyrix 6X86	0.5 micron	3,100,000	392 mm^2
Intel Pentium MMX	0.35 micron	5,500,000	128 mm^2
AMD K6	0.25 micron	8,000,000	68 mm^2
Intel Pentium II	0.35 micron 0.25 micron	7,500,000	131 mm^2
Intel Celeron	0.25 micron	7,500,000	131 mm^2 155 mm^2
Cyrix MII	0.25 micron	6,500,000	119 mm^2
IDT WinChip 2 3D	0.25 micron	6,000,000	88 mm^2
AMD K6-2	0.25 micron	9,300,000	81 mm^2
AMD K6-3	0.25 micron	?	118 mm^2
AMD ATHLON	0.25 micron	22,000,000	184 mm^2
Intel Pentium III CuMine	0.18 micron	28,000,000	106 mm^2
AMD ATHLON "Thunderbird"	0.18 micron	37,000,000 (22 mil. + 15 mil.)	117 mm^2
Intel Pentium 4	0.18 micron	42,000,000	217 mm^2

Intel Pentium 4 Northwood	0.13 micron	42,000,000	116 mm^2
Athlon T	0.13 micron	37,000,000	80 mm^2

Here you see the Intel process generations:

Process generation	Year	Gate length
P648	1989	1.0 micron
P650	1991	0.8 micron
P852	1993	0.5 micron
P854	1995	0.35 micron
P856	1997	0.25 micron
P858	2000	0.18/0.13 micron

Moore's Law

The CPUs have doubled their calculating capacity every 18 months. This is called "Moore's Law" and was predicted in 1965 by Gordon Moore. He was right for more than 30 years. The latest CPUs use internal wiring only 0.25 microns wide (1/400 of a human hair). But if Moore's Law has to be valid into the next century, more transistors have to be squeezed onto silicon layers.

IBM succeeded as the first in making copper conductors instead of aluminum. Copper is cheaper and faster, but the problem was to isolate it from the silicon. The problem has been solved with a new type of coating, and now chips can be designed with 0.13 micron technology. The technology is expected later to work with just 0.05 micron wiring!

Texas Instruments announced on August 27th 1998 that they expect 0.07 micron CMOS processing in the year 2001.

AMD was the first company to mass-produce copper-wired CPU's. This happened in their fab 30 in Dresden, April 2000.

Chip errors

[top]

The following miscalculations occur in 386, 486, and Pentium, when running Excel, Works, or

Pascal, with the numbers 49 and 187:

A1	=1-(1/49*49)	
	A	B
1	1,11E-16	

All CPUs have faulty instructions. Recently flaws have been discovered within the Pentium II and Cyrix 6x86MX.

The Pentium scandal

[top]

Pentium was hit by a scandal in late 1994, when an error in the mathematical co-processor (FPU) became publicly known. It simply miscalculated at a given division. Intel knew of the error from early that summer but more or less kept it secret.

Intel insisted that the error would occur extremely rarely. Compaq immediately modified their production to disable the FPU. Shortly thereafter, IBM announced that they would stop the production of Pentium based PCs. IBM had calculated that the error would occur every 24 days. At the time, IBM was working to extricate themselves from the Intel CPU monopoly. They were moving towards Power PC, Cyrix, and NexGen based PCs. Thus the scandal played right into their hands. You see the error here, where A3 should be equal to A1:

A3	=A1/B1*B1		
	A	B	C
1	4195835	3145727	
2			
3	4195579		

Intel underestimated the significance of the miscalculations, certainly regarding users employing complex mathematical calculations. IBM over dramatized the error for political reasons. This all happened in December 1994, while Intel was running their big TV campaign for Pentium.

That gave birth to a number of jokes: How many Pentium programmers are needed to screw in a bulb? (answer: 1.9990427). Why is Pentium not named 586? Because it would have to be called 585.999983405! In a different vein: How many Apple employees does it require to change a bulb? 7! One to hold the bulb and 6 to design T-shirts. And: how many IBM employees does it require to change a bulb? None! IBM simply announces a new feature called "black bulb."

Intel Owner's Club site

[top]

This site is good if you are interested in the CPUs. Find the Intel Owner's Club, which is a free, easy way for members to:

- get the scoop on the latest Intel technologies
- get info on hot new software and technologies
- interact with Intel & technology experts
- download free software and games
- enter contests.

My membership has helped me to learn how to use the Intel web site, which holds a lot of information. Only Intel's servers can be terrible slow, so you easily get tired from them.

- Next page
- Previous page

Learn more **[top]**

Or continue with the 6th generation CPUs. Click for Module 3e.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com.

Module 3d.1 About Cooling and Over clocking

The contents:

- About cooling
- A clean cooler..

- Next page
- Previous page

Cooling

[top]

All modern CPUs share a common need for cooling. Make sure to include a good cooler. It has to be matched to the size of the CPU.

- It has to be attached properly, either with glue or a clamp, which fits the CPU.
- It must have a substantial size heat sink - the bigger the better.
- The fan must be mounted in roller bearings, to minimize noise.

The bigger the fan and heat sink, the better it is. The CPU will operate more reliably. It will have a longer life span, and it can possibly be over clocked. If you buy Intel CPUs, buy them "in a box". It is a special package, priced slightly higher than just the CPU. They always include a good fan and a three year warranty.

Pentium with fan. Photo taken with Canon Powershot 600. JPG-file 1:30, 32 KB.

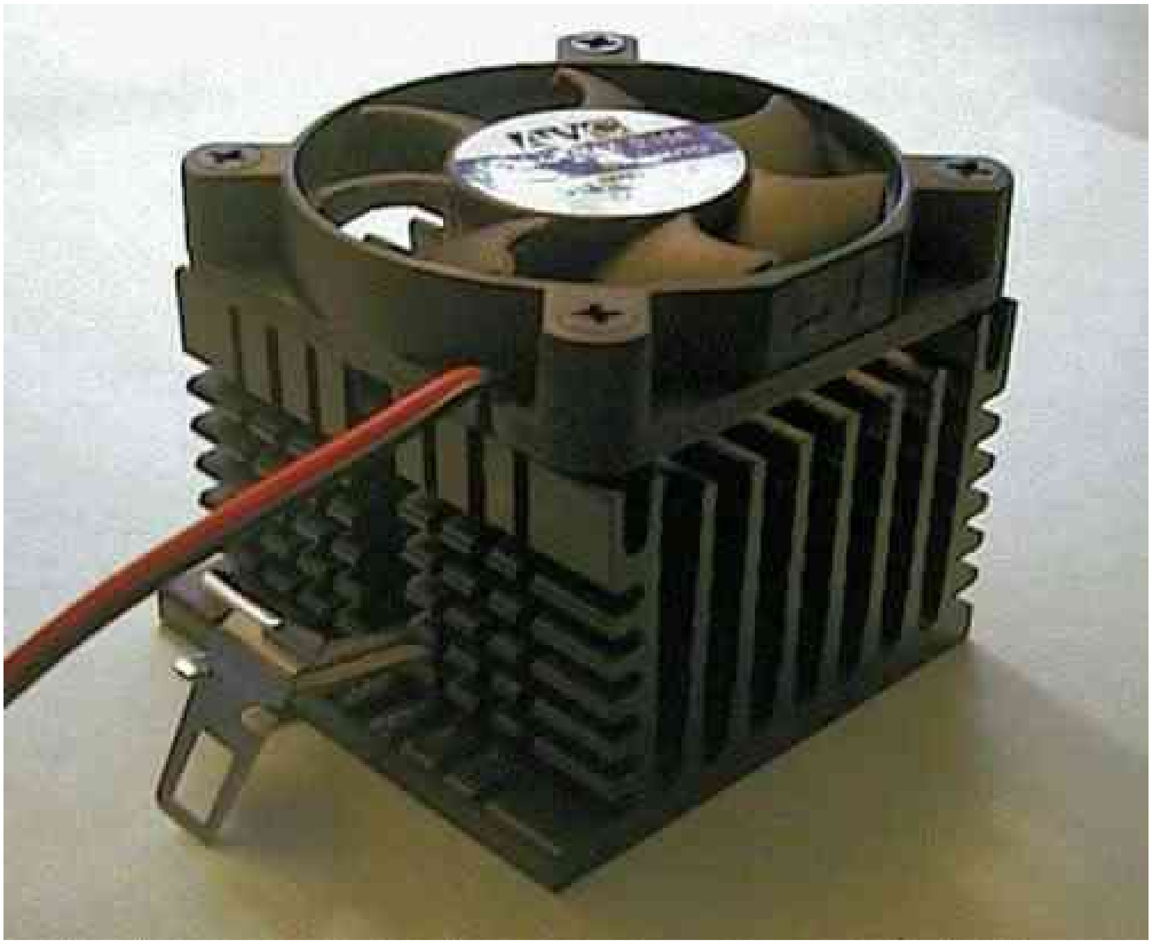
What is a cooler?

[top]

A cooler consists of two parts:

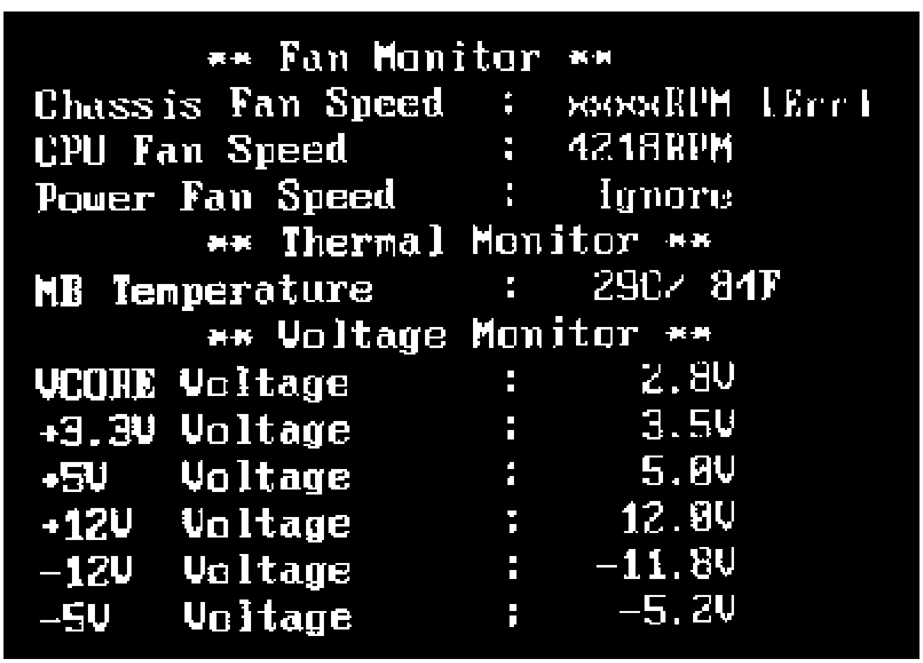
- A fan that needs power supply.

- A cooling element, usually made of metal ribs. The fan is placed on the top of the cooling element, which is fastened very tight to the top of the CPU:

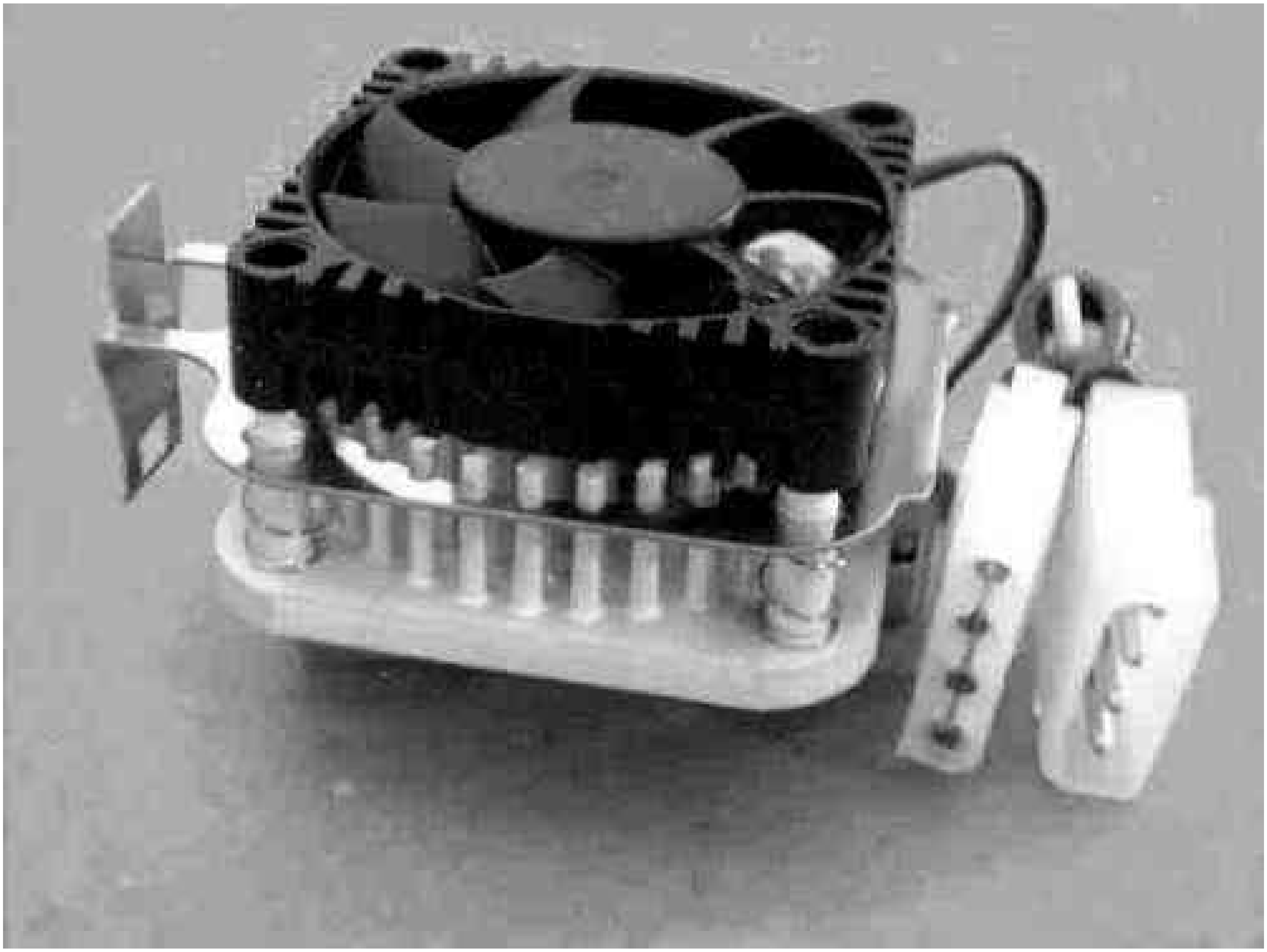


The power supply can be connected two ways:

- From the main power supply of the PC. This is the case in most PCs and all older ones.
- From the motherboard. This way the rotation can be monitored by the BIOS software which then can control the temperature of the CPU. This system is implemented on many ATX-boards. Here you see the BIOS program monitoring the temperature (29 C on my board, right now):



Some coolers use *peltier elements* which give an extra cooling. Look at this one below where you see the white peltier-thing at the bottom. Notice the two-fold power supply:

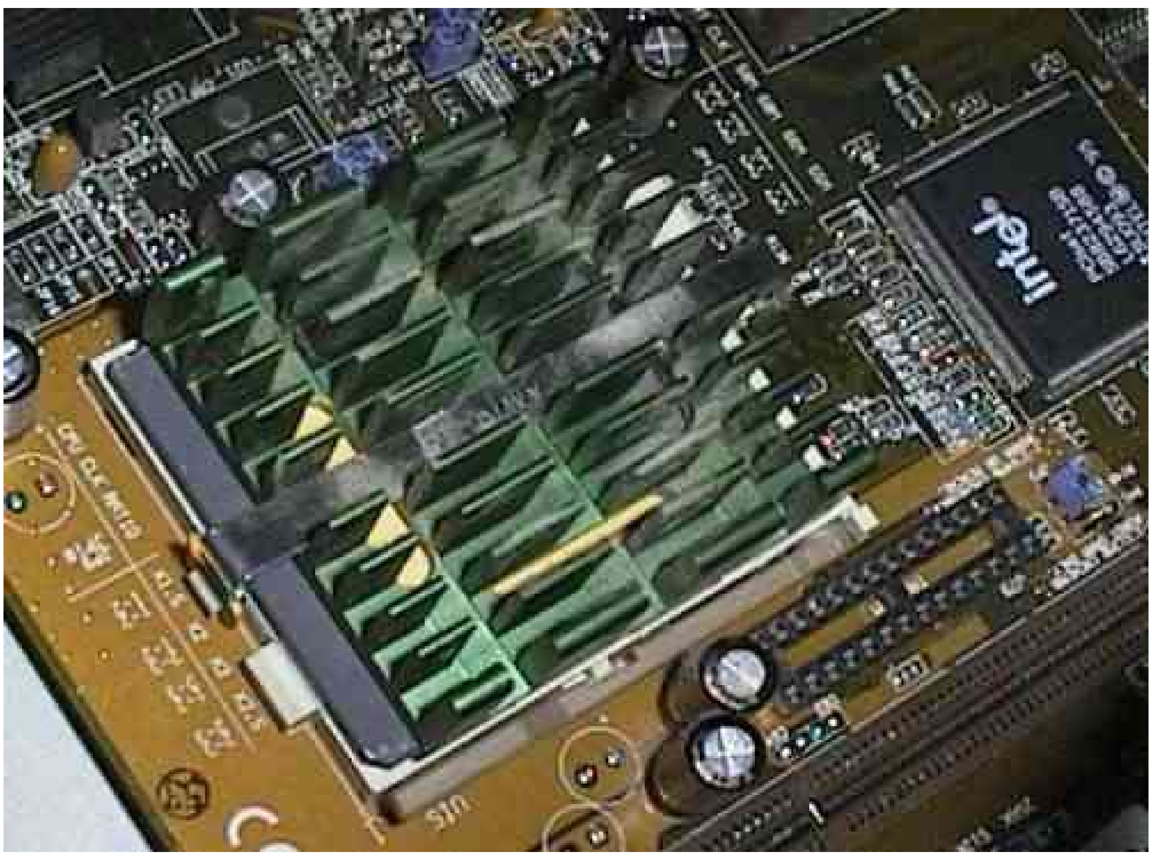


Cleaning the cooler

[top]

Another important thing to take care of is vacuum cleaning the fan on a regular basis. My old Pentium Pro has a very big fan on it. It began giving error messages within Windows . I really could not find out why. Until I discovered that the heating sink was extremely hot. The fan was rotating as it should, but a large amount of dust had gathered just beneath it, so the air did not cool the sink at all!

You should separate the fan from the cooling element to clean it properly. Here is the cooling element alone on the top of the CPU:



What to be learned: Check your CPU fan once a year. Perhaps you have to disconnect the CPU to clean it thoroughly. Take the CPU in your hand and hold the vacuum cleaner close to the sink.

Here you see a powerful cooling device for (over clocking) Pentium IIs. It contains three fans (the third being difficult to see, it's in the middle of the device) plus a peltier element:

- Next page
- Previous page

Learn more

[top]

Read Module 3e - about the latest CPUs.

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the chip sets in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3d.2

About Cooling and Over clocking (continued)

The contents:

- What is clocking?
- Two frequencies to clock on..
- What is over clocking?

- Next page
- Previous page

What is clocking?

[top]

All Pentium CPUs run with clock doubling. That's the way they are built. The PC works with two frequencies, which the user can adjust. The clock doubling is set on small *jumpers* on the motherboard. You simply set a clock doubling factor, to make the CPU work – but who says that you must use the factor listed in the manual?

If you are brave, you try to set your CPU to run *faster* than it is designed to run. Often it works. If you "cheat" the CPU in this manner to work faster, it is called *over clocking.* Over clocking is kind of a PC tuning, which can be fun to fool with – if you are interested in the technicalities of PC-hardware. Otherwise - skip it!.

If you are lucky, you can make a medium speed CPU run as fast as the top of the line version! Please note, I accept no responsibility for the result of your experiments. I will now try to explain the technologies in the over clocking phenomenon. The interesting part is that, like much of the theory I tried to describe in in the modules 3a, 3 b and 3c, it all comes together here in the clock doubling technology.

Two frequencies to clock

[top]

The CPU works on two frequencies: An *internal* and an *external.*

- The external clock frequency (the bus frequency) is the speed between the CPU and RAM. In the Pentium CPUs it is actually the speed between L1 and L2 cache. In the Pentium II it is the speed between L2 cache and RAM.

- The internal clock frequency is the speed inside the CPU, that is between L1 cache and the various CPU registers.

For practical reasons you let these two frequencies depend on each other. In practice you choose a given bus frequency (between 60 and 153 MHz) and double it up a number of times (between 3½ and 8). The latter frequency become the CPU internal work frequency.

Here I show a number of *theoretical* CPU frequencies, resulting form different clock doublings: Many of these frequencies will actually never be used, but they are *possible* because of the system structure:

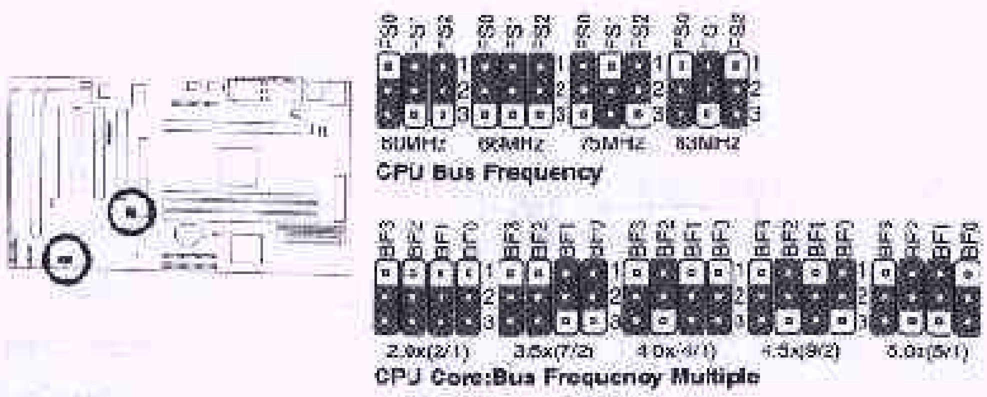
Bus frequencies	Clock doubling factors	Examples of resulting CPU frequencies
60 MHz 66 MHz 75 MHz 83 MHz 100 MHz 117 MHz 133 MHz 153 MHz	1½, 3½, 4, 4½, 5, 5½, 6 6½, 7 7½, 8	233 MHz, 266 MHz, 333 MHz 333 MHz, 366 MHz, 400 MHz, 433 MHz, 466 MHz, 500 MHz, 300 MHz, 338 MHz, 375 MHz, 375 MHz, 416 MHz, 458 MHz, 468 MHz, 527 MHz, 585 MHz 533 MHz, 600 MHz, 667 MHz, 612 MHz, 688 MHz, 765 MHz

Note an important point: The CPU frequency is the result of the the bus frequency multiplied with a factor. If you increase the bus frequency, it affects the CPU frequency, which is also increased.

Look here at a page from the manual to a ASUS P2L97 motherboard. It has a clear instruction about how to set the two values (bus frequency and clock factor). This (old) motherboard accepts bus frequencies up to 83 MHz with a clock factor up to 5:

2. CPU Bus Frequency (FS0, FS1, FS2)
This option tells the clock generator what frequency to send to the CPU. This allows the selection of the CPU's *External* frequency (or *BUS Clock*). The BUS Clock multiplied by the BUS Ratio equals the CPU's *Internal* frequency (the advertised CPU speed).

3. CPU Core:BUS Frequency Multiple (BF0, BF1, BF2, BF3)
This option sets the frequency ratio between the *Internal* frequency of the CPU and the CPU's *External* frequency. These must be set in conjunction with the *CPU Bus Frequency.*



WARNING! Frequencies above 66Mhz exceed the specifications for the onboard Intel Chipset and are not guaranteed to be stable.

Intel Pentium II Processor in a SEC Cartridge
(233-333MHz 256/512KB L2 Cache)

Set the jumpers by the Internal speed of your processor as follows:

CPU Model	Freq.	Ratio	BUS F.	FS2	FS1	FS0	BF3	BF2	BF1	BF0
				(BUS Freq.)			(Freq. Ratio)			
Intel Pentium II	333MHz	5.0x	66MHz	[1-2]	[1-2]	[1-2]	[2-3]	[1-2]	[1-2]	[2-3]
Intel Pentium II	300MHz	4.5x	66MHz	[1-2]	[1-2]	[1-2]	[2-3]	[1-2]	[2-3]	[1-2]
Intel Pentium II	266MHz	4.0x	66MHz	[1-2]	[1-2]	[1-2]	[2-3]	[1-2]	[2-3]	[2-3]
Intel Pentium II	233MHz	3.5x	66MHz	[1-2]	[1-2]	[1-2]	[2-3]	[2-3]	[1-2]	[1-2]

What is over clocking?

[top]

Since clock doubling and bus speed can be freely adjusted on the motherboard according to your desires, you can in principle make the CPU run at 600 MHz. You set the bus to 133 MHz and the clock factor to 4½. Then the CPU runs at 600 MHz – if it runs. The question is whether the chip will tolerate that - and if it will give a stable performance, since clock doubling means more than added heat.

We have now seen that there are two frequencies which can be manipulated, if you want to re-clock the CPU:

- The bus frequency can be increased, let's say from 133 to 153 MHz.

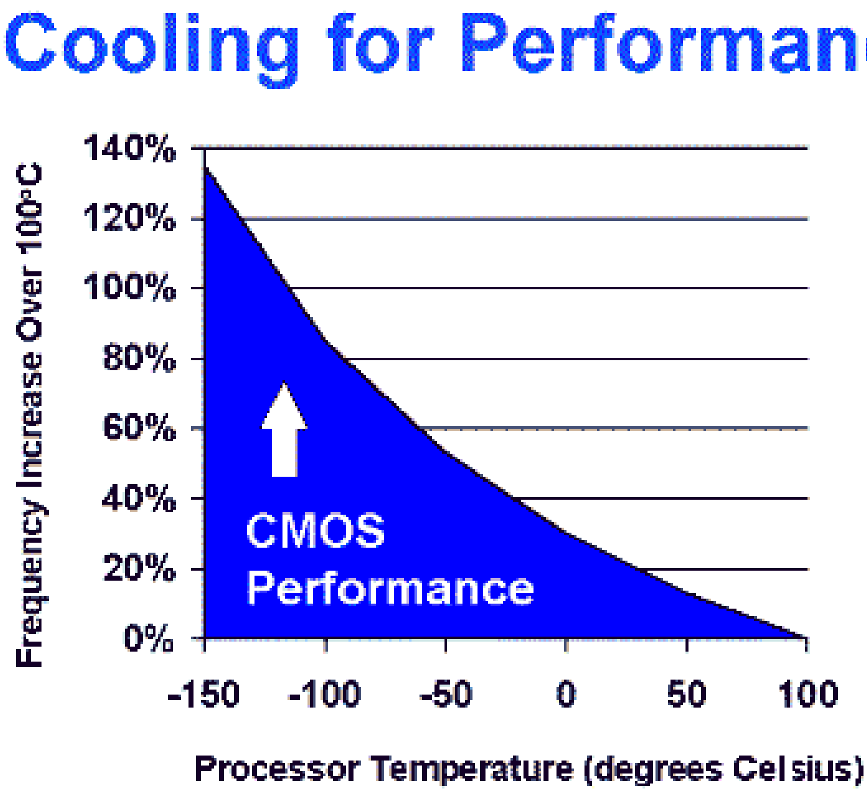
- The CPU frequency can be increased. That can happen as a result of an increased bus speed, which also affects the CPU frequency, or it can happen by using a greater clock factor. The latest is not possible anymore.

Both techniques result in a faster PC. If the bus frequency is increased, it affects all data transport to and from RAM. It will work faster, to the joy of all work done on the PC. However, the RAM has to cope with the increased speed.

When the CPU internal frequency is increased, many applications will be happily affected.

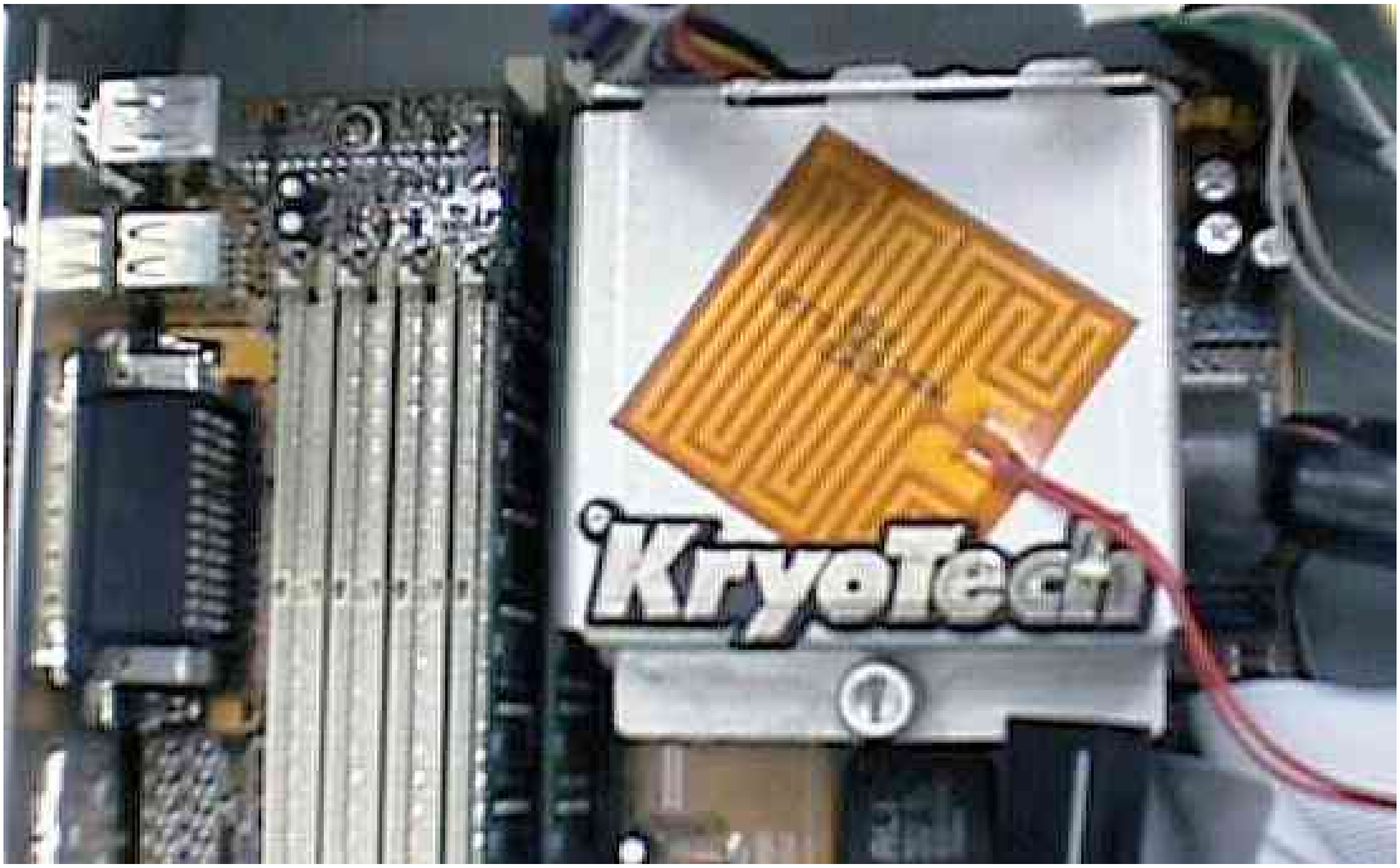
More cooling

The tuning will often work, but it requires good cooling of the CPU, the more cooling the higher you can have the clock frequency. CPUs are built in CMOS technology. That is a type chip which works better the cooler it is. See this relationship between temperature and performance:



You can see that the performance drops drastically with increased CPU temperature.

This problem caused the Kryotech company to manufacture coolers utilizing the Danish Danfoss compressors, just like in refrigerators. See this cooling unit on a CPU:



It is fed from the compressor in the bottom of the cabinet:

This form of cooling is extreme, but it works. Kryotech can make a standard CPU work at 400-700 MHz! But it requires that it is kept constantly cooled to -40 degrees F or C. (it is the same). The Kryotech setup is efficient, but it is expensive, noisy and power consuming.

If you like, look at Kryotech's Home Page http://www.kryotech.com/

Another company in this business is Asetek.

This was to demonstrate that over clocking can be a serious issue....

However, the CPU speed has become less important. To most users it really does not matter whether you have a CPU running at 300 or at 600 MHz.

- Next page
- Previous page

Learn more [top]

Also see Module 3e - about the latest CPUs.

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3d.3

About Cooling and Over clocking (continued)

The contents:

- Which CPUs can be over clocked?
- Risks in over clocking?

- Next page
- Previous page

Which CPUs can be over clocked?

[top]

The first CPUs which were dramatically over clocked were AMD's 5x86 series. That was a 486

CPU, which could be forced up to an excellent performance at 160 MHz.

Since then especially Intel's Pentium CPUs have been over clocked. Many of those seem to be sold with specs far from their optimum performance. Actually it was so easy that as a result many P133s were sold in 1996 as fake P166s. They worked fine, and the users did not know it. But Pentium MMX and Pentium II can also be re-clocked.

It appeared that Intel were aware of this activity, and they don't seem to care. Unfortunately their CPUs came in two groups:

- Clock doubling works.
- Clock doubling does not work - it is *disabled* by the manufacturer.

You cannot guarantee that it always will work. But let me show a couple of examples, which I have made work with good results:

CPU	Manufacturers spec	Tuning result
Intel Pentium	2½ X 60 MHz = 150 MHz	3 X 66 MHz = 200 MHz
Intel Pentium Pro	3 X 66 MHz = 200 MHz	3½ X 66 MHz = 233 MHz
Intel Pentium II	3½ X 66 MHz = 233 MHz	4 X 75 MHz = 300 MHz
Intel Pentium II	4½ X 100 MHz = 450 MHz	4½ X 117 MHz = 527 MHz

Looking at the three examples, number 1 and 3 show the best results, where both bus frequency and clock factors are increased. That simply moved the CPU up one class in performance.

Here is a table of the clock factors, which the CPU's theoretically can accept (according to my studies):

CPU	Clock factor
Intel Pentium (P54C)	1½, 2, 2½, 3
Intel Pentium Pro	2½, 3, 3½, 4
Cyrix 6x86	2, 3
Cyrix 6x86MX (M2)	2, 2½, 3, 3½
Intel Pentium MMX (P55C)	2, 2½, 3, 3½
AMD K5 PR75 - PR133	1, 1½

AMD K5 PR150 and PR166	2
AMD K6-2 and K6-3	4, 4½, 5
Intel Pentium II, Celeron and Pentium III	Up to 8 and 12 (latest models)

Some AMD and Cyrix chips were special, in that they did not always respond to motherboard settings. It is like they determined their own frequencies.

All modern Intel processors are locked at fixed clock factors (*Multiplier Locking*). They only operate with one specific multiply factor.

The Celeron

The original Celeron was a Pentium II without L2 cache. This CPU was very overclocking friendly. There are several reports about 300 MHz Celerons working at 504 MHz without any problems at all.

The Celeron A

The Intel Celeron line starting with models 300A and 333 (both with 128 KB L2 cache on-chip) are both protected against overclocking. They hold a "Multiplier Locking", which locks them to the clockfactors 4.5 and 5.0 respectively.

The Celeron 533 will only work with clockfactor 8, so if you want to overclock it, you have to go for a motherboard with adjustable system bus frequencies. This could be 8 X 100 MHz instead of 8 X 66 MHz increasing the CPU speed from 533 MHz to 800 MHz. Many users have found this in-expensive way to get a higher performance.

Disadvantages and risks in over-clocking?

[top]

Many factors need to be considered, when you start tampering with these system settings. Watch out for:

- Heat. Can the CPU dissipate the heat?
- The L2 cache RAM of old Pentium II, III or Athlon cartridges - how fast can it work?
- RAM speed. Can it keep up with the system bus?

- The I/O bus. Can PCI and EIDE units keep up?

- Will the software still work?

The last two problems are associated with increased system bus speed. This kind of over clocking gives the best results. However those also create the biggest problems, at least in my experience.

The CPU gets hot

First of all the higher CPU frequency causes more wear on the chip. It is said that a CPU can last 10 years. However do not count on that if you over clock it. Actually I am less concerned about the wear. Of course you should not allow the chip to over heat, but I have never heard about burnt out CPUs. In news groups you can read about various monster fans used for cooling of totally over clocked CPUs.

RAM speed

Another problem is in the relationship with the bus frequency. Here we are talking about the system bus, which connects RAM with the CPU. If you increase this speed, RAM must be able to keep up. Here is a guideline table for the maximum bus frequencies with different RAM types:

RAM type	Speed	Maximum bus frequency
FPM	60 ns	66 MHz
EDO	50 ns	75 MHz
SD	10 ns	100 MHz
SD	7 ns	133 MHz

Finally you could say that with cheap CPUs running at 900 MHz and above - you really do not need any overclocking. Most users will not experience any benefit from shifting from say 700 MHz to 1000 MHz.

- Next page
- Previous page

Learn more

[top]

See Module 3e - about the latest CPUs.

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3d.4

About Cooling and Over clocking (continued)

The contents:

- I/O speed and experiments with that
- Side effects
- Fake Pentium IIs
- Jumpers.

- Next page
- Previous page

Experimenting the I/O speed

[top]

The area which has given me the most problems is the increased PCI speed. At 66 MHz the PCI bus runs at half the system bus frequency. At 100 MHz it runs at one third and at 133 MHz one quarter of this frequency. Thus if we increase the system bus, it also affects the PCI bus:

System bus speed	Bus factor	Resulting PCI speed
66 MHz	The half	33 MHz
75 MHz	The half	37.5 MHz
83.3 MHz	The half	41.6 MHz
100 MHz	One third	33.3 MHz
112 MHz	One third	37.3 MHz
133 MHz	One quarter	33.3 MHz

153 MHz	One quarter	38.25 MHz

Side effects

When we increase the PCI bus speed, a number of units are affected. They may not always agree with the faster pace. This includes:

- The EIDE hard disk
- The video card
- The network controller and other I/O cards.

My own experiment with Pentium II

In 1997 I experimented with a very early Pentium II, which was bought as a 233 MHz model.

First I made it run at 3½ X 75 MHz. It worked fine with CPU, RAM (10 ns SD) and hard disk (IBM DHEA). But the net card (a cheap 10/100 Ethernet card) refused. When I copied large volumes of files on the net, it froze up - stopped. It was quite obvious that the problem was in the net card.

I had to accept the traditional 66 MHz. But to soothe the pain, it turned out to run excellently with a clock factor of 4 - thus at 266 MHz.

Within a couple of weeks I was in the mood to experiment again. I now found an adjustment in the setup program. It is called PCI latency. It is not explained anywhere, but it has a default value of 32. I increased it to 36 and increased the bus frequency to 75 MHz – it works. Now the net card runs without problems.

Then I hoped to speed the system bus up to 83 MHz, which should give a significant performance improvement for all RAM transport. My 10 ns SD RAM can certainly handle 83 MHz. But no, it did not work. Regardless of the PCI latency, the PC would not start. This indicates that the PCI latency setting does not work like I expected. Maybe it has nothing to with this - I do not know.

My explanation is, that the video card could not tolerate the 41.5 MHz PCI frequency. Nothing appeared on the screen.

Now the PC runs fine at 4 X 75 = 300 MHz. There can be an occasional unexplained break-down in Windows 95 (that happens under other circumstances also), which I blame on the drastic over clocking. However, the advantages of the significant performance improvement far exceed the annoyance of these small interruptions, which happen far from daily.

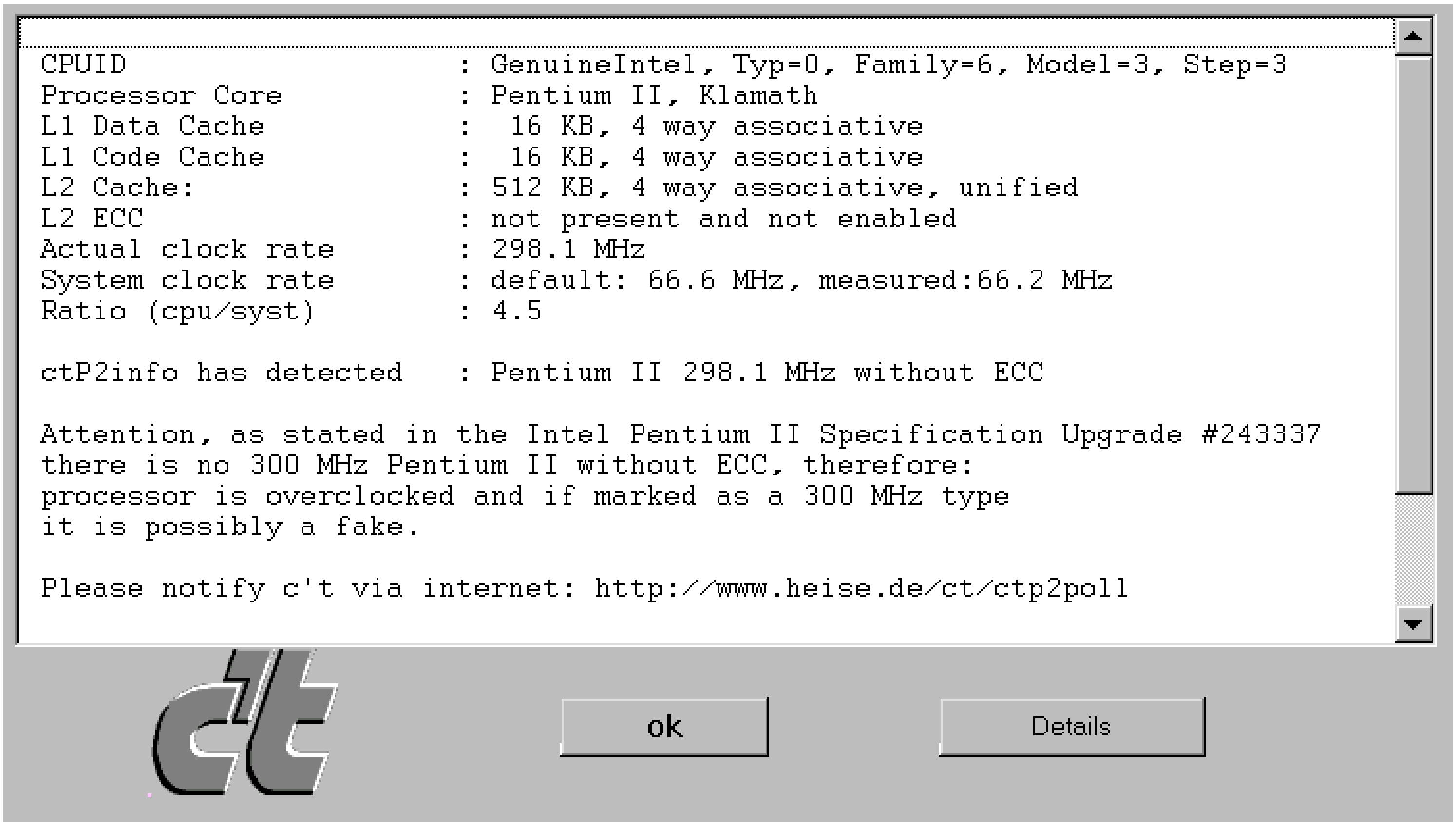
Problems with NT 4.0

Windows NT 4.0 does not install with over clocked CPU. The program tests for "genuine Intel", and seems to register the change in clock frequency. And then it will not work. But if you install NT first, then you can over clock afterwards and NT will work. Actually NT is quite sensitive. One of my friends experienced some peculiar errors. The solution turned out to be moving the RAM module from one socket to another!

Fake Pentium IIs

[top]

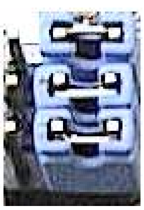
Since some Pentium II-233 perform very well at 300 MHz, they have been sold as such ones. To test your own Pentium II, you can download this test program from C't, which can check your Pentium II. Here is the interface of the Windows 95 version, which correctly detected my CPU to be over clocked:



Jumpers on the motherboard

[top]

To set the clock doubling, some small switches (called jumpers) have to be reset. They are located on the motherboard, as you see here:



You can read in the motherboard manual how to set them. Or you can look at the motherboard! In the picture below you can see some of the printed information *on* the motherboard (this is an ASUS TX97 with a Socket 7).

Here you can read which jumpers to set to select clock doubling 1, 1½, 2, 2½, 3, 3½ and 4 for 6 types of processors:

- P54C and K5
- P55C, K6 and M2 (Cyrix 6x86MX)
- M1 (Cyrix 6x86)

On modern motherboards you may find a software solution to the settings, and that is a lot better.

- Next page
- Previous page

Learn more **[top]**

Also see: Module 3e - about the latest CPUs.

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3d.5

About Cooling and Over clocking (continued)

The contents:

- An example of overclocking
- The SDRAM speed
- Features of the Abit BX6 motherboard.

- Next page
- Previous page

An over-clocked Pentium II

[top]

In the previous pages you can read about the theory behind over clocking. Here I describe a practical case.

In April 1999 we needed a new workstation. It was to used for graphics work and sometimes video editing, so it had to be speedy. We decided to try some over clocking.

Over clocking with Intel - earlier results

[top]

Intel CPUs have always been good for over clocking.

Back in 1997 we had a Pentium Pro designed for 200 MHz. It ran (and still runs) at 233 MHz without any problem at all.

Later we got one of the first Pentium IIs. These processors were very friendly to over clocking, both the frequency of the system bus as well as the clock factor could be changed. A modest 233 MHz version ran (and still does) at 300 MHz.

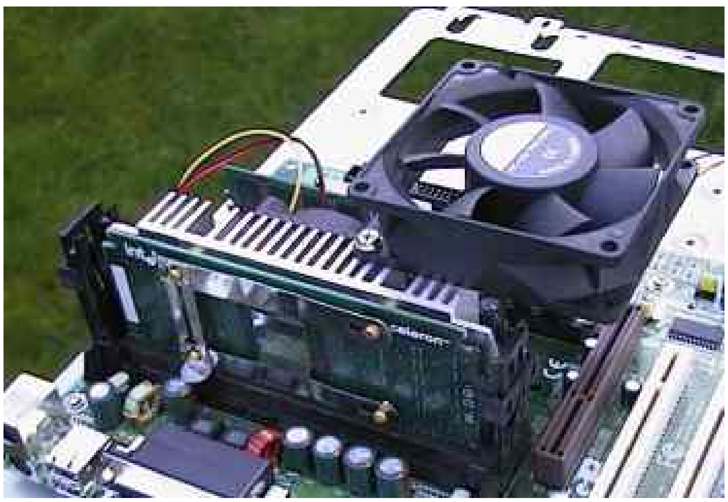
The Deschutes kernel of second generation Pentium II and Celeron was changed, so every CPU only could work with a specific clock factor. This means that you only can over-clock by increasing the bus frequency. This has been the situation with all later Intel processors.

You see our over clocking results as described are not extreme. This has a reason; all our PCs function in a network and they are heavily used for various demanding tasks. So they have to be completely stable, which they also have been. Further over clocking would aggravate the inherent un-stability.

The first attempt

[top]

We started up with the cheapest solution. A 300 MHz Celeron should be doing fine at 450 MHz if the system bus was increased from 66 MHz to 100 MHz. We even added extra cooling, a fan placed above the SEC module:



It never worked. But the motherboard was interesting, so we went for another approach.

Pentium II and Abit BX6

[top]

We then purchased a Pentium II-450 MHz. This processor was the clock factor 4.5 model of the Pentium II you could say.

The motherboard was the newest version (2.0) of the pretty well-known Abit BX6. It is a BX-based board with is capable of delivering a lot of different frequencies. The clock multiplier goes up to factor 8, but since the Pentium II only works with clock factor 4.5, we had these options:

Bus frequency (SDRAM speed)	Clock factor	Resulting CPU frequency	L2 Cache speed
66 MHz	4.5	300 MHz	150 MHz
75 MHz	4.5	338 MHz	169 MHz
83 MHz	4.5	375 MHz	188 MHz
100 MHz	4.5	450 MHz	225 MHz

112 MHz	4.5	504 MHz	252 MHz
117 MHz	4.5	527 MHz	263 MHz
124 MHz	4.5	558 MHz	229 MHz
129 MHz	4.5	581 MHz	290 MHz
133 MHz	4.5	599 MHz	300 MHz
138 MHz	4.5	621 MHz	310 MHz
143 MHz	4.5	644 MHz	322 MHz
148 MHz	4.5	666 MHz	333 MHz
153 MHz	4.5	689 MHz	344 MHz

Of course I could not expect my Pentium II to run at 689 MHz. The values are theoretical.

When you increase the bus frequency it affects a lot of units within the PC. This is due to the architecture, where the system bus so to say is a local bus, with other attached buses and units working synchronously. Increasing the bus frequency influences:

- The CPU clock frequency. Often Intel CPUs are capable of working at a higher frequency than what they are sold for. However, improved cooling is important.

- The L2 Cache of the Pentium II module. It has an upper speed limit as all other RAM types do. Cooling is important for the L2 cache RAM chips.

- The SDRAM speed. The RAM modules have to fast enough to cope with the increased bus frequency.

- The PCI units. The graphics controller, EIDE controller and network controller all have to work at around 33 MHz, otherwise un-stability is the result (at least that is our experiences).

- The AGP bus speed.

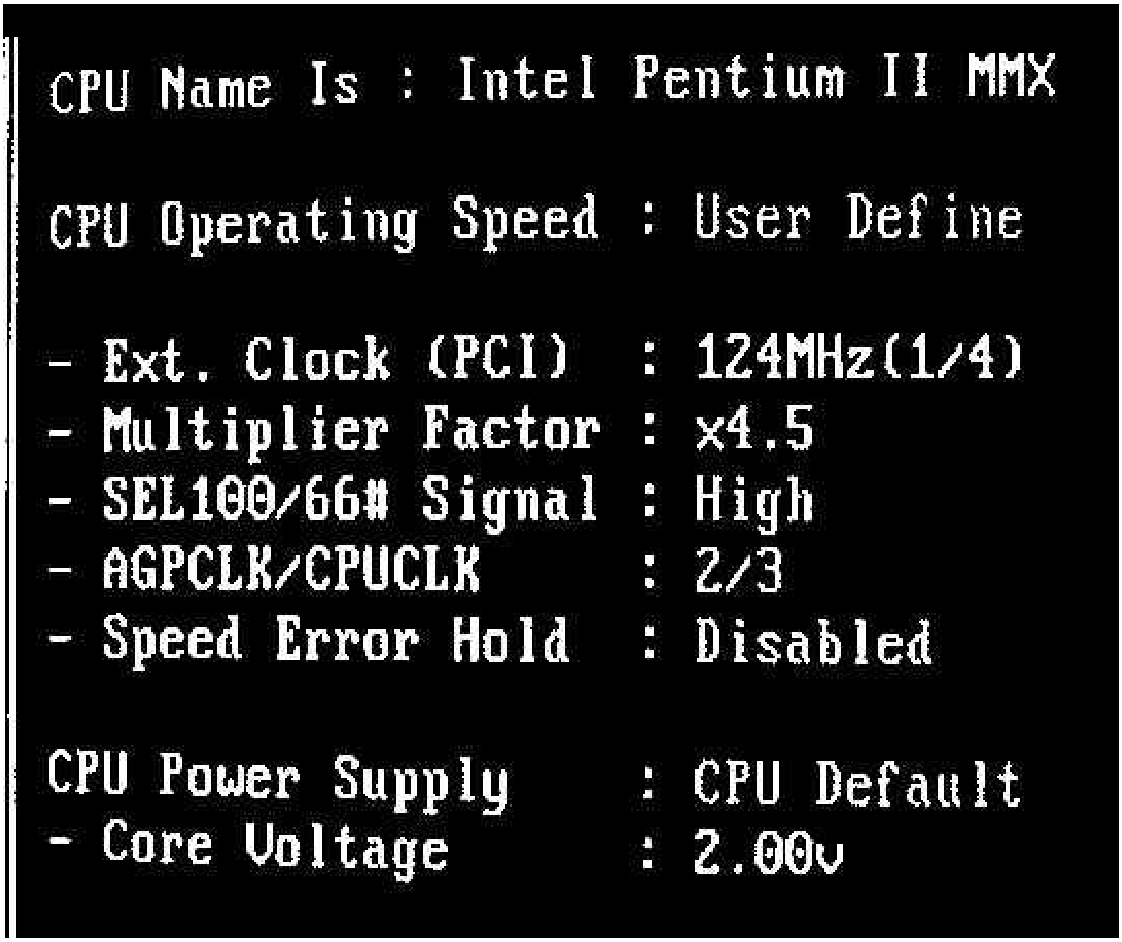
Over clocking a PC is not that simple. All the mentioned units have to be tuned, so they work at right frequencies.

Testing and trying

[top]

One of the biggest problems is to control the speed of the PCI units. Our network (LAN) is a very good tool for testing this. I make a backup of all my documents (> 10.000 files) across the network from harddisk to harddisk, and if this works i am pretty sure that everything is all right with the new PC.

With the Pentium II, I started increasing the bus frequency. Of course everything worked fine at 100 MHz. It should. 112 MHz was completely stable. 117 MHz as well, but at 124 MHz the problems came. Here you see the Soft Menu setting, which is an extremely nice feature of the BX6 board:



The PC seemed to work at 558 MHz, but the file copy-test could not be performed. The PC froze. This probably was due to "slow" SDRAM. With better RAM it might have worked.

SDRAM speeds

[top]

Here is an theoretical calculation of the required SDRAM speed:

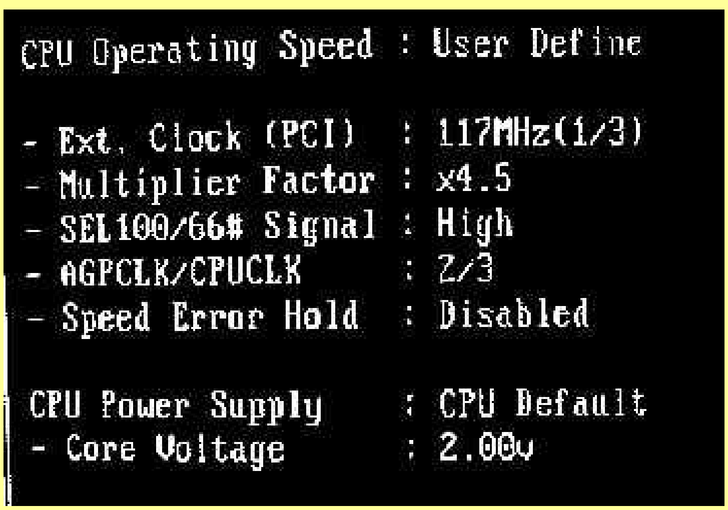
Bus frequency	SDRAM speed (Nano seconds)
66 MHz	15.02
75 MHz	13.33
83 MHz	12.00
100 MHz	10.00
112 MHz	8.93
117 MHz	8.55
124 MHz	8.03
129 MHz	7.75

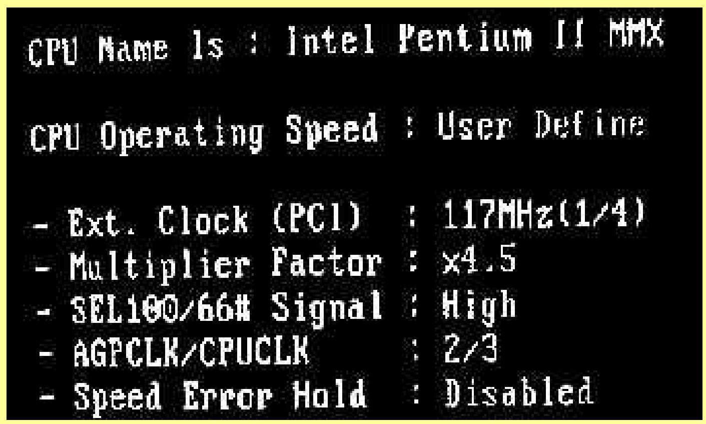
133 MHz	7.52
138 MHz	7.25
143 MHz	6.99
148 MHz	6.76
153 MHz	6.54

The RAM was of PC100 type. But this may be 10, 8 or 7 ns. In our case it was 8 ns, so the 124 MHz setting should have been working, it just didn't.

Two versions of 117 MHz

At 117 MHz I had two options. I could go for a PCI bus at 39 or 29 MHz. These values come out as *one third* or *one quarter* of the 117 MHz bus frequency. Unfortunately 39 MHz was too much for my PCI units:

<table>
<tr><td>Soft Menu setting: PCI 1/3</td><td>Soft Menu setting: PCI 1/4</td></tr>
<tr><td>


</td><td>


</td></tr>
<tr><td>PCI frequency: 39 MHz</td><td>PCI frequency: 29 MHz</td></tr>
<tr><td>System stability: not good</td><td>System stability: 100% all right</td></tr>
</table>

So we ended up with a completely stable Pentium II system running at 527 MHz. That's absolutely OK.

Features of the Abit BX6

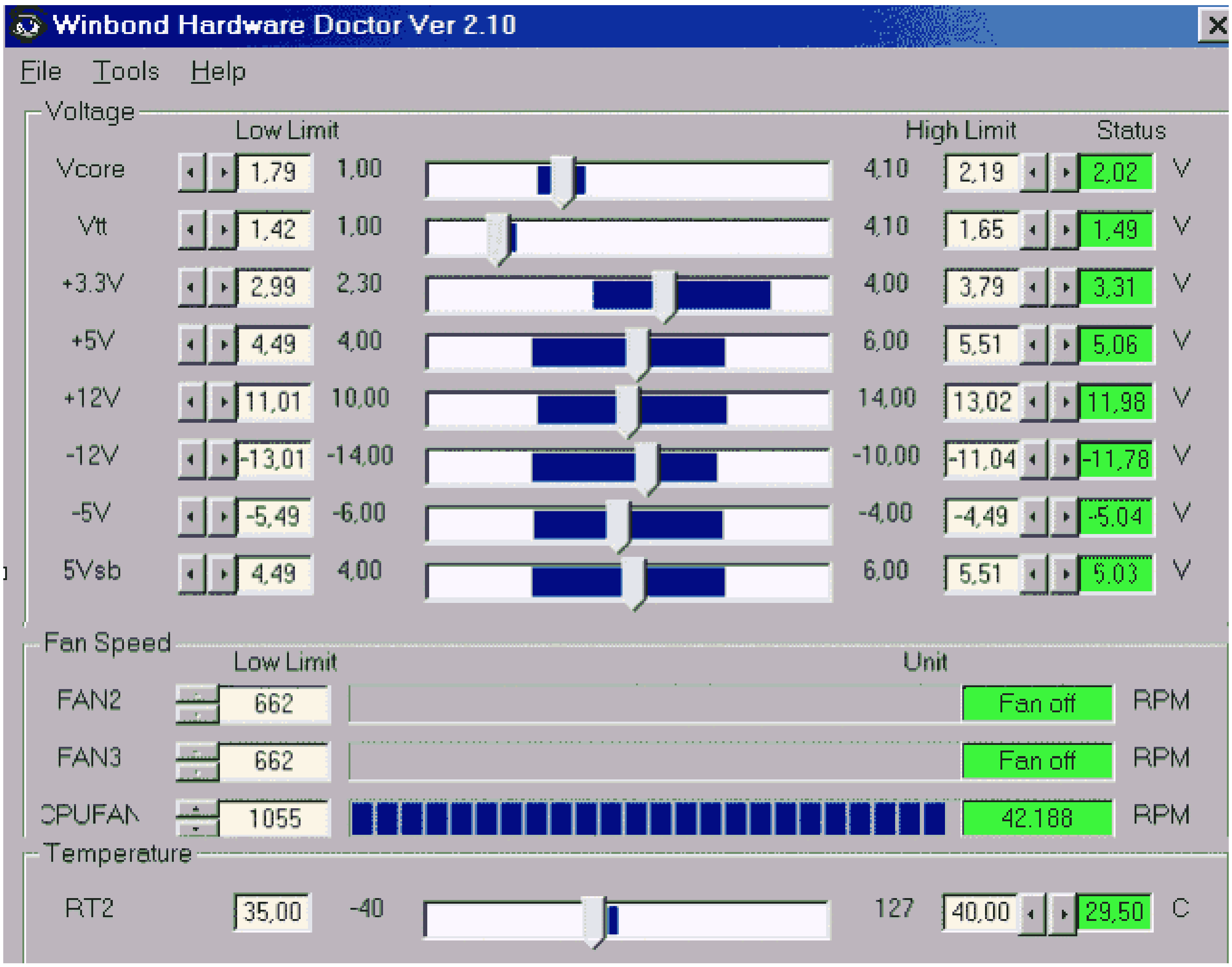
[top]

The Abit board seems pretty cool to me. The manual is OK but not overwhelming impressive. The board has 5 PCI slots which I like. But especially the Soft Menu II is great - a brilliant tool for over-clockers. You do not have to move a simple jumper on the BX6 board, so it is extremely simple to test your CPU and system at various frequencies.

You also get thermistor to detect the CPU temperature:

It is taped to the heat sink and connected to the motherboard.

You get some software, among others this diagnostic tool:



More over clocking?

With better RAM we might tweak the full 689 MHz out of the Pentium II processor. Running with a bus frequency of 153 MHz, the PCI units have to work on 38,25 MHz which I very much doubt they can.

My realistic guess would be that this configuration using 7 ns SDRAM might work:

Bus frequency	CPU frequency	SDRAM speed	PCI frequency
138 MHz	621 MHz	7,25 ns	34,5 MHz

- Next page
- Previous page

Learn more [top]

Read more about the boot process and system bus in Module 2b

Read more about I/O buses in module 2c

Read more about the motherboard chip set in module 2d

Read more about RAM in module 2e

Read about EIDE in module 5b

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e

Module 3e describes the development of 6th generation CPU's. The module is subdivided into the following pages:

01: The Pentium Pro, father of all P6s
02: The first Pentium II
03: The "Deschutes" and the Celerons
04: The P6-like processors: AMD's K6, K6-2, and Cyrix
05: The K6-3
06: The Pentium Xeon
07: The Pentium III
08: The Great Athlon
09: On MMX, 3DNow!, and Katmai
10: On sockets and roadmaps
11: On Intel Itanium (codename "Merced") and IA-64
12: On VIA Joshua
13: AMD Duron
14: Intel Pentium 4

I recommend that you read all the pages one by one. Just follow the links "Next page" to get through the textbook. I hope you find the information useful!

Introduction to the 6th generation of CPUs

The first 6th generation CPU was Intel's Pentium Pro from 1995. However, first from 1997 with both AMD's K6 and the Pentium II the 6th generation performances have been available for us all.

The contents:

- Pentium Pro
- A giant chip
- No DOS with PPro
- Pentium Pro versus Pentium II
- The next module 3e

Pentium Pro was an important CPU, since it became the father to the Pentium II, the Celeron, the Pentium III and made the ground other P6-like processors as K6-2.

Pentium Pro

[top]

Pentium Pro development started in 1991, in Oregon. It was introduced on November 1, 1995.

The Pentium Pro is a pure RISC processor, optimized for 32 bit processing in Windows NT or OS/2. The new hot feature was that the L2 cache is built-in. This is like two chips in one. The new features were:

- Built in optimized L2 cache with 256 KB or 512 KB. This is connected to the CPU itself with a 64 bit back side bus. Thus, the L2 cache runs synchronous with the CPU speed.

- Multiple branch prediction, where the CPU anticipates the next instruction. Data Flow Analysis, which should reduce data dependence. Speculative Execution, where the CPU attempts to anticipate instruction results.

- 5.5 million transistors in the CPU, 15 million for the 256 KB SRAM L2 cache. (6 transistors per bit).

- 4 pipelines for simultaneous instruction execution.

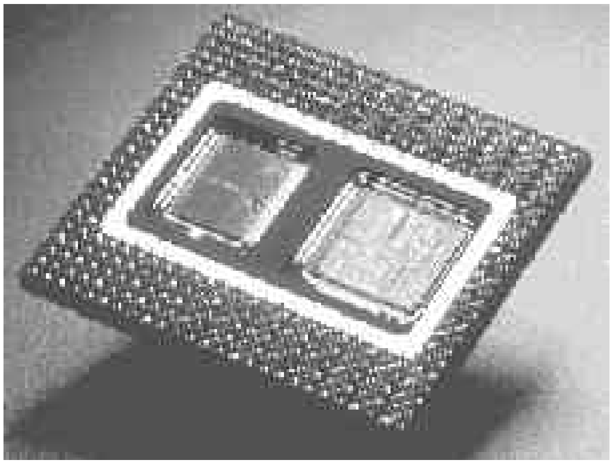
- RISC instructions with concurrent x86 CISC code to MicroOps RISC instructions decoding.

- 2.9 Volt 4 layer BiCMOS processor technology.

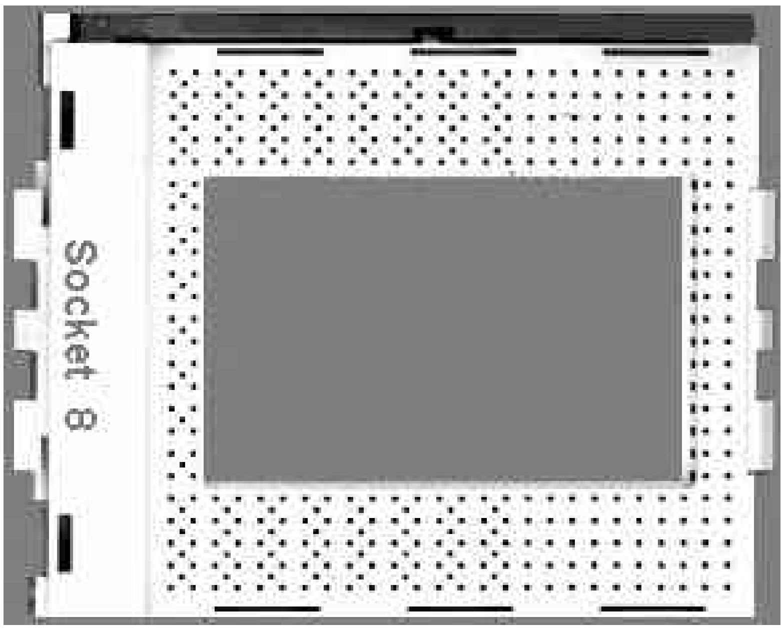
- Patented protocol. Thus, other CPU manufacturers cannot use the Pentium Pro Socket and chip set. This was not to the user's advantage.

A giant chip

Here you see a rectangular chip. The CPU and L2 cache are separate units inside this chip:



It is mounted in a huge Socket 8:



Pentium Pro was not for DOS...

Pentium Pro was primarily optimized to 32 bit program execution. Often you heard about its poor performance executing 16 bit programs. I used a PPro 200 MHz (at 233 MHz) and experienced tremendous power in my Windows 95 environment. However the CPU was aimed at use in servers.

PPro versus Pentium II

After the introduction of Pentium II, the interest in the PPro has declined, and by the end of 1998 it was out of production. However it sold awhile after the introduction of the Pentium II.

Compared to the first generations of this one, the PPro had advantages when used in certain servers:

CPU	Pentium Pro	1. generation Pentium II
Max. RAM	4 GB	512 MB
L2 cache speed	200 MHz	150 MHz
Max. number CPU	4	2

Intel also supplied a Pentium Pro-Overdrive Kit running at 333 MHz. However, with the Intel Xeon CPU the end came to the Pentium Pro.

- Next page
- Previous page

Learn more **[top]**

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.02

The Pentium II

The second 6th generation CPU was Intel's Pentium II from 1997.
The contents:

- Pentium II
- L2 cache out of chip
- The SEC module
- L2 cache speeds compared

- Next page
- Previous page

Pentium II

[top]

Pentium Pro "Klamath" was the code name for Intel's top processor. It ended up as a partially reduced and partially improved Pentium Pro model.

Introduced May 7, 1997, the construction of Pentium II was a little controversial. The features include:

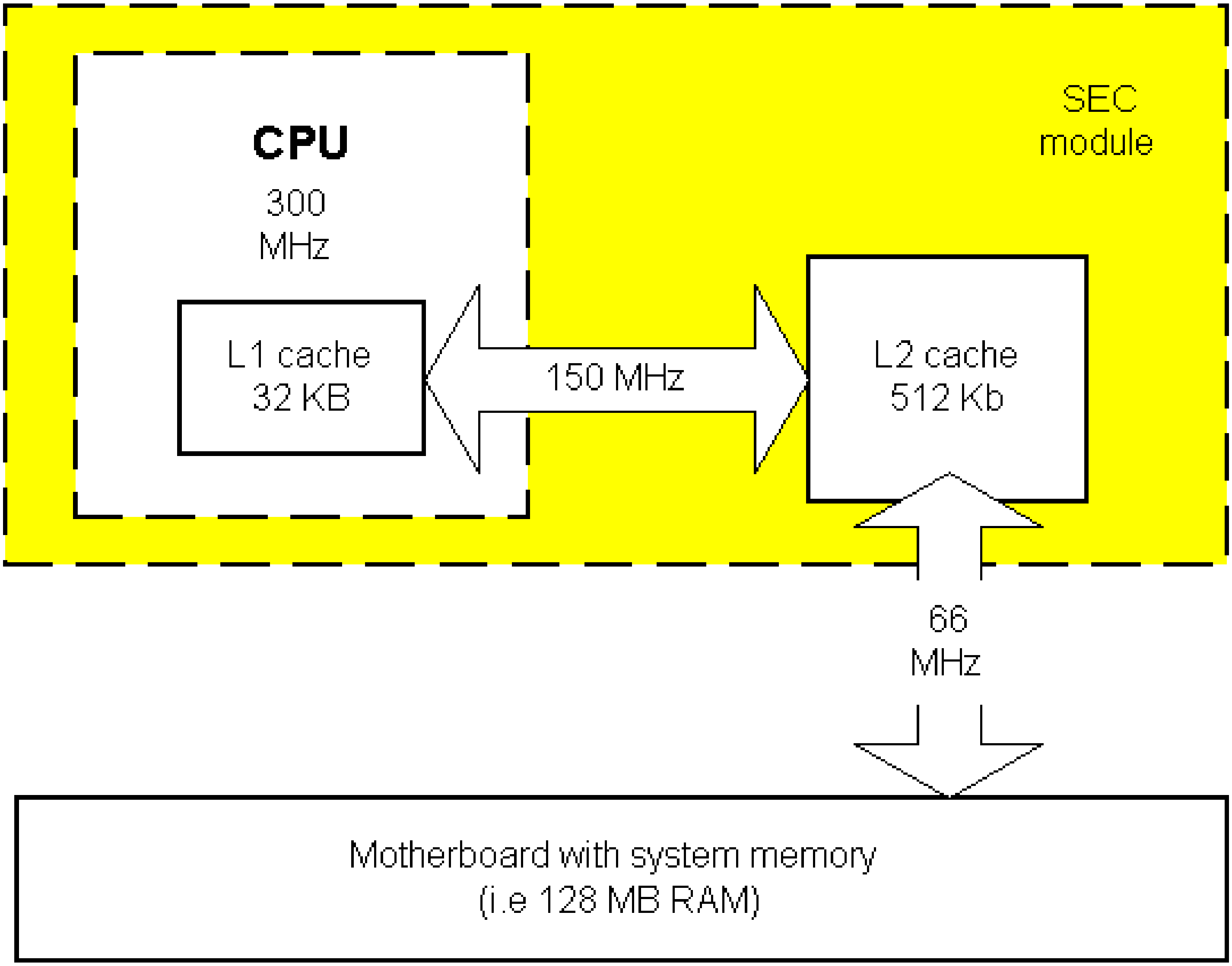
- A CPU mounted together with 512 KB L2 in a SECC (*Single Edge Contact Cartridge)* module
- Connection to the motherboard using the *slot one* connector and the P6 GTL+ bus.

- MMX instructions.
- Improved 16 bit program execution (joy for Windows 3.11 users).
- Doubled and improved L1 cache (16 KB + 16 KB).
- New increased internal speed: from 233 MHz to 300 MHz (later version much higher).
- L2 cache working at half CPU speed.

L2 cache out of chip

[top]

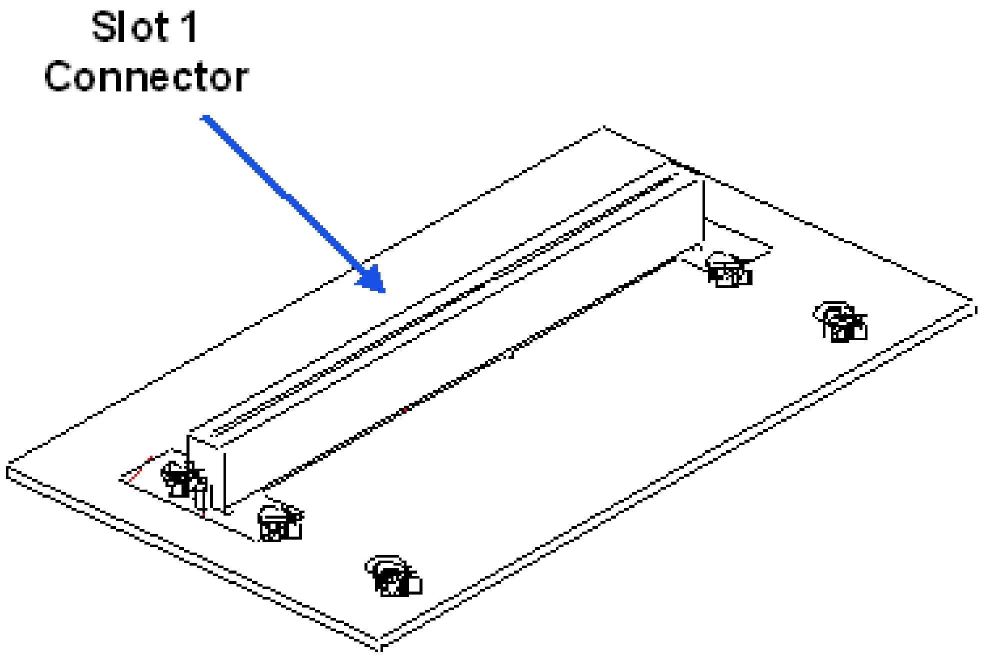
The most interesting change was the separation of CPU and L2 cache. Intel found it too costly to combine them in one chip as in Pentium Pro. To facilitate mass production, cache RAM of a different brand (Toshiba) was used. The cache RAM is marked 7 ns allowing a clock frequency of maximum 150 MHz.



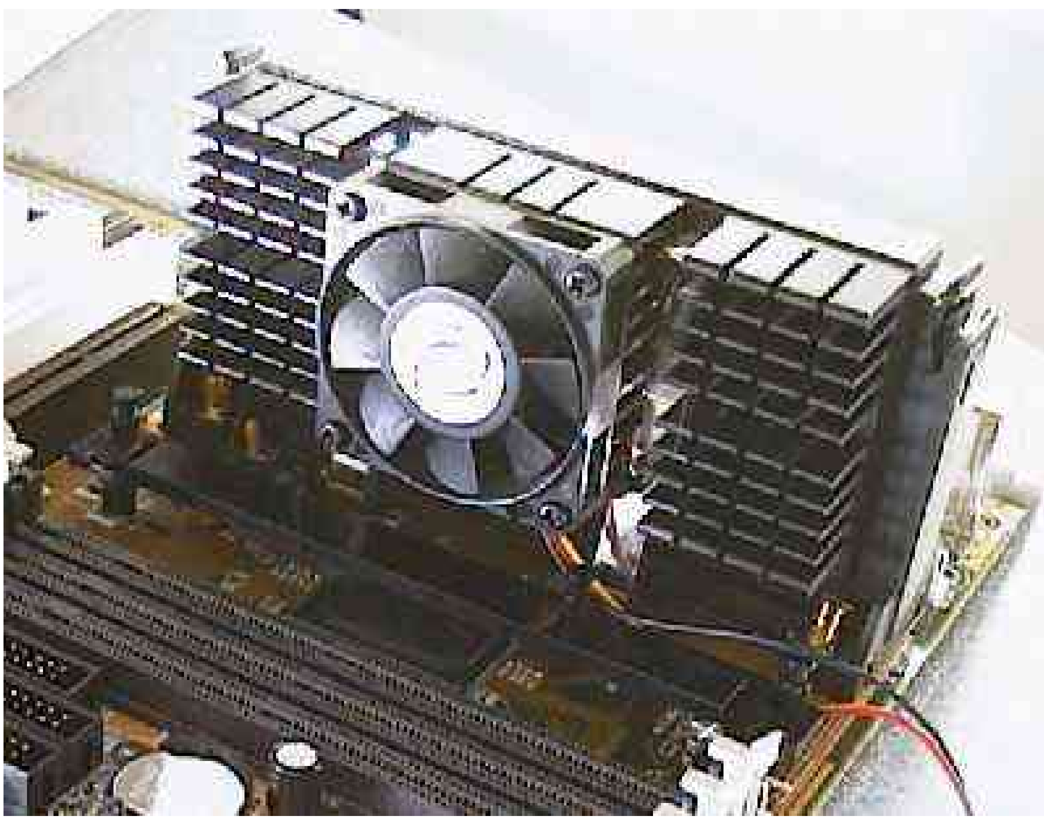
The SEC module

[top]

Pentium II is a large rectangular plastic box, which contains the CPU and cache. There is also a small controller (S82459AB) and a well dimensioned cooling fan. All are mounted on a card. This card with chips weighs about 380 g (13 ounces). It fits in a new 242 pin Single Edge Connector on the motherboard:



Here you see the SEC module mounted in my ASUS board. Note the cooling elements on the cache RAM chips on both sides of the CPU:



L2 cache speeds compared

[top]

With its special design, the L2 cache has its own bus. It runs at half the CPU speed, like 133 MHz or 150 MHz. This is clearly a retrogression from the Pentium Pro, which can run at 200 MHz between the CPU and L2 cache. It is countered by the improved L1 cache, which really zips along! Here you see a comparison:

CPU	L1 transfer rate	L2 clock speed	L2 transfer rate
Pentium 200	777 MB/sec.	66 MHz	67 MB/sec.
Pentium 200 MMX	790 MB/sec.	66 MHz	74 MB/sec.
Pentium Pro 200	957 MB/sec.	200 MHz	316 MB/sec,
Pentium II 266 MHz	1,175 MB/sec.	133 MHz	221 MB/sec.

Pentium II is and has been available in 233, 266, 300, 333, 350, 400, 450, and 500 MHz editions. With the 82440BX and i810 chip sets Pentium II was an excellent performer. Read on for more information on Pentium III.

- Next page
- Previous page

Learn more

[top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

KarbosGuide.com. Module 3e.03

The second generation of Pentium II

The contents:

- The Deschutes
- 100 MHz Front Side Bus
- The Original Celeron
- The Over-clocking
- The Good Celeron (the Mendocino)
- The New Great Socket 370
- Dual Celeron configuration

- Next page
- Previous page

The next Pentium II, the Deschutes

[top]

The third P6 CPU was Intel's Pentium II code named "Deschutes". This new core also lead to the Celerons in various brands.

On January the 26th 1998 Intel introduced the new 333 MHz model of Pentium II.

It was the first of a second generation Pentium IIs known under the code name "Deschutes". The chips are produced with 0.25 micron technology, which reduces the power consumption by more than 50 % compared to the original Pentium II "Klamath" with its 0.35 micron technology. The core voltage is down from 2.8 to 2.0 Volt

100 MHz Front Side Bus

[top]

On April the 15th, 1998 Intel released the next line of Deschutes. The system bus had been increased to 100 MHz. This will internally be multiplied by the clock factors 3.5, 4.0 and (June 1998) 4.5, making the CPU run at 350, 400 and 450 MHz. These CPUs use the new chip set: 82440BX.

So these Deschutes chips use two different motherboards:

- LX-based for the 333 MHz version (5 X 66 MHz)
- BX-based for the 350, 400, 450, and 500 MHz versions
(with clock multipliers of: 3.5, 4.0, 4.5, and 5.0 X 100 MHz).

Fast L2 cache RAM

The L2 cache RAM has to be cooled down and it has to be fast:

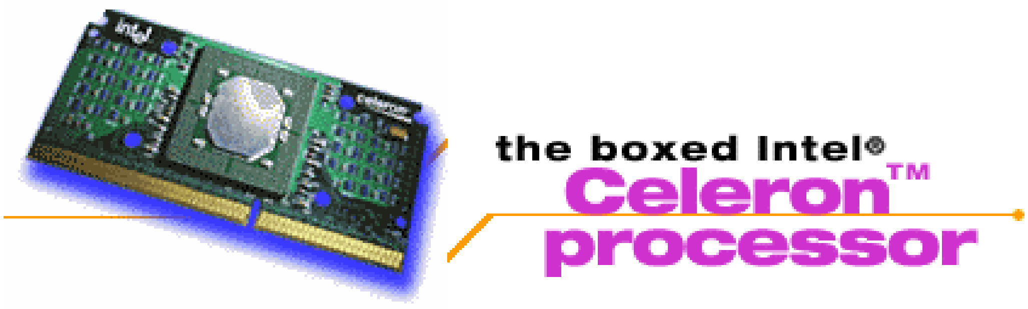
CPU Clock	RAM type	Controller
333 and 350 MHz	5.5 ns	S82459AC
400 MHz	5.0 ns	S82459AD
450 MHz	4.4 ns	S82459AD

The Original Celeron

[top]

Early 1998 Intel was having a hard time with the Pentium II which was pretty expensive. Many users bought the AMD K6-233, which offered very good performance at a moderate price.

So Intel created a brand new CPU called Celeron. It is a Pentium II cartridge except for the L2 cache, which has been chopped away. It uses a 'Covington' core, and we could just as well have called it the Pentium II-SX. In 1998 Intel replaced their Pentium MMX with the first Celerons. Later the design was improved a lot, and Celeron became a very successful product.



This first inexpensive Celeron cartridge fitted into Slot 1 and it ran on a 66 MHz system bus. The internal clock ran at 266 or 300 MHz and delivered good performance for floating point and MMX heavy programs such as certain games. Concerning office applications, the lack of L2 cache was a great disadvantage.

Over-clocking

[top]

The first Celeron were extremely good for over-clocking, since much of the problem here arises from the onboard L2 cache. The L2 cache RAM cannot function at high clock frequencies, but without L2 cache RAM this problem did not occur with the first Celerons.

The Celeron 266 and 300 ran at speeds of 412 MHz and 464 MHz without any problems. However, for non-overclocking purposes the Celeron cartridge could not be recommended. Its lack of L2 cache was too big a disadvantage.

Celeron with L2 cache - the Mendocino

[top]

The next variant of Celeron got the code name Mendocino. First it came in 300 and 333 MHz versions.

The interesting part is that the new cartridge holds 128 KB L2 cache inside the CPU itself. This gives very good performance, since the L2 cache runs at full CPU speed. Here you see a Celeron 300A. A chip on a card:

Integrated L2 cache

The manufacturing price was increased by less than 10%, adding the 128 KB integrated L2 cache, while the performance probably increased 30-40%. The number of transistors were increased from 7.5 million to 19 million due to the L2 cache.

However, on-chip L2 cache is a good technology. In the first 0.25-micron technology, the Mendocino's 128K cache took up about 35 mm2 of die area. It added $10 to the manufacturing costs, but these numbers decreased going into 0.18 micron process technology. And then it is cheaper to produce a big integrated L2 cache than to add the chips to an expensive Slot 1 or 2 module.

These early "Mendocino" cartridges were just as good as the traditional 66 MHz Pentium IIs. The Mendocino-based Celeron cartridge running at 300 MHz was named with an A as suffix to distinguish it from the Celeron 300 without L2 cache.

Also good for over-clocking

Hence, the first two models were the Celeron 300A and 333. They did very well, being priced very low compared to the equivalent Pentium IIs.

In terms of over-clocking they proved successful as well. Here it appears that the 300A was the best. It works fine with a clock doubling of 4.5 X 103 MHz giving 464 MHz. The 333A model "only" runs at 416 MHz (5 X 83 MHz).

Faster with Celeron

On January 4th 1999, Intel introduced a 366 MHz version and a 400 MHz version both working the RAM on a 66 MHz bus. The clock multiplier within the new Celerons goes up to 8.0.

March 15, 1999. The 433 MHz version of the Celeron was launched. A 466 MHz version was released late April.

July 31, 1999. The 500 MHz version of Celeron was launched.

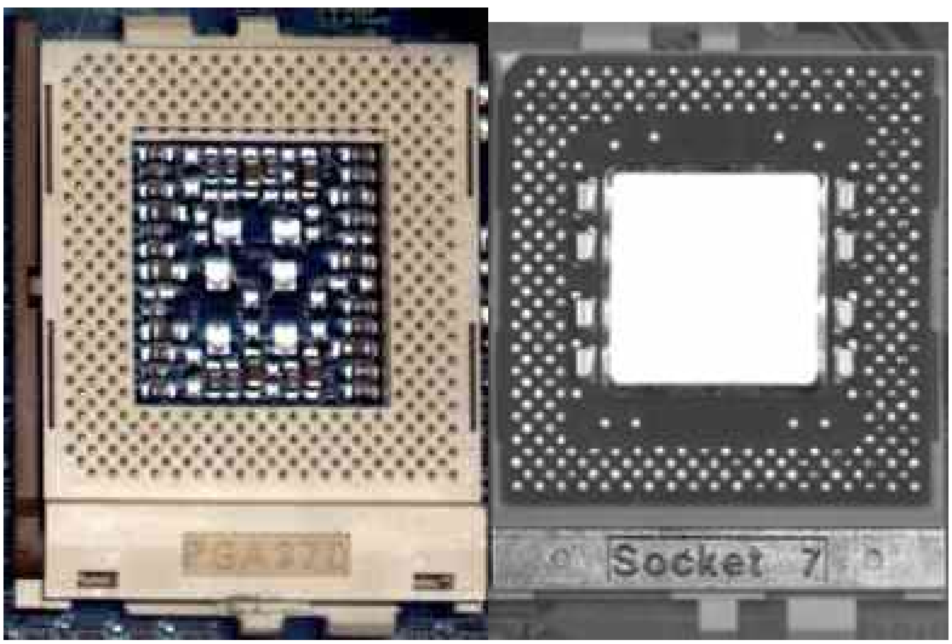
Later 1999 the Celeron came in a 533 MHz version. In 2000 came 566 MHz Celerons produced with 0.18-micron process technology.

New Socket 370 for the Celeron

[top]

The 400 and 366 MHz processors were as all successors available in a plastic pin grid array (P.P.G.A.) form factor.

This PGA370 socket looks quite like a traditional Socket 7. It holds 370 pins:



Both are ZIF (Zero Insertion Force) sockets containing a lever so you can open and close the socket. This makes it very easy to insert the CPU.

However, the PGA uses a different bus protocol (GTL+) than the Socket 7, which also only holds 238 pins. The GTL+ bus is the same protocol as all Pentium II's. Hence, they use the same chip sets.

The socket 370 is cheaper to produce than Slot 1 cartridges, so all Intels mainstream processors will come in this design.

The roadmap for the Celeron looks like this:

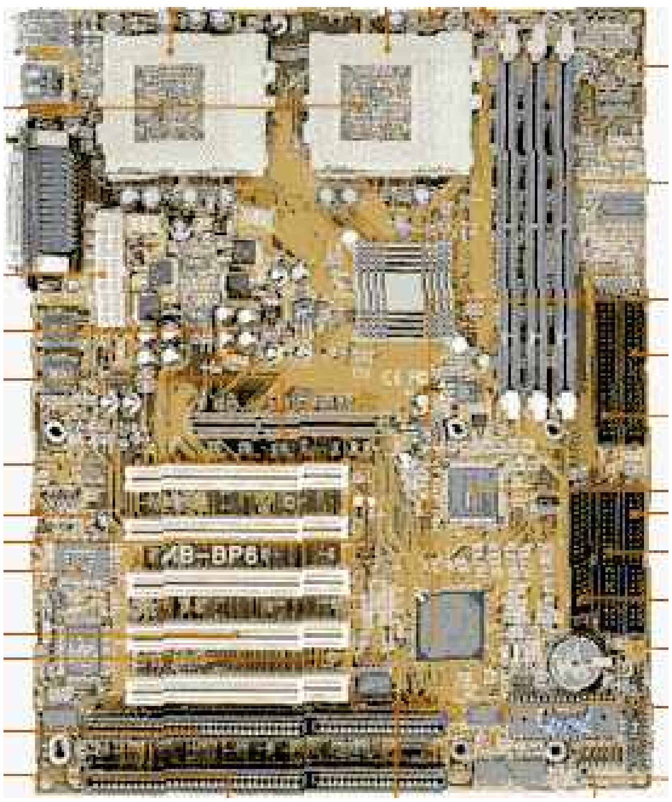
- 66 MHz bus versions up to 800 MHz.
- 100 MHz bus versions introduced in 2001, bringing the chip up to 1200 MHz.

Dual Celeron configuration

Using two Celerons on one motherboard could be a great idea. It would enable people to produce very inexpensive high-powered workstations. Windows NT is capable of using both processors.

During 1998 I heard of several private persons, who made the Celerons work in dual SMP (symmetric multiprocessing) configurations. But in July 1999 two companies produce motherboards for dual Celeron configuration.

Here is a little picture of such a board. You see two socket 370's:



See Abits own homepage on the BP6 board. And see QDI.

- Next page
- Previous page

Learn more **[top]**

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7bon graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.04

About P6-like processors from AMD and Cyrix

The contents:

- AMD's powerfull K6
- AMD's K6-2
- Cyrix 6x86MX
- Pure CISC
- Two brands of 6X86MX
- K6-2

- Next page
- Previous page

K6

[top]

Intel's Pentium II soon got competition from AMD and Cyrix. Both companies have launched several good processors, sometimes giving Intel a hard competiton.

AMD's K6 is from April 2, 1997. In 1996 AMD produced the K5 processor which was not very impressive, however very cheap. The company was put back to business by Mr. Atiq Raza, who brought in the technology from NexGen. This lead to the very successful model K6, which saved AMD from ruin.

The market soon discovered that the K6 performed a lot better than Pentium MMX, which it shared the Socket 7 motherboards with. Here are the data:

- Equipped with 32+32 KB L1 cache and MMX.
- Containing 8.8 million transistors.

K6 is (like K5) compatible with Pentium. Thus, it can be mounted in a Socket 7 on a regular Pentium motherboard, and this soon made the K6 very popular.

BIOS and voltage

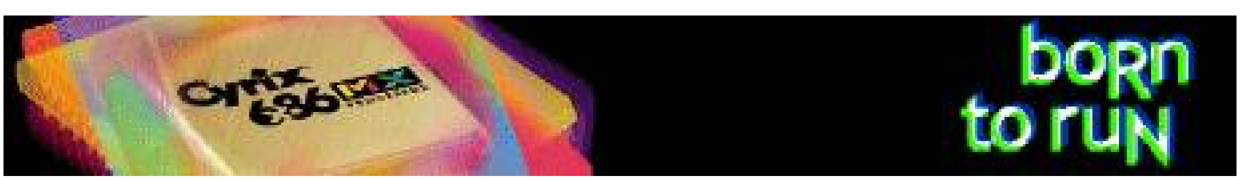
On the older motherboards, it is possible that the BIOS has to be updated to make it work. K6 performs best when the BIOS recognizes the chip, so its full potential can be utilized. That requires a dual voltage motherboards. The K6-200 requires 2.9 volt for its core. The other models require 2.8 volt as the Pentium MMX.

[top]

The K6 model 7 (*Little Foot*) was running at 300 MHz. These high performance K6s were sold at very reasonable prices. The problem seemed in the beginning to be to produce enough chips. The K6 was followed by the K6-2 (and later the Athlon), which gave AMD an enormous success in the late 1990's.

Cyrix 6X86MX (MII)

[top]



Cyrix was a company with another high performance chips, placed somewhere between 5th and 6th generation. The first models were positioned against the Pentium MMX chip from Intel. Later models can be compared to the K6. I have to admit, that I quite seldom saw these processors in my country, but they did exist.

Cyrix powerful P6-classed processor was announced as the "M2". Introduced on May 30, 1997 the name became 6X86MX. Later it has been named MII again. There has always been some confusion about the identification of the Cyrix CPUs.

MMX

This 6X86MX chip is compatible with the Pentium MMX. This gives additional possibilities to assemble PCs on ordinary Socket 7 motherboards.

The 6X86MX has 64 KB internal L1 cache, which is very impressive. Cyrix also utilizes technologies which are not found in Pentium MMX. These chips are named to compare them with genuine Pentiums, although their internal clock speed is lower than corresponding Intel processors.

Pure CISC

[top]

The 6x86MX was unique compared to the other 6. generation CPUs (Pentium II and Pro and K6) since it does not work upon a RISC kernel. 6x86MX executes the original CISC instructions as does the Pentium MMX.

The 6x86MX has plenty of internal registers placing it in company with other 6th generation CPU's:

CPU	Number of 32 bit CPU registers
Pentium MMX	8
6x86MX	32
Pentium Pro	40
K6	48

The 6x86MX had - as all processors from Cyrix - a problem concerning the FPU unit. However, only using standard office applications, this is of no concern. The problem arises when you play 3D games.

Poor performance

The 6X86MX is quite a powerful chip - on the paper. However, there are problems with the supply of them, and also the system bus speed caused troubles. It was difficult to find a motherboard that accepts these speeds. They also lacked good FPU and MMX performance. They did not incorporate the 3DNow! technology. Hopefully this will change as Cyrix has been taken over by VIA.

6X86MX	Internal speed	External speed
PR166	150 MHz	60 MHz
PR200	166 MHz	66 MHz
PR233	188 MHz	75 MHz
PR266	225 MHz	75 MHz
PR300	233 MHz	66 MHz
PR333	255 MHz	83 MHz
PR433	285 MHz	95 MHz
PR466	333 MHz	95 MHz

It was evident that Cyrix intended to continue this line of processors, and this definitely was a positive trend. Intel got competition, and it keept the well tested and inexpensive Socket 7 motherboards in the market. In 2000 the VIA Joshua processor will hold designs originating from Cyrix - ported into socket 370 design.

Two brands of 6x86MX and MII

[top]

The 6x86MX processor was produced by National/Cyrix as well as by IBM. The architecture were the same, but the chips were built at different plants.

On April 14, 1998 the Cyrix MII (*M-two*) version was launched. It was exactly the same chip as the 6X86MX just running at higher clock frequencies. Later the voltage will be reduced to 2.2 Volts.

IBM used a new technology for their PR333 chip. It is patented and called Flip-Chip. The die is soldered directly to the ceramic casing and this causes less induction.

Cyrix MIII

In the year 2000 we were expecting the 3rd generation of the 6X86MX (code name Jalapeno and Mojave). This CPU was to be named MIII, and to come in>600 MHz flavors. The MXi was intended Socket 7 compatible, with a core running on a 133 MHz system bus. It also should include 3DNow! instructions and improved FPU. Read more on this.

AMD K6-2

[top]

The next AMD "model 8" version of the K6 had the code name "Chomper".

This processor of May 28, 1998 was marketed as K6-2, and like the model 7 version of the original K6, it is manufactured with 0.25 micron technology. These chips run on just 2.2 Voltage. They became an immense succes, in many situations competing very successfully with Intel's Pentium II.

Super 7 motherboards and better MMX

The K6-2 is made for a front side bus (system bus) at the speed of 100 MHz. This is to be found with the so-called Super 7 motherboards. AMD made other vendors like VIA produce new chipsets for the traditional socket 7 motherboards, after Intel in 1997 had given up the platform.

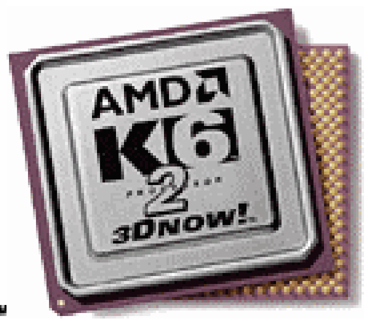
K6-2 is also improved with an MMX performance twofold better compared to the original K6.

3DNow!

The K6-2 holds a new 3D plug-in (called 3DNow!) for better game performance. It consists of 21 new instructions that can be used by software developers giving a better 3D-performance. To benefit from it, you need a graphics driver or a game, which deals directly with the new commands.

The good thing is, that games do not have to include special programming to benefit from 3DNow!. Support is included in DirectX 6.0 (and newer) for Windows . DirectX is a so-called *multimedia API* (in fact a *hardware abstraction layer*) for Windows . It is some programs that can enhance the multimedia performance within all Windows programs.

3DNow! is not compatible with MMX, but the K6-2 holds MMX as well as the 3DNow!. Also Cyrix and IDT launch CPUs with 3DNow!. Read more on 3DNow!



Good and inexpensive power

The K6-2 gave very, very good performance. You can compare the models to the Pentium IIs. A K6-2 350 MHz performed very similar to a Pentium II-350, but was sold a lot cheaper. And you even saved more because of the cheaper motherboard.

100 MHz bus

Not all K6-2s ran with a 100 MHz bus. Here you see some of the versions, which require motherboards with crystals capable of these configurations:

K6-2	Bus	Clock
266 MHz	66 MHz	4.0 X 66 MHz
266 MHz	88 MHz	3.0 X 88 MHz
300 MHz	100 MHz	3.0 X 100 MHz

333 MHz	95 MHz	3.5 X 95 MHz
350 MHz	100 MHz	3.5 X 100 MHz
380 MHz	95 MHz	4.0 X 95 MHz
400 MHz	100 MHz	4.0 X 100 MHz

Two of the CPU's in the table must be the same. AMD calls it a 350 MHz version, but in Denmark e.g it was sold as a 380 MHz version.

K6-2/400 and above

November 15, 1998. The K6-2/400 was introduced. This chip worked on a new core, which should be slightly improved. Hence the performance matched a Pentium II-400.

April 6, 1999. A 475 MHz version of the K6-2 was introduced. The latest version is 533 MHz.

AMD had 39% of the market with K6-2 in 1999!

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.05

The AMD K6-3

The contents:

- K6-3
- K6-2+

- Next page
- Previous page

K6-3

[top]

AMD announced the very powerful K6-3 for quite a long time. It was delayed quite a while, should have been released February 22nd 1999. During the delay Intel had time establish its Socket 370 version of Celeron. But finally it arrived in the summer of 1999.

Using the next version of this chip - model 9 code name "Sharptooth" - you may have three levels of cache!

In the K6-3 you find:

- A slightly improved K6-2 unit.
- An in-chip L2 cache of 256 KB.
- TriLevel Cache design.
- New 133 MHz Front side bus.
- Clock speeds of 400 MHz and 450 MHz.

TriLevel Cache design

Both the 64 KB L1 cache and the 256 KB L2 cache are integrated with the chip. This L2 on-die cache works at full processor speed just like it did in the Pentium Pro, and as it does in the Celeron A and in the Xeon processors from Intel. This will definitely speed up the K6 quite a bit!

Since the K6-3 is to be used in a Super 7 motherboard there is room for another level of cache, the L3 cache. The TriLevel Cache design is constructed to use the existing motherboards with up to 2 MB of cache on-board. This used-to-be L2 cache (on the motherboard) is used as the third level of cache. This happens automatically, and the bigger cache seems to increase performance a lot!

High performance

Tests show performance from the K6-3 450 MHz comparable to the Pentium III 500 MHz processor. This is coming from a Socket 7 motherboard! My prediction is that the K6-3 will be an excellent CPU at a very good price. The problem was to find it, the K6-3 never became very popular.

Still weaknesses in the FPU

Traditionally only Intel can produce a powerful FPU (*Floating Point Unit*). Test confirms that the K6-3 has the same FPU as the K6-2 does. At same clock frequencies it performs 40% under the Pentium III. However the 3DNow! technology, which is supported by the Microsoft DirectX software layer, makes up for the weak traditional FP-performance. In test it ends up 15% under the equaling Pentium III.

The die size

The K6-3 obviously is bigger than the K6-2, due to the integrated L2 cache. However it is a lot smaller than the Athlons :

Chip	Die size
K6-2	81 mm2
K6-3	118 mm2

K7 Athlon 0.25 micron	184 mm2
K7 Athlon 0.18 micron	100 mm2
Intel Pentium III Cumine	106 mm2

K6-2+

In 2000 the K6-3 is expected to disappear from the market. It never became a success. However, the K6-2 has been a great seller, so AMD will launche new versions of the K6-3 as *K6-2+*.

This new chip will have following features:

- 0.18 micron process technology
- Integrated 256 KB L2 cache
- Additional 3DNow! instructions (from Athlon)
- 550 MHz

The maximum speed from this design is expected to be 750 MHz. Hence one can expect the Socket7 platform soon to die out.

- Next page
- Previous page

Learn more **[top]**

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.06

Intel Xeon for servers

The contents:

- Xeon
- A product for servers

- Next page
- Previous page

Pentium II Xeon

[top]

Intel always had an important market supplying CPU's for servers. The original Pentium Pro was for several years used for this purpose. Xeon is a line of CPUs for use in servers.

On July 26th, 1998 Intel introduced the Pentium II cartridge named Xeon. Aimed at servers and perhaps high-end users.

The Xeon is a Pentium II in a new cartridge fitting into a new connector called Slot Two. The module is twice as tall as the current Pentium II, but there are other important innovations and improvements:

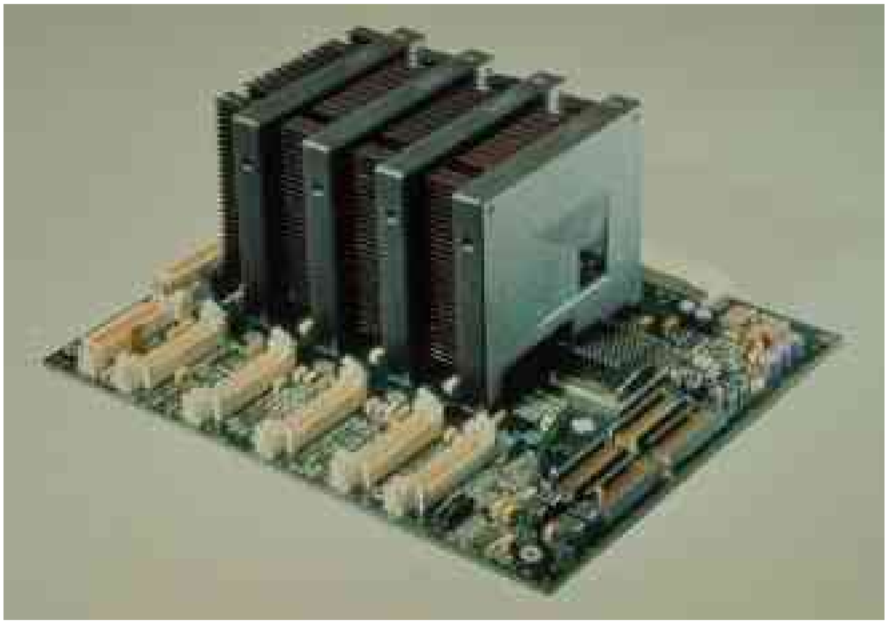
- New type L2 cache RAM chips: CSRAM (*Custom SRAM)*, which runs at full CPU speed.
- Different L2 cache sizes: 512, 1,024, or 2,048 KB L2 RAM.
- Up to 8 GB RAM can be cached.
- Up to four or even eight Xeons in one server.
- Support for clustered servers.
- New chip sets 82440GX and 82450NX.

The new (huge) cartridge fits into a new Slot Two with three layers of edge connectors. The large L2 caches running high speed will use a lot of power, so cooling will be very important. The cartridge is about twice the size of the well known Pentium II.

A server product

[top]

The Xeon chip is for high performance servers. The first top model will hold 2 MB L2 cache on the cartridge, running at full 450 MHz. This chip costs $4,500!



Performance gain from L2 cache at full speed

The L2 cache of the Xeon runs at full CPU clock speed. One could think, that this would have the same performance as the L1 cache. However the interface from L1 to L2 costs some clock ticks in the beginning of each transmission, so there is some latency. But when data is transferred, it runs at full clock speed.

Practical tests only show an increase in the performance of 5-8% comparing Pentium II and Xeon/512 KB, both running at 450 MHz.

Personally, I find the Xeons too expensive. I know companies who have been advised to and bought the modules with 2 MB cache for use in web-servers. Obviously the price does not matter in those cases, and Intel makes a good profit from that. I do not think the performance matches the price.

Tanner

In 1999 the code name “Tanner” chip became known as the Pentium III Xeon.

Later might follow the processor code named "Foster" which should integrate 2 MB of L2 cache in-chip.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.07a

The Pentium III - Katmai and SSE

The contents:

- Katmai New Instructions (KNI)
- Two new features
- The ID number – panic, panic
- SSE
- New registers
- Program support with DirectX 6.1
- My conclusion

- Next page
- Previous page

The first P6 CPU from Intel was the Pentium Pro. Later we got the Pentium II in various flavours (including the popular Celeron). In 1999 the time came for the Pentium III.

Katmai New Instructions (KNI)

[top]

In March 1999 Intel introduced the new enhanced MMX2 set of graphics instructions (70 of them). These are called Katmai New Instructions (KNI) or SSE. They are intended to speed up 3D gaming performance - just like AMD's 3DNow! technology. Katmai includes "double precision floating point single instruction multiple data" (or DPFS SIMD for short) running in eight 128 bit registers.

Katmai New Instructions (KNI) was introduced with the 450 and 500 MHz Pentium III. It was processors very similar to the old Pentium IIs, using Slot 1.

The only new feature was the implementation of Katmai and SSE, which I shall try to describe in this page.

Two new features

[top]

In fact the Pentium III contained two rather different news items, one small and one somewhat bigger. Intel's new top processor is a Pentium II in principle. It is mounted in a BX based motherboard with Slot 1. This processor has some built-in features:

- A rather problematic ID numbering.
- New registers and 70 new instructions.

Finally the clock speed was raised to 500 MHz with room for further increases. Pentium III Xeon (code name Tanner) was introduced March 17th, 1999. It was a Xeon chip updated with all the new features from Pentium III. To utilize it Intel has the 840 chipset.

The ID number – panic, panic

[top]

The ID number PSN (Processor Serial Number), which is unique for each CPU, has caused lots of security discussions. It is a 96 bit value that is electronically programmed into each chip. Actually it was meant as a very sensible initiative, which could make electronic trade and encrypting on the Internet more secure and effective.

The advocates see the ID number as "the permanent cookie."However it turns out that hackers and crackers can easily access the number, which makes the security dubious. At the same time the system is proposing a global registry of Internet users, which can easily be misused for marketing, etc. That will be invasion of privacy, if you no longer can remain anonymous on the Internet.

SSE etc.

[top]

The somewhat bigger news item is a real change in the basic processor architecture. It is actually the first change since 1985, where the x86 architecture was expanded from 16 to 32 bit (compared with the 386 processor).

- New 128 bit registers, which each can handle four floating-point numbers.

- Streaming SIMD Extensions (SSE). 50 new instructions, which enable simultaneous, advanced calculations of more floating-point numbers with a single instruction.

- New Media Instructions. 12 new instructions in this category, which include other instructions for floating-point decimal calculations besides instructions that are designed for coding and decoding of MPEG-2 video streams "on the fly."

- Streaming Memory. Here are 8 new instructions, which improve the interaction between L2-cache and RAM. With optimum utilization of the instructions, it could result in a 20% improvement of the bandwidth on the system bus. These instructions require newly-compiled programs, so it may be some time before we see the effect of this.

The combined new instructions are called KNI (*Katmai New Instructions*) or SSE. The Pentium III construction is an attempt at improving the FPU performance in the processor. The registers are used by the new Katmai instructions.

New registers

[top]

The new 128 bit registers can potentially speed up 3D-graphics and multimedia handling, since the registers can contain four of the important 32 bit floating-point decimal numbers. Since the registers are a new physical creation within the CPU, they require support in the operating system to utilize them. Such support is expected soon in Windows 98.

Program support with DirectX

[top]

Microsoft's DirectX program layer is optimized relative to Katmai. With that, a large number of existing programs should benefit from the additional power of Pentium III (when DirectX 6.1 or better is installed). However the drivers for sound and graphics cards need to be re-written, to enable utilization of the new DirectX edition. Games that are not based on DirectX need to be re-written or expanded with patches, which utilize Katmai. Finally, a patch has been announced for the operating system itself, Windows 98, which should support the new instructions. Whether it will support all or only part of them is yet unknown.

Problems

In April 1999 came reports of *heat problems* with the 550 MHz version. It should be very very hot, so big fans are the issue here...

July 31, 1999. The 600 MHz version was launched. Later the "Coppermine" version was introduced with Socket370, as you will see in the following page.

A provisional evaluation

[top]

It is difficult to evaluate the significance of Pentium III's new registers and instructions. However, it seems that most programs, drivers, etc. are being tuned for the new instructions. Then there is no doubt that the multimedia capabilities have received a great boost. There could be a *doubling* of their performance. So far Adobe have included support for SSE in version 5.5 of Photoshop. This works very well, some very time-consuming processes has been shortened with approx. 40%.

The question is, if AMD will be forced to work in SSE etc. in their top processors or if they will continue the development of 3DNow!. In that case we'll have two systems to work with. Perhaps this is not a problem.

Please read about Cumine and other new chips following Pentium III in the next module.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

KarbosGuide.com. Module 3e.07b

Pentium III CuMine

The contents:

- CuMine
- Many models
- But no copper

- Next page
- Previous page

Coppermine - CuMine..

[top]

Here we shall look at the further development of Pentium III.

600 MHz

31, July 1999 the Pentium III was released in a 600 MHz version. This chip is working on a 100 MHz bus.

27, September 1999, Intel launched two new chips, the 533 and the 600 MHz versions of Pentium III. These chips are both running on a 133 MHz system bus. Unfortunately the new chip set, i820, which was to be launched the same day, was pulled back in the last minute.

The interesting point is that the 820 set with Pentium III Coppermine is supposed to be Intel's answer to the the very successful AMD Athlon.

Some good news: Adobe has updated the graphics program Photoshop 5.5 with support for the SSE set of instructions. It should be very succesfull.

Coppermine

25, October 1999 the next generation of Pentium III processors was released. The new thing here is the process technology and the integrated L2 cache.

The headlines:

- 0.18 process technology with 28 million transistors
- 6 layer aluminium production
- Reduced die size and 1,65 core voltage
- Integrated L2 cache of 256 KB
- New L2 to CPU bus of 256 bits width

The electronic "wires" insides the chip has been reduced from a width of 0.25 micron to0.18, which is 1/500 of a human hairs width... The impact of 0.18 process technology is that the required voltage can be lowered from 2.2 Volt to 1.6 Volt. Hence, the Coppermine chip is developing less heat at the same clock frequency, and it can be produced for higher speeds. The launched topmodel was running at 733 MHz.

Inside the CPU, the architecture has not changed a lot. The die size has decreased, and this way there has become room for an integrated 256 KB of L2 cache. This cache now works at full CPU speed and at a 256 wide bus. This gives a solid increase in performance.

The 28 millions of transistors are loaded into 106 squaremillimeters, which is quite small; the old Pentium III without integrated L1 cache and only 9.3 million transistors took up 128 squaremillimeters.

Many models

The new chips produced with 0.18 micron process technology are labeled with an "E" to distinguish them from older models.

But since the new process technology is also used for Pentium III's running at traditional 100 MHz, the models with 133 MHz are labeled with a "B". This way we will (for a while) have four flavours of 600 MHz Pentium IIIs:

Model	Process	Clock frequency
600	0.25	6 x 100 MHz
600E	0.18	6 x 100 MHz
600B	0.25	4.5 x 133 MHz
600EB	0.18	4.5 x 133 MHz

Obviously, Intel plans to skip the chips produced in 0.25 micron. But meanwhile both process technologies will be sold side by side.

Late we shall have Celerons at 800 MHz and more based on the new CuMine kernel.

But no copper

It was expected this new generation of chips to produced using copper. The name "CuMine" also indicates this. But the first of these Cumines are produced with traditional aluminium wiring in 0.18 microns width. First in 2001 Intel plans to start using copper (in the P860 kernel), and this should lead to much higher levels of speed.

AMD launched GigaHerz version of the Athlon using copper in 2000. Intel also launched GigaHertz versions of Pentium III, but they were only sold in few numbers in 2000.

SpeedStep

Intel launched a new series of chips for notebooks. The first models are 500 MHz versions of Pentium III running on a 100 MHz bus. Producing them in 0.18 techology, Intel has been able to work with the power consumption.

The new chips can work in a "light" mode when the notebook is on batteries. The core voltage is reduced from 1.6 to 1.1 Volt, and the power usage goes down to just 50%! The CPU performance only decreases with 20%.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.07c

The further development of Pentium III

The contents:

- Pentium III
- The Pentium 4

- Next page
- Previous page

CuMine in year 2000

[top]

Here we shall look at the further development of Pentium III.

Athlon was not not good for Intel

In the first months of 2000 it was obvious that Intel had a hard time with Pentium III. From unknown reasons, they were not capable of producing the CPUs the market wanted. In Denmark we could get all the Athlons needed in February, but only very few Pentium IIIs.

At the same time AMD showed up with faster and cheaper versions of the Athlon - all putting a heavy pressure on Intel.

Intel on their side launched a lot of new models. On Jan 11th, 2000 the 800 MHz version was launched running 6 x 133 MHz. In February the company showed a Pentium III running at 1 Ghz without special cooling.

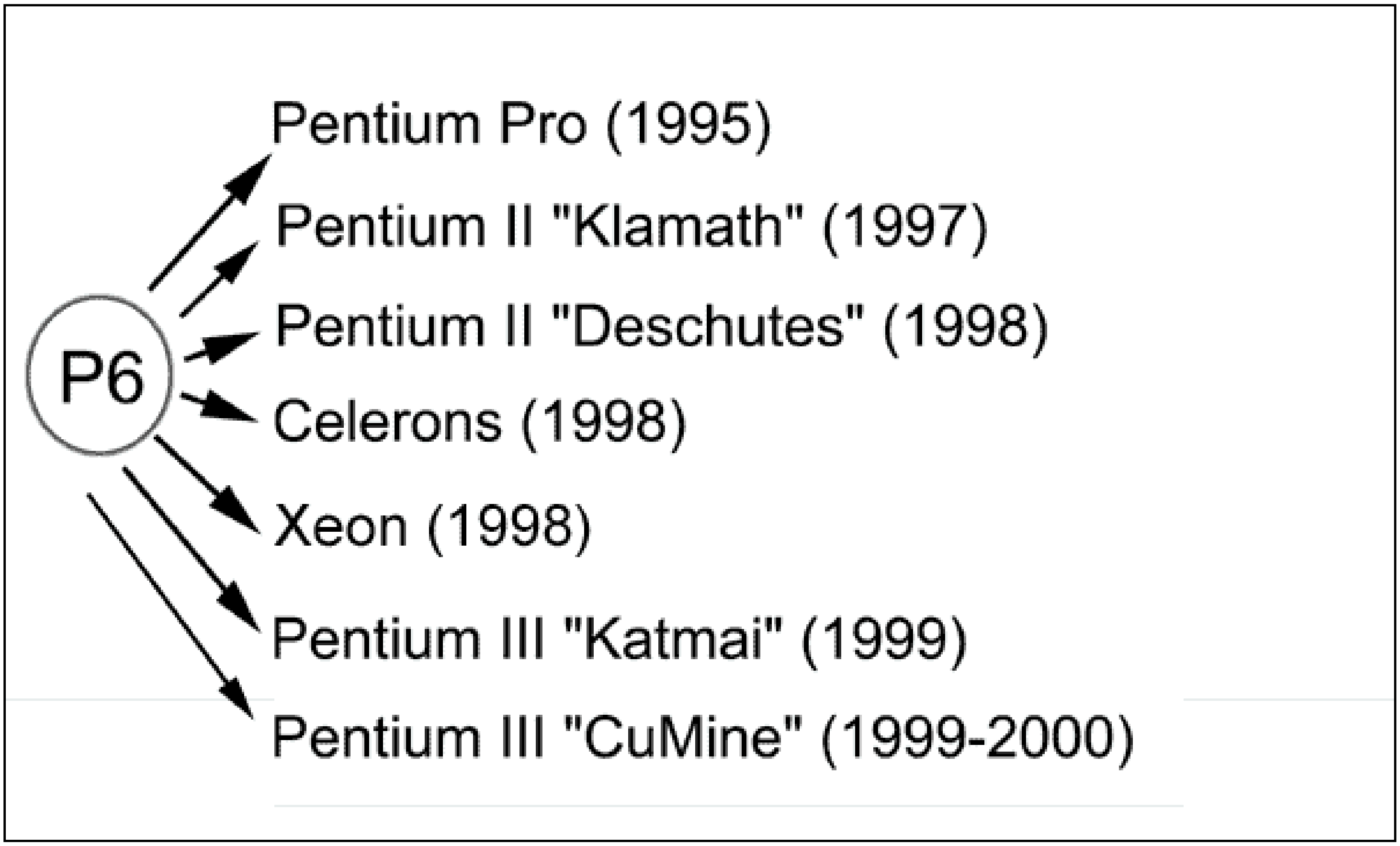
The 850 and 866 MHz were scheduled for February 27/28, 2000. The 933 MHz model came May 27,

2000. This way Intel tried to prove, that they are keeping up with AMD. However in the real world, they were not, being unable to supply the market with processors. In the same period thousands of the competing 650 and 700 MHz AMD Athlons were sold every week.

In the summer 2000, a 1,113 MHz version of Pentium III was taken out of the market due to unstability, and it appears that 1000 MHz is going to be the topmodel of Pentium III. Pentium 4 is heading for 2 GHz in 2001.

The last Pentium III

The interesting issue is, which version of Pentium III to be the last one. If we look at the scheme, we see that Pentium III is bulding on pretty old technology, namely the P6 core:



It was expected, that the P6 line of processors would end with CuMine. However, Intel seems to have decided to continue the line one more year:

The "Tualatin" core

In August 2001 Intel launched the 0.13 micron version of Pentium III, which has been known as codename "Tualatin" (Intel's codenames is from rivers in the pacific Northwest of the USA). This processor uses copper interconnects.

The new processor is a Pentium III "Cumine" with 256 or 512 KB L2 cache integrated. Hardware data prefecth is a new feature, which gives an 8-10% increase in performance.

It runs on 1.475 Volt and is mounted in a Socket370. The Tualatin comes in 1.2 and 1.13 GHz versions, both using a 133 MHz Front Side Bus.

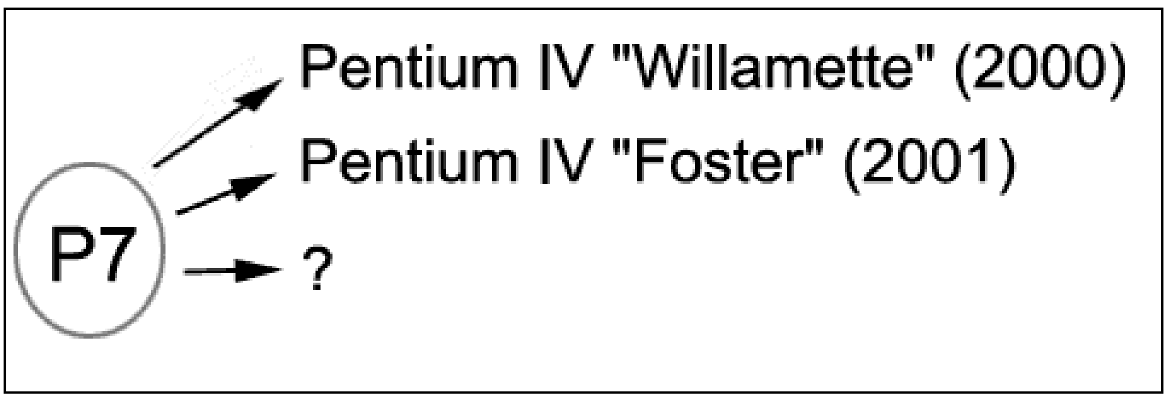
The Intel i815EEA chipset is designed to this processor as well as VIA Apollo Pro 266T. The latest gives support for DDR RAM.

Marketing ...

This the last Pentium III processor has not been advertised very much. Probably Intel does not want it to compete with the 1.4 GHz Pentium 4. Many people was looking forward to the 1.2 GHZ model with 512 KB cache. However this one is only sold for use in servers and in mobile PCs - not for desktop use!

The Pentium 4

The roadmap Intel made several years ago was to abandon the P6 core in favour of a completely new core. The processor codenamed "Willamette" should be the first of a new line of IA32 processors, which should be marketed side-by-side to the IA64 Itanium ("Merced"):



Read about the Pentium 4 in a following module.

- Next page
- Previous page

Learn more [top]

The Pentium 4

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.08a

AMD "The Great" Athlon.

The contents:

- An introduction to K7 Athlon
- The background
- No system bus

- Next page
- Previous page

An overview of Athlon

[top]

The Athlon is a powerful CPU, the first 7th generation CPU in my opinion. It was expected June 1999 but was delayed until August 1999. Intel's response (code name Willamette) was scheduled October 2000.

Athlon was designed using technologies from DEC Alpha 21064 and 2162 RISC processors. Their "farther" Dirk Meyer came to AMD and brought in an engineering team who succesfully developed the Athlon, which ended up being an enormous success to AMD.

Within the first months, the markets response to the Athlon was very positive. It seemed (as expected) to outperform the Pentium III at same clock frequency.

Let us look at what Athlon has to offer:

Raw data

- Mounting in a Pentium II like module, which is entirely AMD’s own design. The socket is called slot A.
- A clock speed of 500 MHz in the first versions.
- Up to 8 MB L2-cache (minimum 512 Kb, without extra TAG-RAM).
- 128 KB L1-cache.

- 22 million transistors (the original Pentium III had 9.3 million).

New bus type

- A brand new system bus type, which in the first versions will work at 200 MHz. An increase to 400 MHz is expected later. The bus is ready for new fast RAM types.

- Independent backside bus, which connects the L2 cache. Here the clock speed can be 1/4, 1/3, 2/5 or identical with the internal CPU frequency. That is the same system which is used in the P6 systems where the L2 speed is either half (Pentium II and III Katmai) or full CPU frequency (at Celeron, Xeons and Pentium III CuMine).

Heavy decoding and FPU

- Three instruction decoders, which translate the X86 program's CISC instructions to the effective RISC instructions, ROP's, where up to 9 can be executed simultaneously. The first test show a decoding of 2.8 CISC-instruction per clock cycle. This is roughly 30% better than Pentium II and III.

- Can handle and rearrange up to 72 instructions (ROP out of order) simultaneously (Pentium III can do 40, K6-2 only 24).

- Enormous FPU performance with three simultaneous instructions and one GFLOP at 500 MHz (1 billion floating-point number operations per second) with 80 bit floating-point numbers. Two GFLOP with MMX and 3DNow! instructions. That at least equals Pentium III's performance with full utilization of Katmai. The 3DNow! engine has even been improved comparing to the K6-3.

The first tests show this FPU performance:

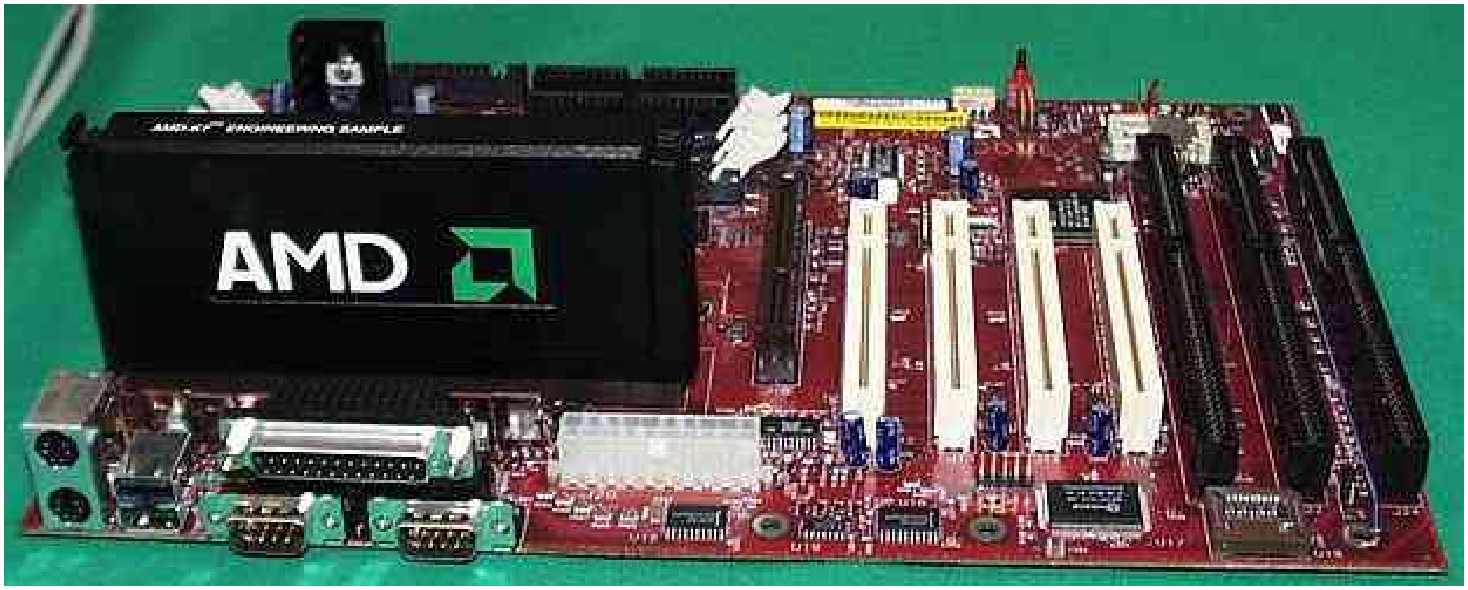
Processor	FPU Winmark
Intel Pentium III/500 MHz	2562
AMD Athlon /500 MHz	2767

The background

[top]

AMD has no license to use the *Slot 1* architecture, so the controlling logic comes from Digital Equipment Corp. It is called EV6 and was designed for the 21264 Alpha CPU. AMD developed the first chip sets (750) themselves, but the architecture is royalty free to use. It will be AMD's first processor using a motherboard and chip set specially designed by themselves. VIA has developed a series of chip sets for the Athlon.

The use of the EV6 bus gives a lot more bandwidth than the Intel GTL+. This means that the Athlon has the capacity to work with new RAM types such as RDRAM. Also the use of 128 KB L1 cache is pretty heavy. The L1 cache is important when the clock speed increases and 128 KB is twice the size in Pentium II's.



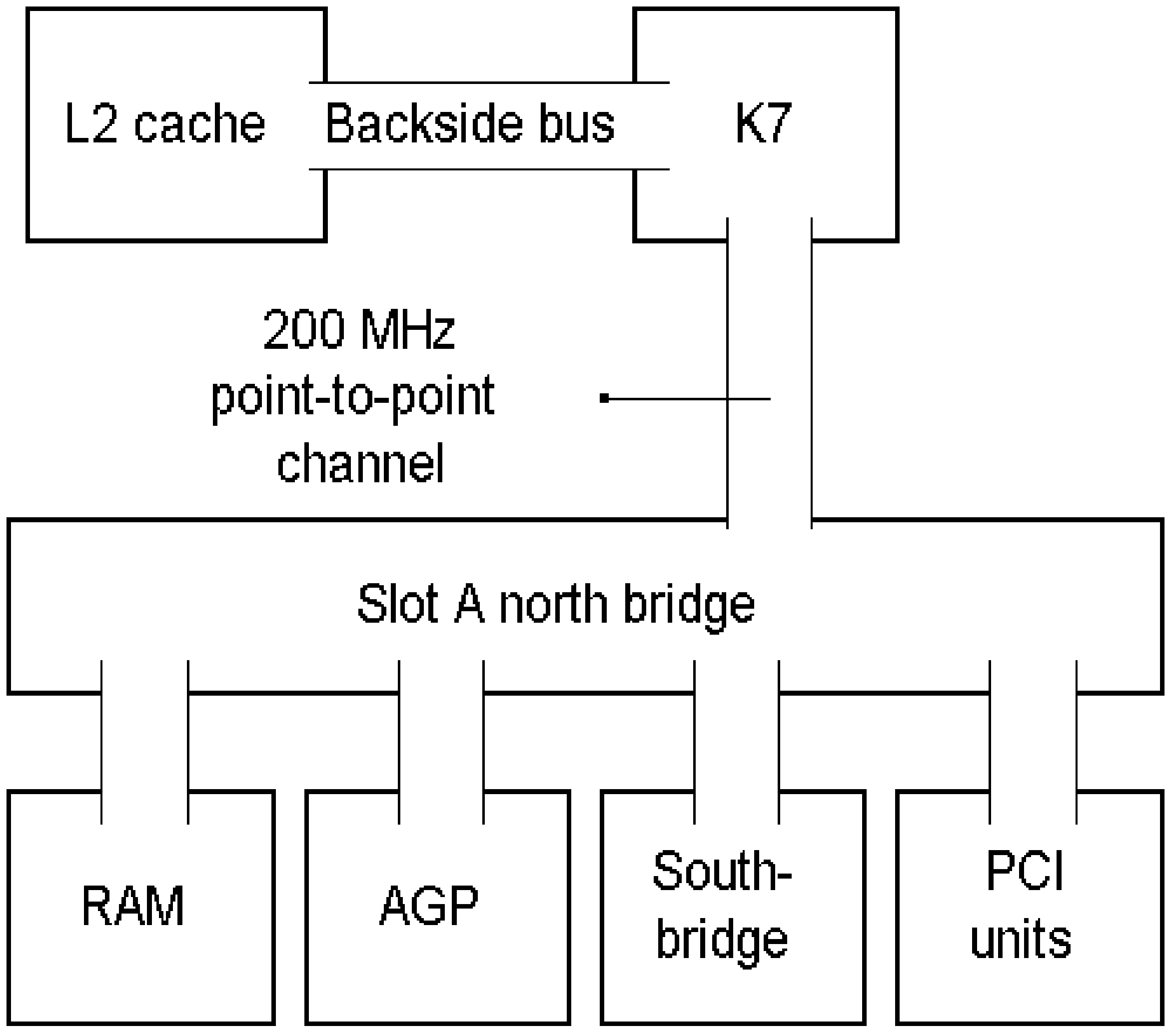
The Athlon came in several versions. The "slowest" ones will have the L2 cache running at one third of the CPU speed, where the best ones (like "Thunderbird") work at full CPU speed (as the Xeons do). The Athlon was intended to give Intel competition in all segments including servers, where the topmodels are being compareable with the best Intel Xeon processors.

No system bus

[top]

Since the Athlon is not installed in the same way as Pentium II and III, AMD could develop a brand new architecture. This means that there really is no system bus. The Athlon module is connected directly to chipset's "North bridge" in the first edition through a 200 MHz data channel. In a multi-processor system, each CPU will have its own 200 MHz channel.

That channel connects only two units: the CPU and the chipset. In the P6 systems the CPU, L2-cache, RAM, PCI units, the AGP unit and the chipset are all connected to the system bus. In Athlon the traffic is split. North bridge comes first, then come RAM, AGP, the PCI units and the South bridge.



Better bandwidth

By eliminating the system bus and replacing it with the new system, Athlon achieves access to a much bigger bandwidth. In theory the bandwidth in a 200 MHz connection is:

200,000,000 x 64 bit/ bit/second = 1.6 GB per second.

That was significantly better than Intel's present systems:

System	Maximum total bandwidth
Intel 100 MHz	800 MB/sec.
Intel 133 MHz	1064 MB/sec.
AMD Athlon, 200 MHz	1600 MB/sec.
AMD Athlon, 400 MHz	3200 MB/sec.

The new architecture opens up for new RAM interfaces. We will see support for 100 and 133 MHz SDRAM, for DDR SDRAM and for RDRAM.

It is also likely that we later can choose from 64, 128 or 256 bit wide RAM access.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.08b

AMD K7 "The Great" Athlon.

The contents:

- The Athlon Tbird
- The Sledgehammer

- Next page
- Previous page

The Thunderbird Athlons

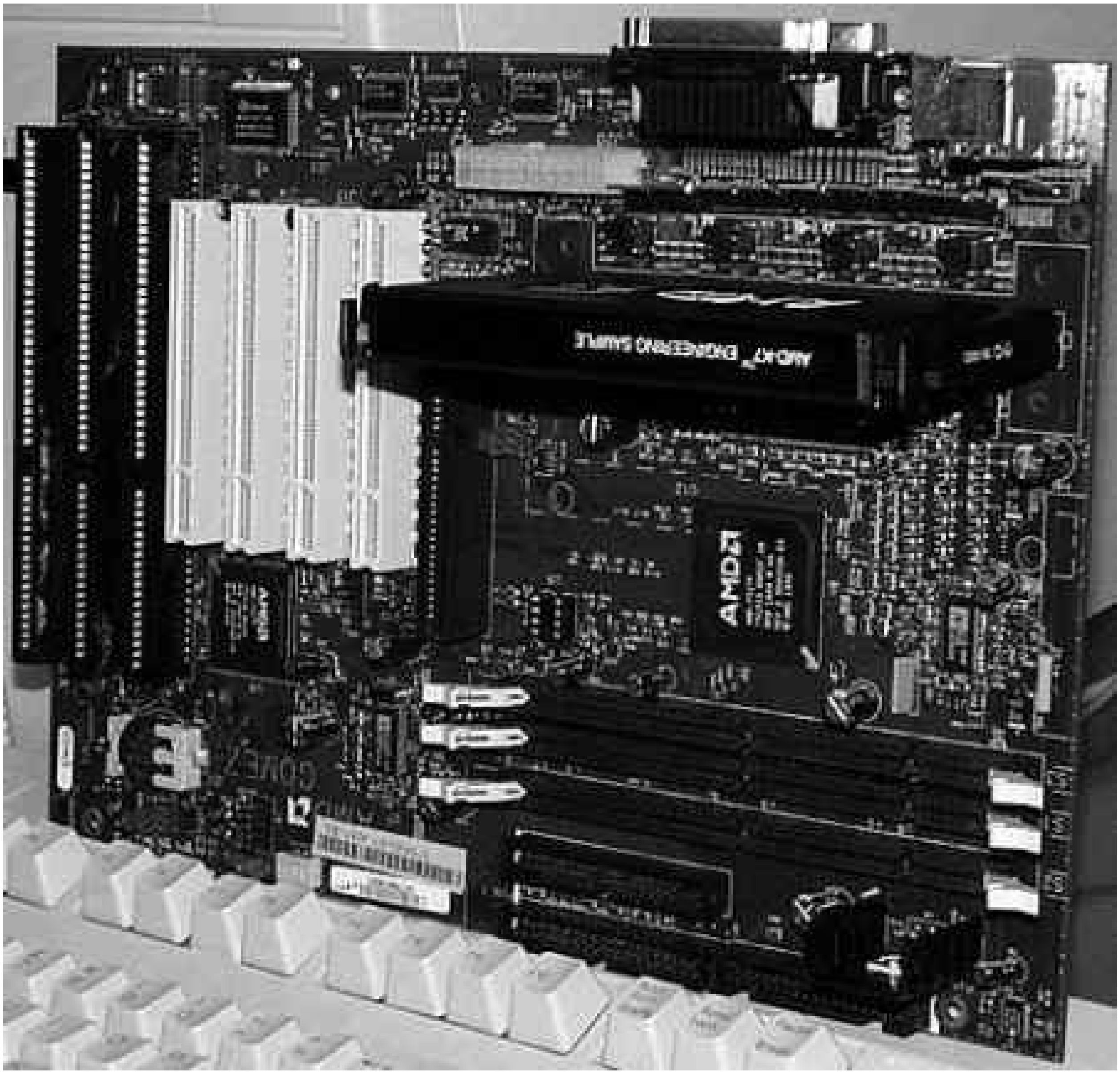
[top]

The Athlon processor has a very large potential. Back in 1999 some analytics believed that Athlon will be the most important and dominant processor in 2001. And they were almost right. All major vendors - except Dell - are using Athlons in high-end PCs.

The Athlon has an enormous bandwidth that obviously not will be needed in the immediate future.

The point is that the architecture is looking ahead. There is lots of room for technological advances in the coming years, such as significantly faster RAM, hard disk etc.

At the same time AMD signals that this is the future server architecture, since especially larger network servers will have a need for the large bandwidth.



The Thunderbird core

AMD has a new fabrication unit (*fab 30*) in Dresden, Germany. From this the new "Thunderbird" core was shipping June 2000.

The first Thunderbirds:

- 750 MHz to 1 GHz versions using 0.18 micron copper technology.
- 256 KB L2 cache integrated.
- New 462 pin socket A.

- Both copper and aluminium chips available.

The Thunderbird is a very powerful chip with its 37 million transistors. It competes directly with the Pentium III "Cumine".

The "old" Athlon design using a Slot A-based cartridge suffered from poor L2 cache performance. The 512 KB of L2 cache was placed outside the CPU. This gave a connection to the CPU working at only a half or a third of the processors clockspeed.

Integrating the L2 cache with the processor, the 256 KB is accessed at full processor speed, as it should. On-die cache gives the best performance. The reduction in the L2 size from 512 to 256 KB is of less importance; the full clockspeed has an enormous effect.

The Thunderbird chip performs just as good as or slightly better than Pentium III running at the same clock frequencies. With the new on-die L2 cache of 256 KB in combination with the original 128 KB L1 cache, AMD indeed has a very powerful product.

Narrow L2 cache to CPU pipeline

Still Pentium III Cumine has one advantage to the Thunderbird. When Intel decided to integrate the 256 KB of L2 cache with the processor, they gave it a 256 bit wide bus to work with.

When the L2 resides outside the CPU you have to stick to a 64 bit bus between CPU and L2. This restriction comes from the number of CPU pins you want to allocate to the L2 connection.

If the L2 cache is integrated with the CPU there is no need for this limitation. Intel wisely went from 64 to 256 bits width. This AMD has not done. For some reason, the Thunderbird core still only connect to the L2 cache on a 64 bit wide bus.

Copper or alu?

The new Thunderbirds are being produced two fabs:

- At fab25 in Austin, Texas (0.18 micron aluminium)
- At fab30 in Dresden, Germany (0.18 micron copper)

AMD told that there should be no difference between the two chips.

Copper is the most sophisticated material since it opens up for much higher clock frequencies than aluminium, due to the better electrical conduit. However, at sub-GigaHertz frequencies aluminium works fine, and there should be no difference between the chips coming from different fabs.

Chip sets

The original Athlon chip set AMD 750 works fine with the Thunderbird processor. However, AMD is soon introducing the 760 chipset for use with Thunderbird. It is expected to support DDR RAM.

The popular VIA KX133 chip set had problems with Thunderbird. Therefore VIA produced the KT133 chipset specially designed for Thunderbirds and Durons. This chipset was at first introduced as "KZ133" which was a very unwise choice in naming. KZ was the Nazi-German abbreviation for concentration camp - the camps in which millions of Jews and other Europeans were murdered. VIA wisely renamed the chip set when the historical significance of the two letters KZ came to their minds.

Another brand of Athlon is the "Spitfire" core, which was launched as "Duron" for cheaper PCs - Celeron-killer so to say. Please see module3e13 on this chip.

Sledgehammer

The successor to Athlon is codenamed "Sledgehammer". It sounds interesting:

- Multi-core technology with several complete microprocessors working parallel within the same CPU.
- The IA32 set of instructions are beeing extended to include a 64 bit mode.
- More powerfull FPU.

The value of 64 bit instructions is disputeable. We already have 64 bit and even 128 bit instructions within SSE and 3DNow!. Here it means new 64 bit instructions, registers, busses and memory addresses. To benefit fully from this 64 bit power, all software have to be recompiled. But AMD claims that the processor will run all existing 32 bit software at full speed as well.

From my humble viewpoint, "Sledgehammer" (what a name) sounds far more interesting than Intel's Itanium. Backward compatibility has always been extremely important.

Lots of RAM

One of the limitations of the 32 bit architecture is the amount of RAM. A 32 bit processor can "only" address 4 Gb of RAM. This is not enough for the biggest systems.

With 64 bit addressing you can use 18 Exebytes of RAM. That's a lot.

Sledgehammer will be introduced in 2002.

Please also see the article on die sizes here.

- Next page
- Previous page

Learn more **[top]**

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.09

On MMX, 3DNow!, and Katmai

The contents:

- An introduction
- The FPU
- Working with 3D graphics
- MMX
- 3DNow!
- Katmai

- Next page
- Previous page

Multimedia, MMX and Katmai

[top]

With the Pentium MMX we had the first of several improvements of the microprocessor's set of instructions. Later, we got 3DNow! and Katmai. What does all this mean?

In 1995 the Pentium processor was expanded with the so-called MMX instructions. That was announced as a multimedia expansion with 57 new instructions.

Today the emphasis in multimedia is especially in 3D graphics. Here the most important operation is the so-called *geometric transformations*, which deal with floating-point numbers. Let us take a look at these issues.

FPU

[top]

FPU stands for Floating-point Unit. That is the unit in the processor, that handles floating-point numbers. It is difficult for the CPU to manipulate floating-point numbers, since it requires lots and lots of bits to perform an accurate calculation. Math with integers is much simpler, and is done with hundred percent accuracy each time.

The FPU works with floating - point numbers of various bit length, depending on the desired degree of accuracy. The most accurate type has a bit length of 80!

All the modern P6 processors have 8 FP registers, each of which has a bit length of 80. So there is room inside the CPU itself for 8 numbers each of 80 bit length or, for example, 16 numbers each of 32 bit length. Read more...

Working with 3D graphics

[top]

When you draw people and landscapes, which are altered in 3D graphics, the figures are built up from small polygons (usually triangles or rectangles).

A figure in a PC game can typically be built from 200-1500 such polygons. For each change in the picture these polygons have to be re-drawn in a new position. This means that each corner (vertex) in every polygon has to be recalculated.

Floating-point number operations

To calculate the placement of the polygons, you need to use floating-point numbers. Integer calculations (1, 2, 3, 4 etc.) are not nearly precise enough. Instead, you use decimal numbers such as 4.347. These numbers are *single precision*. They are 32 bits long. There are also 64 bit numbers (having more decimal places). They are called *double precision* numbers, which are useful for even more demanding calculations.

However the 32 bits numbers are sufficient to design 3D objects. When the figures in a 3D landscape move, you need to make a so-called matrix multiplication to calculate the new vertices. If a figure consists of 1000 polygons, it requires up to 84,000 multiplications, each with two 32 bit floating-point numbers.

It is quite a hefty piece of math, for which the traditional PC is not well equipped. Actually, the largest spreadsheet available to the finance ministry is a drop in the bucket compared to Quake II, as far as number crunching ability is concerned.

What assists the 3D execution?

The CPU can easily run out of breath when it comes to work with 3D movements across the screen. So what assistance can it get? That can be provided in different ways:

- Generally speaking, the faster the CPU, the higher the clock speed, the faster the traditional FPU performance will be.

- Improvements in the CPU's FPU with pipelines and other acceleration. We see that in each new CPU generation.

- New instructions for more effective 3D performance. Instructions which can be called by the programs, 3DNow! and SSE, are examples of this.

- 3D accelerated graphics cards.

MMX

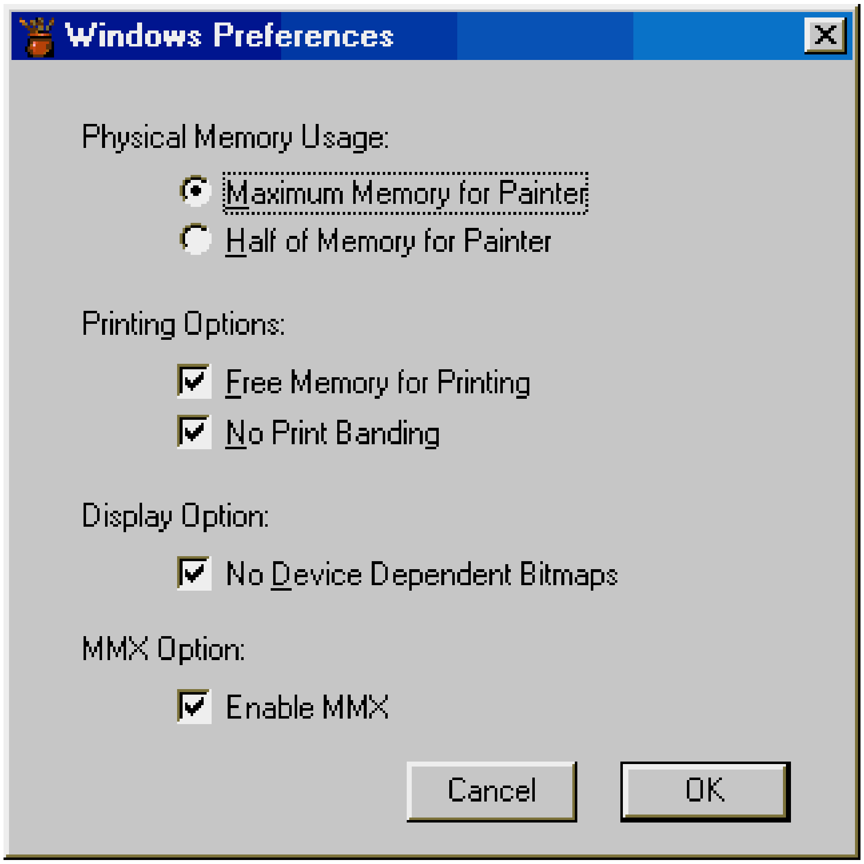
[top]

The Pentium MMX processor was a big success. However, that was not because of the MMX instructions. Many regard them as a flop.

The point is that MMX only works with integers. Furthermore the system is so weak, that it can only work with either MMX or with FPU, not both simultaneously. That is because the two sets of instructions share registers.

The MMX instructions can be of assistance in other tasks in the redrawing of 3D landscapes (the surface etc.), but for all the geometry you need much more *umph!*

Here you see the MMX enabling in a program. It is "Painter Classic" a great drawing program, which is bundled with Wacoms drawing tablets. The program utilizes MMX:



3DNow!

[top]

During the summer of 1998 AMD introduced a new collection of CPU instructions, which improve the 3D execution.

- 21 new instructions.

- SIMD instructions, which enable handling of more data portions with just one instruction.

- Improved handling of numbers, especially the 32 bit numbers, which are used widely in 3D games. 3DNow! became a big success, since the instructions soon became integrated in Windows , in different games (and other programs) and in the driver programs from the hardware producers.

The instructions use the same registers, as do MMX and traditional FPU. So they have to share them. Since the registers are 80 bits wide, they can hold two 32 bit numbers simultaneously.

Katmai

[top]

Katmai (SSE) is Intel's way to improve 3D execution in Pentium III. Read also the description in module 3e7. The problem with Katmai is that the instructions require software support, and that will take some time to get in place.

In principle Katmai is significantly more powerful than 3DNow! The 8 new 128 bit registers can actually hold four 32 bit numbers at a time. But to take advantage of this, the FPU pipeline should also have been doubled, so each multiplication or addition pipeline could receive four numbers at a time.

However that was not done in Pentium III, since it would have delayed its introduction. So the pipelines can still handle two 32 bit numbers at a time. In that way the full potential of Katmai is not reached within the actual Pentium III design.

With the current FPU unit Pentium III can perform twice as many 32 bit number operations per clock tick as can the other P6 processors (Pentium II and Celeron). That is the same performance as we find in the 3DNow! processors. But Pentium III is scheduled for future editions with a four-fold increase in FPU performance as far as the 32 bit numbers are concerned.

SIMD

SIMD stands for Single Instruction Multiple Data. This technique was introduced in the MMX processors, where more than one *integer* could be processed simultaneously. In Pentium III this technique was given another lift, so now it can handle more than one *floating-point number*. Multimedia handling especially will benefit from this, since many floating-point number operations are handled in sound and video programs.

With the introduction of Pentium 4, the SIMD instruction set was further improved with144 new instructions.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.10

On CPU sockets

The contents:

- The CPU Sockets and chip sets
- Pentium II road map
- Three lines of Intel CPUs

- Next page
- Previous page

The CPU Sockets and chip sets

[top]

To get an overview on all the different Intel CPUs, you may take a look at the various sockets, that are used to mount the CPU. Each socket is working together with specific chip sets. Let us finally also look into the future...

There are many different CPU sockets in use for the various CPUs. Here you see a handfull of them:

Socket	Fits CPU	Intel Chip set	Number of pins
Socket 7	Pentium, MMX, K5, 6x86, K6, IDT WinChip, 6x86MX, K6-2	82430TX	321
Socket 8	Pentium Pro	82440FX	387
Slot One	Pentium II	82440FX 82440LX	242
Slot One	Pentium II (100 MHz system bus), Pentium III (100 and 133 MHz)	82440BX 82440JX	242
Slot One	Celeron	82440EX	242
PGA370	Socketed Celeron Pentium III	82440BX 82440LX 82440EX 82440EZ i810 i815	370
Slot Two	Pentium II Xeon, Tanner	82440GX 82450NX	330
PGA423	Pentium 4	i850	423
Slot M	Merced	?	?

Only Socket 7 may be copied freely. The other ones are Intel's patents. They may be manufactured by others on license from Intel. Cyrix is expected to produce Slot 1-compatible modules.

A road map to Intel CPUs

[top]

Putting all the 6th generation's CPUs from Intel together, we get a picture like this:

CPU name/ code name	Year	CPU/ bus MHz	L2 cache	Socket	Process techno- logy	Extra instructions
Pentium Pro	1995	233/66	512-1024 full speed	Socket 8	0.35	None
Pentium II "Klamath"	1997	300/66	512 KB half speed	Slot 1	0.35	MMX

Pentium II "Deschutes first"	1998	300/66	512 KB half speed	Slot 1	0.25	MMX
Pentium II " Deschutes second"	1998	400/100 450/100	512 KB half speed	Slot 1	0.25	MMX
Celeron	1998	266/66 300/66	No	Slot 1	0.25	MMX
Celeron A "Mendocino"	1998/ 1999	300/66 333/66 366/66 400/66	128 KB full speed (on-die)	Slot 1	0.25	MMX
Celeron socketed	1999	366/66 400/66 433/66 466/66 500/66 533/66	128 KB full speed (on-die)	Socket 370	0.25/0.18	MMX
Xeon	1998	400/100	512-2048 full speed	Slot 2	0.25	MMX
Pentium III	1999-2000	500/100 533/133 600/100 600/133 650/100 700/100 733/133 ... 1133/133	256 Half speed	Slot 1/ PGA370	0.25/0.18	MMX SSE
Pentium III Xeon (Tanner)	1999	550/100	512-2048 full speed	Slot 2	0.25	MMX SSE
Coppermine	1999/ 2000	733/133	256 KB full speed (on-die)	Slot 1 and PGA370	0.18	MMX SSE
Pentium 4	2000	1400	256 KB full speed (on-die)	PGA423	0.18/0.13	SIMD2

"Northwood"	2001	1600	512 full speed (on-die)	PGA478	0.13	?
"Tualatin"	2001	1500	512 full speed (on-die)	Socket 370	0.13	?
Celeron 2	2001	900 - 1200	128 full speed (on-die)	Socket 370	0.13	?

Three lines of Intel CPUs

[top]

If you study the scheme above, you see the development from Intel coming in three "lines". Each line will be developed independently:

Market segment	Processor line	CPU speed	Bus Speed/ Front Side Bus speed	Number of CPUs in system
The consumer	Socketed Celeron	566-1200	66/100 MHz	1
The professional	Pentium III Pentium 4	733-1400 MHz	100/133 MHz or 400 MHz	1 or 2
The server	Xeon Tanner (Cascades)	600 - 1200 MHz	100/133 MHz or 400 MHz	4

- Next page
- Previous page

Learn more [top]

Read about drives in module 4a

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.11

A guide to Intel Itanium.

The contents:

- An introduction to Intel's Itanium/Merced
- The specs
- The perspective for IA 64

- Next page
- Previous page

The Itanium/Merced

[top]

Merced was the code name for a completely new CPU, which Intel has developed together with HP, who has a vast experience in the manufacture of high end CPUs (RISC).The chip is due 2000 and will be launched under the name Itanium.

The chippen will cost around $4000.

Itanium is a IA-64 processor. This means that it is targeted for a completely different type of programs than those we are used to.

The specs

This is what I know of the Itanium:

- A 64 bit CPU with IA-64 architecture.
- Starting clock frequency: 800 MHz.
- 25.4 million transistors.
- "Massive hardware units": 128 integer and 128 floating point registers with multiple integer and floating point units all working in parallel.
- 0.18 micron technology at 1.8 Volt.
- Slot M cartridge.
- 4 MB of Level 3 cache, holding 320 millions of transistors.
- The L3 cache runs on a 12.3 GB per second bus.
- VLIW design

The first chip set for Itanium should become 460GX, which allows four Itaniums on the same motherboard and 64 GB of RAM.

Later it should be possible to construct super computers holding 512 Itaniums (in clusters of four).

DDR RAM

The Itanium chip sets should be designed to use DDR RAM and not Rambus

Problems with heating

Heating problems have been reported. The Itanium is extremely power hungry and runs very hot. It has been using up to 130 watts in some tests, and this appears to be a really serious problem. The problems should arise from the choice of VLIW design, which should not be suitable for a general-purpose CPU as Itanium as some articles indicate. I am no expert on these issues, and it sounds weird if Intel should choose thewrong architecture.

The perspective for IA 64

There is no doubt that the Itanium is going to be a heavy processor. But it will not end up on many desktops. It is too expensive, and the design is 100% intended the server market.

There have been speculations about a lousy IA-32 performance. All the programs we use (including Windows 2000) are of 32bits design. This corresponds with the P6 processors (like Pentium III etc.) which also are of 32 bits architecture.

Now Intel comes with a 64 bit processor. It has to *emulate* the 32 bit instructions, to execute 32 bit programs like Windows . An emulation is costly, it takes power from the processor. This is also the case with the Itanium; it has to translate each of the IA-32 instruction. Some magazines have claimed that the Itanium will be terrible slow executing IA-32 programs.

Some articles even claim that Intel wants to dump the Itanium and go for the successor 'McKinley'.

Anyway, the Itanium will require a new 64 bit operating system - could it be Windows 2064? Linux 64 and NT 64 should be upcoming.

- Next page
- Previous page

Learn more [top]

Read about drives in module 4a

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.12

The VIA "Joshua" processor

The contents:

- An introduction
- Compared to Athlon

- Next page
- Previous page

Joshua or Cyrix MIII

[top]

Cyrix was working on a brand new processor nucleus (code name Jalapeño) before the company was sold to VIA. The processor was expected to be marketed in the MIII in the middle of year 2000, but I do not know what happened to the design.

The features were:

- Clock frequencies starting at 533 MHz (PR533).
- Built in 64 KB L1 cache, which works at full clock speed.
- Built in 256 KB L2 cache, which works at full clock speed.
- Built in 3DNow! 3D graphics acceleration.
- Socket 370 with 133 MHz bus for RAM.

From the original MIII design the following rests: Built in hardware coder for MPEG. Use of Direct RDRAM (Rambus RAM). Powerful memory controller, which should permit transmission at 3.2 GB per second.

VIA3 later introduced a Cyrix III processors based on IDT's WinChip technology.

Compared to Athlon

Joshua was not a high end processor like AMD's K7 Athlon. However it is intended to be a powerful low price CPU with integrated sound and graphics controller. If we compare the design with the Athlon, we see a slightly lower performance:

	Joshua	Athlon
Number of program instructions which can be executed superscalar (simultaneously)	2	3
Number of internal operations per clock cycle	6	9
Pipelines to floating-point number operations (FP, MMX, 3DNow!)	2	3

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.13

The AMD "Duron" processor

The contents:

- An introduction
- Compared to Celeron

- Next page
- Previous page

The Duron

[top]

In the summer 2000 AMD introduced a new and very powerful low-end chip, formerly known as codename "Spitfire".

The Duron is a Athlon in new design:

- 64 KB L2 cache on-die
- 128 KB L1 cache on-die
- Exclusive L2 cache design
- Socket A for in-expensive motherboard design
- Clock frequencies from 600 to 950 MHz

The Duron is the first CPU to have an internal L2 cache smaller than the L1 cache.

The exclusive cache design means that you never find the same data in L1 as in L2 cache. This increases the efficiency of the cache.

A Celeron killer

The Duron processor is designed for the lower end of the market. Here we find the Intel Celeron processor an AMD's own K6-2, which is on its way out of production.

Compared to the Celeron, Duron has a more powerful layout:

- Bigger cache
- Choice of PC100 or PC133 RAM
- A better processor architecture

The Morgan Kernel

Late in 2001 new 1200 MHz versions of the Duron processor was introduced. This processor holds a new kernel of same generation as the Palomino kernel in AthlonXP.

Here you find SSE support and other news similar to those in the AthlonXP.

The first tests of the new 1200 MHz Duron showed a very convincing performance almost compareable to a Pentium 4 working at 1500 MHz!

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 3e.14

The Intel Pentium 4 processor

The contents:

- An introduction til Pentium 4
- SSE2
- The Execution Trace Cache
- "Northwood"

- Next page
- Previous page

The Pentium 4

[top]

In November 2000 Intel introduced the new and very powerful high-end chip Pentium 4, formerly known as codename "Willamette". It was (as expected) delayed.

The Pentium 4 is a completely new processor holding several new designs. Here is a highlight:

- 400 MHz Front Side Bus of 128 bit width
- Execution Trace Cache
- 20 KB L1 cache and 256 KB L2
- The ALU (Arithmetical Logic Unit) runs at twice the clock speed
- A new socket for simple motherboard design
- Clock frequencies from 1500 MHz
- 20 stages pipeline
- SSE2 and 128 bit MMX
- 42 millions of transistors
- A new 423 pins socket design
- Dual Rambus memory channel with i850 chipset
- Only single processor mode available.

NetBURST

Intel uses the term NetBURST to describe some features in Pentium 4:

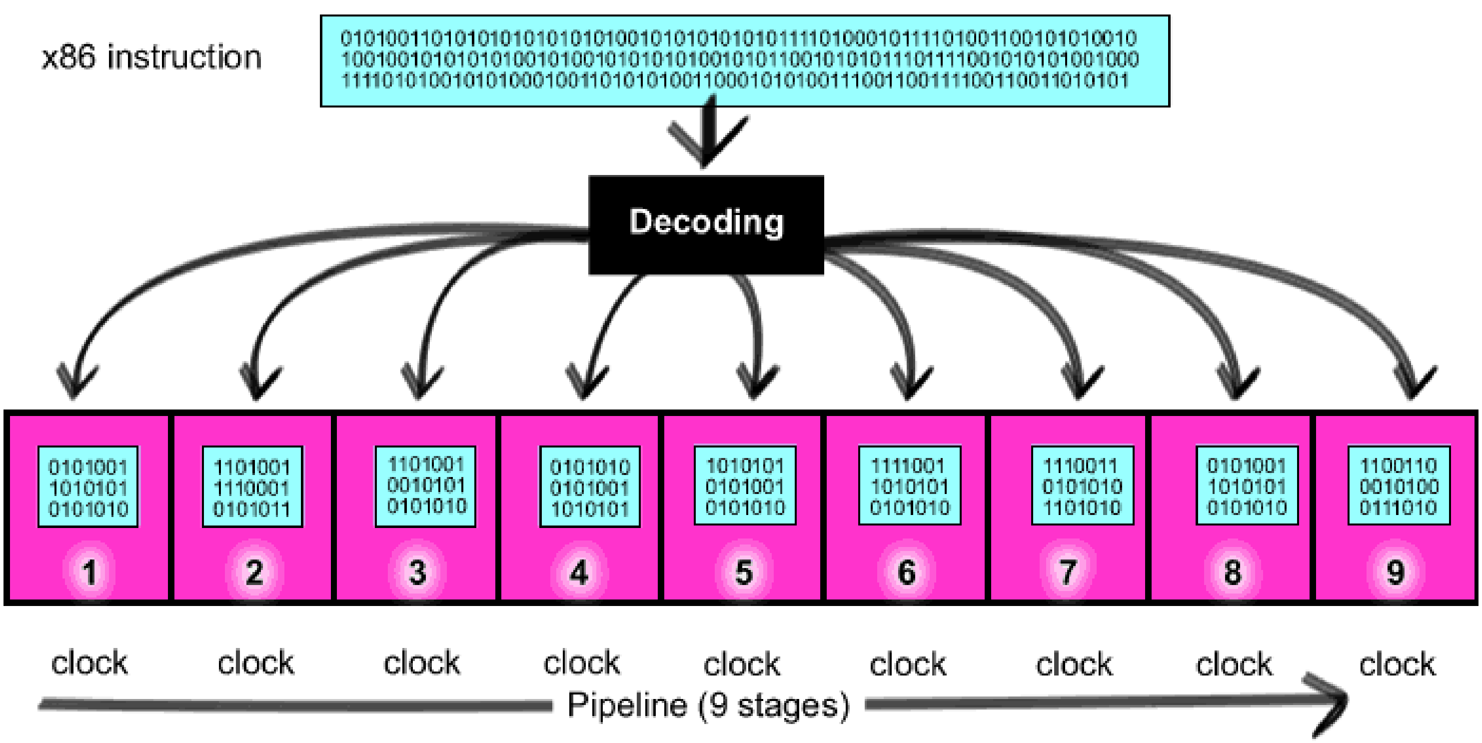
- Advanced Dynamic Execution
- The Rapid Execution Engine

Advanced Dynamic Execution means that the processor may execute up to 6 instructions simultanously.

Using *Rapid Execution Engine* certain instructions may be executed at twice the normal speed.

20 stages of pipeline

The pipeline is a execution unit which takes in decoded micro-instructions. The X86 instructions are first decoded and then sent to the pipeline. The longer the pipeline is, the quicker an instruction can be executed. Each stage executes a minor part of the work and by spreading the work on "more hands", the efficiency is increased.



In Pentium III the pipeline was of 10 stages. In Pentium 4 it has been increased to 20 stages.

The problem with the long pipeline is, that it takes to longer time load the instruction. And if the instruction should not have been loaded at all, the error is most costly (in time) the longer the pipeline becomes.

With many instructions being executed simultanously you cannot avoid loading the wrong instruction (called *misprediction*) from time to time. And a shorter pipeline is quicker to recover this error - fewer stages has to be cleared and reloaded.

The analytic work preventing mispredictions is done by the Branch Prediction Unit. This has, according to Intel, been improved with a 30% better performance compared to the Pentium III.

Stalling is another phenomen. Normally the data to be used is located in the cache. But if, for some reason, the data is missing in the cache, it has to be loaded from RAM. This takes a lot of time, and the longer the pipeline is the longer time it takes. .

The benefit from a pipeline increased from 10 stages to 20, is to open up for new higher clock frequencies. When each instruction is executed in more stages, it can be done a lot quicker.

At lower frequencies this gives no advantage. In fact all reports indicate that a 2000 MHz Pentium 4 is slower than a 1600 MHz AthlonXP. This is due to the difficult prediction of the order of the instructions. Wrong predictions gives wasted clock cycluses and a poorer performance.

However, a longer pipeline is required for processor speeds above 2 GHz. Intel expects this new NetBurst core to live three to five years. Hence we may expect Pentium 4 successors to reach 5 GHz or more.

SSE2

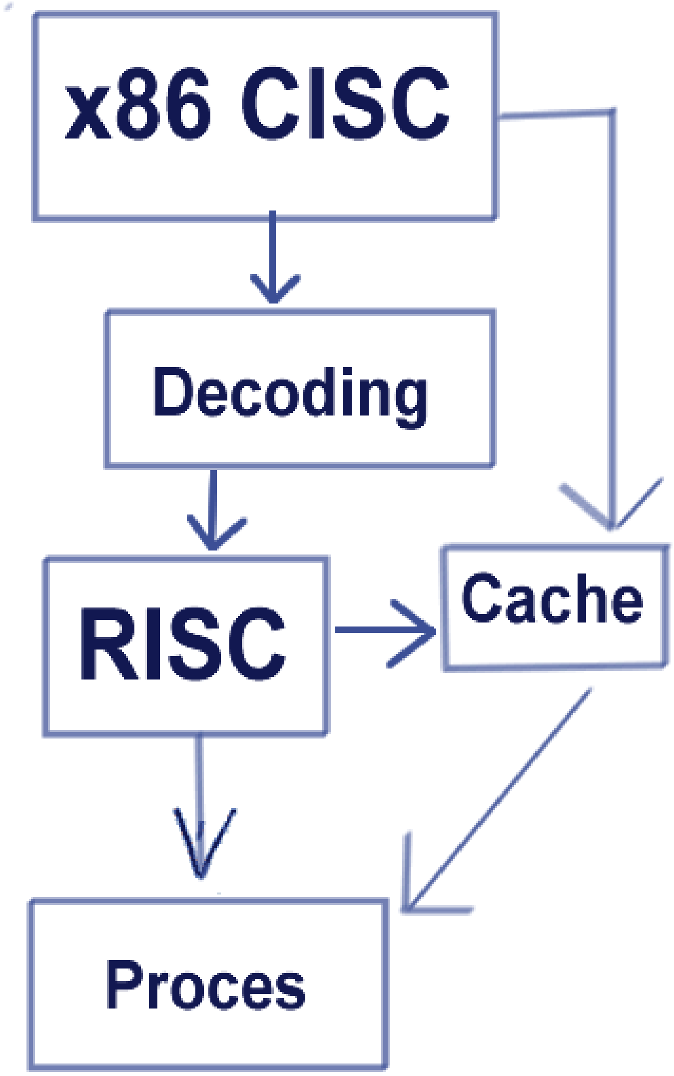
The Streaming SIMD (Single Instruction Multiply Data) Extensions known from Pentium III has been improved. The data width has doubled from 64 to 128 bit.

Also 144 new instructions in SSE2 gives more parallel execution. Now four Internet/Multimedia-based operations can be executed simultanously. The new design appears to have been accepted by software developers, and it will probably be very useful within programs like Photoshop.

The Execution Trace Cache

The Pentium 4 is the first CPU to have a "code cache". All instructions are translated inside the CPU. This happens in all modern x86 processors. They receive x86 instructions from the software. These instructions are "crunched" into smaller instructions which then are executed natively.

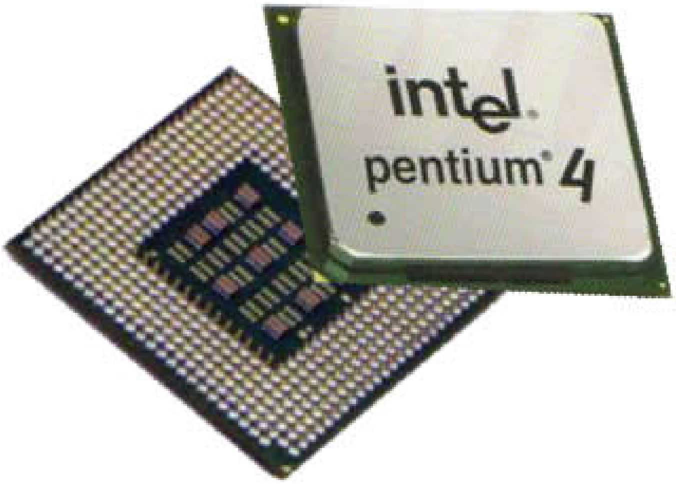
The new thing in Pentium 4 is that these translated instructions are cached and reused. The logic of this setup may look like this:



This new cache, being a part of the L1 cache holds 12 KB. Together with 8 KB of data cache, the L1 cache consists of 20 KB.

Heavy hardware

The first Pentium 4s were giant chips with a die size of 247 mm^2. The 42 millions of transistors uses 60 watts and requires heavy cooling. However, Intel has done a lot to provide sufficient cooling using new materials and design.



The sockets

The first Pentium 4s came with a 423 pin design. These CPUs could only use RDRAM, which only requires few pins interface.

Later came a 478 pin design with support for SDRAM.

The chipsets

The chip set i850 ("Tehama") is using a dual RDRAM bus. This is also heating up the MCH (north bridge) chip, so additional cooling is required.

The Intel i845 chipset was introduced August 2001. It gives an interface to standard 133 MHz SDRAM. This chipset is found in many cheap Pentium 4 computers.

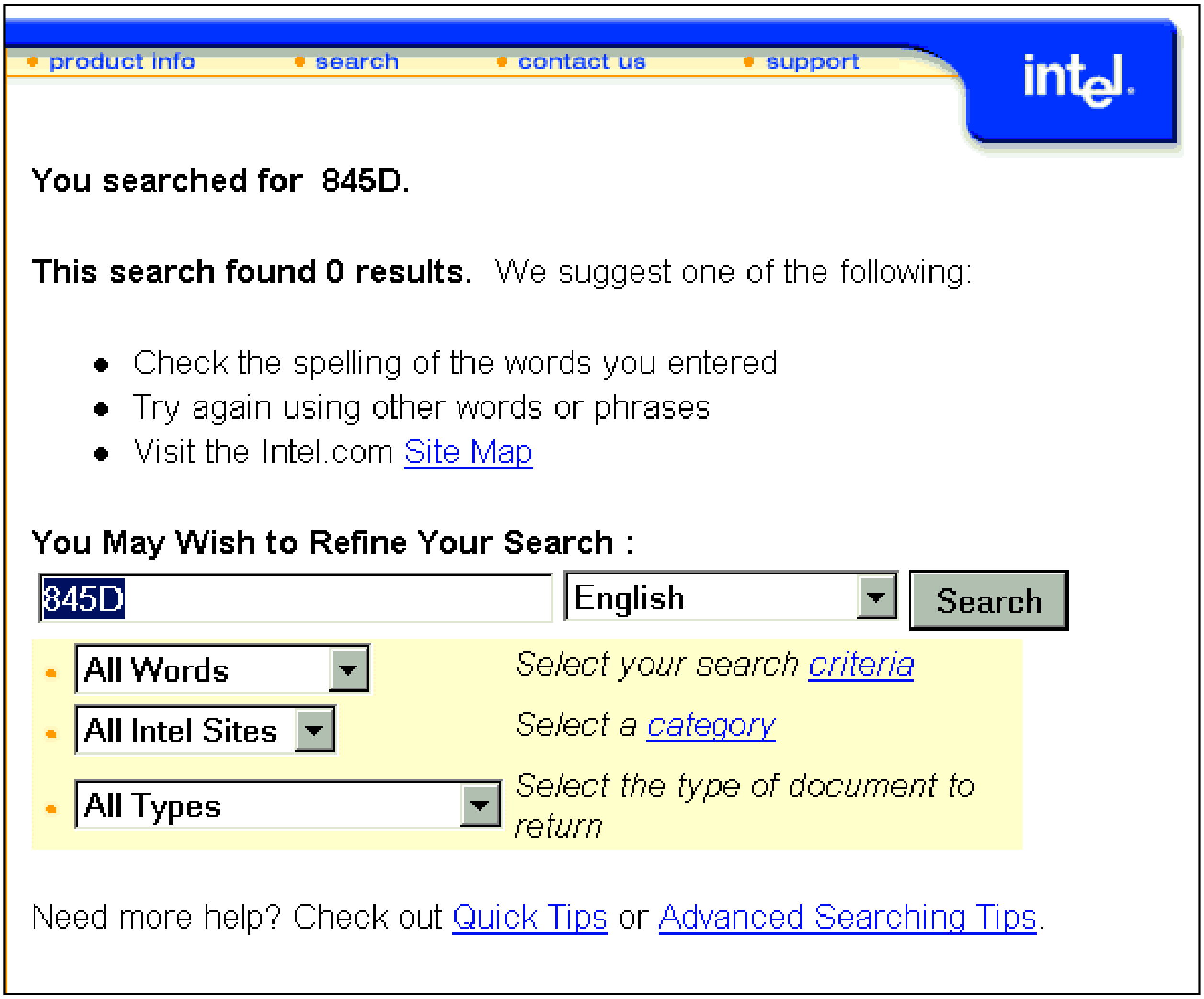
Support for 266 MHz DDR RAM is found in the VIA P4X266 and P4X266A chipsets to be used with Pentium 4 (against Intels will; they don't like VIA at all).

SIS 645 is a chipset for Pentium 4 with support for both 266 and 333 MHz DDR RAM.

The i845D chipset

December 2001 Intel suddenly introduced a new flavour of the 845 chipset. It is designed to use with DDR-RAM!

There was no press releases about this product. Searching Intel's web for info on this chipset gave no result (December 27th 2001):



However the i845D chipset was found in computers from Dell and others.

Intel has been under a hard pressure from AMD, VIA, and the Taiwanese motherboard companiess in all 2001, and it was fine to see, that they finally had to adapt to common sense. DDR is the RAM type to use - almost nobody wants RAMBUS! And Intel's dominating position in the market will promote better DDR RAM products.

Northwood

On January 7th 2002 Intel introduces a new 2.2 GHz version of the Pentium 4.

This processor comes with the new Northwood kernel:

- L2 cache doubled from 256 KB to 512 KB
- 0.13-micron process

New techniques should also improve the clock circle/instruction execution ratio.

Due to the new design, the performance of this Pentium 4 was increased with 30% compared to the 2 GHZ version - where one should expect only 10%.

- Next page
- Previous page

Learn more [top]

Read about chip sets on the motherboard in module 2d

Read more about RAM in module 2e

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP and module 5c about Firewire.

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 4a.

- Next page
- Previous page

Drives are storage media

A drive is the name for several types of *storage media.* There are also storage media, which are not drives (RAM, Tape Streamers), but on these pages, we will discuss the drives.

Common to drive medium is:

- A *file system* can be assigned to them.
- They are recognized by the operating system and they are assigned a *drive letter*.

During start up, drives are typically recognized by the PC system software (ROM BIOS + operating system). Thus, the PC knows which drives are installed. At the end of this configuration, the appropriate drive letter is identified with each drive. If a drive is not "seen" during start up, it will not be accessible to the operating system. However, some external drives contain special soft-ware, allowing them to be connected during operation.

Some examples of drives

[top]

Storage media	Drive letter
Floppy disks	A: B:
Hard disk	C: D: E:
CDROM/DVD	F:
MO drive	G:

Network drive	M:
RAM disk	O:

On this and the following pages, I will describe the various drive types, their history and technology. The last two drive types in the above table will not be covered.

Storage principles

[top]

Storage: Magnetic or optic. Data on any drive are *digitized.* That means that they are expressed as myriads of 0s and 1s. However, the storage of these bits is done in any of three principles:

The physical drive principle	Disk types
Magnetic	Floppy disks Hard disk Syquest disks Zip drive LS-120 disks
Optic	CD-ROM DVD
Magneto optic	High end drives

Interface

[top]

Individual drives are connected to other PC components through an interface. The hard disk interface is either IDE or SCSI, which in modern PCs is connected to the PCI bus. Certain drives can also be connected through a parallel port or the floppy controller:

Interface	Drive
IDE and EIDE	Hard disks (currently up to 40 GB) CD-ROM
SCSI	Hard disks (all sizes) and CD-ROM

ISA (internal)	Floppy drives CDROM and super floppies connected through parallel port

Let us start evaluating the drives from the easy side:

The traditional floppy drive

[top]

We all know diskettes. Small flat disks, irritatingly slow and with too limited storage capacity. Yet, we cannot live without them. Very few PCs are without a floppy drive.

Diskettes were developed as a low cost alternative to hard disks. In the 60s and 70s, when hard disk prices were exorbitant, It was unthinkable to use them in anything but mainframe and mini computers.

The first diskettes were introduced in 1971. They were 8" diameter plastic disks with a magnetic coating, enclosed in a cardboard case. They had a capacity of one megabyte. The diskettes are placed in a *drive*, which has read and write heads. Conversely to hard disks, the heads actually *touch* the disk, like in a cassette or video player. This wears the media.

Later, in 1976, 5.25" diskettes were introduced. They had far less capacity (only 160 KB to begin with). However, they were inexpensive and easy to work with. For many years, they were the standard in PCs. Like the 8" diskettes, the 5.25" were soft and flexible. Therefore, they were named *floppy disks*.

In 1987 IBM's revolutionary PS/2 PCs were introduced and with them the 3½" hard diskettes we know today. These diskettes have a thinner magnetic coating, allowing more tracks on a smaller surface. The track density is measured in TPI (tracks per inch). The TPI has been increased from 48 to 96 and now 135 in the 3.5" diskettes.

Here you see the standard PC diskette configurations:

Diskette size	Name	Tracks per side	Number of sectors per tracks	Capacity
5.25" Single side	SD8	40	8	40 X 8 X 512 bytes = 160 KB
5.25" Double side	DD9	40	9	2 X 40 X 9 X 512 bytes = 360 KB
5.25" Double side High Density	DQ15	80	15	2 X 80 X 15 X 512 bytes = 1.2 MB
3.5" DD	DQ9	80	9	2 X 80 X 9 X 512 bytes = 720 KB

3.5" HD	DQ18	80	18	2 X 80 X 18 X 512 bytes = 1.44 MB
3.5" XD (IBM only)	DG36	80	36	2 X 80 X 36 X 512 bytes = 2.88 MB

Diskette drives turn at 300 RPM. That results in an average search time (½ revolution) of 100 ms.

The super floppy drives are described in module 4d.

The floppy controller

[top]

All diskette drives are governed by a controller. The original PC controller was named NEC PD765. Today, it is included in the chip set, but functions like a 765. It is a programmable chip. It can be programmed to handle all the various floppy drive types: 5.25" or 3.5" drives, DD or HD etc.

The controller has to be programmed at each start up. It must be told which drives to control. This programming is performed by the start up programs in ROM (read module 2a). So you don't have to identify available drive types at each start up, these drive parameters are saved in CMOS RAM.

The floppy controller reads data from the diskette media in serial mode (one bit at a time. like from hard disks). Data are delivered in parallel mode (16 bits at a time) to RAM via a DMA channel. Thus, the drives should be able to operate without CPU supervision. However, in reality this does not always work. Data transfer from a diskette drive can delay and sometimes freeze the whole PC, so no other operations can be performed simultaneously.

- Next page
- Previous page

Learn more [top]

Module 4b about hard disks.

Read Module 4c about optical media (CDROM and DVD).

Read Module 4d about super diskette and MO drives.

Read Module 4e about tape streamers (which *are not* drives).

Read Module 5c about SCSI.

Read Module 6a about file systems.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 4d.

Other drives

Here we look at other drive types not previously mentioned:

- Super floppies
- Zip
- LS 120
- Sony HIFD
- MO-drives

- Next page
- Previous page

Super floppies

With the increasing performance in magnetic hard disks you may expect a similar development in the production of the floppy drives. However,it took many years to get alternatives to the 1.44 MB floppy disk. Now we have three drives to choose from, all using special 3½ inch media.

All of them perform very well and are stable and pretty fast:

Drive	Capacity	Comments

Iomega Zip	100 MB	10 million sold units makes it the most compatible drive
LS120	120 MB	Read and write on 1.44 MB floppies as well
Sony HiFD	200 MB	Read and write on 1.44 MB floppies as well

Use the Zip, the LS120 or the HiFD in new PCs - they are cheap and good for backup. You also find portable PC Card versions of these drives.

The Zip drive

[top]

The Zip drive uses a kind of diskette, which can hold 100 MB. In my opinion, the Zip drive works excellently. They are stable, inexpensive, and easy to work with. The drives are not the fastest.

I and many others have used Zip drives since they came on the market. This provides us with a common standard to move large files and to make back-ups. For example, you can use this drive to install Windows 95/98 on a computer without a CDROM and avoid having to insert numerous floppy disks.

The 100 MB Zip disk is borderline size. However, compared to the work I had to do previously, compressing files with PKZIP onto multiple diskettes, these are very practical.

Two types of interface

The Zip drive exists in different versions:

- Internal and external
- For SCSI and to floppy/parallel port

The SCSI model is by far the fastest. That is really good. If your SCSI controller is installed with Windows 95/98, you just have to install the drive with two screws and two cables and you are in business.

The parallel port version is good, because it can be connected to *any* PC. I have a boot diskette, which includes a driver plus the program GUEST.EXE. I connect the drive to a parallel port, and boot with the diskette. Then it is ready to run.

I have the quite fast SCSI version installed in my stationary PC. I use the somewhat slower parallel port version "in the field."

My latest information is that almost 20 million Zip drives have been sold (October '98) and that they sell 1 mill. a month. This speaks for itself and makes it a de facto standard.

The BIOS manufacturers AMI and Phoenix include the floppy version of the drive in their programs as a boot device. That will eliminate the need for other drivers, and you will be able to boot from the Zip disk.

LS120

[top]

Since 1993, we have heard about the LS120 drive and now it is available. It is a 120 MB standard designed by the company Imation. LS120 is supposed to replace the regular floppy drives. At the same time, they read the traditional 3½" floppy diskettes (DD and HD) much faster than the ordinary floppy drives.

The LS120 ought to have become the new floppy standard, but it has come too late with all the installed Zip drives.

The drives, coming from Imation, use EIDE interface, and they are comparable to the Zip drives.

However the SCSI-version of Zip is several times faster than the LS120 drive.

HiFD

[top]

Sony has a super diskette drive called HiFD (High Floppy Disk) holding 200 MB on a 3½" floppy disk.

Like the LS120, the HiFD disk drives can read and write old 1.44 MB floppy disks in addition to the new high density disks. However, it should be a lot faster than the LS120 drive.

- 3.6 MB/s maximum transfer rate (read)
- 1.2 MB/s maximum transfer rate (write)
- 3,600 rpm rotational speed with dual discrete gap head (flying head type) and a high speed head actuator with VCM (voice coil motor)
- Recording Capacity: 200 MB (formatted), 240 MB (unformatted).
- Disk Diameter: 86 mm
- Track Density: 2822 TPI (111 t/mm)

The HiFD drive is the best performing of the three super floppies mentioned in this article. However, compatibility is often more important than performance. In that case Zip is the winner.

MO drives

[top]

Magnetic Optic drives represent an exciting technology. The medium is magnetic, yet very different from a hard disk. You can only write to it, when it is heated to about 300 degrees Celsius (The Curie point)

This heating is done with a laser beam. The advantage is that the laser beam can heat a very minute area precisely. In this manner the rather unprecise magnetic head, can write in extremely small spots. Thus, writing is done with a laser guided magnet. The laser beam reads the media. It can detect the polarization of the micro magnets on the media.

MO disks are fast, inexpensive, and extremely stable. They are regarded as almost wear proof. They can be written over and over again forever, without signs of wear. The data life span is said to be at least 30 years. There are many MO drive variations, but all are very expensive. The only mainstream use of the MO-technology is found in Sony's recordable MiniDisc.

All other drives are very expensive (>$2000). For example:

Sony SMO-F551 MO Drive

- Continuous/Composites (ISO/IEC 15286) format
- Direct overwrite magneto optical drive with 5.25-inch double sided disk
- 5.2 GB (2,048 Bytes/sector), 4.8 GB (1,024 Bytes/sector), 4.1 GB (512 Bytes/sector)
- 3,600 rpm/ 3,300 rpm

Maxoptix T6-5200

HIGH CAPACITY, MULTI-PURPOSE 5.2 GB READ/WRITE OPTICAL DISC DRIVE

Near-field recording

A new magneto optical technology with flying heads and solid immersion optical lenses is called near-field recording. It promises 20 GB high density magnetic storage on 5.25" plastic media. Check www.terastor.com for further information.

- Next page
- Previous page

Learn more **[top]**

Module 4e about tape streamers (which not are drives).

Module 5c about the SCSI interface

Module 6a about the file systems.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 4e.

About tape streamers

This subject is not very interesting, and the text in this module reflects this point of view. Tape streamers are used for data *backup*. They come in different types and price ranges, but have this in common:

- The tape streamer does not work like a *drive*. You cannot retrieve any particular file. The data must be read using special back-up software.

- Data are stored *sequentially* on the tape. This means that you can not, contrary to disks or CDROMs, read in random fashion. You must wind the tape to the desired location.

The advantage of Tape streamers is their low cost. They contain lots of data on inexpensive tapes. They are available in different types:

- Next page
- Previous page

Tapestreamers - variations	
Installation	• *Internal* in the PC • *External* in a box connected to the PC
Interface	• On floppy controller or parallel port • EIDE (rather new) • SCSI
Tape types	• QIC 40/80/Wide • Travan TR1/TR3/TR4 • DAT

Tapes		
Standard	Capacity	Typical interface
QIC	120 MB	Floppy/parallel
Travan TR1: TR3: TR4:	 800 MB 1,6 GB 4 GB	 Floppy/parallel Floppy/parallel EIDE or SCSI
DAT	2/4 GB	SCSI

- Next page
- Previous page

Learn more [top]

Module 5a about adapters.

Read Module 5c about SCSI.

Read Module 6a about file systems.

Last revised: February 20, 1997.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

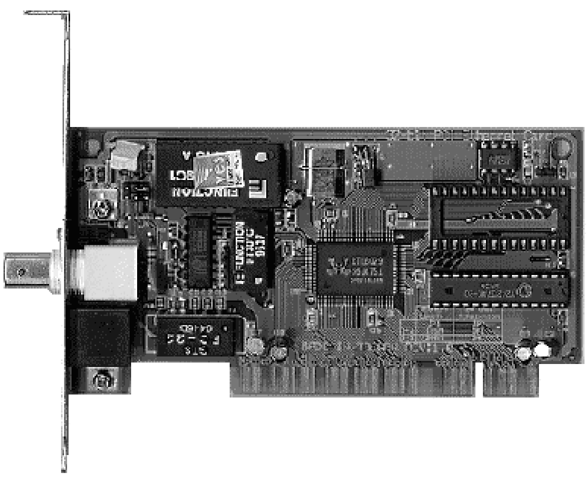
KarbosGuide.com. Module 5a1a.

About the PC I/O system. Expansion cards. Adapters, etc.

The contents:

- Intro to I/O
- A model

- Next page
- Previous page



Intro to I/O

[top]

This page should preferably be read together with module 2c, 2d, 5b and 5c. The first two describe the I/O buses and the chip sets. Here we will look at the other end of the I/O buses, the "exit."

There are four I/O buses in the modern PC architecture and each of them has several functions. They may lead to internal and external ports or they lead to other controlling buses. The four buses are:

- ISA, which is old, slow, and limited, compared to the alternatives listed below. We hope that it is replaced by the following interfaces:

- PCI, which is the newer high speed multifunction I/O bus.

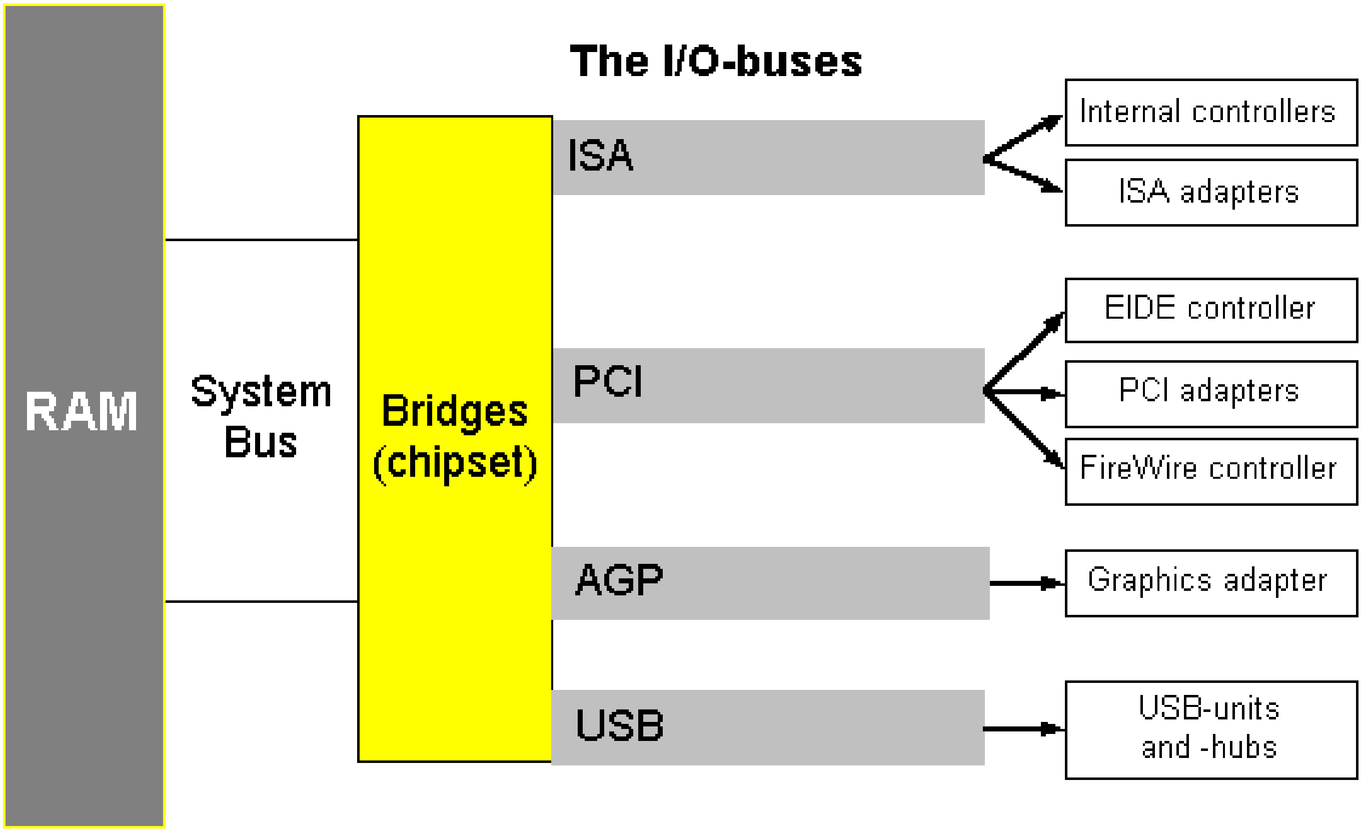
- AGP, which only is used for graphics adapter.

- USB, which is the new low speed I/O bus to replace ISA.

The ISA and the PCI bus both end up having to exits:

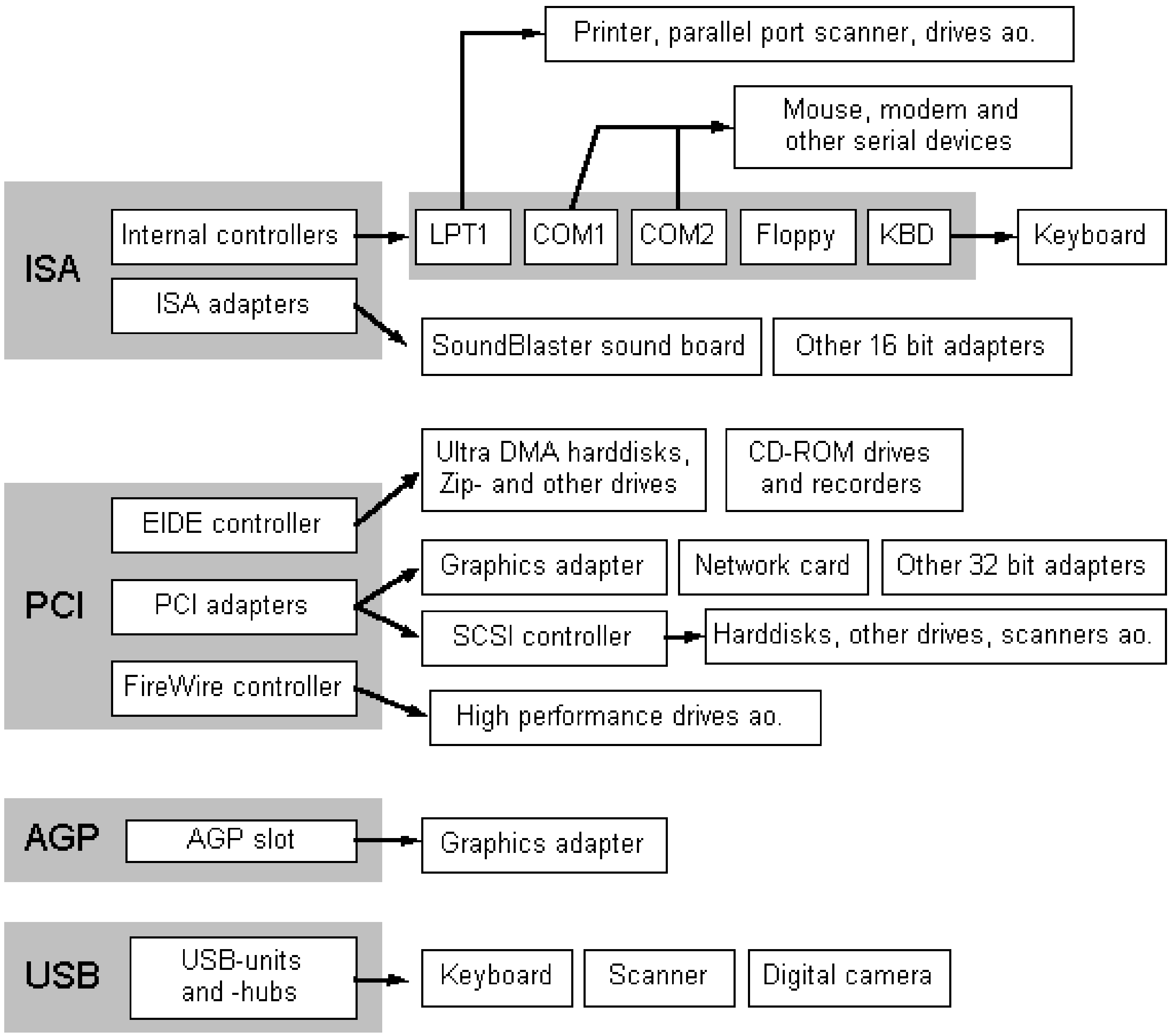
- Internal I/O ports (LPT, KBD, COM1, COM2, EIDE etc.)
- Expansion slots in the system board, in which we can insert adapters.

If you look at this illustration you will see the overview of this architecture:



A model

If we focus on the right end of the illustration we approach the I/O units. Here you get a closer look at that:



As you see, there is room for a lot of units to be connected to the PC.

The PCI bus is the most loaded of all the buses. It is used for so many purposes that the output for the graphics adapter has been isolated on its own AGP-bus.

But still the PCI bus is heavyly loaded, connecting the system bus to the network controller and the various EIDE- and SCSI drives. Because of the high bandwidth of the FireWire bus, overall throughput of both interfaces would be improved by separating these. We hope to see a separate FireWire interface in future motherboard architectures.

- Next page

- Previous page

Learn more [top]

Read: Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about SCSI, USB etc.

Read A little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

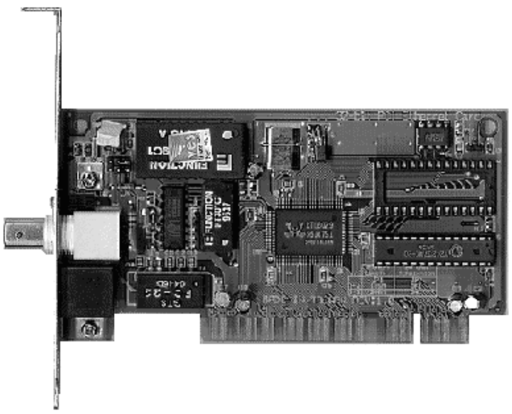
KarbosGuide.com. Module 5a1b.

About I/O units, continued

The contents:

- The Internal I/O ports and units
- The serial ports

- Next page
- Previous page



The internal I/O ports

[top]

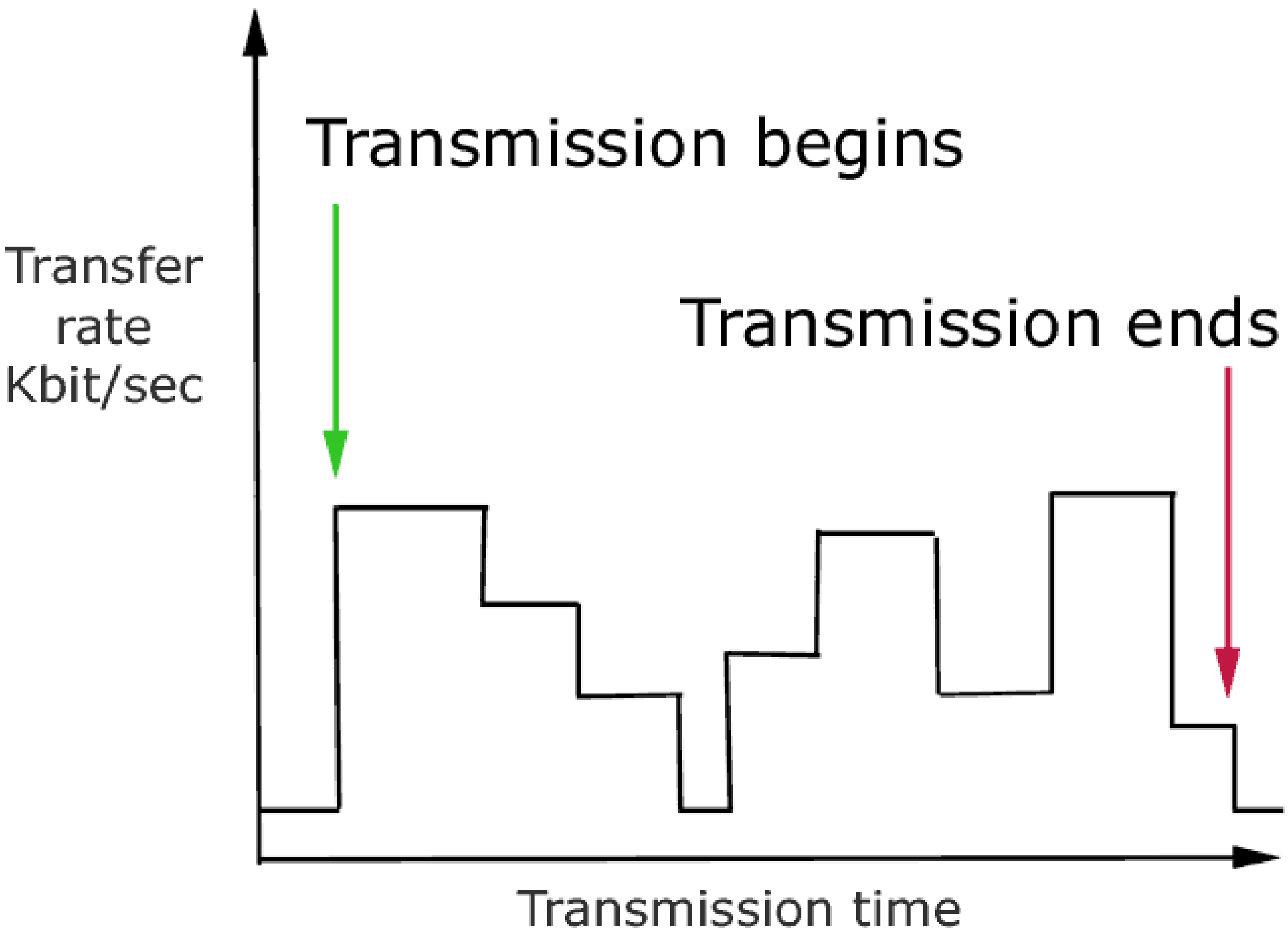
As mentioned, the USB is going to become the main bus for low speed devices. But so far we still use the internal "face" of the ISA bus for a range of purposes. At any PC motherboard you find these:

- The floppy controller
- The serial ports
- The parallel port(s)
- The keyboard controller

They all occupy IRQs which is a central part of ISA architecture and a pain in the a... Let us take a moment to look at these ports and controllers.

The serial ports

Serial transmission means to send data from one unit to another one bit at the time. The PC architecture traditionally holds two RS232 serial ports. The RS-232 standard describes an *asynchronous* interface. This means that data are transmitted only when the receiving unit is ready to receive them:



Synchronous/asynchronous

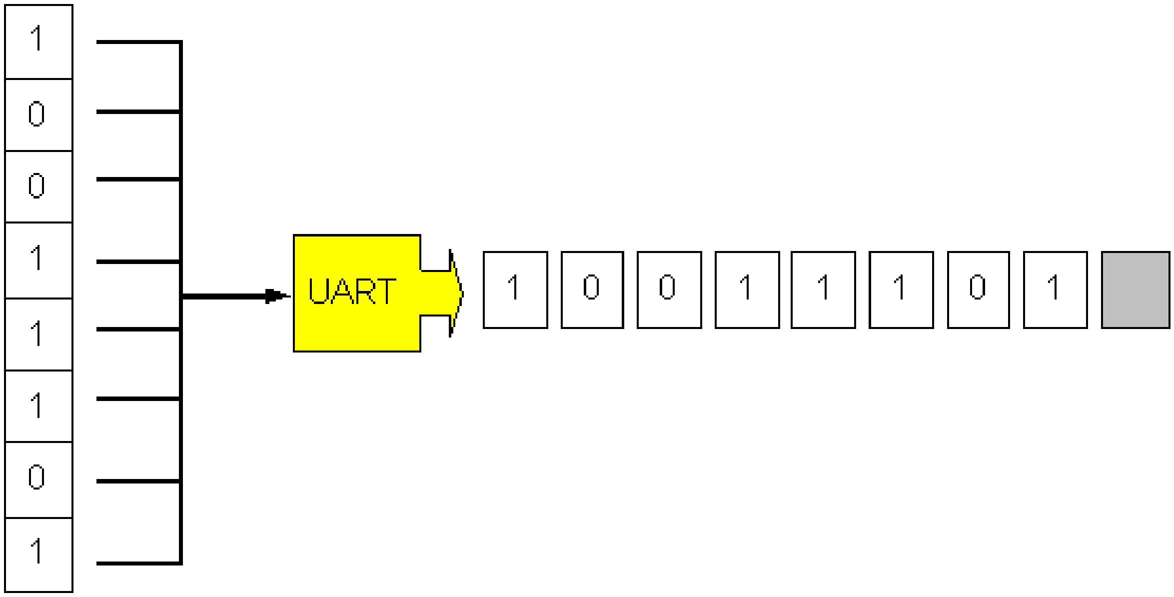
In the synchronous transmission you need two seperate cables. With every clock pulse (i.e. the positive going edge of the clock) one data bit is transferred.

In the asynchronous transmission clock and data are transferred with only one cable.

The clock has to be reconstructed from the mixed signal in the receiver: After a "1" start bit come 8 data bits and then a "0" stop bit (or two "0" stop bits) and so on.

The UART chip

The serial ports are controlled by an UART chip (*Universal Asynchronous Receiver Transmitter)* like 16550 AFN. This chip receives bytes from the system bus and chops them up into bits. The most common package is called 8/N/1 meaning that we send 8 bits, no parity bit and finally one stop bit. This way one byte occupies 9 bits:



The serial transfer is limited to a speed of 115,200 bits per second. The cable can be up to 200 meter long. The serial ports can be used to connect:

- The mouse and digitizers
- Modems
- ISDN adapters
- Printers with serial interface
- Digital cameras
- ....

These units are connected to the serial ports using either DB9 or DB25 plugs.

In modern PCs most of these devices connect to the USB bus instead. This gives a much higher transfer speed.

The parallel port

Parallel transmission means that data are conducted through 8 separate wires - transmitting a full byte in one operation. This way the parallel transmission is speedier than the serial, but the cabling is limited to 5-10 meters. The cable is fat and unhandy, holding up to 25 wires and the transmission is controlled according to the Centronics standard.

Most printer manufactures use a 36-pins Amphenol plug, where the PC's parallel port holds a 25-pinned connector. Hence the special printer cable. To the left you see the 25 pin connector, to the right the 36-pin:

Then parallel port represents the most uncomplicated interface of the PC. It is always used to connect the printer, but with the bi-directional parallel port (EPP/ECP), other devices have found their way to this interface. Today you find:

- ZIP drives
- Portable CD-ROM drives
- SCSI adapters
- Digital cameras
- Scanners all using the parallel port to connect to the system bus.

The EPP/ECP ports

Today we operate with Enhanced Parallel Port/Enhanced Capability Ports. This method for bi-directional (half duplex) parallel communication offers higher rates of data transfer (up to 1 megabyte per second) than the original parallel signaling method. EPP is used for non-printer peripherals, where ECP is for printers and scanners. You find the settings for the printer port in the setup program on the motherboard.

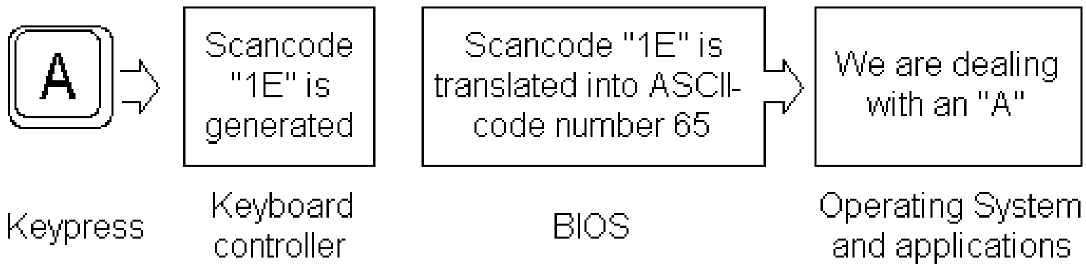
Both port types are parts of the IEEE1284 standard, which also includes Centronics.

To get the best results all the involved hardware and the operating system has to be EPP/ECP compatible. Windows supports IEEE1284 in its parallel plug-and-play feature. It also supports ECP if you have a printer and a parallel port with ECP. The printer cable has to be complete with all 25 wires connected.

The keyboard

Traditionally the keyboard is connected using a DIN or PS/2 mini DIN plug. Soon we shall have USB keyboards but the old ones connect to the internal ISA bus occupying an IRQ.

The keyboard operates with scan codes, which are generated each time a key is pressed and released. The scan codes are translated into ASCII values, which are translated according to the code pages (see module 1a and 1b). Here you see a simple illustration of the system:



This system is quite flexible because it allows for arbitrary remapping of the keyboard codes. This is especially useful if you find the placement of the Caps Lock and Control keys awkward. It's a simple matter to remap them to swap places.

Each key generates a unique scan code. This happens completely independent of the typeface that is printed on the plastic key.

At the other end, the code pages represents a programmable interpretation of the key press; you can assign any type to any key as you want it. Languages like German and French use different keyboard layouts as well as many other languages.

- Next page
- Previous page

Learn more [top]

Read: Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about SCSI, USB etc.

Read A little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

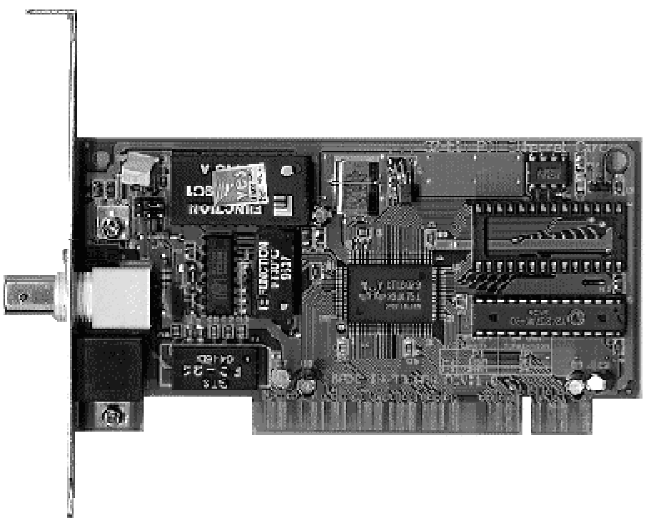
KarbosGuide.com. Module 5a2.

About adapters

The contents:

- Adapters
- The modular PC design

- Next page
- Previous page



Adapters

[top]

In a stationary PC, adapters are typically printed circuit boards called expansion boards or expansion cards. They form a link between the central PC unit and various peripherals. This is the so-called open architecture.

- Typically, adapters provide functions, which are *separated* from the system board.
- Adapters provide *expansion* capability to the PC.

There are PCs without expansion slots. In that case all functions must be built into the system board. You could easily include chips for graphics, ethernet, SCSI, and sound on the system board. This is not common in stationary PCs. Portable, *laptop PCs* have nearly all electronics on the system board. This is called closed architecture.

A traditional PC has a system board which contains all standard functions (except the graphics chip). To this system board you can add various expansion cards, which *control* one or more peripheral units:

The system board	Expansion boards
Standard functions incl. control of keyboard, COM and LPT ports. and four EIDE units.	Video card Network controller Sound card SCSI card 3D graphics controller (for 3D games)

Other expansion board types:

- Internal modem (in lieu of external modem)
- ISDN adapters
- Extra parallel ports
- Video editing boards
- Special graphics cards, which supplement the usual (3D and MPEG)
- TV and radio receivers.

The integrated hard disk controller

[top]

In the Pentium based PC, the hard disk is connected to an EIDE controller, which is integrated on the system board. Likewise, the serial and parallel ports are connected directly to the system board. This is new. On the 386 PCs, you had to install special controller cards (I/O cards) to handle these functions. They are included in the modern chip sets on the system board. Other functions are not integrated. That includes:

The video controller

[top]

You have to install a video card to make the PC function. It would be illogical to assemble a PC without a video card. You would not be able to see what you are doing, since the video card governs data transmission to the monitor.

The advantage of this design is, that the user can choose between numerous video cards in various qualities. A discount store may offer a *complete Pentium based PC* (without printer) and with the cheapest video card for $669.-. If the buyer is quality oriented, he would want to spend an additional $40 to get a much better video card.

The modular PC design

[top]

In this way, various expansion boards provide flexibility in assembling a customized PC. At the same time, various electronics manufacturers are specializing their production:

ASUS and Tyan are good at making system boards. Others, like S3, Matrox, and ATI specialize in making graphics chips and expansion boards. Olicom make only net boards. Adaptec make only SCSI controllers and Creative Labs make SoundBlaster sound boards.

This variety of manufacturers offers the consumer wide choices. Your PC can be customized and configured according to your needs and wallet size.

About the electronics

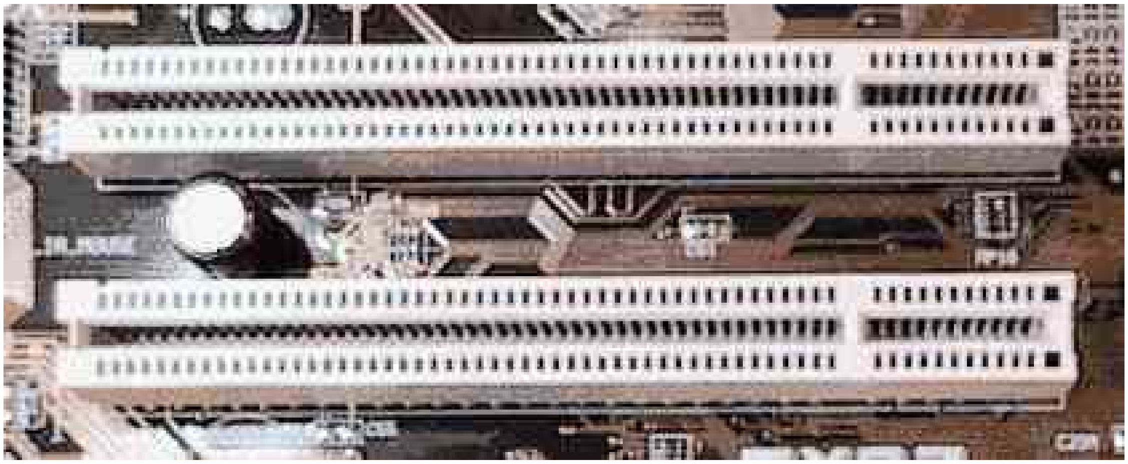
[top]

The adapter is a printed circuit board. They have an *edge connector*, so they can be inserted in expansion slots in the system board. The expansion slots connect to the I/O buses. Since the Pentium system board has two I/O buses, it has two types of expansion slots:

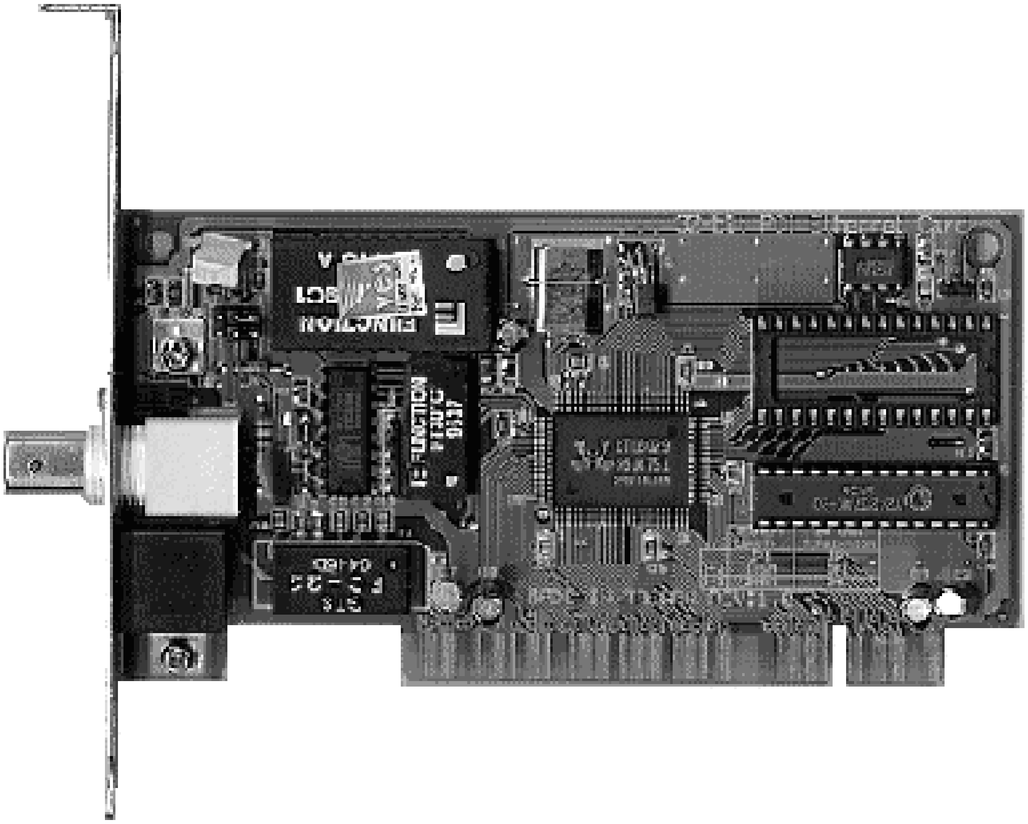
- ISA slots
- PCI slots

Typically, on a regular Pentium system board there are three or four of each type. That gives a total of 7 expansion slots. One expansion board can be installed in each of these. You simply press the edge connector of the expansion board into the expansion slot. Now it is connected to the bus.

Here you see two PCI slots open for video cards, network controllers and others:



Below, you see a *network adapter.* It is an ethernet card with PCI interface, so it fits in a PCI slot in the Pentium. This inexpensive board allows your computer to join a network with other net board equipped PCs. Please compare the edge connector at the button of the card with the sockets above. They fit together!



- Next page
- Previous page

Learn more **[top]**

Read: Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about USB.

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5a3.

About IRQs, DMA, bus mastering etc.

The contents:

- IRQ
- DMA
- Bus mastering

- Next page
- Previous page

IRQs

[top]

When you install an expansion board in a slot, it gets connected to the I/O bus. Now the board can send and receive data. But who regulates the traffic? Who gives clearance to the new controller to send data? It would appear that data traffic could soon be chaotic.

To control data traffic on the I/O bus, the concept of IRQ (*Interrupt ReQuest)* was created. Interrupts are a fundamental principle in the PC design. There are two types of interrupts: Software Interrupts are used to call any number of BIOS routines. Hardware Interrupts are the subject of this page.

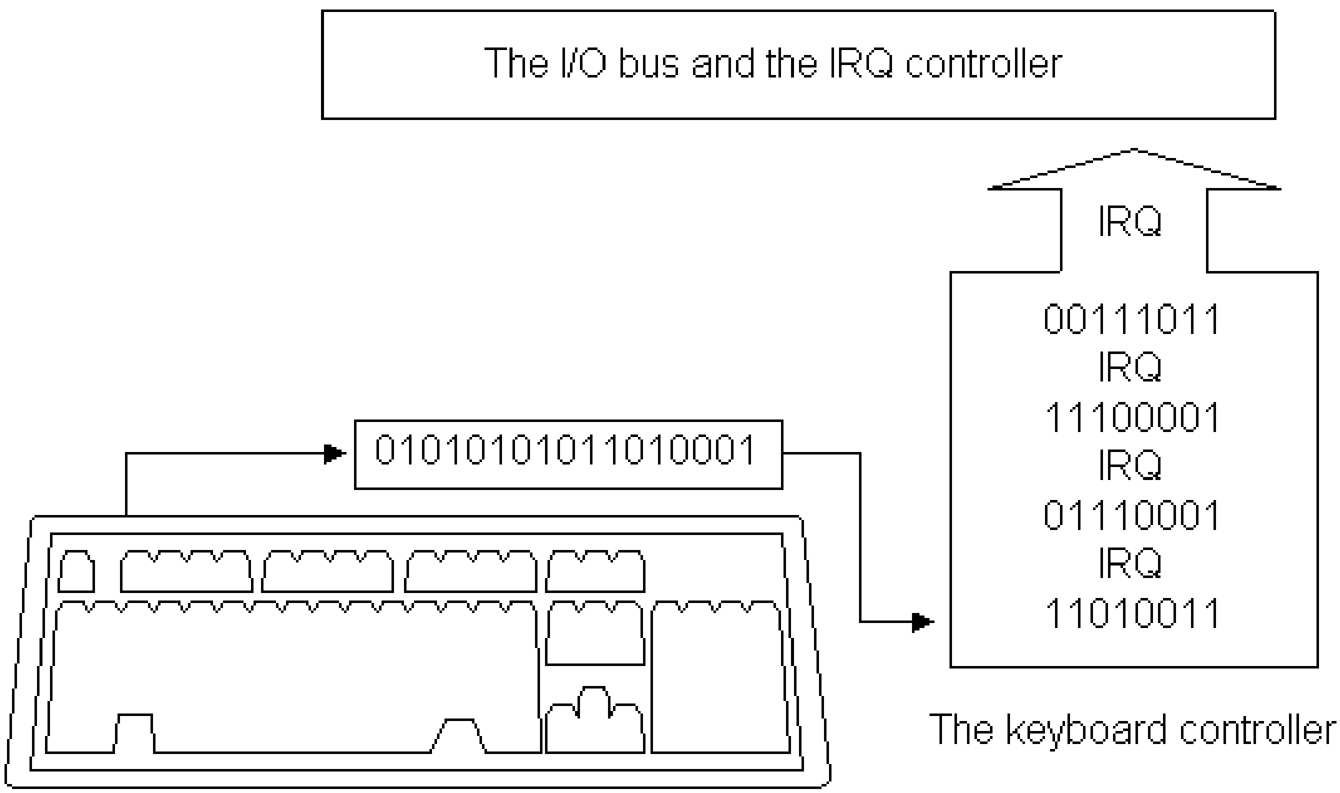
Hardware Interrupts

[top]

The adapter or unit on the I/O bus uses the interrupt to signal request to send or receive data. An interrupt signal is like a door bell. The unit signals by applying a voltage to one of the wires in the bus - an IRQ. When the CPU acknowledges the signal, it knows that the unit wants send or receive data, or is finished.

The advantage of IRQs is that the CPU can manage other tasks, while an adapter "massages" its data. When the adapter has finished its task, it will report to the CPU with a new IRQ.

As an example, let us see how keyboard data are handled. The keyboard send bits, serially, through the cable to the keyboard controller. The controller organizes them in groups of 8 (one byte). Every time it has a byte, it sends an IRQ to the I/O bus. The IRQ controller asks the CPU permission to use the bus, to send the byte to wherever. The IRQ controller reports back to the keyboard controller, giving clearance to send the next character (byte):



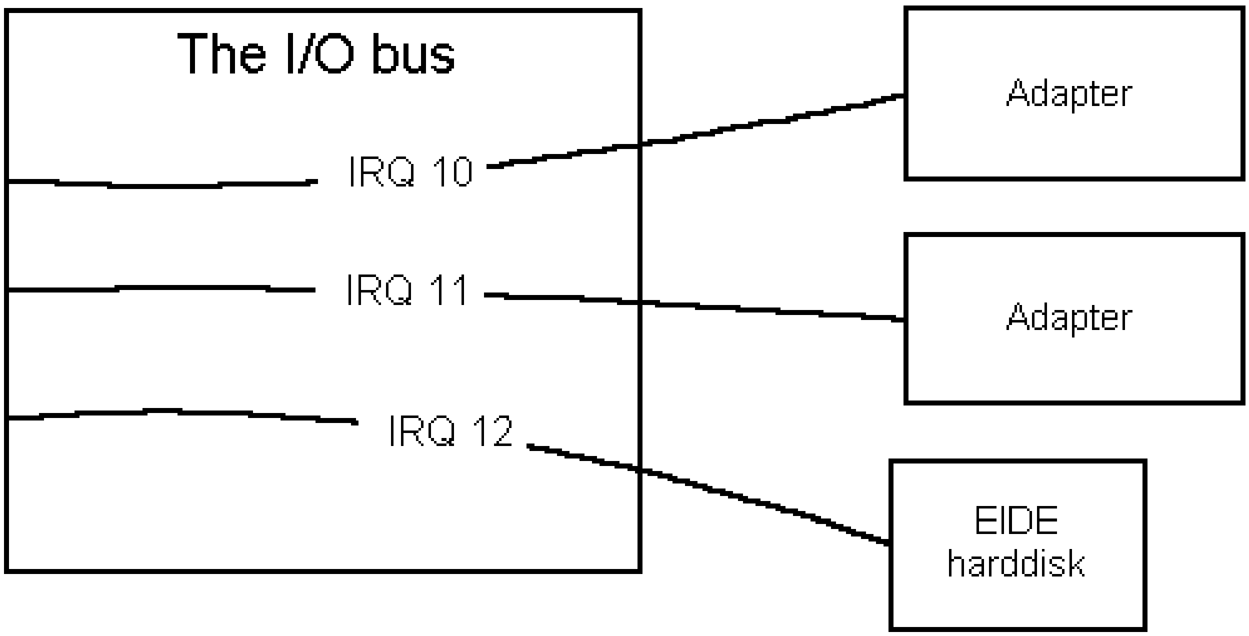
IRQ wires

[top]

Physically, the IRQ is a wire on the bus. This wire connects to all expansion slots. Therefore, regardless of in which slot you install an adapter, the adapter can communicate with an IRQ.

In the pristine PC design (the original PC/XT bus) you found 8 IRQ's. The more recent PC is "born" with 16 IRQs, but five of them are internal, and cannot be used with I/O cards, and one of them connects the lower IRQ's with the higher (IRQ2/9).

We find 10 accessible IRQs on the I/O buses. Each of those consist of a circuit board wire, which goes through the entire bus. When you install an expansion card in a vacant slot, one of the IRQs is assigned to it.



When a signal arrives on an IRQ channel, that is a message to the CPU. It is told that a unit wants to get on the bus. *Which* unit is to be identified through the *IRQ number*.

Next the unit is admitted to the bus, to send or receive data. When the transaction is completed, another signal is transmitted to the CPU to indicate that the bus is vacant.

The IRQs have different *priorities*, so the CPU knows which IRQ have priority, if two signals are sent simultaneously.

The IRQ system is guided by a controller chip, like Intel 8259. It can handle 8 IRQ signals and couple two of them together, via IRQ 2 or 9. All PCs with ISA bus include two 8259 chips.

MSD (Microsoft Diagnose System)

[top]

Let me show an image of the MSD diagnostic program, which you can run in Windows 95/98. It shows the use of IRQs on a PC:

IRQ	Address	Description	Detected	Handled By
0	CC00:0000	Timer Click	Yes	Unknown
1	0A2C:08D2	Keyboard	Yes	KEYB
2	F000:FF47	Second 8259A	Yes	BIOS
3	F000:FF47	COM2: COM4:	COM2:	BIOS
4	F000:FF47	COM1: COM3:	COM1:	BIOS
5	F000:FF47	LPT2:	No	BIOS
6	0956:009A	Floppy Disk	Yes	Default Handlers
7	0070:0465	LPT1:	Yes	System Area
8	0956:0035	Real-Time Clock	Yes	Default Handlers
9	F000:F7C9	Redirected IRQ2	Yes	BIOS
10	F000:FF47	(Reserved)		BIOS
11	F000:FF47	(Reserved)		BIOS
12	0956:00E2	(Reserved)	Not Detected	Default Handlers
13	F000:2930	Math Coprocessor	Yes	BIOS
14	0956:00FA	Fixed Disk	Yes	Default Handlers
15	F000:FF47	(Reserved)		BIOS

MSD shows the IRQs of the PC, where the program is run. There are a total of 15 IRQ channels and each IRQ is assigned to a unit. However, it is not always possible to utilize IRQ 9. It functions like a bridge between two parts in the IRQ system.

In the above illustration, IRQ numbers 5, 10, 11, 12, and 15 appear vacant.

IRQ numbers 2 and 9 show the linking between those two IRQ controllers.

Some IRQs are reserved for various internal units, which must also be able to disconnect the CPU. Those are IRQ numbers 0, 1, 2, 8, and 13, as you can see in the illustration above. They are not available for other units. In principle, the remainder are available for expansion boards and EIDE units.

IRQs are assigned during the PC start-up. An ISA expansion board is assigned a given IRQ during start-up. That IRQ is used every time that expansion board uses the bus.

Shared IRQs

The modern I/O buses MCA, EISA and PCI permit shared IRQs. Thus, two adapters can share one IRQ. When the IRQ is activated, the drive programs for the two adapters are checked, to identify which is on the bus.

IRQ and conflicts on the ISA bus

[top]

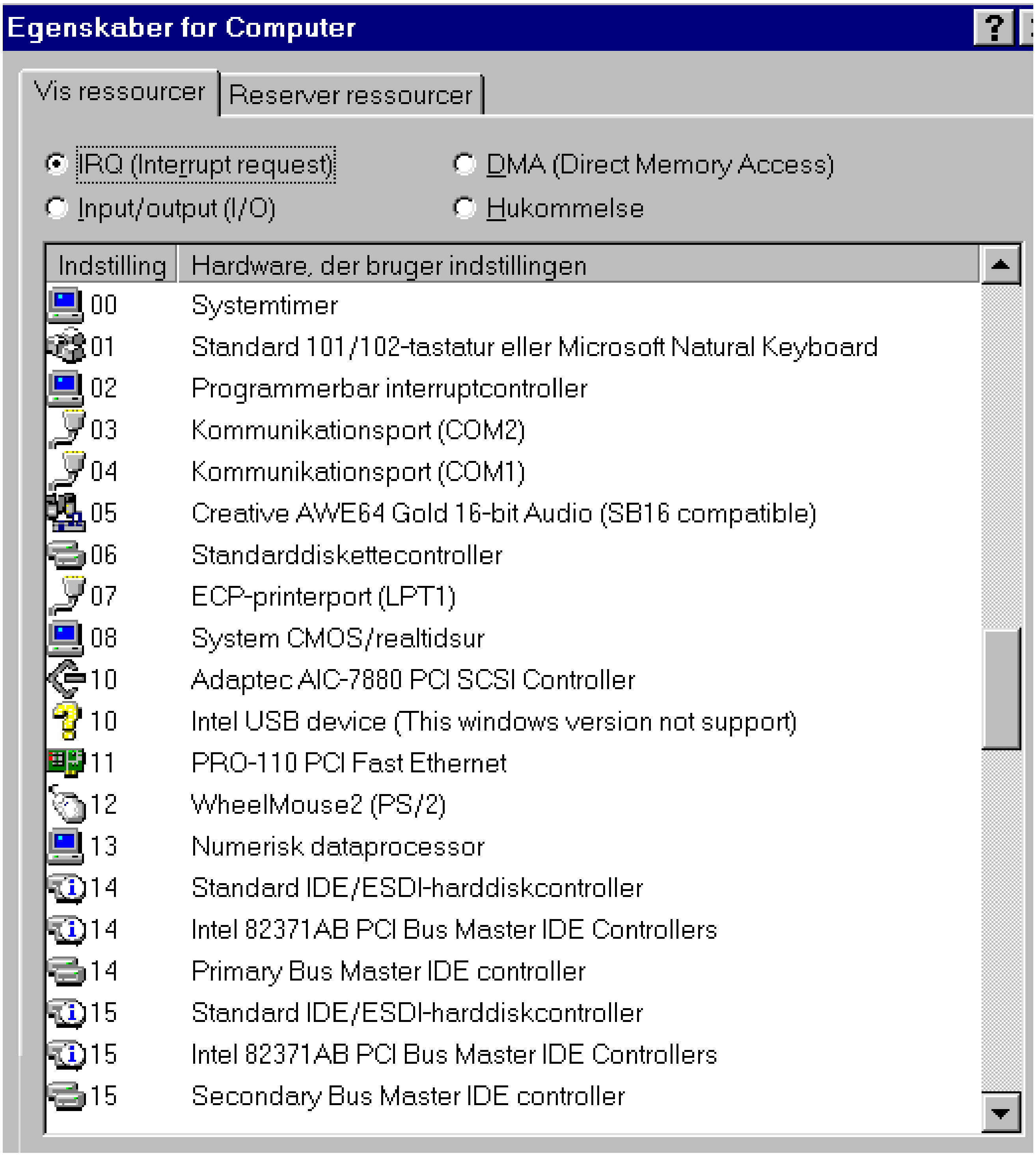
The IRQ system can cause some problems on the unintelligent ISA bus. When bus and adapters are referred to as unintelligent, it implies that they are unable to organize the IRQ distribution on their own.

In order to function, an ISA network controller must be assigned an IRQ. The manufacturer could preset it to work with IRQ 9, 10, 11, or 12. One of these values, let us say IRQ 11, is preset as the *default* value. When the customer installs that board, during start-up it will try to access the bus as IRQ 11. If no other units are connected to IRQ 11, it should work. If IRQ 11 is occupied, we have a problem. Those two units would get in a *conflict*. Often, the PC will not start at all and panic erupts.

The solution is to change the IRQ of the adapter. The manufacturer has designed the board to work on IRQ 9, 10, 11, or 12. Number 11 was the default. If that does not work, you must adjust to another. This can be done with the accompanying software, or by resetting a little *jumper* - an electric contact on the board, which has to be reset. The manual for the board will include instructions about how to do this.

These IRQ problems can be a terrible nuisance. If both sound and net boards had to be installed in ISA slots in the same PC, sometimes I had to give up.

In Windows 95 (System, Computer, Properties) you can find an excellent overview of the IRQs. Here it is from my Danish version:

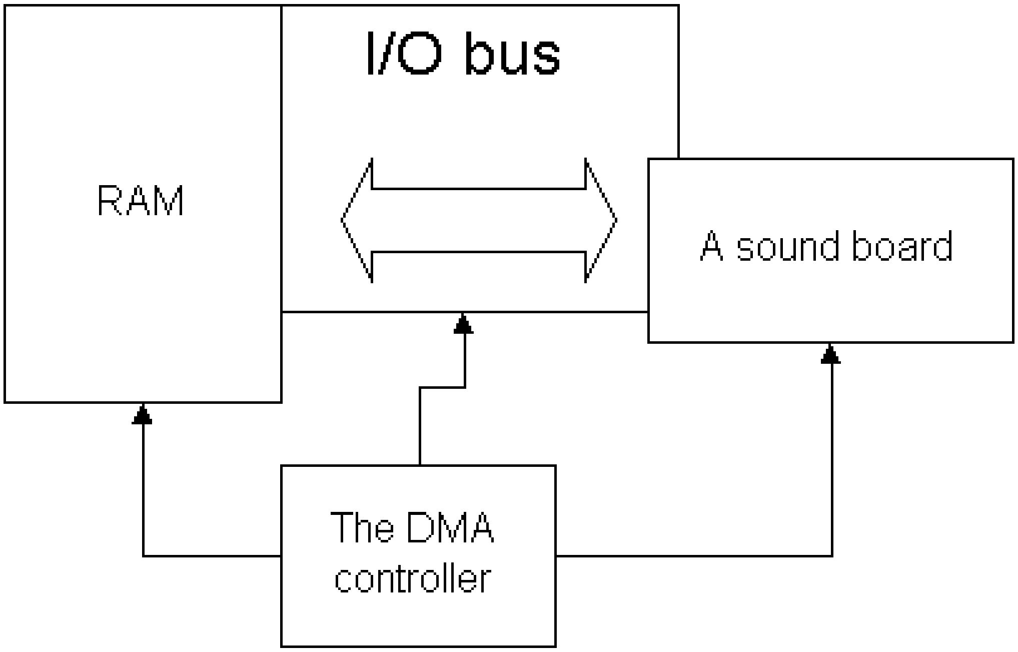


DMA

[top]

IRQs are only one of the problems with ISA boards. The other one is DMA (Direct Memory Access). That is a system which allows an adapter to transfer data to RAM without CPU involvement.

Normally, the CPU controls all bus activities. With DMA, this "intelligence" is assigned to a DMA controller on the system board. This special controller chip (Intel 8237) has clearance to move data to and from RAM, via the I/O bus, without burdening the CPU.



You can implement a number of DMA channels, which can be used by the ISA boards. Each channel has its own number and one controller can be in charge of four channels. Each ISA unit can occupy one of these channels, if so designed. Diskette drives utilize DMA.

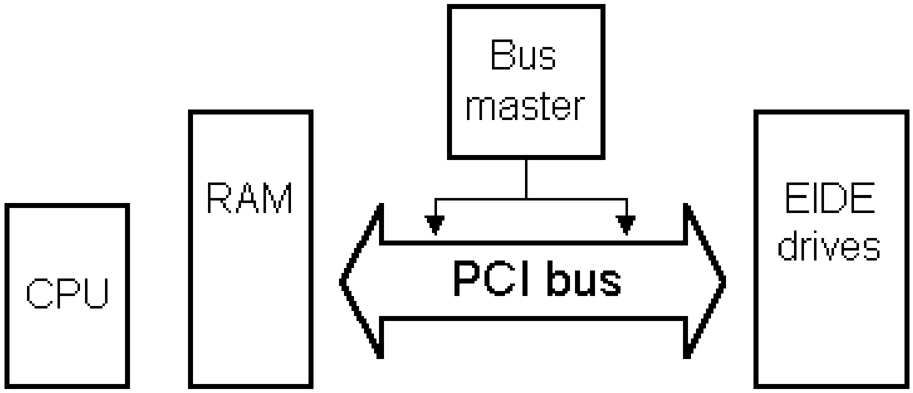
The DMA system can result in conflicts between two units on the bus, if they have requested the same DMA channel. As an example, on ISA sound boards you have to reset both IRQ and DMA number.

It is important to enable DMA for your harddisk in Windows . See how in this Windows tip.

Bus mastering

[top]

There are no DMA channels on the PCI bus. Instead *bus mastering* is employed. It is a similar system, where special controller functions allow adapters to control the bus. Thus, they can deliver their data directly to RAM, minimizing the workload on the CPU. It does not need to keep track of the transactions, the bus master takes care of that.



This allows the PC to multitask, handle more than one task at a time. The hard disk can pour streams of data to RAM, while the CPU handles some other task. The bus mastering system works *fairly well* with EIDE hard disks. However in this particular area, the SCSI controller is far more advanced. EIDE bus mastering is rather new and we will see further developments in this area.

Bus mastering version	Chip set	Year
DMA mode 2	82430FX	1995
Ultra DMA	82430TX	1997

The latest version Ultra DMA/66 is described in module 5b.

- Next page
- Previous page

Learn more **[top]**

Read: Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about SCSI, USB etc.

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5a4.

About I/O port, Plug and Play, and PC Cards.

The contents:

- I/O addressing
- Plug and Play (PnP)
- The physical connector on adapters

- Next page
- Previous page

I/O addresses

[top]

Finally, we need to mention how the CPU finds all these units - adapters, ports. etc. They all have an address - an I/O port number.

Each unit can be reached through one of many I/O ports. Each port is a byte port. That means that 8 bits (one byte) can be transmitted simultaneously - parallel mode.

If the unit is on the ISA bus, it handles 16 bits at a time (*words*). Then you link two consecutive ports together, to make a 16 bit channel. If we talk about a 32 bit PCI unit, we link four byte ports together to get 32 bits width (called *dword).*

The PC has a built in listing of all I/O units, each of which has their own "zip code" - a port address. Since the PC is basically a 16 bit computer, there are 2 at the 16th power possible addresses (65,536) - from 0000 to FFFFH. They are described in the hexadecimal number system as 5 digit numbers. Hexadecimal is a 16 digit number system. Digits go from 0 to 9 and continue with 6 letters A - F. Let me show you some examples of I/O addresses:

Unit	I/O ports
CMOS RAM	0070H
Keyboard	0060H ... 0063H
Serial port 1 (COM 1)	03F8H ... 03FFH
Parallel port 1 (LPT1)	0378H ... 037FH

Fortunately, you do not have to adjust port addresses too often. Some adapters give room to adjust to user option I/O addresses, but you have to have bad luck to encounter any conflict in this area.

Plug and Play

[top]

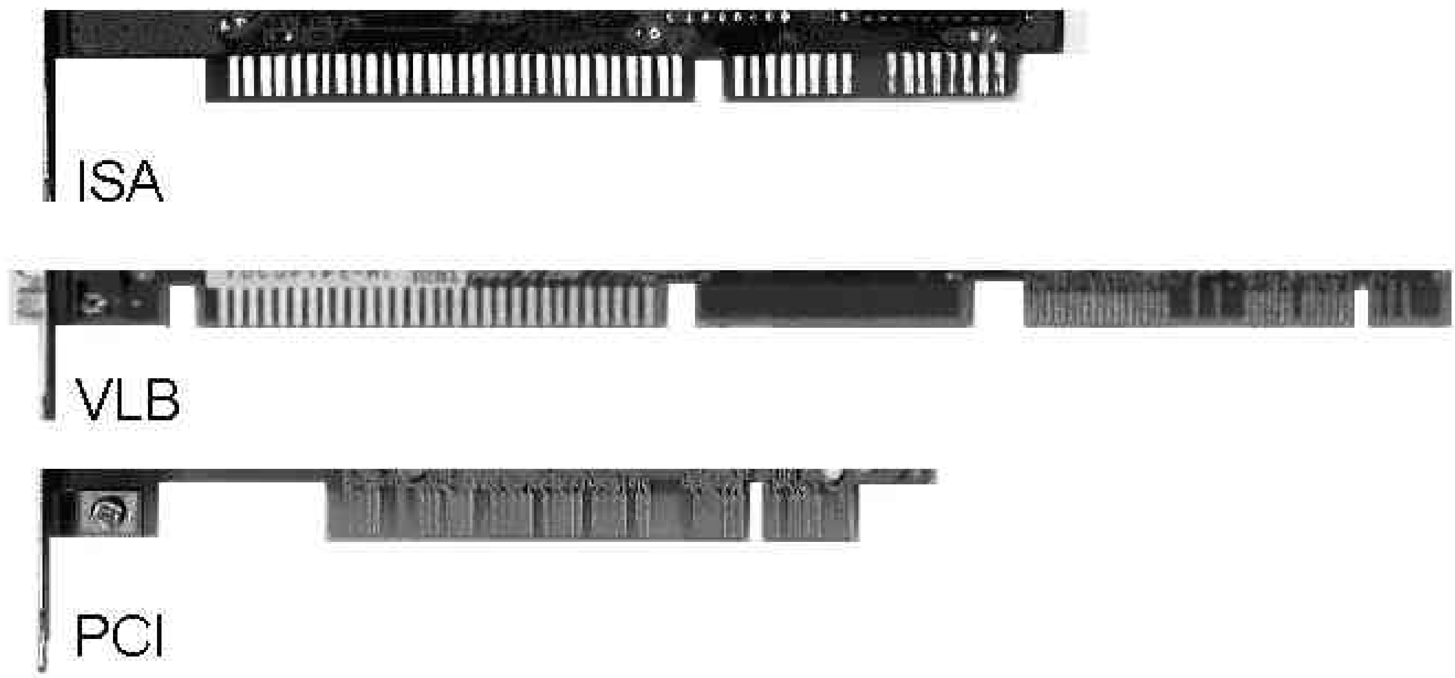
Plug and play (PnP) is an industry standard for expansion boards. If the board conforms to the PnP standard, the installation is very simple. The board configures itself automatically. These are the minimum requirements:

- The PC system board must be PnP compatible.
- The operating system must be capable of utilizing PnP, as Windows is.
- The adapter must be able to inform the I/O bus which I/O addresses and IRQs it can communicate with.
- The adapter must be able to adjust to use the I/O address and the IRQ, which the I/O bus communicates to the adapter.

The physical connector

[top]

The different I/O cards each fit with a particular I/O bus. The different buses each have their own system board slot configuration. That is a socket in the system board, in which you press in the expansion board. Here you see three different edge connectors fitting into each their own type of socket. The ISA bus has a total of 98 prongs (31+18 on each side).



- Next page
- Previous page

Learn more **[top]**

Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about SCSI, USB etc.

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5a5

About PC Cards.

The contents:

- PC Card
- PC Card II

- Next page
- Previous page

The PC Card

In portable PCs, the adapter is usually a PC Card. This is a little tiny box which fits into a special slot. The PC Card used to be called a PCMCIA card, but this obviously was a little difficult to remember.

The first generation of PC Cards were technically connected to the ISA bus. The newer ones are PC Card32 working with the CardBus. They are internally connected to the 32 bit PCI bus.

Here you see a network controller, as a PC Card. It is about the size of a credit card, but slightly thicker:

The PC Card is placed in a special socket, where it can be inserted or removed, while the PC is operating. Actually, each socket acts like an I/O unit, regardless of whether there is a PC Card in it or not. When the card is inserted, it is automatically configured with I/O address, IRQ, etc. Windows 98 provides by far the best support for PC Cards.

I use two PC Cards myself: The network controller you see above connects my laptop to my network. And my digital camera (Canon Powershot 600) uses a PC Card with 4 MB Flash RAM. Having taken the photos I just move the PC Card from the camera to the laptop. Here it instantly becomes a D-drive and I use Explorer to move the photos to a folder on the server. The operation takes less than a minute.

PC Card II

The latest standard (PC Card II) makes it possible to use harddisks within the PC card. The IBM Microdrive is very handy when using digital cameras. All modern cameras of good quality can hold such a harddisk of 340 MB. This gives room for more than 300 pictures in a very high quality.

The PC card, holding Compact Flash Card, SSFDC (SmartMedia) or a MicroDrive, can be read by a USB-connected reader like these:

If you use PC Card, you should invest in a reader for USB. It is very handy and speedy.

CardX

The new high speed serial ports USB and FireWire will also become available for portable PC users. CardX is a new PC Card interface built on FireWire and allowing transfers of up to 400 megabit per second.

Two function adapters

[top]

Integrated adapters with more than one function are space savers. Especially, the ASUS company has introduced dual function boards to stationary PCs, since they utilize both the ISA and PCI bus to share a slot:

- Graphics + sound
- SCSI + sound There are also two function PC Cards for portable PCs:

- Ethernet network controller + modem
- Token Ring network controller + modem

- Next page
- Previous page

Learn more **[top]**

Module 5b about EIDE, Ultra DMA and AGP.

Read Module 5c about SCSI, USB etc.

Read Module 6b with a little about Windows 95/98.

Read Module 6c about the relationship between BIOS, OS and hardware

Read Module 7a about the videosystem

Read about video cards in Module 7b.

Read about digital sound in Module 7c.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b1.

About the interfaces EIDE, Ultra DMA and AGP

The contents:

- What is EIDE?
- Four units controlled by the motherboard

- Next page
- Previous page

On the following pages:

- The EIDE cable
- The Promise FastTrack EIDE controller
- Transfer speeds and protocols
- What does Ultra DMA offer?
- Looking at a good harddisk
- ATA/66
- Configuring your EIDE hard disk
- What is AGP?

What is EIDE?

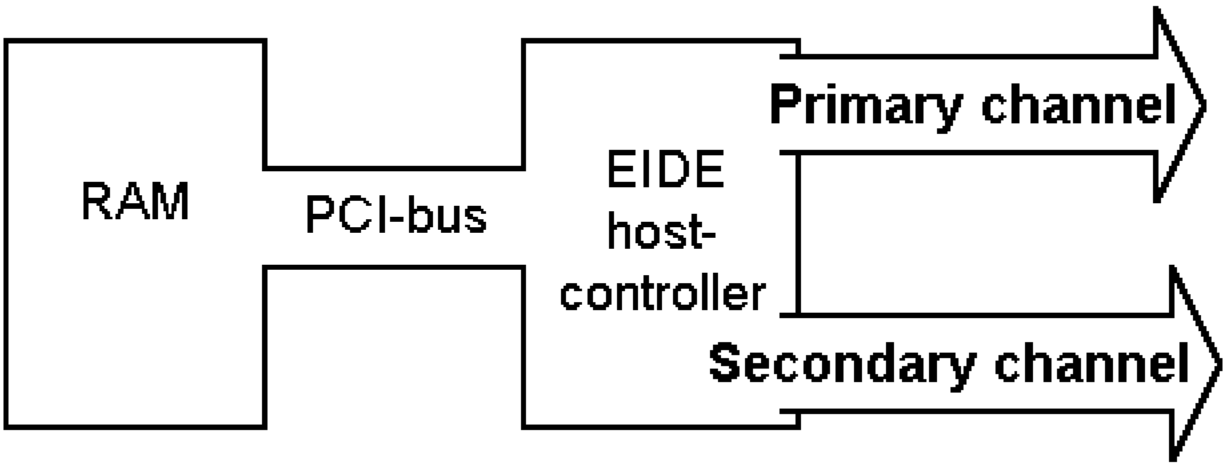
[top]

EIDE is the current standard for inexpensive, high performance hard disks used in PCs.

EIDE stands for Enhanced IDE and it is registered name own by harddisk manufacture Western Digital. They also own the name "IDE".

Other companies like Seagate, IBM, Quantum and Maxtor Uses the term ATA, which stands for Advanced Technology Attachment. But it is all the same. However there are many differant protocols behind the terms.

You can think of EIDE as a *bus* - which is a *host controller* - which controls it, and you can connect up to four units. Here you see the controller and its two channels:



All Pentium system boards since 1995 have this EIDE controller built into the chip set. That allows the hard disk and other EIDE units to be connected directly to the system board.

Improvements

The EIDE standard is a great improvement over the old IDE. Here are some examples:

- The hard disk can exceed the 528 MB IDE limit. Currently the largest EIDE disks are of 37 GB and this number keeps increasing. IBM has promised harddisks of more than 100 GB before year 2001.

- The hard disk's interface is moved from the ISA bus to the high speed PCI bus.

- Four units can be connected to the system board, which has two EIDE channels. Each channel can be connected to a *master* and a *slave* unit.

The most important feature is the interface directly on the PCI bus. This has given EIDE transfer speeds and disk capacities, which far exceed older controller principles. Concurrently, there is a continual development of the *protocols,* which are needed for the connection between the units and the EIDE bus.

Four units controlled by the motherboard

The EIDE interface is not designed for hard disks only. There are four channels, which can be connected to four independent units:

- Hard disks
- CD-ROM drives
- CR-RW drives
- DVD drives
- LS 120, Zip or HiFD drive
- Tape streamers

EIDE is thus designed as an inexpensive all-round interface, which can be connected to all kinds of storage media.

Auto detect

The BIOS on the system board has a neat auto detect feature, which often allows EIDE units to be connected directly and work immediately. The PC start up program automatically finds the necessary information about the drive via the auto detect function.

Sometimes you have to assist the hard disk installation by activating the auto detect in the CMOS Setup program, but often it runs by itself. You definitely do not have to key in information about cylinders, etc., as you had to with earlier IDE units.

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

Click for Module 7d about digital music, MP3s

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b2a.

About the EIDE interface - continued

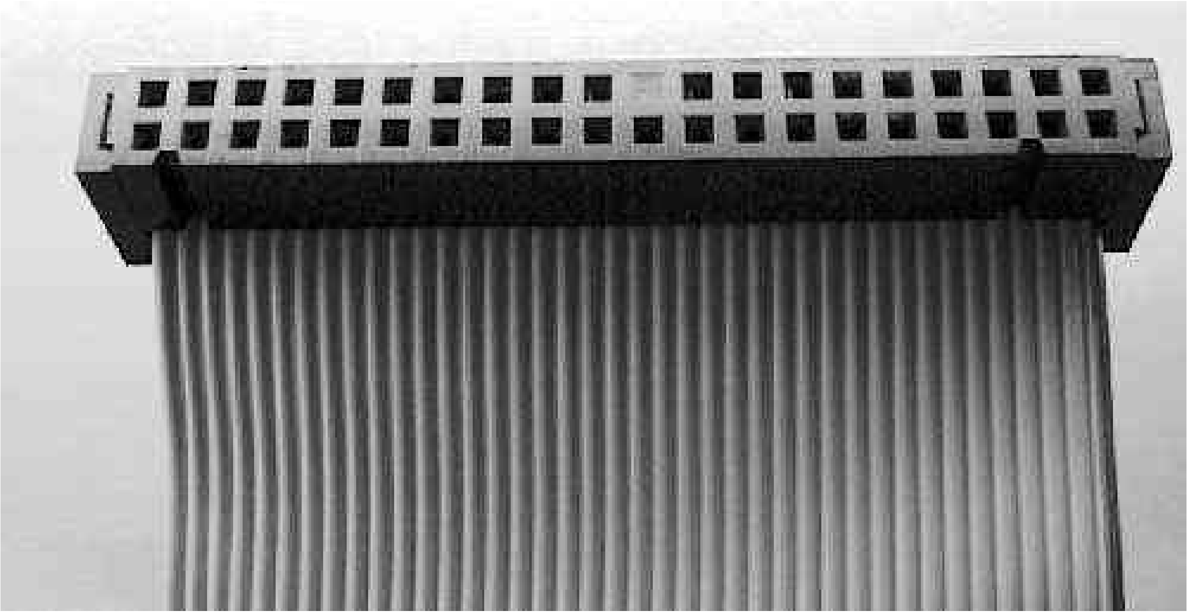
The contents:

- The EIDE cable
- Problems with assigning two EIDE harddisks
- The Promise FastTrack EIDE controller

- Next page
- Previous page

The EIDE cable

Here you see a typical EIDE cable:

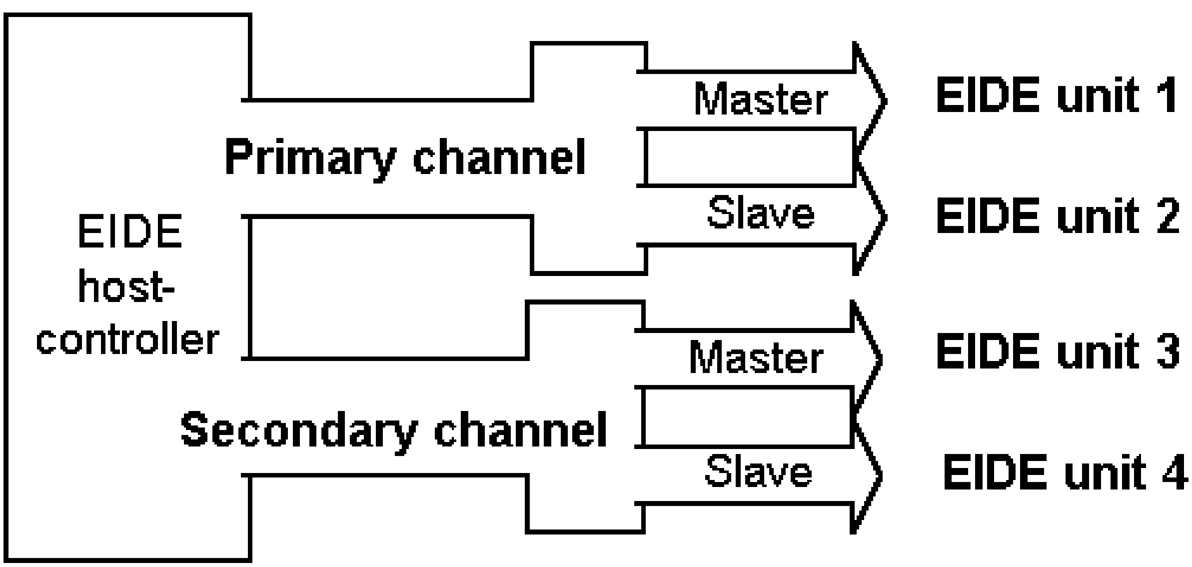


These cables are diffficult to handle, if you want to make a nice PC. The are so big, that they disturb the air circulation and hence the cooling inside the PC cabinet.

Since each channel can handle two units, there are two of these connectors on the system board. Note the blind hole in top center. Note also the stripe (it is red) in the far right edge of the cable. It tells you that lead number one is on this edge. Both of these features help prevent incorrect installation of the cable.

Using primary and secondary channels

The system board has sockets for two EIDE cables. Each EIDE cable (primary and secondary) has sockets for two units (master and slave).



Settting up four units

If we have to use all four connections, it causes some problems. The setup may look like this:

Primary, master	Hard disk 1
Primary, slave	Hard disk 2 or CD-ROM
Secondary, master	CD-RW DVD drive
Secondary, slave	ZIP/LS120 diskette drive

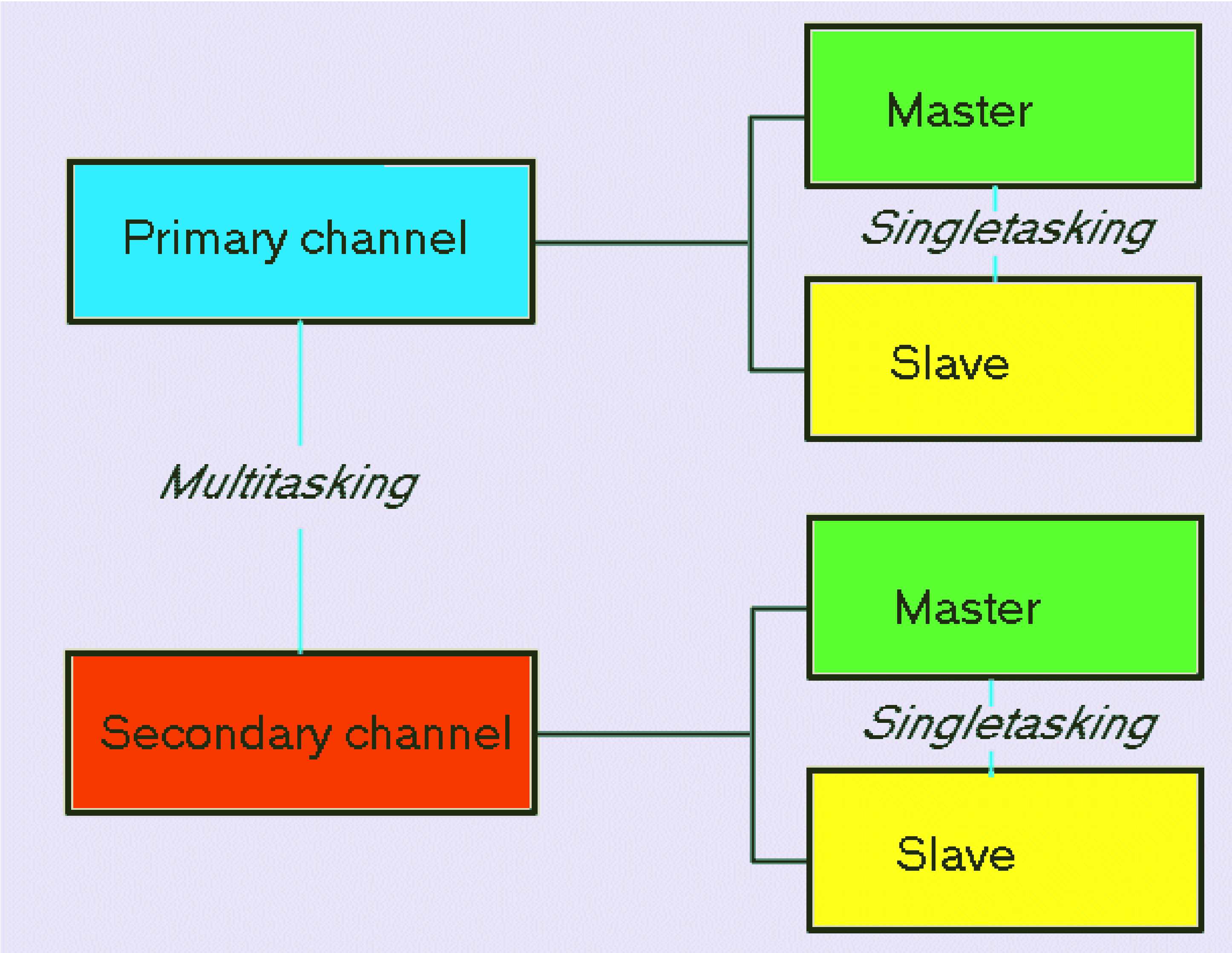
Typically, a PC has two EIDE units connected: the hard disk and the CDROM drive. However, as you can see, other units can be connected as well. The hard disk must must be on the primary EIDE channel. On some system boards, this has the greatest transfer capacity.

Using two EIDE harddisk

In the table above I indicate that it is not good to assign two harddisks to the primary channel. They should be assigned to the masters of the primary and the secondary channel.

This is due to the fact, that the two main controllers (primary and secondary) are capable of multitasking. The two channels can process data simultaneously and independently, at the same time.

The two sub-channels (slave and master) do not multitask; here only one operation is processed at the time, be it on the master *or* on the slave channel.



So for best performance, the two harddisk have to be assigned one at the primary EIDE channel and the other at the secondary EIDE channel. This leaves us with the problem of the CD-ROM drives, which also have to be placed on EIDE channels.

My conclusion is that if you have to use two EIDE harddisks (and many of us do), the motherboard should be enhanced with further EIDE channels. Please continue the reading...

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b2b.

RAID/EIDE controllers

The contents:

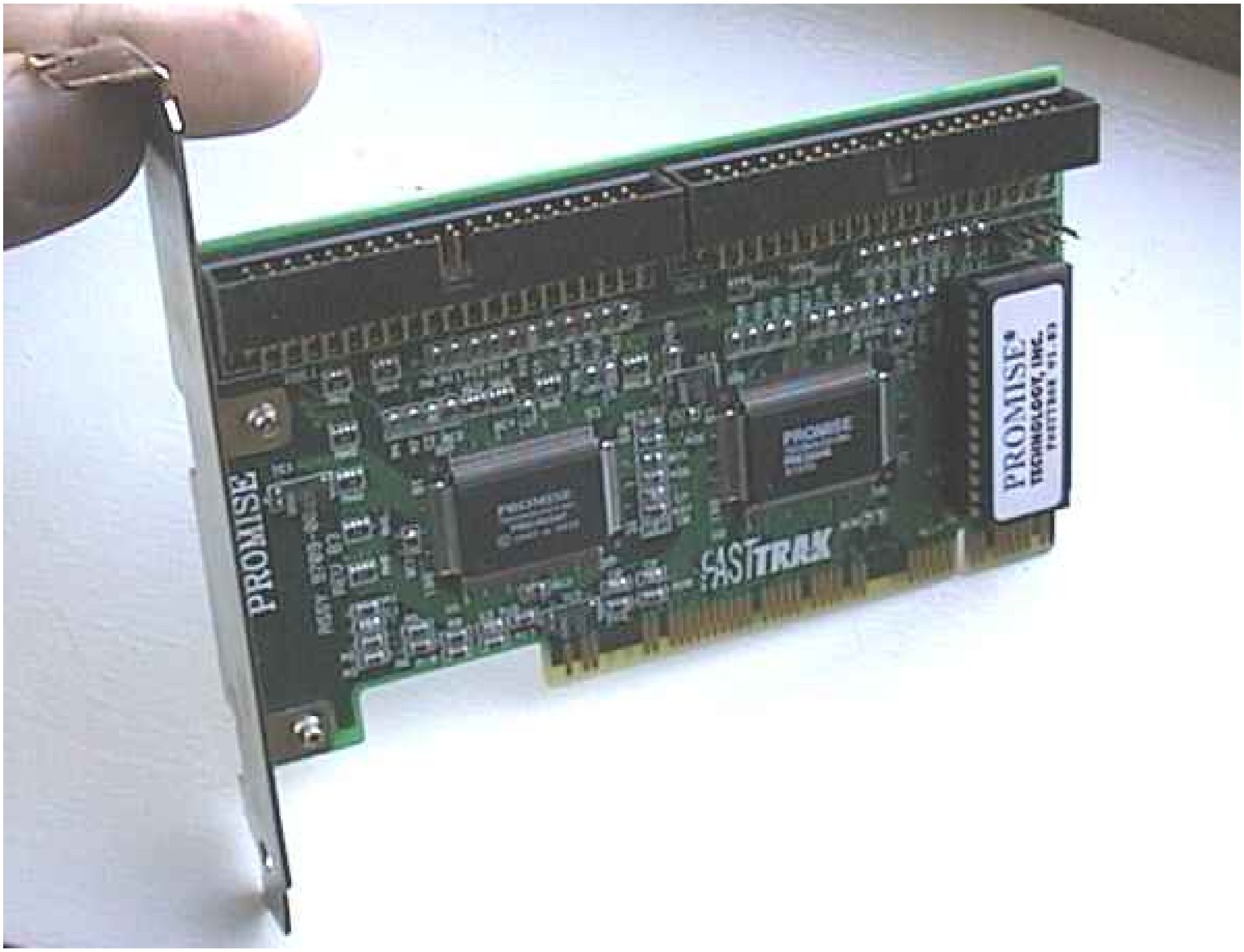
- The Promise FastTrack EIDE controller
- Adaptec UDMA RAID
- Highpoint UDMA RAID onboard

- Next page
- Previous page

The Promise FastTrack EIDE controller

You may expand your system with a little PCI-based EIDE-controller from Promise.

This controller can be connected to two or four EIDE hard disks. And it works side by side with the existing EIDE-controller on the motherboard! I have tested it, and it works fine:



The card is very powerfull. You can use it in multiple setups:

- Connecting four EIDE hard disk to your PC; the onboard EIDE-controller can be used for CD-ROMs, ZIP or other ATA-based drives.
- Increased performance (2X or 4X) by using the RAID 0 functionality.
- Increased capacity since two or four disks can be assigned to one big virtual drive.
- Increased security using RAID 1 mirroring.

The most impressive thing is that the controller holds its own BIOS. It works complete independent of the hosting PC.

Striping

The FastTrak controller can double or quadruple the transfer speed if you want it.

Using RAID 0 striping, the data are striped over more than one drive. You use two or four identical drives, which all are unified into one big volume.

When you write data to the drive, it is simultaneously written to all two or four drives in the array. This is the principle of striping: interleaving during the read/write process.

The stripe size (blocks to be striped) can be set from 1 KB to 1024 KB. The default value for business application use is stripe size 8 KB. For working with bigger files (for instance sound and video editing) the recommendation is 64KB.

Mirroring

Using RAID 1 mirroring, the data are duplicated to several drives. This gives an increased security. The system is fault tolerant; if one drive is damaged, the data can be read from another.

My own use

I was very glad to use the FastTrack PCI RAID controller. I used it just to improve the capacity of the EIDE system.

The PC looked like this regarding EIDE units:

- Motherboard primary EIDE, master: A 28 GB Seagate hard disk, 7200 RPM.
- Motherboard primary EIDE, slave: Nothing.
- Motherboard secondary EIDE, master: A 40X CD-ROM drive.
- Motherboard secondary EIDE, slave: A HP 8100plus CD-RW drive.
- FastTrack primary EIDE, master: A 30 GB Maxtor hard disk.

The two hard disks (each formatted into just one big partition) cooperated very well. I could copy between them at 10 MB/sec.

With the system described, I had room for expansion, since I could add another disk to the FastTrack secondary master. I could also have opted for stiping, if I'd wanted to - it only requires two identical disks. I have tried it, and it works fine.

The FastTrack adapter can stripe disks for better performance, but as described here, it is also great just for adding versatility to your PC system!

Adaptec UDMA RAID

The name Adaptec has for many years been synonymous with SCSI and SCSI-based RAID.

The AAA-UDMA RAID controller supports up to 4 ATA/66 hard disks in RAID 0,1, 0/1 and RAID 5 (fault tolerance) arrays. The RAID 5 array is interesting for use in servers, giving the option of hot-plugging a substitute disk in case of diskfaults. All data is reconstructed on the fly.

This is the first time that this is possible using EIDE disks, and it shows how the quality of the EIDE/ATA standards is improving nowadays.

The maximum output of the RAID controllers is 66 MB per second due to the limitation of the ATA/66 protocol. With ATA/100 and later Serial ATA we shall see much more powerful systems!

RAID controllers on the motherboard

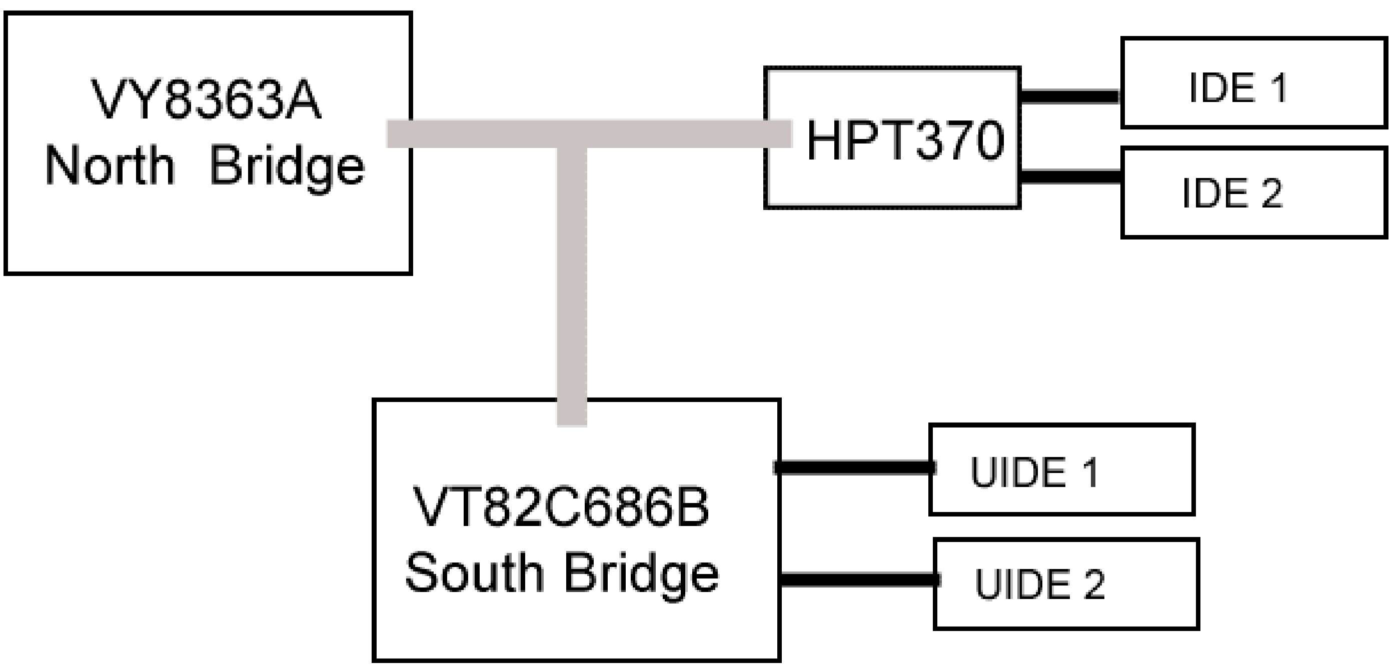
In 2001 we saw several motherboards including ATA/100-based RAID controllers onboard.

I bought an Epox 8KTA3+ board for my 900 MHz AMD processor. Overall, it is a very nice board, but the best feature is the ATA/100-based RAID controller which is integrated.

The RAID controlling logic comes from HighPoint; their HPT370 chip is located on the motherboard:

Here it cooperates with the traditional "south bridge" from VIA:

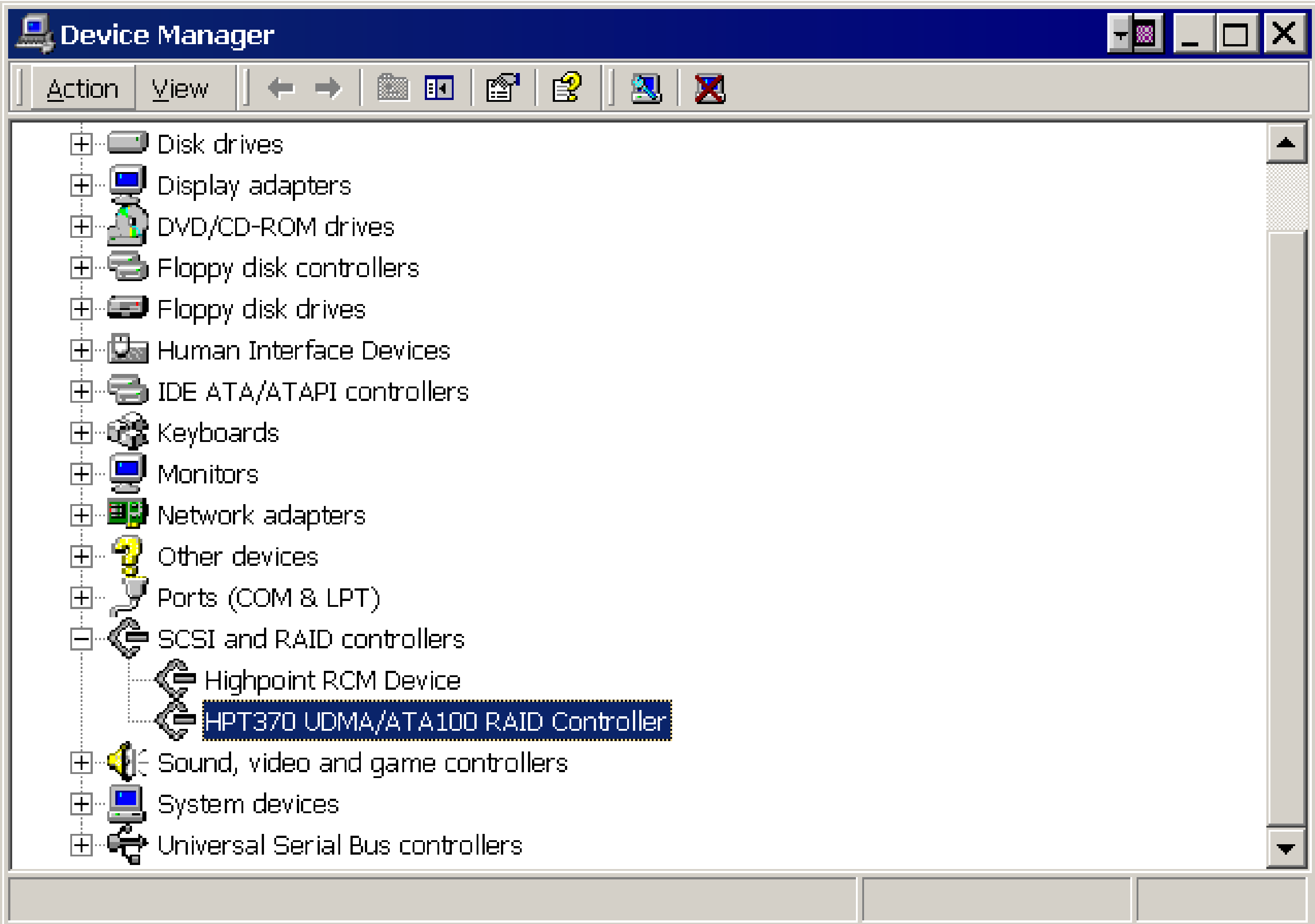
Together they control four EIDE-channels all ATA66/100 compatible:



This way you are free to assign up to 8 IDE units to the motherboard. Most people would probably have two hard disks and two CD drives. With this board they can be mounted as master on each channel. That is really good!

With Windows , you always get RAID controllers recognized as SCSI controllers. This is also the case here. Installing Windows 2000 or XP using the HighPoint controller, you have to use the driver diskette included with the board. It's a little bit weird why these Windows NT-based OS's cannot detect a RAID controller and install a driver automatically.

However the driver installation works perfectly. You see the controller under SCSI/RAID controllers:



- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b3.

About the EIDE interface - continued

The contents:

- Transfer speeds and protocols
- Improving PIO protocol
- The max. disk size

- Next page
- Previous page

Transfer speeds and protocols

[top]

EIDE exists with different protocols, like PIO (Programmed Inpu/Output) 3, PIO 4, UDMA/33, and UDMA/66. They are backwards compatible, therefore always choose the latest.

The protocol is significant to the transfer speed, since it sets the standard for the drives external speed. The protocol controls the interface between drive and motherboard.

An overview

Each EIDE unit communicate according to a specific protocol. Here you see the four best known:

Protocol for EIDE interface	Year	Maximum theoretical transfer
PIO 3 or Multi-word DMA Mode 1	1993	13.3 MB/second
PIO 4 or Multi-word DMA Mode 2	1994	16.6 MB/second
Ultra DMA (*ATA/33*)	1997	33.0 MB/second
Ultra DMA (*ATA/66*)	1999	66.0 MB/second
Ultra DMA (*ATA/100*)	2000	100.0 MB/second

Improving PIO

Traditionally PIO data transfers rely heavily on the CPU to do all the work, processing each and every little task of reading or writing data.

With the ATA-2 specification, including DMA-support solved this problem. Hence data are transported between RAM and hard disk without the supervision of the CPU.

ATA-2 was to be known as EIDE.

The max. disk size

[top]

Another problem solved by this new protocol was the 2 GB limit of disk size. LBA (Logical Block Addressing) was the new setup. LBA is an addressing and translating scheme that replaces the CHS system (Cylinder Head Sector). This new scheme enables the BIOS to address up to 8.4 GB hard disk using 24 bit long addresses.

The 8.4 GB limit was later broken by the new FAT32 file system.

Only one protocol per channel?

The two main channels (primary and secondary EIDE) can work with each one protocol. As an example, the primary channel may host UDMA/66 drives, while the secondary channel hosts PIO4 based CD-ROMs.

However, sometimes the slave/master channels only have room for one protocol. Be aware of this potential problem. Is there only room for one common protocol? In that case the "winner" will invariably be the slowest of the ones connected.

Harddisks always on the best protocol

[top]

It is important that you connect your hard disk to an EIDE channel, which runs ATA/100 (or better).

Ultra DMA requires the installation of drivers. Windows 95 before OSR2, due to age, did not recognize Ultra DMA, while Windows 98 naturally does.

The motherboard vendor ASUS (as well as others) provides an excellent, simple patch program on CD. You run it just once. Then the drivers are stored in the right locations. After one or two re-boots everything works.

Big disks

Another problem can arise if you connect two hard disks to the system board. Despite the suppliers assurance that "it is very simple," it does not always work. Therefore, it is important to start with one sufficiently large hard disk.

Please also read the article Problems with assigning two EIDE harddisks

---

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read module 7a about monitors, and 7b on graphics card.

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b4.

About the EIDE interface - continued

The contents:

- What does Ultra DMA offer?
- Looking at a good harddisk

- Next page
- Previous page

What does Ultra DMA offer?

[top]

The real EIDE improvement was accomplished with the introduction of the Ultra DMA or Ultra ATA (UDMA). It is an interface patented by Quantum but supported by all motherboard and disk drive manufacturers.

New protocol

The technology involves an improvement of the interface - the governing electronics which deliver the hard disk data to the system board. Quantum succeeded in eliminating the bottle neck in existing electronics to deliver data to the EIDE hard disks.

The UDMA hard disk is no faster it self, but the data paths have been optimized. Within the new protocol, the speed is doubled by allowing twice the data transfer per clock cycle.

Is 33 MB per second realistic?

Introducing ATA/33, it sounded great but exaggerated to talk about a 33 MB per second transfer rate, well knowing that no EIDE hard disk at this time could deliver more than 7 MB per second. Then PIO4, which can move 16 MB per second, should suffice? No, not so - the secret is in the EIDE host controller.

Overheads

The host controller, among other things, must retrieve data from the drive and deliver them to the PCI bus. Or it must retrieve data from the bus and deliver them to the disk.

The host controller has certain administrative duties to handle between reading to/from the disk. And they take some time. One clock cycle in the EIDE controller lasts 400 micro seconds. Of these, 275 are spent on "administrative overhead" - handling protocol commands, etc. The remaining 125 micro seconds are used to read from the hard disk. Therefore, a maximum transfer rate of 33 MB per second is necessary to keep up with the hard disk's capacity.

40% better

Actual measurements show that Ultra DMA disks yield up to 40% better performance than comparable PIO 4 disks. That was a clear improvement - even though the disks could not deliver the advertised 33 MB per second.

Another new feature in Ultra DMA was the CRC (*Cyclical Redundancy Check*) - automatic error correction for better data protection and verification.

EIDE on the motherboard

The system board and with that the chip set, must be set up for Ultra DMA in order for you to utilize such a disk. As always, check the chip set, when you buy a new PC. Since it provides solid performance improvement at no extra cost, it is important that it supports Ultra DMA.

Note: With the increasing magnetic density on harddisk platters, we use the bandwidth of the UDMA ATA/66 or Serial ATA interface.

Conclusion

For the EIDE hard disk to function in the Ultra DMA protocol, the following conditions must be met:

- The hard disk must be the Ultra DMA type.
- The system board must have a chip set which supports Ultra DMA with the latest protocol.
- BIOS must "log on" the hard disk with Ultra DMA protocol. You can verify that in the start up screen.
- Drivers for the chip set must be installed in Windows .

A good Ultra DMA disk

[top]

Here you see my (old) Maxtor harddisk (Diamond Max 91728D8 from 1998/99). It holds 17 GB. It rotates with a modest 5400 RPM, which makes it noise-free and it does not produces a lot of heat.

Yet it is quite fast due to the magnetic density, which is 4.32 GB per platter. This thing holds 4 platters:

The harddisk is configured as a slave unit on the primary EIDE channel as described later.

More modern harddisks from Maxtor take exactly the same physical dimensions; internally, they just hold 30, 45 or 60 GB.

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

KarbosGuide.com. Module 5b5.

About the UDM/66 interface

The contents:

- Ultra DMA ATA/66
- Chipset support to UDMA/66
- ATA/100
- Serial ATA

- Next page
- Previous page

Ultra DMA ATA/66

[top]

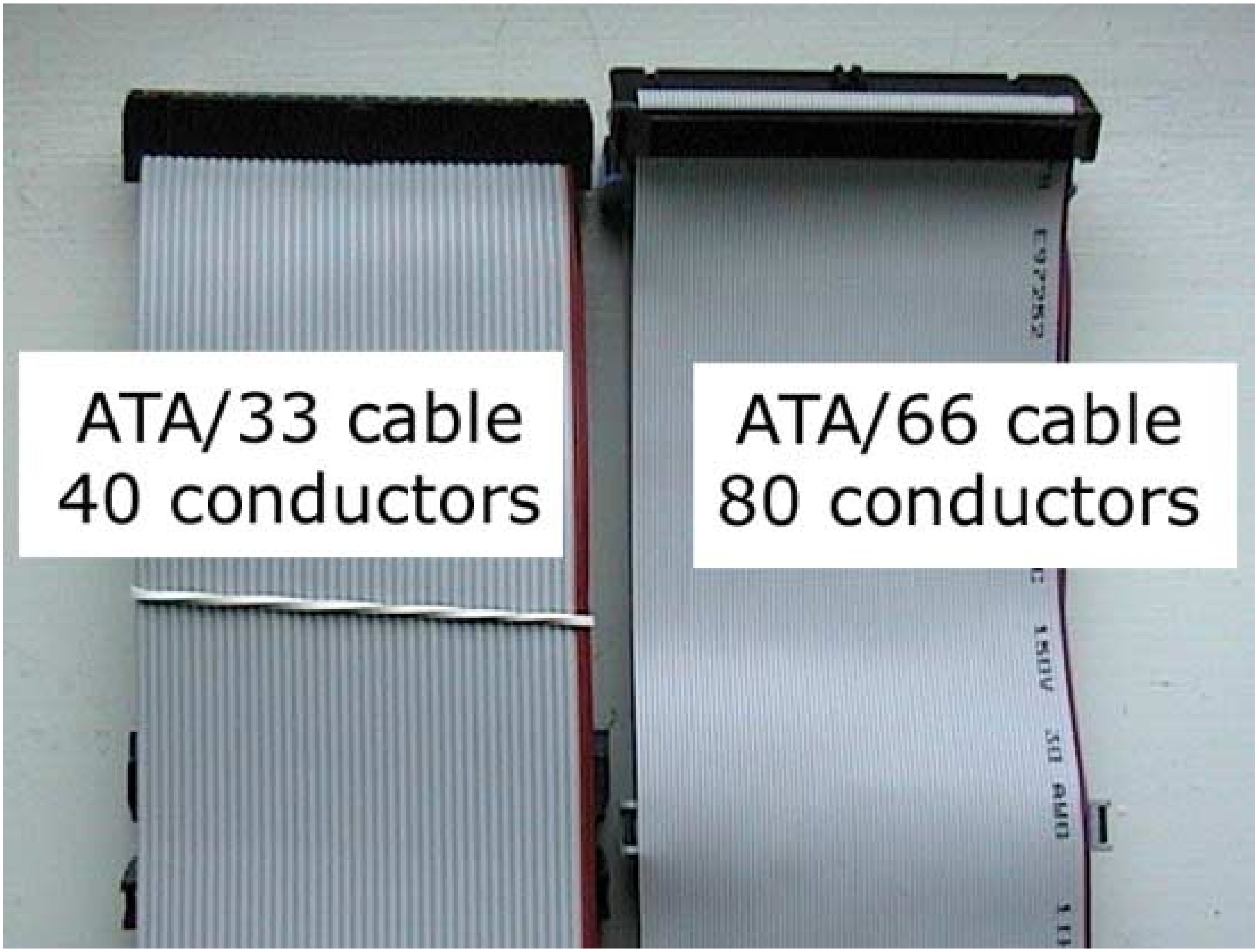
In 1997-98 Intel and Quantum created the new Ultra DMA standard called ATA/66. This gives a theoretical bandwidth of 66 MB/sec.

The new system requires a new cable with 80 conductors. The 40 new conductors are used for grounding. In the old version, only 7 cables were used for grounding. The new and very effective grounding removes the so-called crosstalk (i.e. noise remaining in the cable after a transmission).

In ATA/66 the controller had to wait for noise in the cable to disappear before the next transmission. With the new cables the noise is dramatically reduced, the transmissions can follow one by one without delay, and the bandwidth goes up.

The ATA/66 protocol is fully compatible with ATA/33. You may use both type of drives on motherboards either with ATA/33 or ATA/66. Of course you only get ATA/33 performance using a ATA/66 drive on a ATA/33 motherboard. You may upgrade your motherboard with an PCI-based ATA/66 adapter. This is quite cheap.

The new cables also use the same old 40-pin plugs:



If you use a ATA/66 system with a 40-pin cable, the protocol will automatic switch to ATA/33.

What is required?

According to Western Digital (who had some of the fastest UDMA/66 drives), to use the ATA/66 technology a PC system must have:

- Ultra ATA/66 compatible logic either on the system motherboard, or on an adapter card

- Ultra DMA compatible BIOS
- DMA-aware device driver for the operating system
- Ultra ATA/66-compatible IDE device such as a hard drive or CDROM drive
- 40-pin 80-conductor cable

See this white paper from WD

ATA/66 was very neccessary for further development of the EIDE harddisks. With the increasing data density the media data transfer rates are going up and up. Therefore the host data transfer rates also must increase. All the time the controlling logic must have a better transfer rate than the media or else performance is reduced. The new EIDE disks coming out from IBM and other vendors delivered such a powerful data output that the old UDMA ATA/33 standard could not cope with it.

Chipset support to UDMA/66

VIA's Apollo Pro+ chipset fully supports UDMA/66. The best performance should be gained from the T82C686A super south bridge controller.

Intel supports it with the 810 (Whitney), 820 (Camino) and 840 (Carmel) chip sets.

Other setups include special logic chips to includefull UDMA/66 performance tofor instance a BX-based motherboard.

Promise FastTrak66

Promise produces a PCI-based controller called FastTrak66 that does the job. It controls UDMA/66 disks at full speed, and it even allows doubling or quadroubling the speed using RAID techniques. See the description of the older Promise FastTrak controller.

ATA/100

In spring 2000 the new IBM disks became so fast, that ATA/66 was out of business. The disks use a new protocol called ATA/100, being developed by Quantum, who holds the Ultra ATA patents.

The ATA interface started in 1996 with ATA/33, which in 1998 was upgraded to ATA/66. Two years later the ATA/100 was released.

A kind of ATA/66 Second Edition

Where ATA/33 gave a very powerful boost in the bandwidth between controller and harddisk, the ATA/66 gives a minor gain in performance. On the other hand it solves a lot of compatibility problems by improving timings and other parameters in the specification.

As a result, the ATA/100 is reported being more simple to implement in the chipset logic. It is cheaper to produce and fully compatible with both ATA/33 and ATA/66.

There is an upper limit of disksizes at 137 GB in the ATA/100 interface. However there has been made some workarounds to this problem in some comtrollers.

100 Megabyte per second

The ATA/100 interface have theoretical bandwidth of 100 MB/sec. This is more than any harddisk can deliver at present. However the harddisk technology is improving in very high speed, so the disk soon will reach this limit. Hence the *Serial ATA* will be needed.

Another technology we will see more and more is ATA-based RAID. Using two or four cheap ATA-disks you can have a very powerfull disk system with doubled bandwidth. This requires a RAID-controller like the Promise FastTrak or a motherboard with onboard controller. Theserequires higher bandwiths in the ATA protocols to show real powerful performances.

ATA/133

In 2001 some hardware vendors introduced ATA/133 as new version of the interface.

Serial ATA

Another more interesting new technology is the *Serial ATA*. Intel, Dell, IBM, Maxtor, Quantum and Seagate and other partners are about to replace ATA/100 with a faster drive interface.

The new Serial ATA interface, can pump out 160 MB per second in the first version (Serial ATA 1x or SA1X). Later version promises bandwiths up to 528 MB per second. This will give us headroom for the next five years of harddisk technology improvements.

Even more promising is the new cable design of Serial ATA. Instead of 40/80 conducters the cables only holds *four* conductors. This thinner cabling is great news for everyone putting together his own PC. I also expect the number of onboard ATA-channels to increase from 4 to 8.

Serial ATA probably will kill the remaining hope for generel use of the IEEE 1394/FireWire interface for PC's harddisks. This was never really supported by Intel.

Hopefully the new interface will operate with a command queue, which has been a great lack in ATA-design compared to SCSI.

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b6.

Practical hints to setting up a harddisk

The contents:

- Setting up the disk
- The BIOS

- Next page
- Previous page

Configuring your EIDE hard disk

[top]

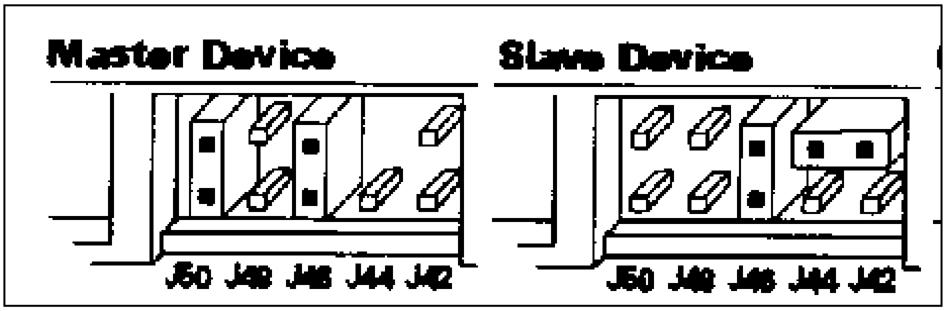
Often setting up a hard disk is very easy. But sometimes it teases you - sometimes even a lot. Let me here describe a few checkpoints:

- The cabling. Your hard disk has to be connected to the master EIDE channel, and the connection has to be ensured.
- The jumpers - is the drive a slave or a master? Read below.
- The BIOS setup has to be configured. Read below.
- Run FDISK to partition the hard disk. Boot from a diskette with fdisk.exe on it. Read about file systems in module 6a.

Slave or master

If you connect EIDE harddisk number 2, you probably have to connect as a slave unit on the primary EIDE channel. This is not very good, as I describe it in the article Please also read the article "Problems with assigning two EIDE harddisks".

Anyway, I'll use this setup as example. You have to make sure the jumper setting is correct. Here is the text from a Maxtor manual, which tells how to set the unit up as a master or a slave. It is very easy - and it works:



The jumpers are on the back side of the drive, between the cable and power connectors:

The BIOS

BIOS is a low-level layer of system software. The BIOS has to identify the hard disk at the Start-up. If BIOS does not have the proper values for the specific hard disk, it will not function. There will be no access to the hard disk.

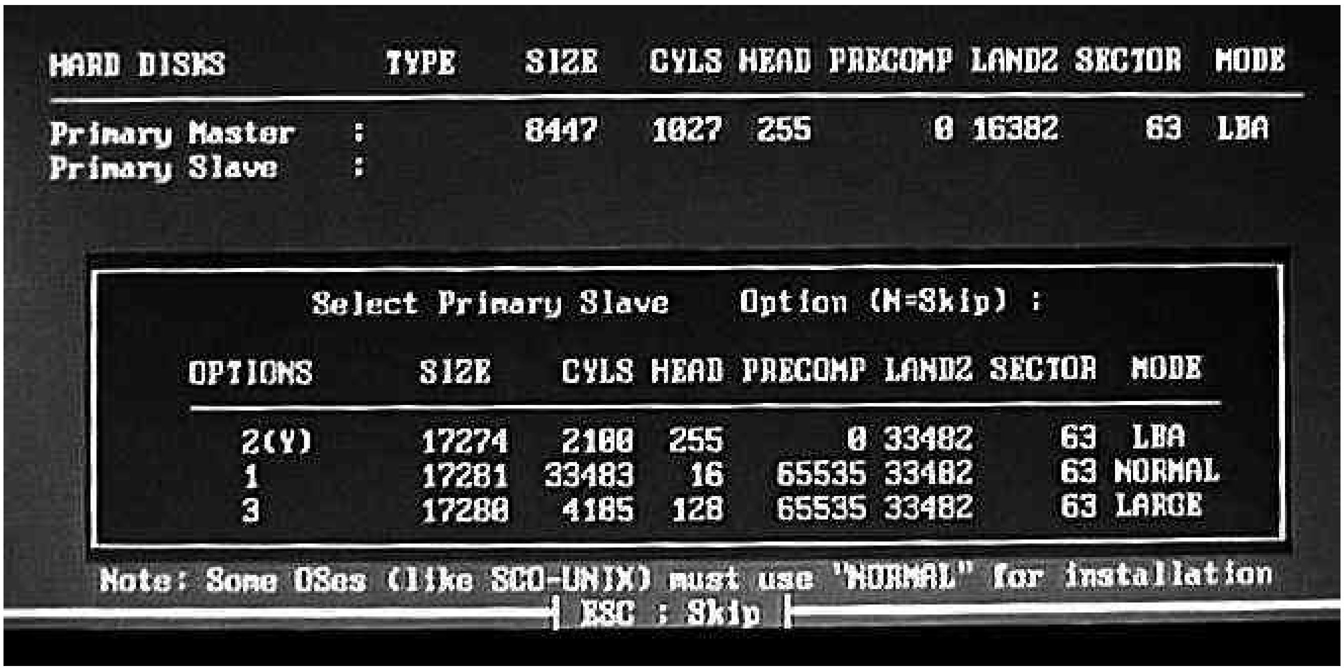
The hard disk is recognized by BIOS on certain parameters like *number of cylinders, number of heads, sectors* etc. These values are to be stored in the CMOS memory. But often the PC cannot identify these values without a helping hand. I here describe the installation as I know it (the ROM software is from Award). It is not complicated at all, but it involves two menus from the Setup utility.

If you have mounted your hard disk and ensured that the cabling is correct, it is time to boot the PC. Then you hit the [Delete]-key to enter the Setup program. First you choose "IDE HDD AUTO DETECTION" as here:

Then you have to let the program detect each of the four EIDE channels. On the first one you should find your hard disk. If not - then something definitely is wrong.

When you find the hard disk you usually get three options concerning the protocol assigned the drive. Here we find the harddisk on Primary Master channel. Choose the prompted one, here it is number 2 (LBA):

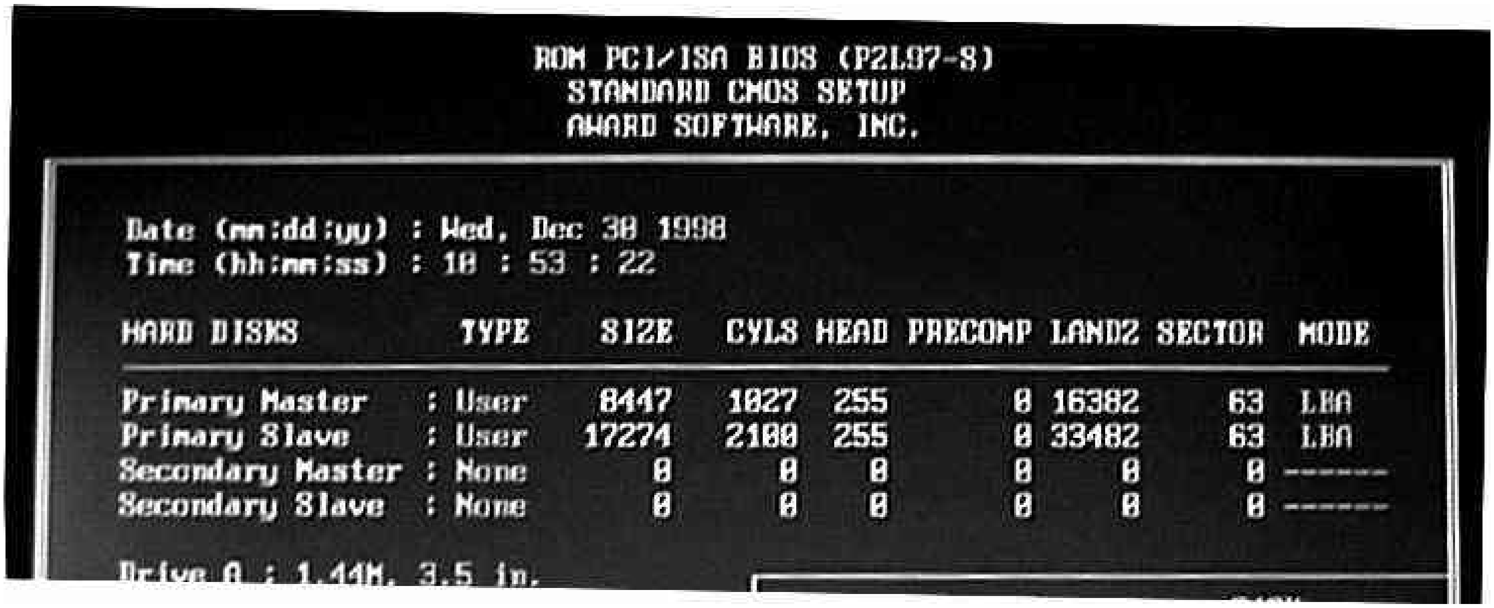
Continue with other channels. Here I install harddisk number 2:



Sometimes a CDROM drive is identified here, other times it isn't, but it functions anyway.

When the hard disk and other EIDE units are identified by the auto detect utility, you return to the main menu of the Setup program. Choose "STANDARD CMOS SETUP" as here:

Here you should find the hard disk listed. In my case it looks like this, using two harddisks:



Be aware that the list above has to correspond directly to the physically installed units - especially if you reconfigure your system. Otherwise it won't work. This has teased me when I installed hard disk number two and changed the master/slave setup. It is not enough to run the auto detect, also the standard CMOS setup menu has to be updated.

When all this is in place, you have to save the changes and reboot.

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c about SCSI, USB etc.

Read module 7a about monitors, and 7b on graphics card.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5b7.

About the interface AGP

The contents:

- Introducing AGP?
- The technology

- Next page
- Previous page

What is AGP?

[top]

A new bus was introduced in 1997. It is called AGP (*Accelerated Graphics Port)*, and it is exclusively designed for video cards.

Introduction

AGP was designed with two purposes:

- To relieve the PCI bus of work with graphics data. Hence, it can concentrate on other demanding transport duties, like transfer to and from network adapter and disk drives.

- To have better bandwidth within the video system.

AGP was introduced with the Pentium II processor and Intel 82440LX chip set. Intel hoped to lift more of the CPU market away from the Socket 7 compatible CPUs by designing a completely new motherboard layout and including a new powerful bus for the graphics card. However both Ali and VIA soon introduced chip sets for Socket 7 motherboards including AGP. So today AGP is found on most motherboards.

The technology

AGP includes several techniques:

PCI

The PCI bus in version 2.1 is the basis. It runs with 66 MHz bus frequency. That gives a doubling of transfer speed compared to traditional PCI, from 133 to 266 MB/sec.

Clockdoubling

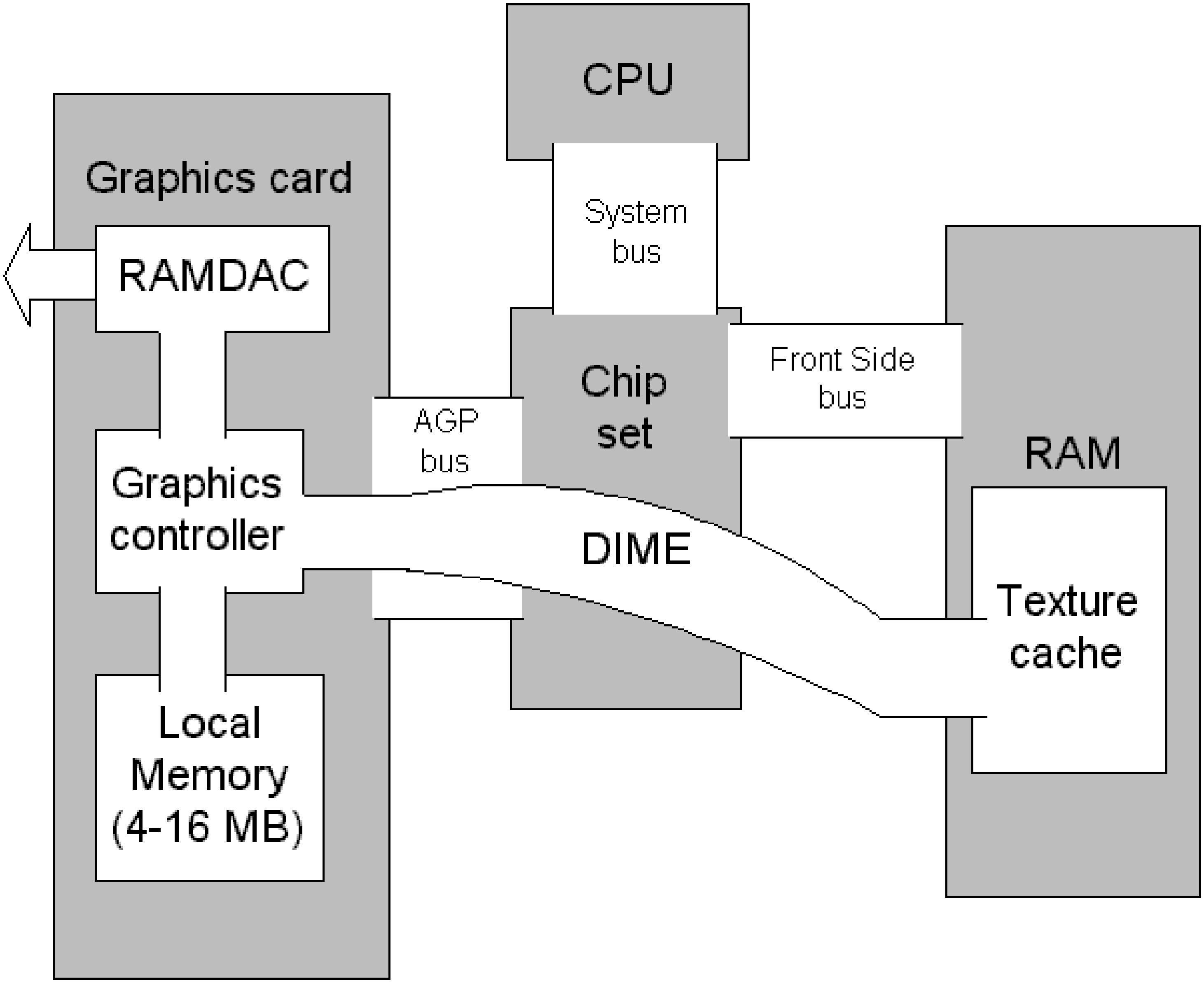
A kind of clock doubling in the 2X mode, where the bandwidth is expanded to 533 MB/sec.

Texture cache

This is a method to utilize system board RAM for *texture cache*. A technology that expands the memory used by the graphics card, utilizing the ordinary RAM in the PC.

The data for texture (backgrounds) need not be processed by the graphics controller, so it is just being loaded from the RAM. This technology is called *DIME* by Intel (for Direct Memory Execute).

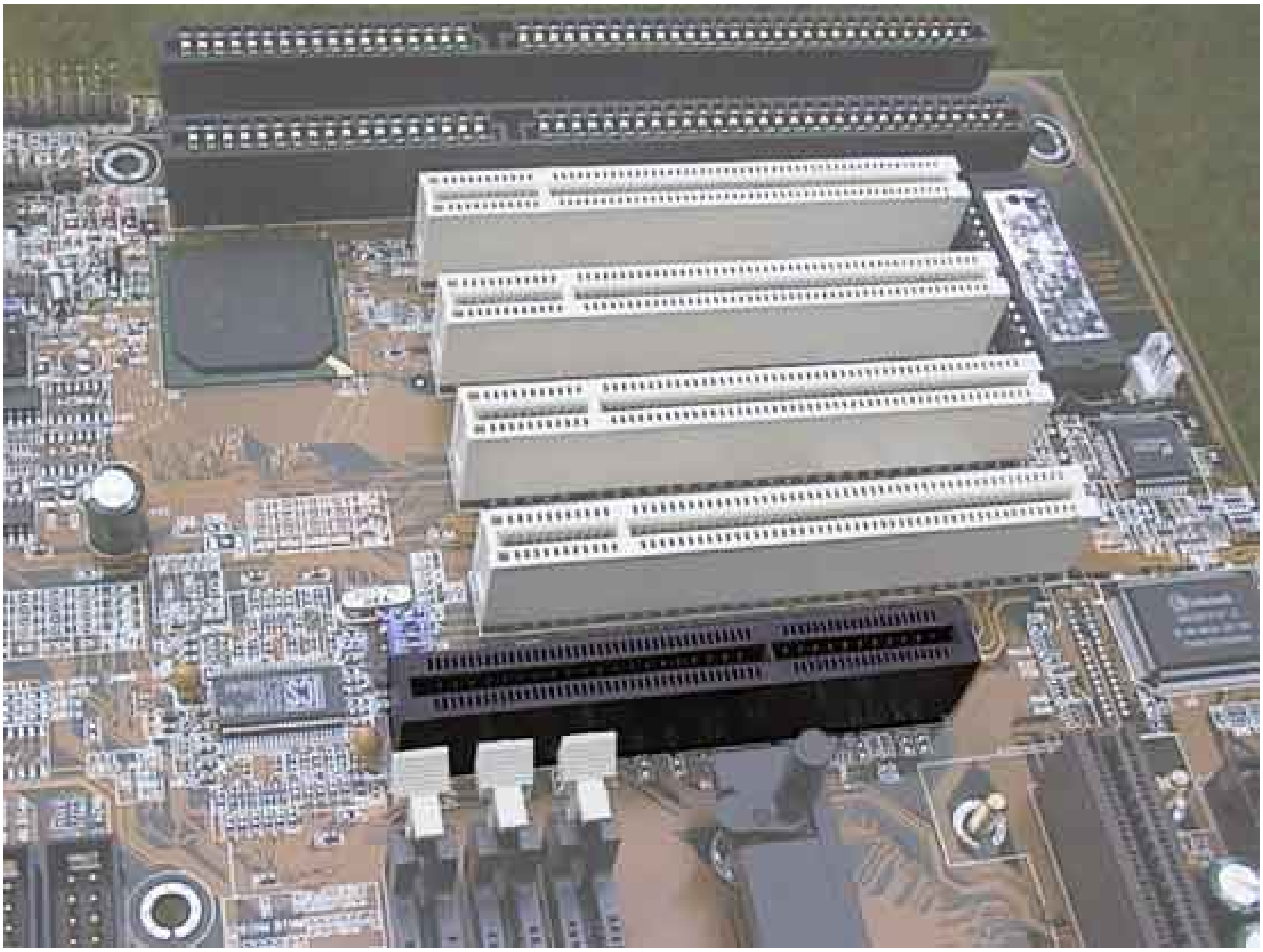
Here you see the architecture involving chip set, main memory and AGP:



The socket

Below you see the AGP socket at the bottom. It looks like a PCI socket, but it has been placed in a different position on the board.

In the top you see two (black) ISA sockets. Then four (white) PCI sockets, and then the brown AGP socket:



At a first hand glance, the AGP socket very much look like a PCI socket. But it is placed in a different position, so you cannot plug an AGP card into a PCI socket.

AGP 4X and above

In 2000 we saw first AGP4X chip sets (like i820) and motherboards implementing this feature, which many graphics controllers already have been made for. With AGP the bandwidth to the video subsystem is up at 1066 MB per second.

Late in 2001 the AGP 8X standard is expected to hit the market. It will increase the bandwidth to more than 2 GB/sec. To use this power, we have to have more powerfull chipsets with high-speed RAM. Intels i850 with dual Rambus channels will deliver the necessary RAM bandwidth.

Try www.agpforum.org for more info.

- Next page
- Previous page

To learn more

[top]

Read more the boot process and system bus in Module 2b

Read about file systems in module 6a

Read about I/O buses in module 2c

Read about the motherboard chip set in module 2d

Read about RAM in module 2e

Read Module 5c2 about USB.

Read module 7a about monitors, and 7b on graphics card.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 5c1a.

About SCSI

The contents:

- The host adapter
- 7 units in a chain On the following page:

- SCSI is intelligent
- About the SCSI standards
- What do you gain with SCSI?
- Next module: USB
- IEEE1394 FireWire
- Device Bay

- Next page
- Previous page

Introduction

[top]

SCSI (Small Computer System Interface) is high end technology. It is a technology which provide means for data exchange among hardware devices such as drives, tape streamers and scanners. SCSI is especially used in high end PCs such as network servers or just powerful workstations.

SCSI might be compared to the EIDE interface, which also uses a host adapter controlling drives. However SCSI has two major advantages over EIDE:

- A SCSI host controls 7 or 15 devices (using only one IRQ).

- The SCSI system holds its own computer power, thus freeing the CPU from workload.

If you are critical about your PC power, the SCSI would be worth considering.

The host adapter

[top]

A SCSI system is built around a central, intelligent controller called the *host adapter* . A host adapter can control several SCSI units:

- Many units on the same host adapter.
- Many types of *drives* : Hard disks, CD and DVD drives, Zip drives, MO drives, etc.
- Tape streamers (DAT and others).
- Scanners.

The host adapter has its own BIOS separate from the PC's. When you boot the PC, you will see the adapter communicating with connected SCSI devices.

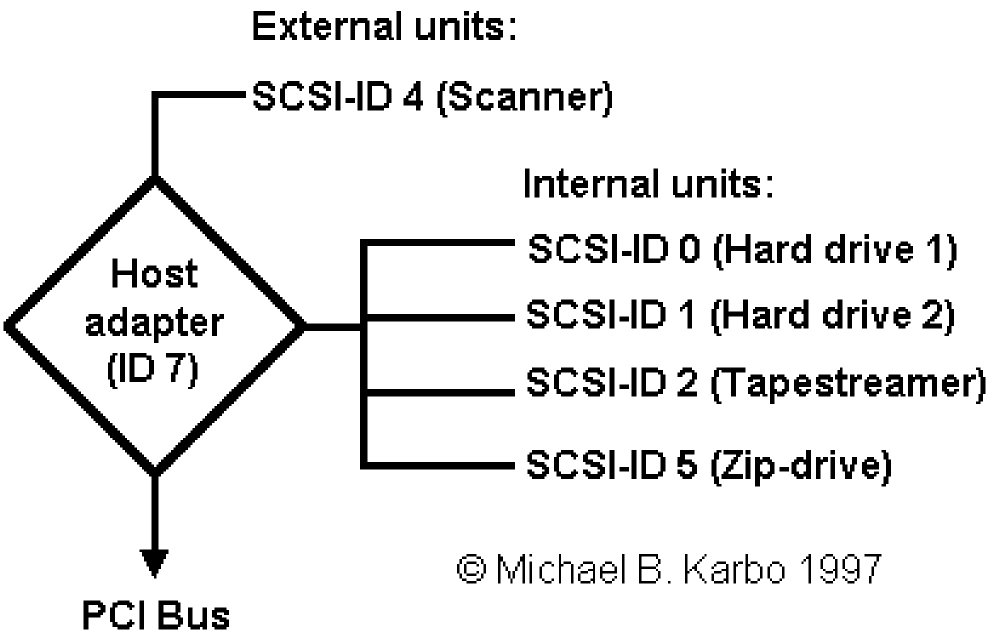
The adapter is rather expensive. Currently, the best for ordinary use is called Adaptec 2940 U2W (priced at around $200). It is PCI based, so you could use it in your next PC too. I have had good experiences with ASUS motherboards in versions with onboard SCSI controller. That is the most easy solution - to have the controller onboard.

7 units in a chain

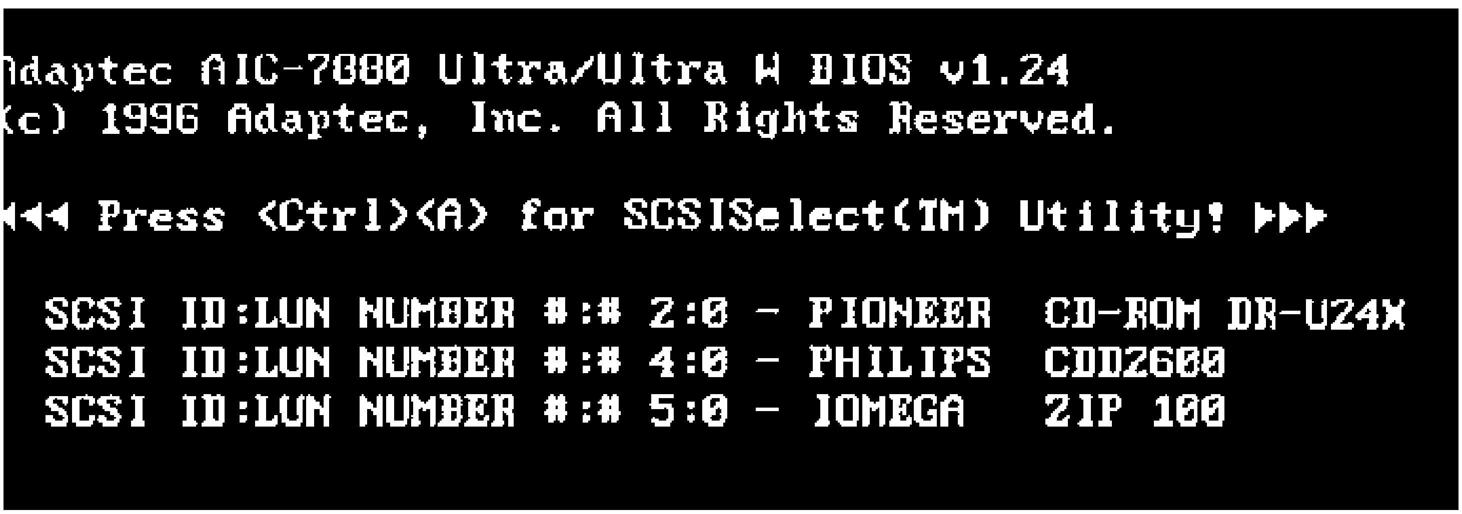
[top]

The regular SCSI 2 system can handle 8 devices including the adapter itself. SCSI Wide handles 15 devices. Each device has to be assigned a unique number going from *ID 0* to *ID 7.* The SCSI devices can be *internal* (installed inside the PC cabinet) or *external* . The host adapter is a device itself. Typically, the host adapter will occupy ID 7.

Here is an illustration of a SCSI string with host adapter (ID 7) and five units (ID numbers 0, 1, 2, 4, and 5):



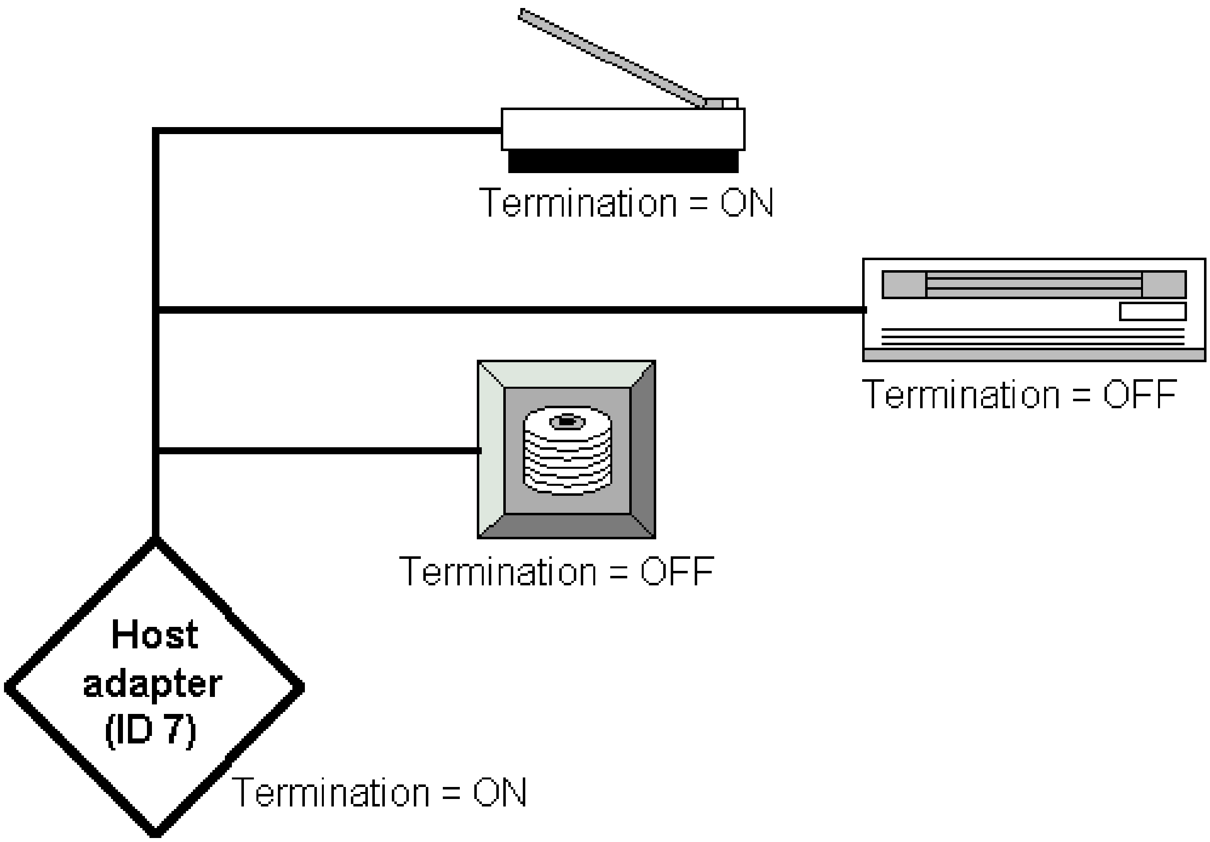
Here you see the BIOS report of a found SCSI chain:



Terminators in both ends

The last unit in both ends of the SCSI chain must be terminated. This means that there must be resistors (jumpers or switches) attached to two of the units.

If you only use two devices, you do not have to worry about it. The host adapter is one end of the chain and the other device is the other end. With three or more units you have to take care of termination:



- Next page
- Previous page

Learn more **[top]**

Read about FireWire in module 5c3

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

KarbosGuide.com. Module 5c1b.

About SCSI - continued

The contents:

- SCSI is intelligent
- About the SCSI standards
- What do you gain with SCSI?

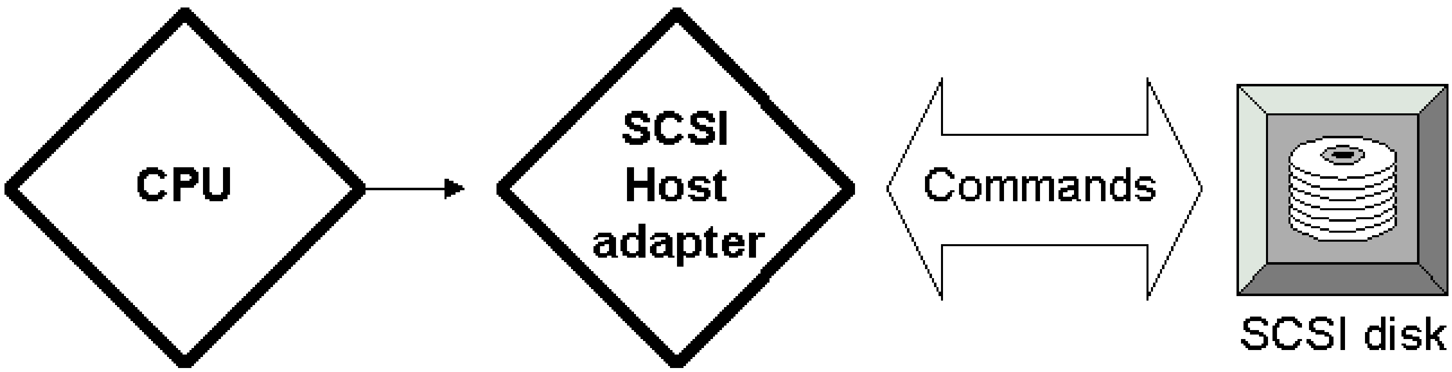
- Next page
- Previous page

SCSI is intelligent

[top]

SCSI is remarkable in having an intelligent protocol, which assures maximum utilization of the data transfer. The basis of SCSI is a set of *commands* . Each individual device holds its own controller, which interprets these commands.

All commands within the SCSI system are handled internally, meaning the CPU does not have to control the process:



While the read/write head moves across a SCSI disk, the host adapter as well as the CPU can handle other jobs. Therefore SCSI is well suited for multitasking environments.

About the SCSI standard

[top]

SCSI stands for Small Computer System Interface. It is intended as a universal interface, defined and designed in 1982 by NCR and Shugart Associates. It exists in numerous variations. Here you see some of the more significant editions:

Standard	Year	Bus speed	Bus width	Max. bandwidth
Standard SCSI	1986	5 MHz (Asynchronous)	8 bit	5 MB/sec
Fast SCSI Narrow	1990	10 MHz (Synchronous)	8 bit	10 MB/sec
Fast SCSI Wide	1992	10 MHz (Synchronous)	16 bit	20 MB/sec
Ultra SCSI	1994	20 MHz (Synchronous)	16 bit	40 MB/sec
LVD Ultra2	1996	40 MHz	16 bit	80 MB/sec

Today there are many SCSI standards. Among others, you can come across SCSI-20 and SCSI-40, which refers to the bus speed. The SCSI standard seem to have its own life with plenty of new development.

LVD Ultra2

The latest version of SCSI is called LVD (*Low Voltage Differentiale).* You also find the term SCSI Ultra 2 - there have always been so many terms... LVD is an improvement to SCSI-3.

LVD gives twice the bandwidth of the ordinary SCSI-3. Another improvement is the cabling which works up to 12 meters. Traditional SCSI only works within 3 meters. LVD has to compete with FireWire, which also has a powerful bandwidth.

Adaptec have a SCSI-controller delivering up to 160 MB/sec. This unitcalled Adaptec SCSI Card 3950U2 uses two independent Ultra2 SCSI buses in one card. It connects up to 30 SCSI devices!

What do you gain with SCSI?

[top]

Expensive but good. SCSI makes the PC a more expensive, but more versatile. The advantages are, that on the same PC you have free access to use *many* and *good* units:

- It is easy to add accessories as DAT streamers, CD-ROM recorders, MO drives, scanners, ZIP- and DVD-drives etc.
- You can use SCSI hard disks.
- You can use CDROM drives on SCSI, which may give a better performance.

The advantages of SCSI hard disks

SCSI hard disk are generally of higher quality than other disks.Typically, good SCSI disks come with a 5 year warranty. Traditionally they are faster than the EIDE disks. At 10,000 or 14,000 RPM they have shorter seek times. They also have a bigger cache.

Another advantage is the large number of accessories, which can be attached. If you buy a 18 GB SCSI disk today, you will guaranteed need additional disk storage in a few years. Then you just add disk number two to the SCSI chain, and later number three. The system is more flexible than EIDE, where you can have a maximum of four units incl. CD-ROM.

The SCSI hard disks can also adjust the sequence in the PC's disk read commands. This allows reading the tracks in an optimal sequence, enabling minimal movements of the read/write head. Quantum calls this technology ORCA (*Optimized Reordering Command Algorithm*). It should improve performance by 20%.

Finally, the SCSI controller can multitask, so the CPU is not locked up during hard disk operations, which you can experience with IDE.

SCSI hard disks can achieve substantially larger transfer capacity than the IDE drives, but they have the same bottle necks: the serial handling of bits in the read/write head, where the capacity is highly dependent on the rotation speed.

SCSI is for servers

However, today the importance of SCSI is decreasing except for use in dedicated servers. Modern CD-ROM and CD-RW drives work just as good on EIDE as on SCSI. USB has taken over when it comes to controlling units like scanners, cameras and Zip drives.

Finally, modern EIDE-based harddisks have an extremely high quality compared to the products we had five years ago. Hence, there is no reason to prefer SCSI-based harddisk to the more inexpensive EIDE drives.

But for servers SCSI still has a market.

Booting from SCSI disk

If the hard disk has to be booted, traditionally it has to be assigned ID 0. If the SCSI controller has to control the hard disk, then the PC CMOS setup must be modified, so the (IDE) hard disk is not installed if not both types of hard disks are installed.

The operating system will find the host adapter after start up and BIOS will be read from the hard disk through the adapter. New BIOSs allow a choice of booting from either IDE or SCSI disk.

Fast and Ultra Wide

The newest generation of SCSI hard disks are both fast, ultra and wide. Therefore, the best advice is to buy an adapter like Adaptec 2940UW2, which can handle the newest disks.

IBM disks

Allow me to advertise IBM's SCSI disks. They are fantastically good. Unfortunately, not many people know about them. I have had a few of them. They excel in high quality at reasonable prices. The physical construction is very appealing: The electronics are integrated in very few components. Everything exudes quality! And they are very quiet. You simply cannot hear them.

32 bit problems in Windows 3.11

Windows 32 Bit Disk Access has given problems with SCSI disks. For a long while, it was impossible to install a 32 bit driver in Windows 3.11 to the SCSI disk. This was solved in 1995 and there have been no problems with Windows and SCSI since then.

Links

About SCSI: SCSI Pro and DPT offer some information.

- Next page
- Previous page

Learn more **[top]**

Read about FireWire in module 5c3

Read about chip sets on the motherboard in module 2d

Read Module 4d about super diskette and MO drives.

Read module 5a about expansion cards, where we evaluate the I/O buses from the port side.

Read module 5b about AGP

Read module 7a about monitors, and 7b on graphics card.

Read module 7c about sound cards, and 7d on digital sound and music.

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

Karbosguide.com

A Few Software Tips

The contents:

Which are the advantages of Windows 98?
About the swap files and RAM.
About the disk cache.
About temporary files (1).
About temporary files (2).
About file types - show only some of them!
Use the desktop for favorite Internet addresses.
Permanent folder for download in Internet Explorer.
Choose a start page.
Color changes in menus - a option in Windows 98.
Replace opening screen in Windows.
Upper case letters in folder names.
Single click in Explorer - smart idea.
TWEAK UI - the hidden tool in Windows 98.
FAX program - what happened to that?
Windows - autotexts in any program with ShortKeys.
Windows - permanent, global and local macros..
Running out of space on my hard disk...
Enabling DMA on the harddisk...
Use MSConfig to alter the Windows start-up.

- Next page
- Previous page

Karbosguide.com. Software Tip 1

What are the advantages of Windows 98?

- Next page
- Previous page

In my opinion and experience Windows 98 was an excellent operating system when introduced. At its introduction in the spring of 1998, some papers made the comment "it is not worth the money" and "there is not much new compared to Windows 95." I did not agree with them.

Here are some of the advantages compared to earlier versions of Windows:

Generally better performance

Generally Windows 98 utilizes the PC resources better:

- Memory management has been completely changed. It finally works.
- The file system is better integrated into the operating system, which gives new functionality.
- Program loading can be up to four times as fast.
- Hardware support is significantly improved with a new driver model, and all new chipsets etc. are supported. However these conditions may change.

These improvements are sufficiently significant to justify an upgrade.

Better user interface

[top]

Superficially the Windows 95 user interface has not changed much. But you need not dig down very deep to see many novelties. I am talking about small items with better adaptations in the Start menu, new tool bars, etc. But these small items are really very smart when you need to set the user interface.

Better system tools

The system tools are significant for the more demanding user, who really wants to know and be in command of the PC. A number of new tools have been added to Windows 98. They improve surveillance facilities. All of this will be thoroughly described in my upcoming "Windows 98 and hardware" booklet (or whatever the title will be).

Stability

Many will experience that previous *instability* is just gone. The PC can be left on for weeks on end without going down a single time. Many may laugh at this - "why should we pay to correct Windows 95 errors." That may be true, but consider the wasted time with PC's which fail and need to be restarted, etc. Cut the mustard and get 98 on your machines - then it works. Life is too short for lousy software!

Windows 98 is not good enough

In the years after the introduction of Windows 98, we saw new and faster hardware coming extremely frequent. We got faster CPU's, the clock frequency increasing 3 to 5 times in few years. Also harddisks became faster and RAM as well.

Having a moderne PC with plenty of fast hardware, Windows 98 or Me (the later version) is not good enough. You need Windows 2000 or XP to benefit from the hardware. This is a fact. Just look at the way Windows 98 manages memory - it does not work using more than 128 MB of RAM. And that is a petty. If yOutlookwork with graphical applications like Photoshop or FireWorks, you will see a great performance using 512 MB RAM or more - but not with Windows 98/Me.

- Next page
- Previous page

Software Tip 2a

About RAM and swap file

- Next page
- Previous page

In Windows 95 and 98 it is important to understand the of the relationship between:

- Amount of RAM in the PC
- Size of and control of Disk cache
- Free memory
- Size of the swap file

Windows are in all versions (as all Microsoft software) a very resource demanding operating system. Then you might ask, why bother to use Windows? We all know the answer: The Microsoft Office packages are undoubtedly the finest, most user friendly and most thoroughly planned office programs on the market - no question about that. They can work satisfactorily on your PC, but it requires some hardware. A lot of hardware indeed.

The processor should be fast, as all modern processors are. Plenty of RAM and a roomy and fast hard drive is also very desirable for running Windows.

The need for RAM

Windows gobbles up memory. Therefore, sufficient memory is essential for its satisfactory performance. Try to check how much you really need - you will be surprised. The memory comes from two locations:

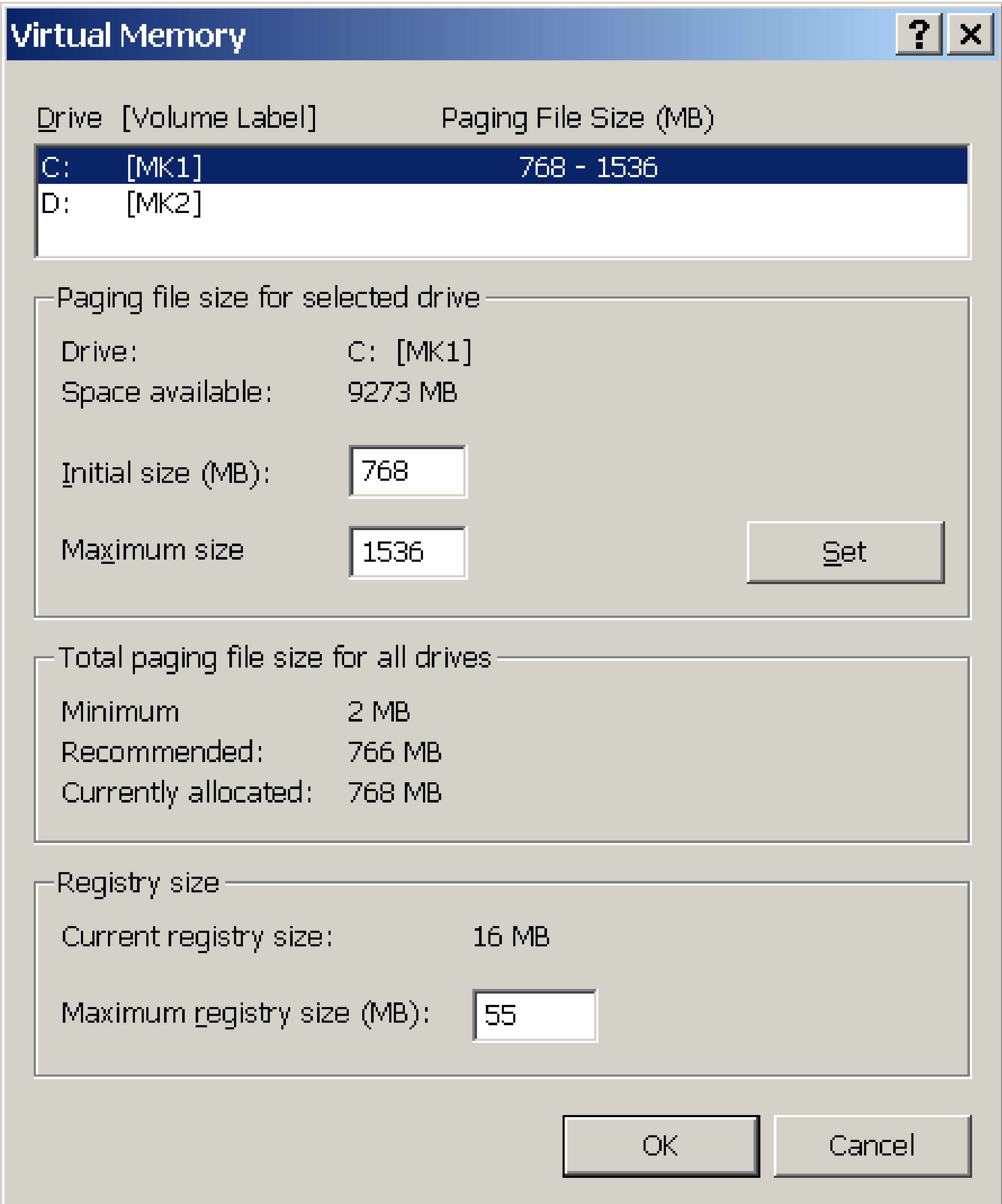
- The installed RAM
- The swap file, which is created automatically, when you run out of RAM.

Windows is clever using the swap file . It "extends" its RAM to the hard disk. If you only have 64 MB RAM in your PC, you can be assured that you have a sizable swap file on your disk.

Controlling the swap file

You may choose which drive, you place the swap file. Some experts prefer to place the swapfile on a separate partition, which only is used for the swap file. That way, the swapfile does not interfere with the other disk data, which become more easy to defragment.

You deside the placement and size of the swapfile using the System Properties dialogue box. Here you see it from Windows 2000:



We recommend that you limit the swapfile to a size of 512 MB using Windows 98/Me. If you use Windows 2000 (which is working a lot better than 98/Me) you should leave Windows to deside the size of the swap file.

Anyway, you need to keep an eye on the swap file. In Windows 95 many breakdowns originated in swap file use. But luckily Windows have improved a lot; Windows 98 is is better at controlling RAM and swap file than Windows 95 is.

Windows 98 has a better algorithm to control RAM etc. The swap file is still there, and it is big - but that does not have to be a problem. Windows only reads to and from the swap file, while no work is done on the PC. In that way we do not even notice that there is a swap file.

In the Windows versions 2000 and XP there is no need to worry about memory management, it works fine (but please use 512 MB RAM).

No swap file?

Some experts recommend if possible to eliminate the swap file in Windows 98. It sounds great but is not not very smart in practice. The problem arises from the extremely lousy memory management you find in Windows 98. Any onboard RAM above 256 MB find no use! Even upgrading from 128 to 256 MB gives almost no benefit; Windows still runs out of memory all the time.

- Next page
- Previous page

Index:

3D graphics
3DNow!
5X86
6X86
6X86MX (M2)

8086-compatibility
810 chip set
815 chip set
820 chip set
820E chip set
82430HX
82430TX
82430VX

82440BX
82440EX
82440FX
82440LX
82440GX
82450NX
82440ZX

Adapters
AGP
AMD K5 and K6
ASCII
ATA/66, ATA/100
AthlonXP
ATI
ATX

BIOS
Bits & Bytes
Boot record
Boot sequence
Bus mastering
buses

EPP/ECP
EIDE
EDO-RAM
EISA bus

FAT16
FAT32
FDISK
FireWire
Foster
FPU

G400

Hard drives
HiFD

i740
IDE
IEEE1394 FireWire
IBM drives (Deskstar etc..)
IBM compatibility
Introduction to Click & Learn
IRQs
ISA bus

K6
K6-2
K6-3
K7 Athlon

Katmai

LCD display
LS120

Neumann, John von

Over-clocking CPUs (tuning)

PC100 RAM

PC133 RAM
PC-Card
PCI bus
Pentium
Pentium II
Pentium III
Pentium 4
PentiumPro
Pipelines
Plug and Play (PnP)
POST (tests)
Powersafe, in BIOS

RAID 1, 2.

RAMDAC
RDRAM
RISC instructions
ROM chips

S3
SCSI
SDRAM
Setup program
SiS (Chip sets)
SIMD
SIMMs
Slot One
SPD
SSE2
System bus

Camino
Celeron
CD-ROM
CD-RW
Chip set
CISC instructons
CMOS
CPU-cache (L1 and L2)
CuMine

Cyrix DDR RAM
DirectX
Diskcache in W95
Diskette drive

Dixon
DMA
Drives
Drives, Interface
Dolby AC-3
Drivers
Dual Voltage
DVD drives
DX4

LVD

M3 (Cyrix CPU)
Matrox
MCA bus
Merced
MIDI
MMX
MO-drives
MPC-3
MP3
Motherboard
Multi-read CD-drives

Tapestreamers
.Tmp files
TNT2
Trinitron
Tseng

USB
Ultra DMA

VC133 RAM
Vesa Local Bus
VIA chip set
VIA Apollo MVP3
VIA Apollo+

Wave table
Wait states
Willamette
Windows optimizing

Xeon
ZIF
Zip-drives

KarbosGuide.com. Module 2d.06

Intel's i810

The contents:

- Introduction
- The Accelerated Hub Architecture
- The Graphics Memory Controller Hub of 810
- The AC97
- Karbo's conclusion

- Next page
- Previous page

Intro to Intel 810

[top]

With i810, Intel has launched the first chip set of a new generation.

In late April 1999 the 810 "Whitney" chip set was introduced. This set is new in several aspects.

Intel® 810 Chipset

- A new type of memory controller with built-in graphics technology.
- Support for up to 512 MB SDRAM.
- Built-in audio-codec controller.
- No ISA bus!

810 is an inexpensive chip set built on the BX technology. However, the new memory bus will come in other chip sets as well. The built-in audio-codec controller enables software audio and modem implementations. This meens that no sound card or modem is required. And finally we see the first attempt to produce modern PC's without the old ISA bus.

Chips in 810

The chip set consists of three chips:

82810	Graphics Memory Controller Hub	421 Ball Grid Array (BGA)
82801	Integrated Controller Hub	241 Ball Grid Array (BGA)

82802	Firmware Hub	32-pin PLCC or 40-pin TSOP

The Accelerated Hub Architecture

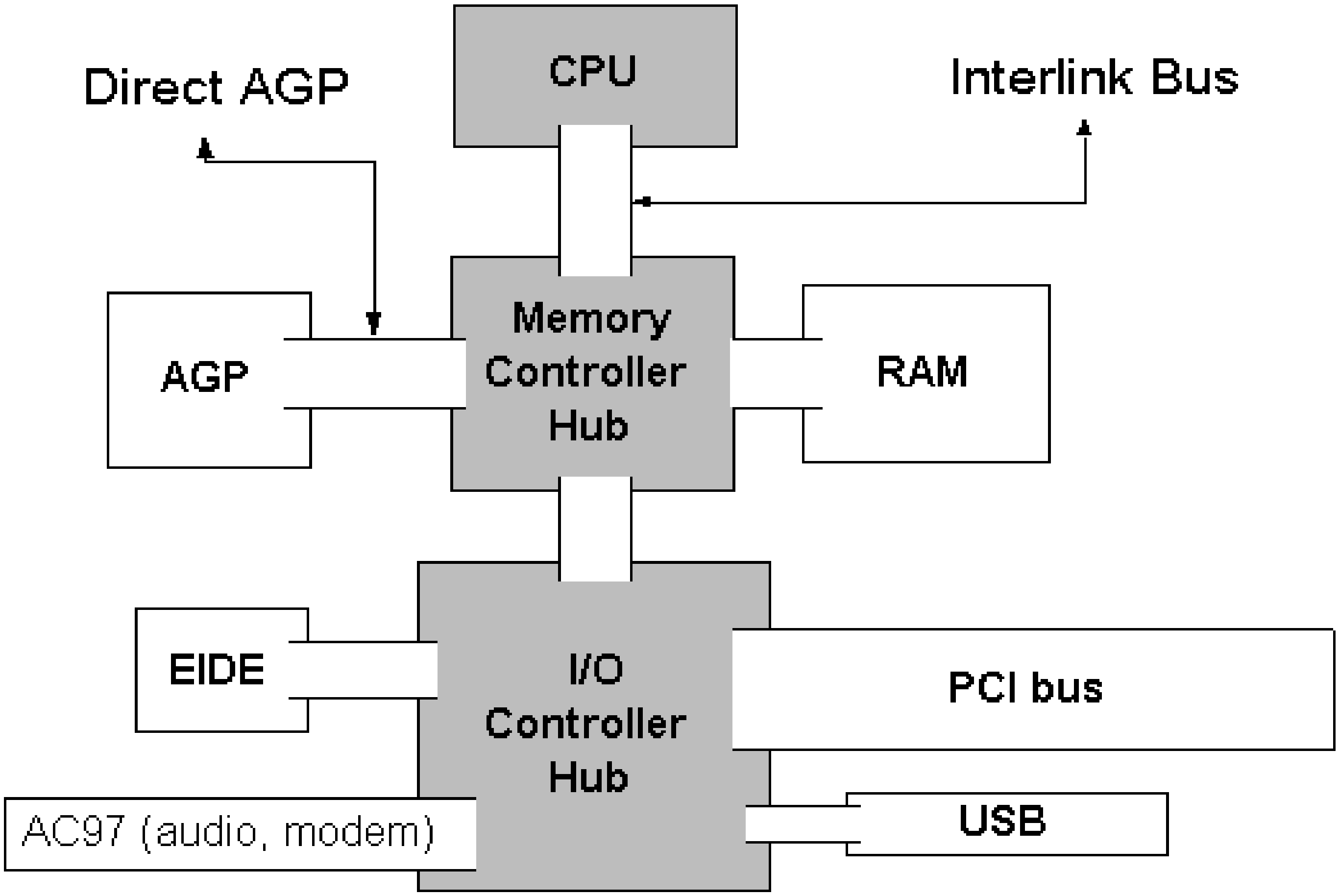
[top]

Usually we talk about north and south bridges in chip sets. These refer to the two controllers a chip set usually consists of. Intel replaces these terms with "hubs".

The new thing in this hub architecture is, that the two controllers not are connected by the PCI bus. Instead they connect via a new Interlink dedicated bus. This is a high speed bus, currently with twice the bandwidth of the PCI bus. This architecture resembles the new K7 Athlon point to point channel.

266 MB/sec

The interlink bus operates at 133 MHz in 2X mode. Being 64 bit wide this gives a bandwidth of 266 MB/sec (2 X 133.000.000 X 8 byte).



Also see the MCH below.

Graphics Memory Controller Hub

The 82810 Graphics Memory Controller Hub (GMCH) is a MCH "north bridge" including a graphics controller and using Direct AGP (integrated AGP, where the graphics controller is directly connected to the system RAM) operating at 100 MHz.

The 82810 chip features a "Hardware Motion Compensation" to improve soft DVD video and digital video out port for digital flat panel monitors. The graphics controller is a version of Intel's new model 752. Optional, the chip set can be equipped with a display cache of 4MB RAM to be used for "Z-buffering".

Dynamic Video Memory Technology (D.V.M.T.) is an architecture that offers good performance for the Value PC segment through efficient memory utilization and "Direct AGP". A new improved version of the SMBA (*Shared Memory Buffer Architecture*)used in earlier chip sets as VX. In the 810 chip set 11 MB system RAM is allocated to be used by the 3D-graphics controller as frame buffer, command buffer and Z-buffer.

82801 I/O Controller Hub

This "south bridge", the 82801 (ICH), employs an accelerated hub to give a direct connection from the graphics and memory to the integrated AC97 (Audio-Codec) controller, the IDE controllers, the dual USB ports, and the PCI bus. This promises increased I/O performance.

82802 Firmware Hub (FWH)

The 82802 Firmware Hub (FWH) stores system BIOS and video BIOS in a 4 Mbit EEPROM. In addition, the 82802 contains a hardware Random Number Generator (RNG), which (perhaps and in time) will enable better security, stronger encryption, and digital signing in the Internet.

AC97

The Integrated Audio-Codec 97 controller enables software audio and modem by using the processor to run sound and modem software. It will require software, but using this you need no modem or soundcard.

This feature is smart if you do not use audio or modem on a regular basis. It adds a heavy work to the CPU, which has to act as a modem and as a sound card beside its regular tasks.

Karbo's conclusion

[top]

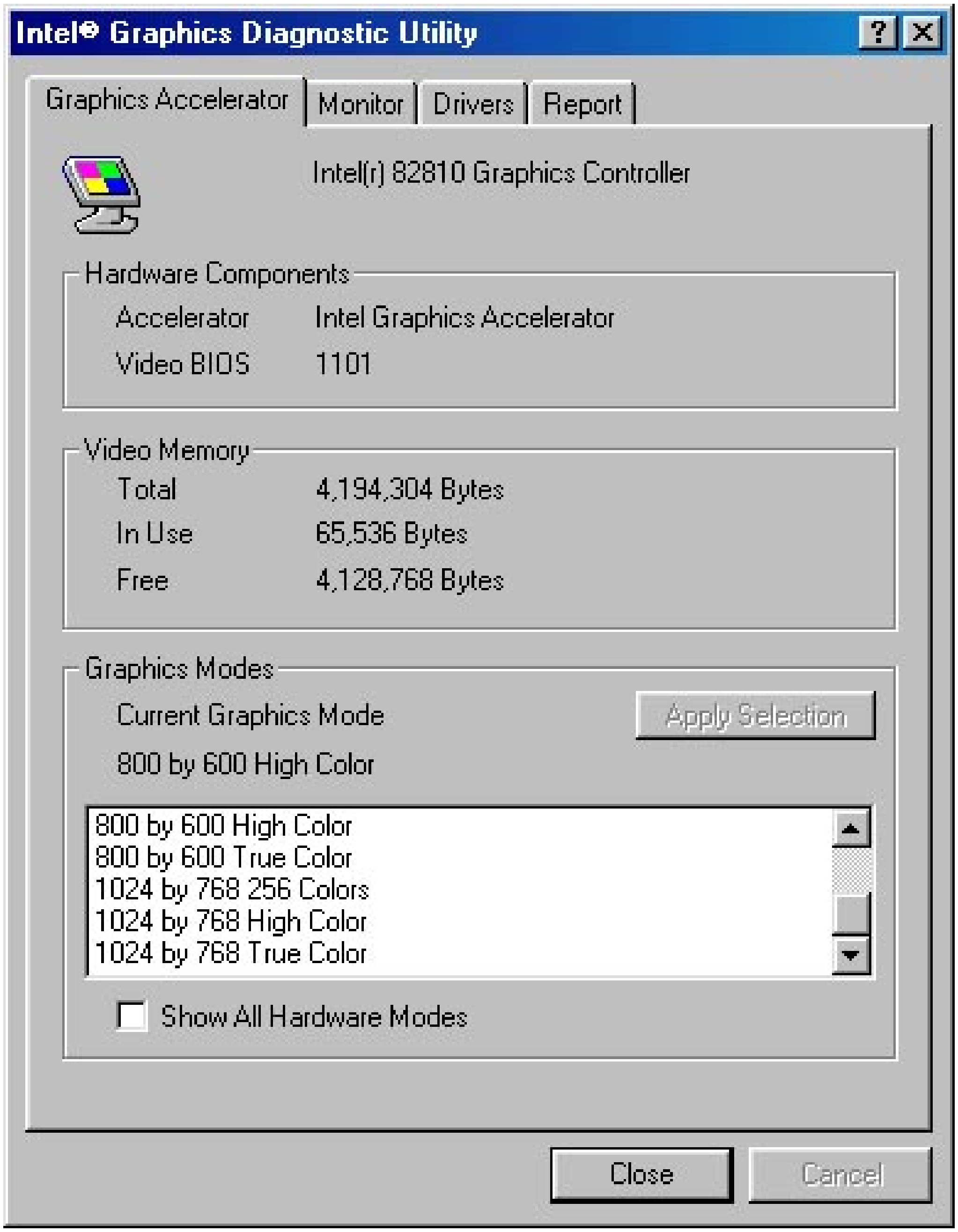
The 82810 controller represent a new generation of low-priced chip sets. I find these aspects interesting:

Integration of a powerful graphics accelerator

The RAMDAC is of 230 MHz giving a max. 2D-resolution of 1280 X 1024 pixels with 24 bit color depth and a refresh rate of 85 Hz. The graphics controller offers 3D acceleration with both DirextX and OpenGL support.

I found the performance to be quite OK for non-game use. The visual quality of the screen images seemed to match the out put from mid-range graphics adapters from ATI and Matrox. I could live with this graphics without any problems.

Here you see a dump from the Windows -driver that goes with the chip set:



Many users will not like that you cannot disable the graphics controller. So for gamers this chip set is no good. It was never really accepted by the motherboard manufactures, nor by the press. However, I liked it ...

The new support for software-based sound and modem

Will this work, and what are the consequences going to be?

No ISA bus. This is good.we shall soon see a lot more USB-based devices. And it will become very easy to built small, inexpensive, and elegant PCs using all the integrated hardware and only connecting external units using the very handy USB cabling.

100 MHz support

The 810 chip set is made for Intel Celeron processors. But so far these processors only work with a system bus frequency of 66 MHz. Why does the chip set then support 100 MHz? The obvious reason is that Intel planed to move the Pentium III processors to a Socket 370 platform. And the Celerons comes operating at a 100 MHz bus frequency. This is good news, and it all happened in 2000 and 2001.

Go for i815E

After the arrival of i815E in June 2000, that is the chip set to go for. It holds all the nice features from i810 plus a lot of great news.

- Next page
- Previous page

Learn more **[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.07

The intel i820 "Camino" chip set

The contents:

- Introduction
- The chips in i820

- Next page
- Previous page

Intro to Intel 820

[top]

In 1999 the new generation of high-end Intel chip set was code named "Camino".

This i820 "Camino" chip sets was originally set for debut in May or June, but were delayed. The Rambus technology was problematic.

The i820 chip set was finally to be launched September 27th, 1999 but was delayed again. This time motherboards with more than two SDRAM sockets did not work. And there have been so many problems with this chip set, which soon became a nightmare for Intel. In the press the situation was described as "Caminogate"

Anyway, we have to look into the architecture. We find:

- New hub-based architecture
- 133 MHz FSB
- Rambus
- Up to 1GB RAM
- AGP4X
- ATA66

The chipset was designed for high-end use with Pentium III processors.

The chips in i820

The chip set consists of two main controllers:

- The 82820 Memory Controller Hub.
- The 82801 I/O Controller Hub (ICH)

The (MCH) provides the CPU interface, DRAM interface, and AGP interface. This chip is found in two versions: A single processor (82820) or a dual processorchip (82820DP).

The ICHmakes a direct connection from the graphics and memory to the integrated AC97 controller, the ATA66 controller, dual USB ports, and PCI add-in cards.

Besides the two main controllers you also find:

- 82380AB PCI-ISA Bridge
- 82802 Firmware Hub

- Next page
- Previous page

Learn more **[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.08

Intel i820 "Camino", continued

The contents:

- The 82802 Firmware Hub and BIOS updates
- The Memory Controller Hub (MCH)

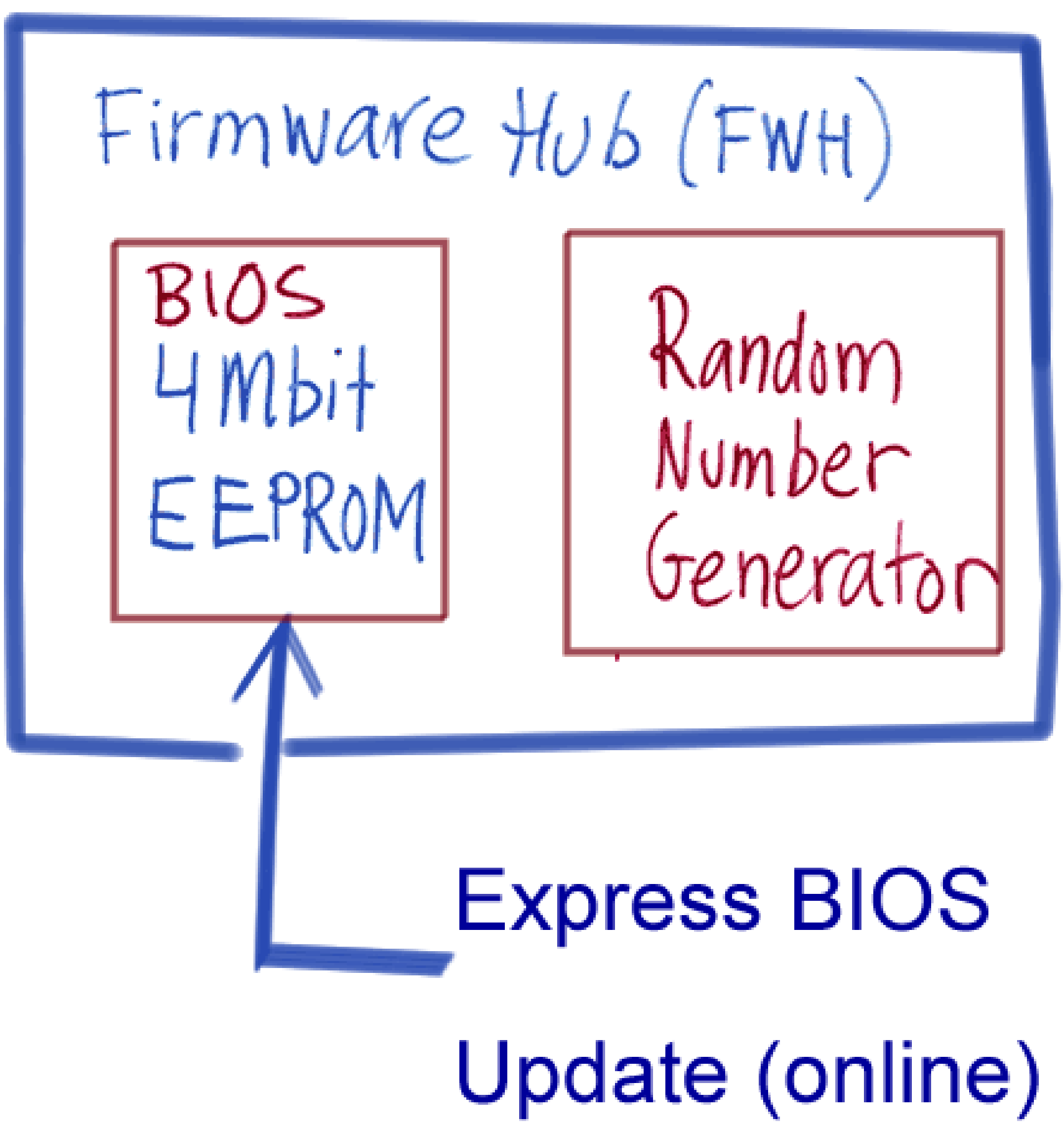
- Next page
- Previous page

The 82802 Firmware Hub and BIOS updates

The 82802 Firmware Hub (FWH) stores motherboard BIOS in a 4 Mbit EEPROM. In addition, the 82802 contains a hardware Random Number Generator (RNG), which (perhaps and in time) will enable better security, stronger encryption, and digital signing on the Internet.

Intel has succeeded in setting up a fine system for BIOS updates, using the FWH. Traditionally BIOS was updated using a boot diskette, but since many modern PC systems do not have a floppy disk, it has become a problem to update BIOS on new machines.

Intel choose to place their BIOS-Update-Patch on the Internet. You download the 1.2 MB file "Express BIOS Update" and execute it under Windows . After re-boot, your i820-based motherboard is updated with new BIOS. This is really smart!



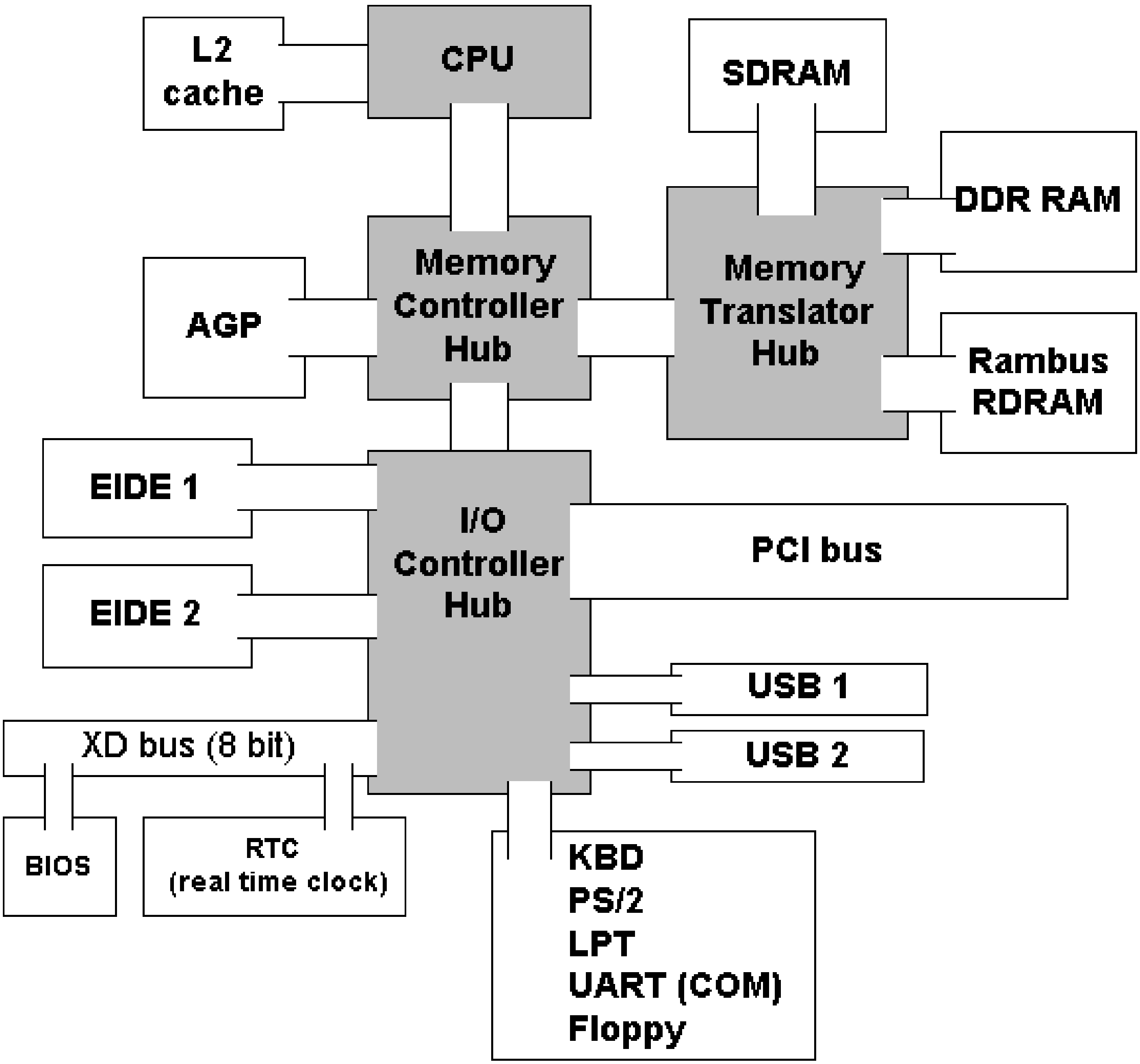
The new BIOS include a much-wanted feature: Rapid BIOS Boot (RBB). It speeds up the POST sequency radically, hence reducing the boot time with some 15 - 30 seconds. This is especially designed to work with Windows ME.

However, the first versions of new BIOS was a failure - soundscards did not function after the update ... Intel really has had a hard time with this chip set.

The Memory Controller Hub (MCH)

Central in the chip set is the Memory Controller Hub. This device controls the data flow to and from RAM. The idea is to assign maybe two or four RAM channels for higher bandwidth.

Here is my early guess on the design:



The idea of using a Memory Translator Hub as you see above, was that it would enable Intel to produce boards using PC133 RAM as well as RDRAM. A sound idea, indeed. DDR was never planned, since Intel is not allowed using this type of RAM, according to their agreement with Rambus (covering the years 2000-2002).

- Next page
- Previous page

Learn more	**[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.09

Intel i820 "Camino" continued

The contents:

- Caminogate: No PC133 RAM, no MTH
- The hub-based architecture

- Next page
- Previous page

Caminogate: No PC133 RAM, no MTH

Today we know the i820 chip set is used only with RDRAM; Intel failed to produce a reliable MTH (or *Memory Conversion Hub* (MCH), as it later was named).

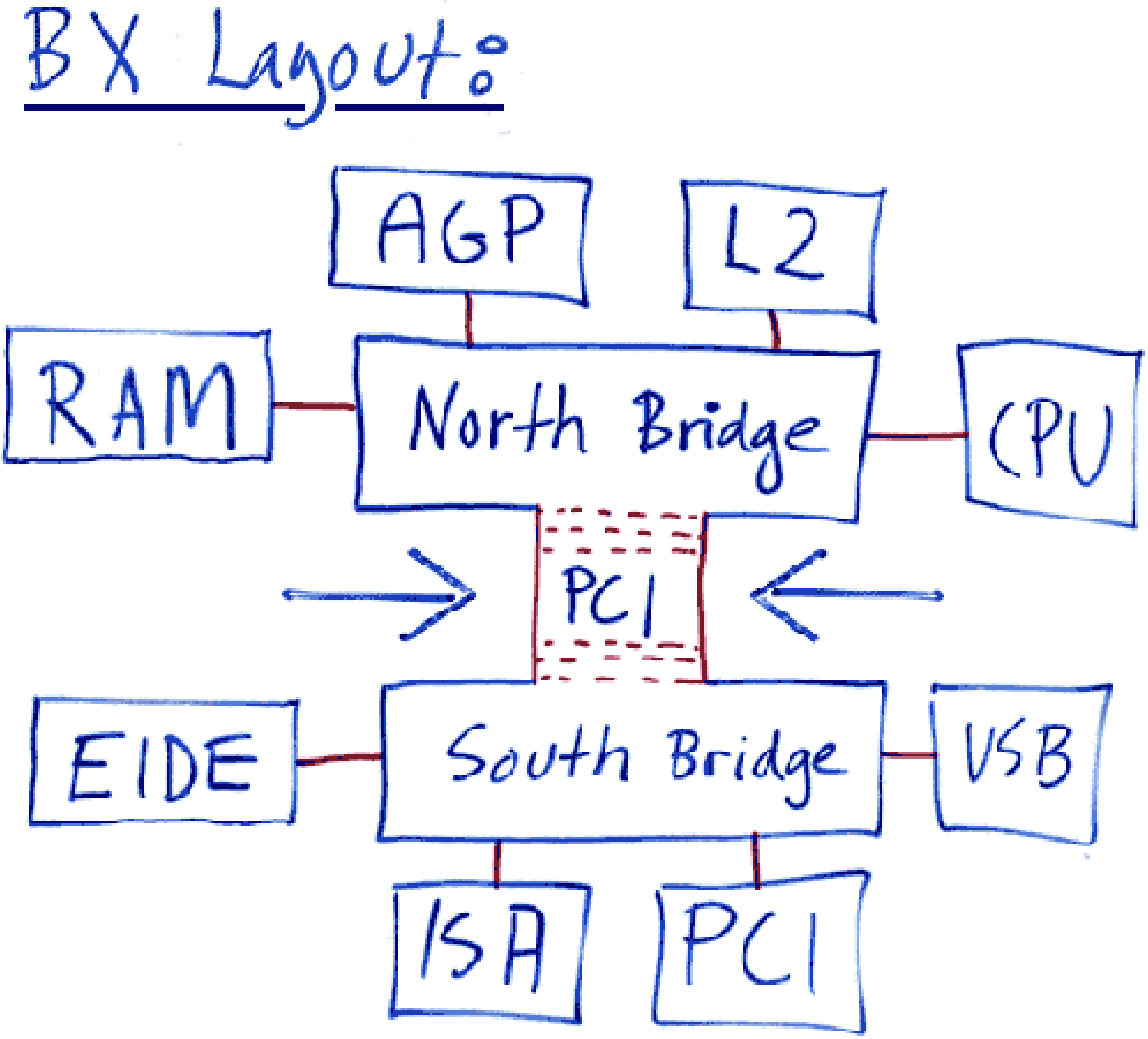
One of the problems appeared to be that the Serial Presence Detect (SPD) chip included on SDRAM DIMMS was missing in some modules. This SPD was crucial to the MCH. A solution was to add a 150 ohm resistor between the Memory Translator Hub and the SDRAM ...

Wisely Intel finally gave up all this business. In the end they had to recall a million of Intel motherboards and had to give away RDRAM in large numbers to unlucky customers among buyers of other motherboard brands using the ill-faited i820 chip set. Here the ITH caused sudden reboots when PC133 RAM was installed.

However you may say that without support for either PC133 or PC2100 RAM, there is no big use for the i820 chip set.

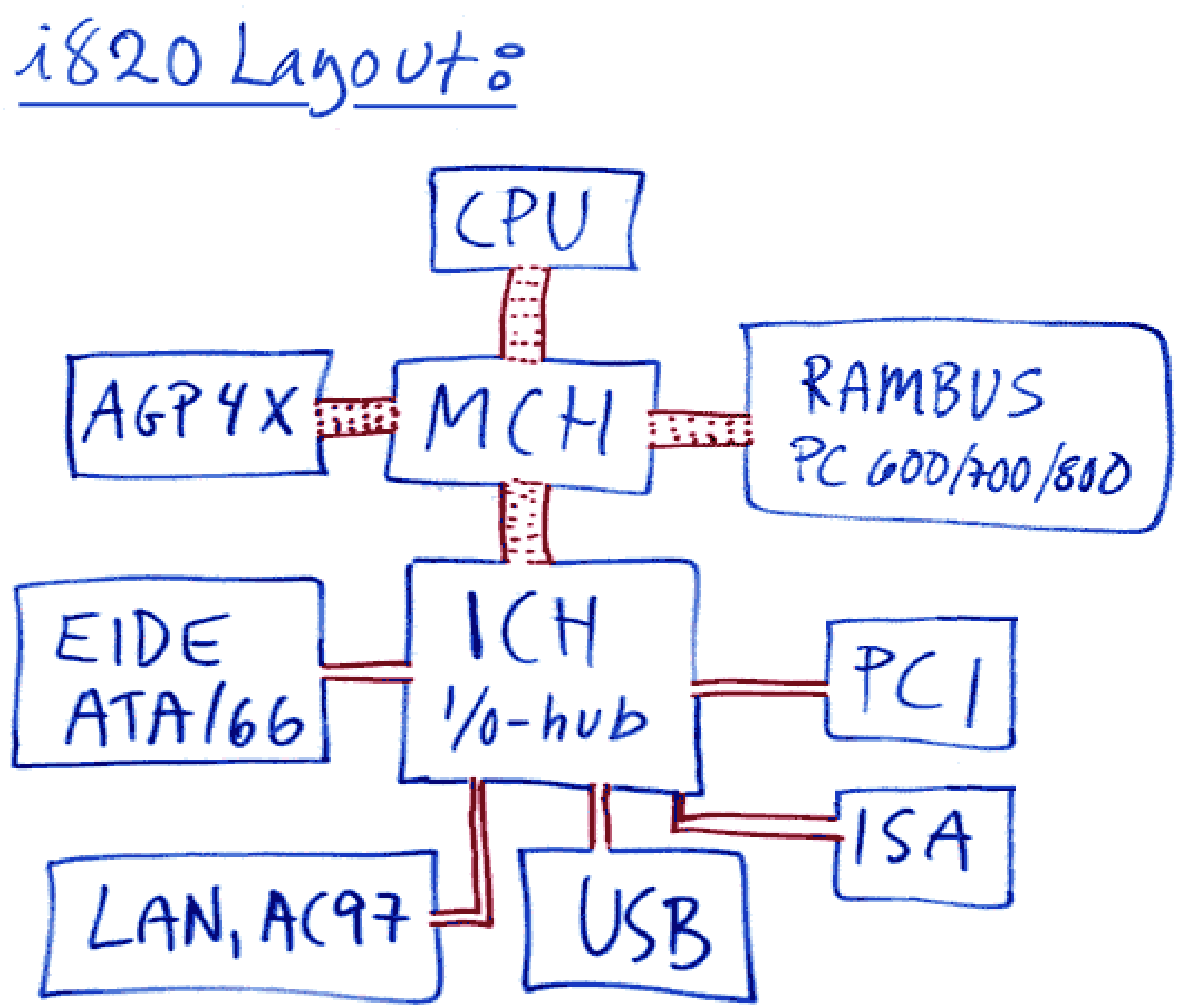
The hub-based architecture

In earlier designs (like BX) you had the two controllers united by the PCI channel:



This design but a heavy strain on the PCI bus, having a 133 MB/sec bandwidth. All data to between RAM and disks, network adapters and other I/O boards such as PCI-based graphics controllers had to pass through the PCI bus.

In the new design, we first saw within the i810 chip set, we have "hubs" instead of "bridges":



- Next page
- Previous page

Learn more **[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

KarbosGuide.com. Module 2d.10

Intel i820 "Camino" continued

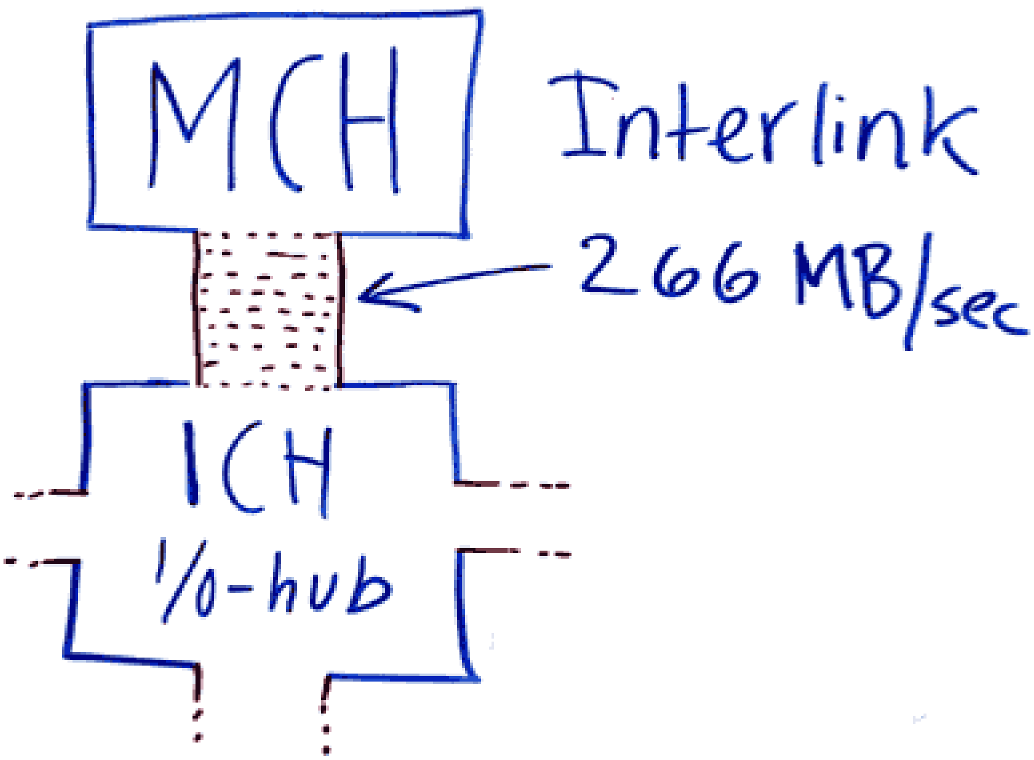
The contents:

- The Interlink channel
- AGP4X
- RDRAM

- Next page
- Previous page

Interlink

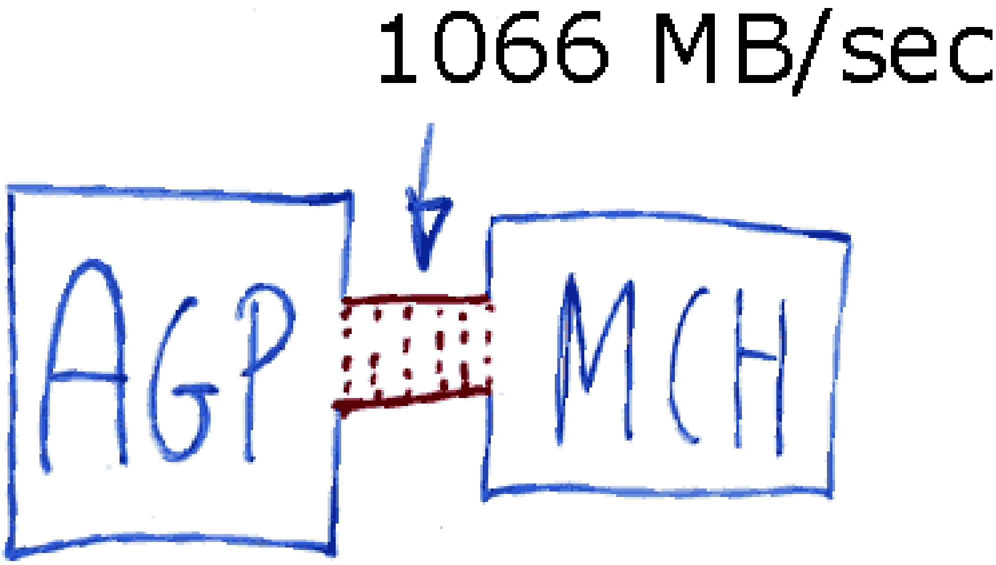
The two controllers are united by a new "Interlink" channel. It runs at 266 MB/sec:



The interlink bus operates at 133 MHz in a 2X mode making it 128bit wide. This gives a bandwidth of 266 MB/sec (133.000.000 X 128 / 8).

The 4X AGP

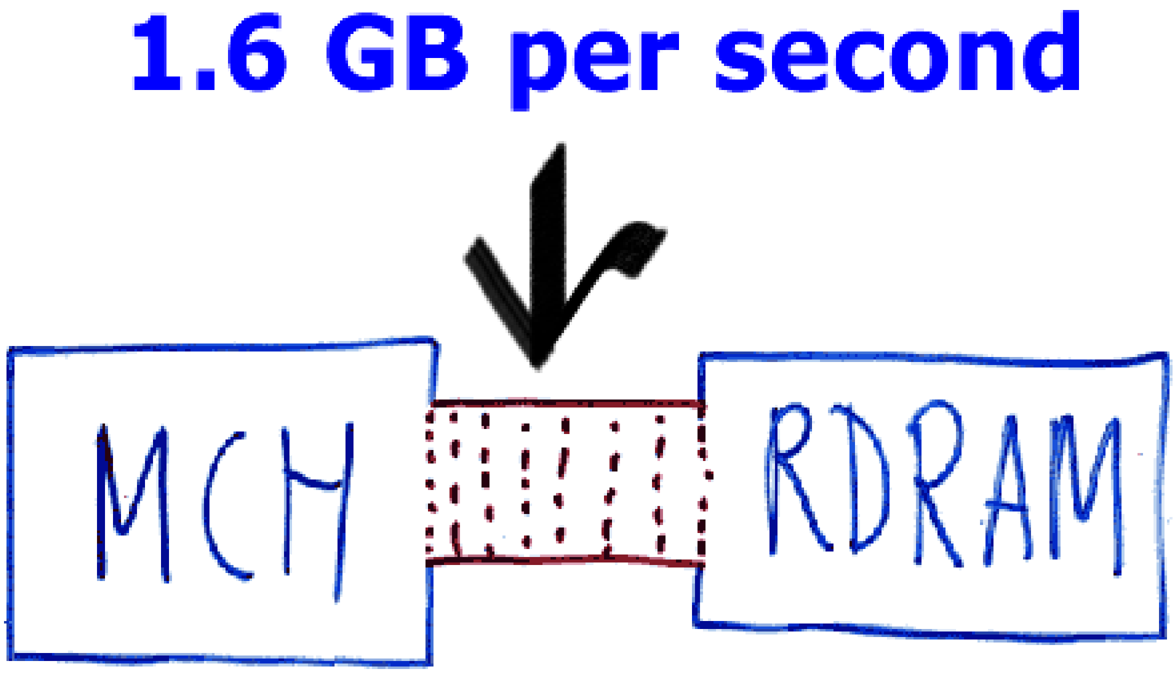
Intel was one of the first companies to implement AGP 4X in the chip set. Using AGP 4x, the bandwidth to the graphics subsystem has doubled from 533 MB/sec in AGP 2X to more than 1 GB/sec:



This is good for all gamers. 3D gaming needs a powerfull channel to RAM to produce high quality screen frames.

The RDRAM channel

The use of Rambus should provide the memory optimal bandwidth.The RDRAM supports PC600, PC700, and PC800, delivering 1.6 GB/s of memory bandwidth in the PC800 - twice the peak memory bandwidth of 100MHz SDRAM systems.



However the price of RDRAM was extremely high in the first year of i820. Therefore, i820 never became popular.

- Next page
- Previous page

Learn more **[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.11

Intel i820 "Camino" continued

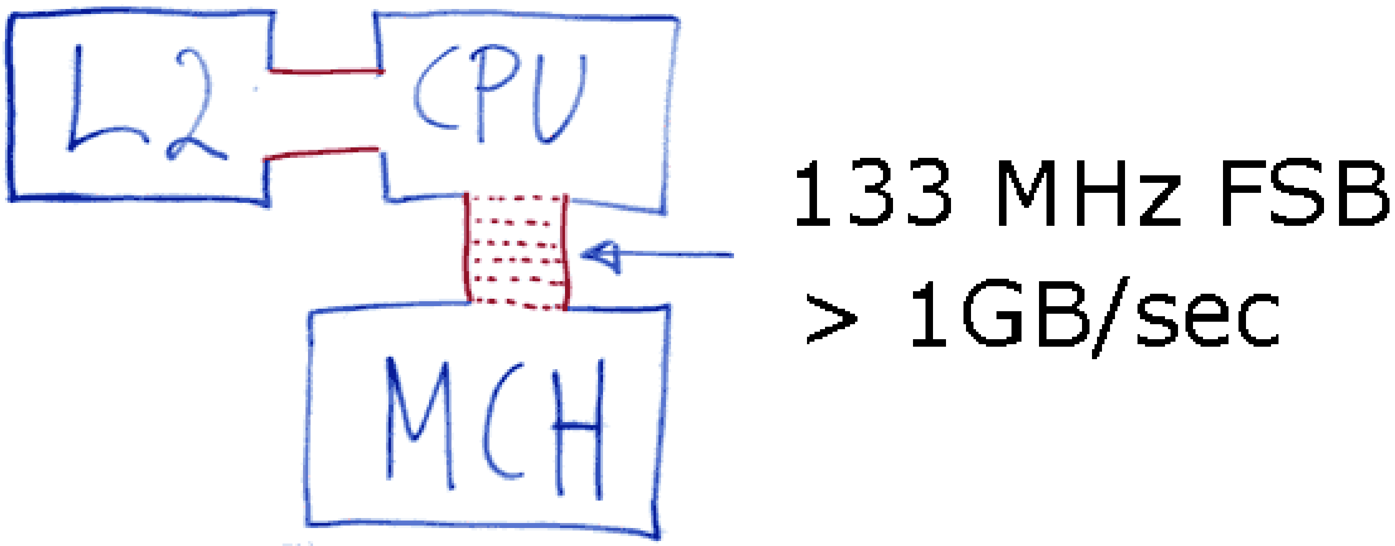
The contents:

- The 133 MHz FSB
- Karbo's conclusion

- Next page
- Previous page

The 133 MHz FSB

The Front Side Bus is the bus connecting the CPU to the MCH. In older systems this bus was the system bus. The FSB is dependant on the CPU; the older Pentium IIIs ran on a 100 MHz FSB. But the i820 chip set was intended to be used with the newer Pentium III "Coppermine", which operates at a multiply of 133 MHz.



The increase from a 100 to a 133 MHz FSB is not as important as it sounds. This is due to the fact that the greatest work is performed inside the CPU and between its L1 and L2 cache, where a powerfull bandwidth really is essential.

The data intensity between CPU and RAM is less demanding. However a 133 MHz FSB will give better performance when working with lagre data amounts (using Photoshop for instance).

A conclusion

We have not tested a 820-based board ourselves. However, we did like the 810-based board we used for a while earlier. Not for gaming - but office use, it was and it is a fine little chip set.

The i820 thing has been a disaster for Intel. First it was discovered, that you only could use 2 out of 3 RIMM-sockets. Then the MTH did not work. Today some analysts believe, that Intel will stop producing chip sets after all this chaos. I understand them, but it will be a pity. Intel used to produce excellent chip sets, and they should continue.

The i820 chip set should have been brought out of circulation a long time ago. The new i815 chip set could probably take over, so the venerable BX set could retire.

We see two points in which Intel has failed, and this could easily have been avoided:

First of all the company should not commit themselves to an uncertain technology like RAMBUS as they did. Intel produces great CPUs. The customers might use them with Rambus or SDRAM or DDRRAM or EXP3RAM or whatever the industry might come up with.

We believe that Intel trusted to much in own powers; they wanted to "force" the market into a certain behavior. We do not like that; it is against the free will and intelligence of users all over the world. Last time Intel tried this attitude was in 1997 when skipping Socket 7 in favour of Slot 1. A clupmsy design, which they now have abandoned themselves.

The second lesson is that Intel never again should market *untested products*. We have seen this several times during the "Caminogate" affair. Maybe Intel has felt threathend by AMD's succesfull Athlon project. But there is no excuse for marketing lousy untested products. Both Intel and Microsoft should learn from this.

- Next page
- Previous page

Learn more [top]

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.12

Intel i820E "Camino II"

The contents:

- I820E

- Next page
- Previous page

820e

In June 2000 Intel's revamped i820E chipset started shipping.

Featuring a brand new ICH2 I/O controller hub (i82801BA), the new chip set holds:

- Four USB ports using dual controllers
- Integrated LAN controller
- Dual Ultra ATA/100 IDE controllers
- Dolby surround-capable six-channel audio.

The i820e uses the same RDRAM Memory Controller Hub (MCH) found on the original 820.

This indicates that there will still be a maximum of two RIMM slots on the motherboards. However other reports mention 4 RIMMs

The new ICH2 offers a bandwidth of 2,4 MB/sec across four USB ports. This is a good thing as the ATA/100 support is!

The enhanced AC'97 interface should support 6 channelsfull surround-sound for DVD Dolby Digital audio.

A future for this set?

It is too early to see if there should be a hope for the 820E chipset. If you have read all the previous pages, you know about the troubles Intel and millions of users have had with the I820 set. We doubt that the market will forget this.

Our advise should be to forget 820 and go for the later i815! However, Intel have planned a new version, "Camino III", scheduled to launch early in 2001.

Another chipset to come after i820 is the i850 "Tehama" which is to be used with Pentium 4. This is also a RAMBUS-only solution, using dual RIMM channels for better bandwidth.

- Next page
- Previous page

Learn more [top]

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.13a

Intel i815E "Solano"

The contents:

- Introduction
- The chips in i815
- The features

- Next page
- Previous page

Intro to Intel 815

The Intel 815/815E chipsets from June 2000 are great products. After all the troublesome affairs with i820 it seemed that Intel finally was back in the chip set business.

Intel i815/815E is an update of the succesfull 810 chip set. It holds an integrated graphics adapter as well as a lot of new functions. It is intended to replace the workhorse BX chip set, which is to be phased out late in 2000. In i815 Intel finally supports PC133 RAM!

The only thing I do not understand is why Intel come up with two flavours of the chip set.

The i815E is the only one, they should sell. Why market both a fine and a lousy product?

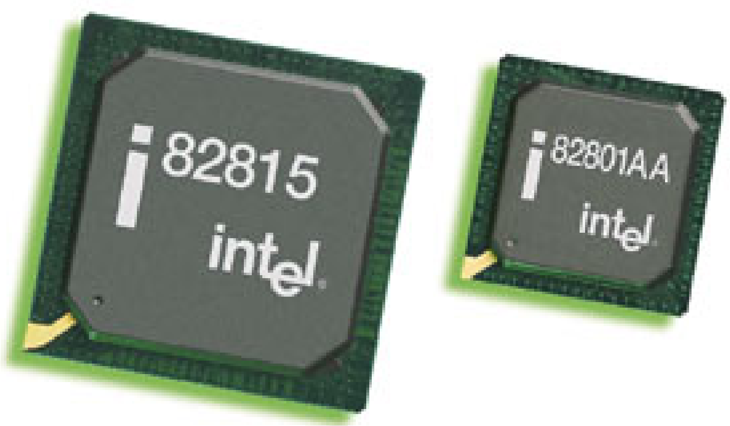
The chips in i815

The chip set is designed using the same hub-based layout as the former i810 and i820 chip sets.

As in i810 we find a memory controlling hub with integrated graphics (GMCH:"Graphics and AGP Memory Controller Hub") and an I/O-hub (the ICH2):

- 82815 (GMCH)544 Ball Grid Array (BGA)
- 82801BA (ICH2)360 Enhanced Ball Grid Array (EBGA)

The ICH2 is a chip also used in the I820E chipset. Here you see the two chips:



The features

The main idea with the 82815 Graphics and AGP Memory Controller Hub (GMCH) is to give an rather inexpensive motherboard with integrated graphics and high power memory management. The ICH2 on the other hand is fully updated with the lates improvements forming a sophisticated I/O hub. Totally we see:

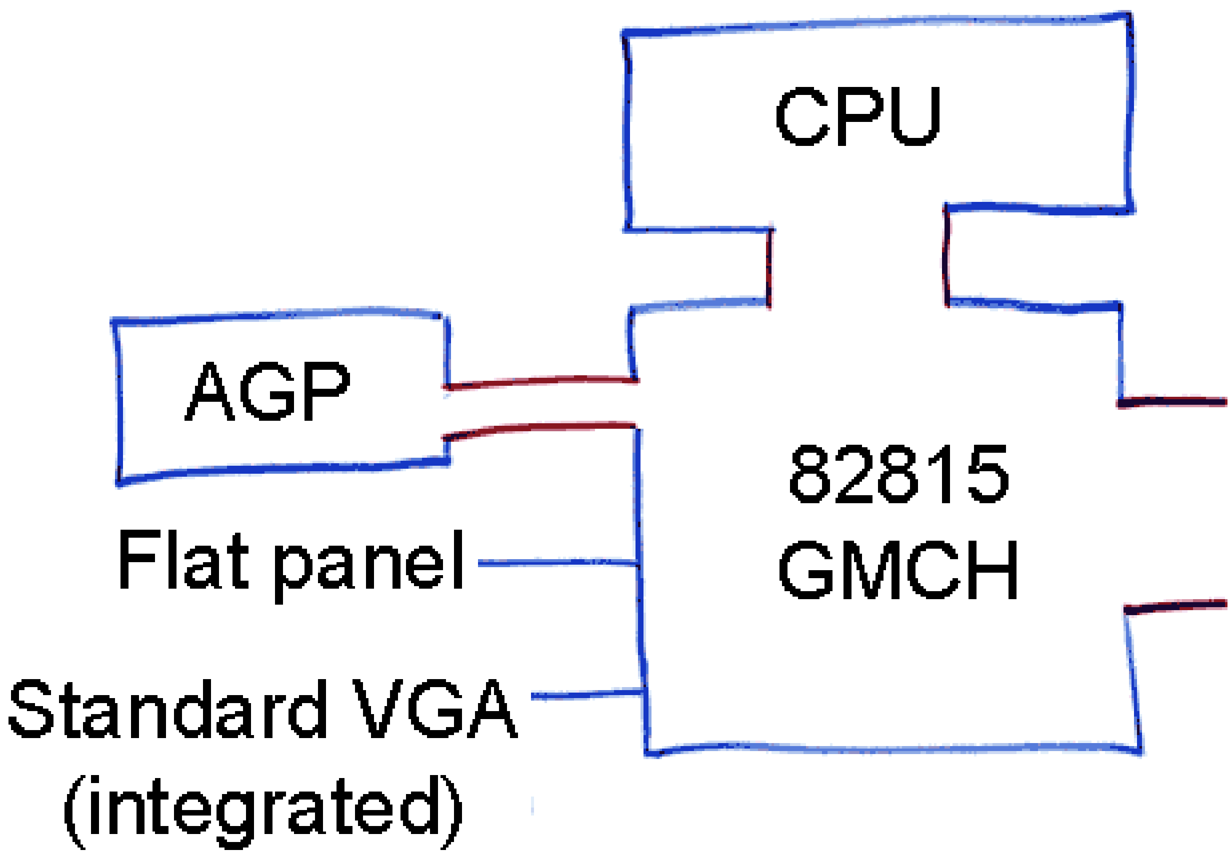
- Integrated VGA graphics
- Up to 512 MB of PC100 or PC133 SDRAM
- Asynchronous FSB
- ATA/100 interface

- Dual USB root hub

The graphics part

We have not tested the chip set. However we were (contrary to many others) very pleased with the 2D graphics performance of the i810 chipset. It works very well, and we expect the same to be the case with this chip set.

One big improvement has been made compared to the i810; you can choose to disable the graphics engine and install your own adapter in the AGP slot. The chip set also supports TV out and digital out for flat panels:



You may choose to upgrade the integrated graphics by adding more RAM using a AIMM (AGP Inline Memory Module). However, do not expect to do 3D-gaming using the integrated graphics of i815. Tests shows that high-end graphics boards work at 6-8 times speedier than this one!

The enhanced 82801BA I/O Controller Hub (ICH2)

This chips is also found in the i820E chipset. It delivers twice the I/O bandwidth as traditional bridge architecture, using an interlink connection to the GMCH chip as in i810.

The two USB controllers double the bandwidth to 2.4 MBps across four USB ports.

AC97 audio supports full surround sound with up to 6 channels and a soft modem implementation.

The integrated LAN is used for three networking environments (1Mbps, 10/100Mbps LAN and managed 10/100Mbps LAN). This means that we do not need an ethernet adapter, it is all in the chipset!s

ATA/100 is a great new interface giving a noticeable improvement to disk performance if the disk is build to ATA/100. Otherwise the interface is fully downwards compatible.

- Next page
- Previous page

Learn more	**[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 2d.13b

Intel i815E "Solano", continued ...

The contents:

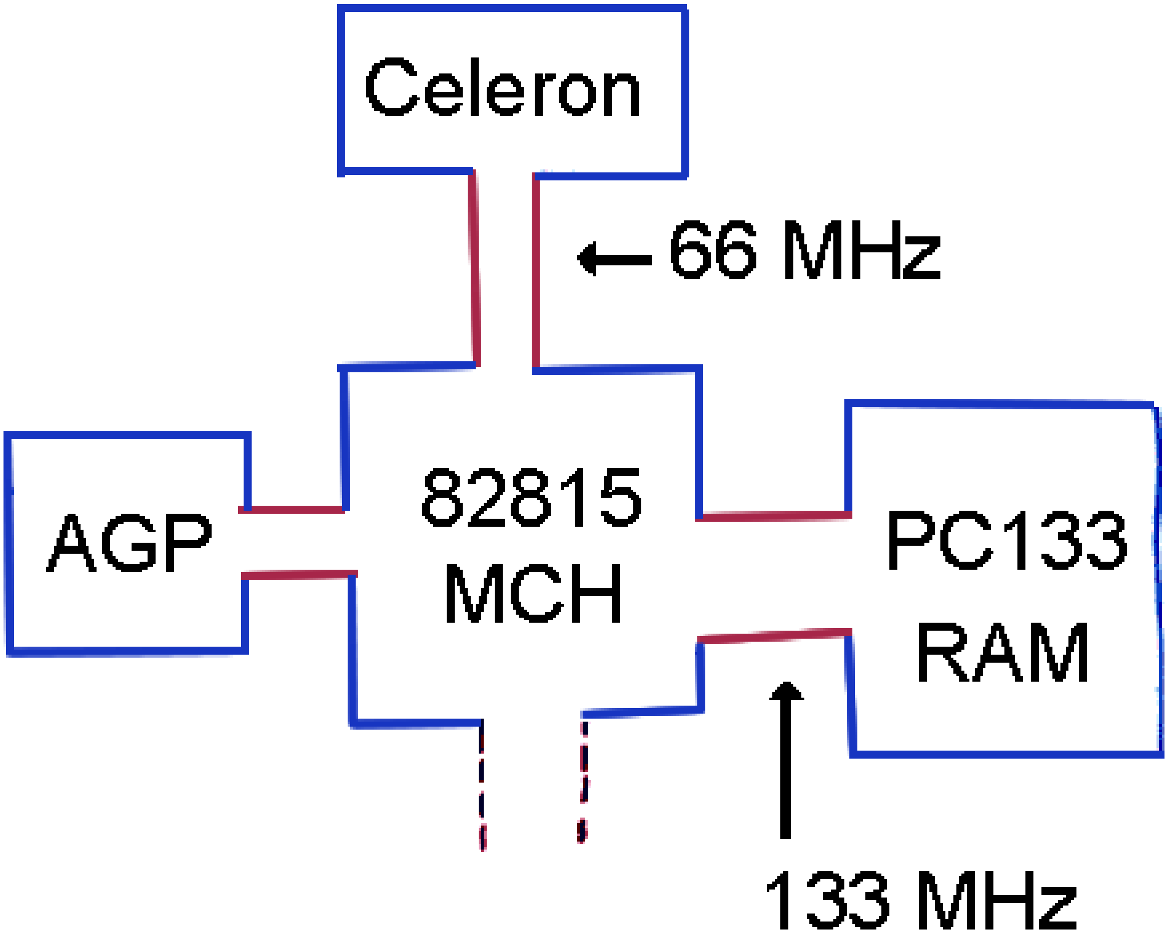
- Asynchronous FSB
- A conclusion
- Next chip set: "Almador"

- Next page
- Previous page

The Asynchronous FSB

A nice new detail in the i815E chip set is that the clocks of the FSB and the RAM can operate independently.

This way you can use PC133 RAM together with a Celeron with its mere 66 MHz bus:

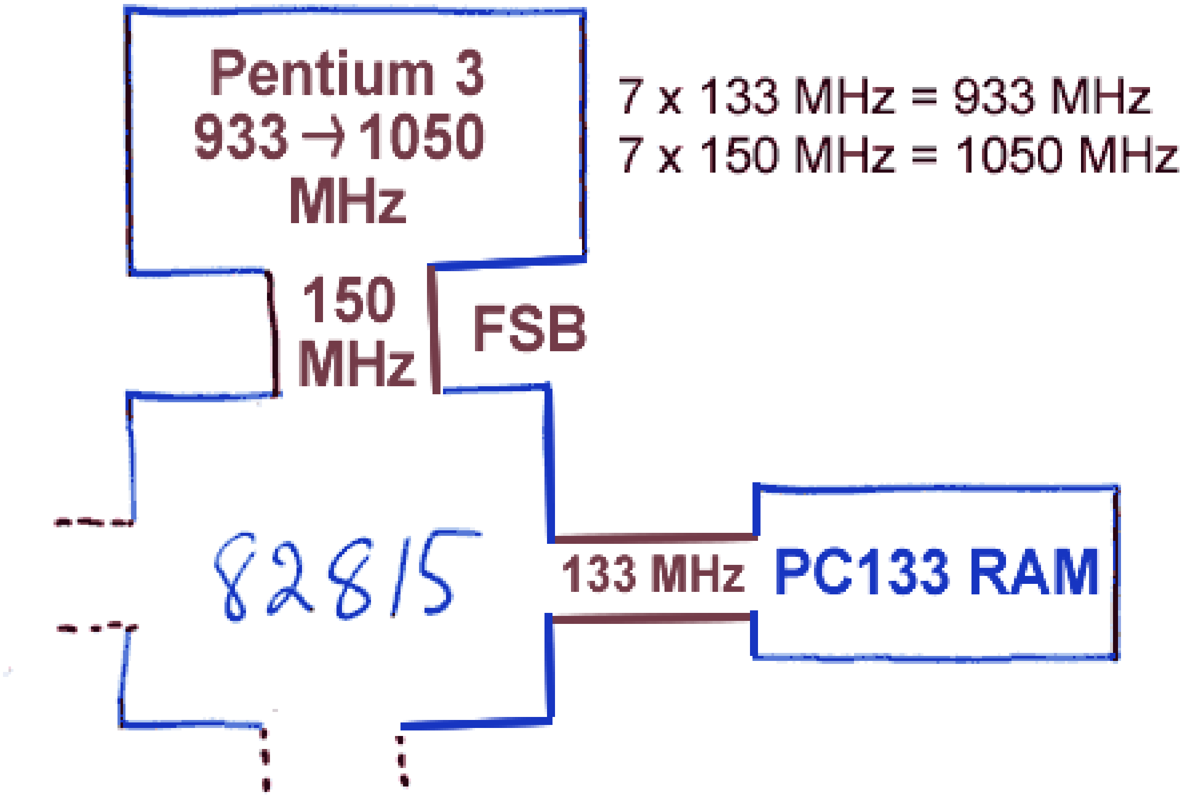


You may also reuse old PC100 RAM for a while before updating it to PC133 RAM.

Overclocking

We do not overclock our PCs anymore in this company. However, looking at the asynchronity of FSB and RAM, it brings into one's mind, that this must be very fancy for overclocking.

With a i815-based motherboard, you can heat up the FSB without any impact on RAM (or PCI) speed. With proper cooling I am pretty sure that a socketed Pentium III-933 CPU will function at 1.05 GHz:



A conclusion

From the details described here, we firmly believe that Intel is on the right track with the i815 chip set.

The BX set is worn out, and details like ATA/100 and asynchronous FSB/RAM will be very much appreciated!

It appears that Intel have to do a revision of the graphics drivers. In the first version, no monitor would work at higher refresh rate than 60 Hz, if the monitor did not have its driver installed in Windows ! This is completely nuts. Often a monitor works fine with the standard Super VGA driver within Windows .

PC133 RAM is a good product at the time, and the first tests show that i815 with SDRAM performs better than i820 with RDRAM!

We strongly hope that Intel will redirect their strategy into using DDR in a future revision of this chip set (be it code name "Almador" or what ever).

The "Almador" set

The successor to i815 is codenamed "Almador" (maybe i817?). The intersting issue is that Intel has announced support for SDRAM.

Using DDR RAM, Intel will have a very powerfull platform for Pentium III. Also the integrated 3D graphics engine will benefit from the better RAM bandwidth.

The south bridge of this set will be the new ICH3 chip with these features:

- ATA/100
- Six USB ports of version 2.0
- Integrated LAN etc. like the ICH2 of i815.

- Next page
- Previous page

Learn more **[top]**

Read about the Pentium in module 3c

Read about the Pentium II's etc. in module 3e

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7a.4

Refreshing the screen image

The contents:

- Electronic beams
- High refresh rate
- The horizontal scan frequency

- Next page
- Previous page

Electronic beams

In traditional CRT monitors, the electron gun continually sends out very precisely aimed beams of electrons, moving from pixel to pixel. The beam actually flickers, as it sweeps the screen. Each dot on the screen receives a quick flash of electrons, before the beam moves on to the next dot. And the beam intensity is varied from dot to dot.

The phosphor coating on the screen has the peculiar ability to light up, when hit by electrons. But the light quickly fades away. In practice, the electron beam "visits" again, before there is any visible fading of the light.

The result is that the it looks to us as a steady screen image. But actually the pixels of the image flickers every time the electron beam hits the phosphor coated dots.

The screen works overtime

Typically, each pixel is hit 60, 70, 75, or 80 times per second. Thus, the electron gun must move extremely fast to make 18 million or more hits per second. If the image is refreshed 75 times per second, we talk about a refresh rate of 75 Hz.

The video card issues the refresh signals, thus controlling the refresh rate. Thus, the video card has to match the monitor, so the two units can interface with a suitable electronical signal.

Let us think of a monitor with a resolution of 1280 x 1024 and a refresh rate of 75 Hz. That requires the electron gun to make 98 million pixel hits per second! That screen works at a very hectic pace – which can sometimes result in beam contamination.

High refresh rate

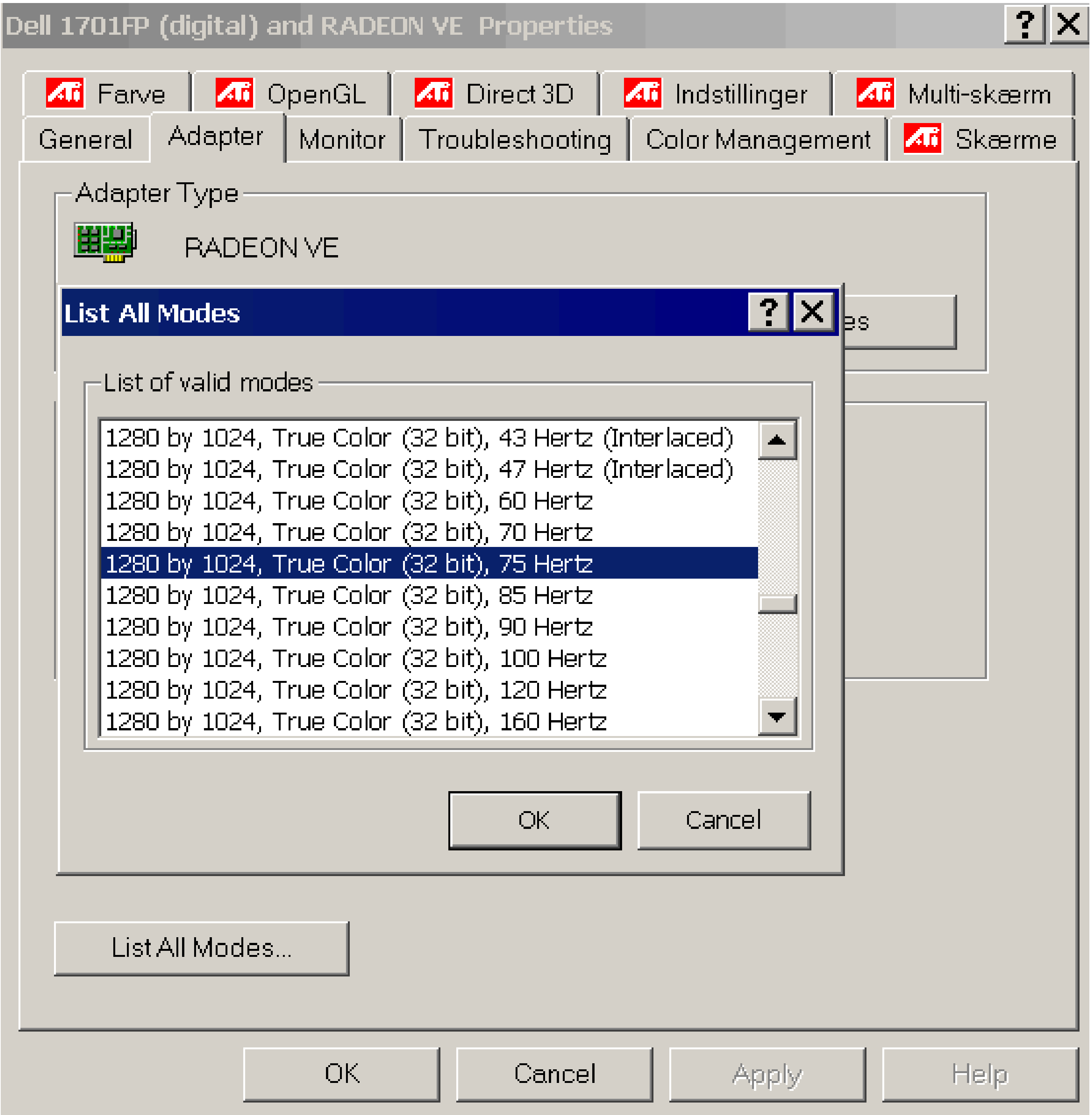
Top

The screen image appears more steady, the higher the refresh rate. You see the same in TV, where traditional sets have a refresh rate of only 50 Hz. Some manufacturers now produce TV sets with 100 Hz refresh rate. Some claim that they cannot notice the difference. However, once you have been used to 100 Hz refresh rate, it is uncomfortable to return to 50 Hz. Similarly with PC monitors, only here we have more options.

Older and inferior screens can only work at 60 Hz, which produces a low quality, flickering image which is not suitable for Windows . The general consensus is that 70 Hz produces an acceptable image.

I find 75 Hz acceptable, but 80 or 85 Hz may be better when you have to work many hours daily in front of the screen. You have to try these rates to find the best on your gear. Not unoften 75 Hz is the best refresh rate.

Here you see a dump from settings of a ATI Radeon graphics controller. It can deliver 11 different refresh rates (from 43 Hz to 160 Hz) in the the 1280 x 1024 resolution:



Note: refresh rate is also called vertical frequency or vertical refresh rate, but I have chosen to use the term refresh rate.

The higher the refresh rate, the better quality monitor you need. If you want both high resolution and high refresh rate, you will need both a high quality monitor and a high quality video card. The bigger the screen, the more it must be able to produce.

Screens can always run with higher refresh rates in lower resolutions. Here are three examples, showing how the screen performance drops with resolution.

CRT Screen	800 x 600	1024 x 768	1280 x 1024	1600 x 1200
Standard 15"	75 HZ	70 Hz	60 Hz	-
15" Trinitron	90 Hz	80 Hz	75 Hz	-
17" Trinitron	110 Hz	100 Hz	90 Hz	85 Hz

For the screen to deliver images at the desired refresh rate, both screen and video card must be matched to the correct specifications. Normally the CRT monitors have a feature called multisync. This means, that they automatically adapt to the signal coming from the video controller.

A good monitor usually is expensive. Cheap monitors may function at high refresh rates, but the image may not be good. Always check a new monitor visually before buying it.

And please remember: You will have the monitor for an average of 5 years. It will serve more than one PC, so buy quality!

More about screens

Top

Let us take a closer look at the monitors. If you read ads for monitors, you might see many hard to understand technical terms. They may mention many frequencies and dots and pitch?

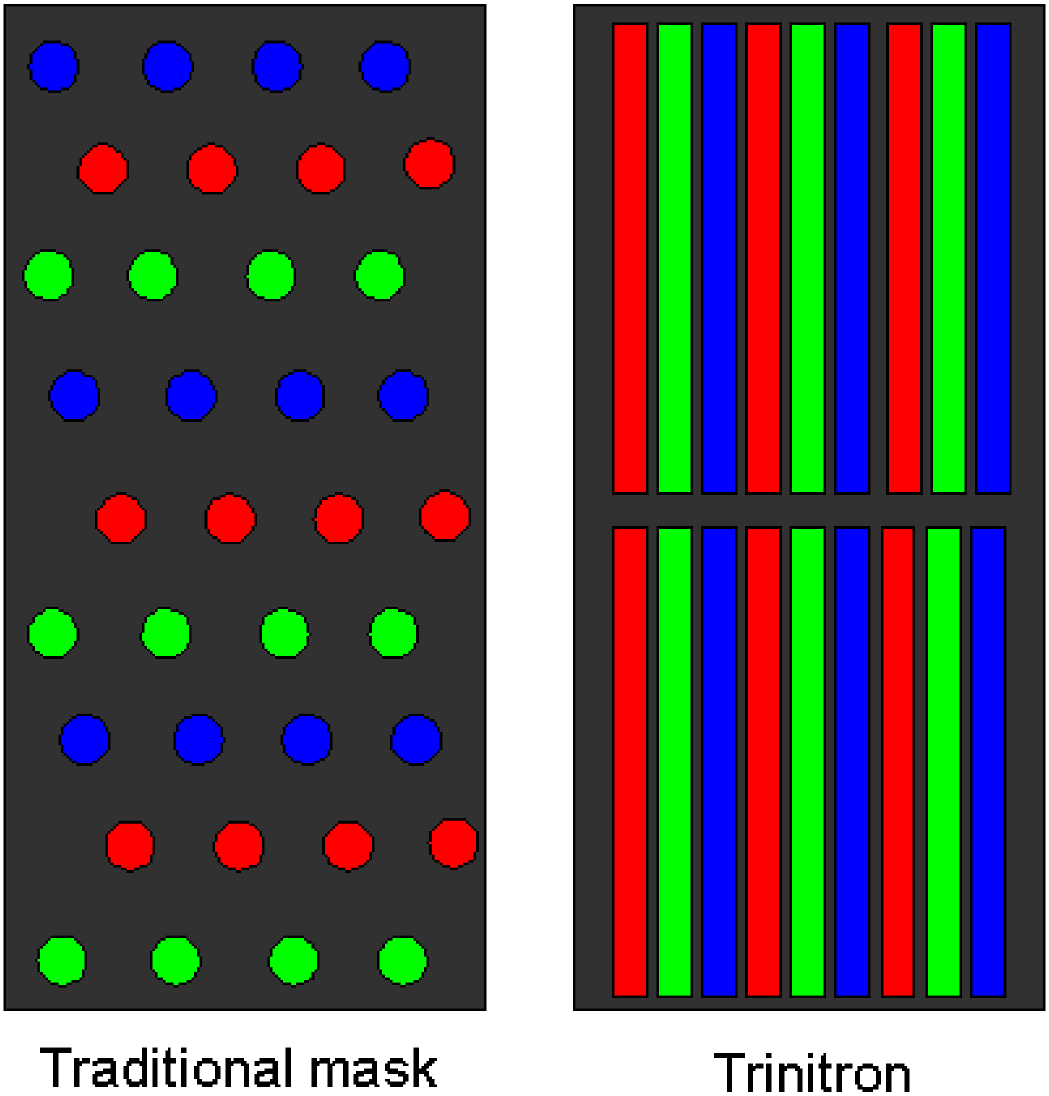
Note: In many ads, these terms (frequencies, etc.) can appear mixed and unclear. Therefore, be critical when you read monitor data.

Trinitron or Invar

Top

When we talk about traditional CRT monitors, there are two primary types of tubes. The best use the so called Trinitron tube. That is a technological principle, which was patented by the Sony company. Since the patent has expired, there are now some clones (ChromaClear, SonicTron etc.).

In the Trinitron screens, the light sensitive pixels on the inside of the tube are placed in a vertical grid, while traditional screens have round masks for the color dots. With the grid mask, you can achieve denser coverage and thus more color saturated images. Here is an attempt to illustrate the difference between those masks:



The Trinitron screens are generally very high quality. Since the Trinitron tube is more expensive than the traditional Invar tubes, manufacturers also include better control electronics in the Trinitron tubes. That increases their price somewhat, but that money is well spent!

The only disadvantage of the Trinitron (besides price) is the thin lines, which run across the screen. They are visible wires, which contain a grid. In daily work, you will not notice them, but rather enjoy the pleasure of an extremely fine and sharp image.

Invar for contrast

The traditional screen can provide more contrast than the Trinitron screens, which is important in some technical applications. But for ordinary use – in home and offices, where you would typically choose 17" or 19" screens – the Trinitron screens are an obvious choice. Of course, they cost a little more than traditional types, but there is a marked difference in the visible quality. You will experience a much better screen image with a Trinitron tube, no doubt about that! But the best is a TFT display, as I'll show you later.

The horizontal scan frequency

Top

The most important factors are maximum resolution and refresh rate. The screen must be able to deliver an image in a suitable resolution (depending on screen size) and at a good refresh rate (75 Hz or more). The screen can display many different image types – in various resolutions and refresh rates. The interesting point is the *maximum* refresh rate at different resolutions.

These data are often reported together in a number, called the *horizontal scan frequency.* The number is measured in KHz and it is very important. Basically, the horizontal scan frequency is calculated from resolution and refresh rate. As an example, an 800 x 600 resolution at 75 Hz gives a horizontal scan frequency of 60 KHz. You cannot calculate the number yourself. Also it varies slightly from screen to screen.

Here are examples of horizontal scan frequency. As I said, the numbers can vary slightly from screen to screen, but they are still in the same ball park:

Resolution	Refresh rate	Horizontal scan frequency
640 x 480	60 Hz	31.5 KHz
640 x 480	72 Hz	37.8 KHz
800 x 600	75 Hz	46.9 KHz
800 x 600	85 Hz	53.7 KHz
1024 x 768	75 Hz	60.0 KHz
1024 x 768	85 Hz	68.8 KHz
1152 x 864	85 Hz	77.6 KHz
1280 x1024	75 Hz	80.0 KHz
1280 x 1024	85 Hz	91.2 KHz

Ususally you get the best performance using the highest refresh rate available. The resolution depends on screen size and user habits. In all cases, it would be foolish to run the screen at 31.5 KHz.

NOTE: Using a digital interface for a TFT monitor, there is no horizontal scan frequency to concern about!

- Next page
- Previous page

To learn more

Top

Read about video cards in Module 7b .

Read about sound cards in Module 7c .

Read about digital sound and music in Module 7d .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7a.5

Adjusting the monitor

The contents:

- Multi Sync
- Color adjustments
- Screen savers
- Environmental standards

- Next page
- Previous page

The multisync screen with digital control

Top

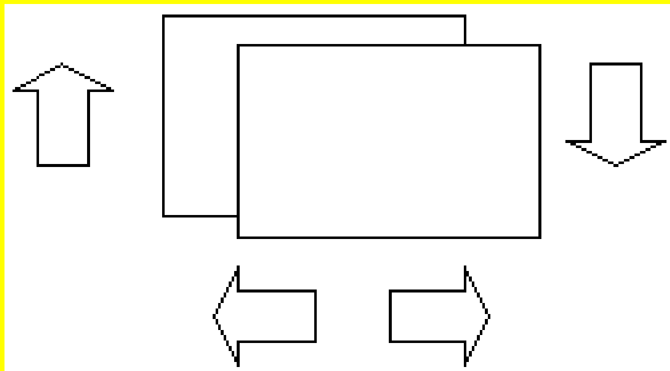
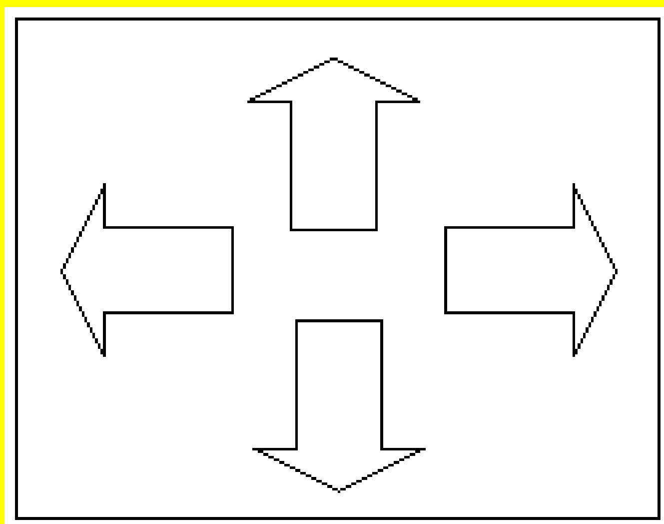
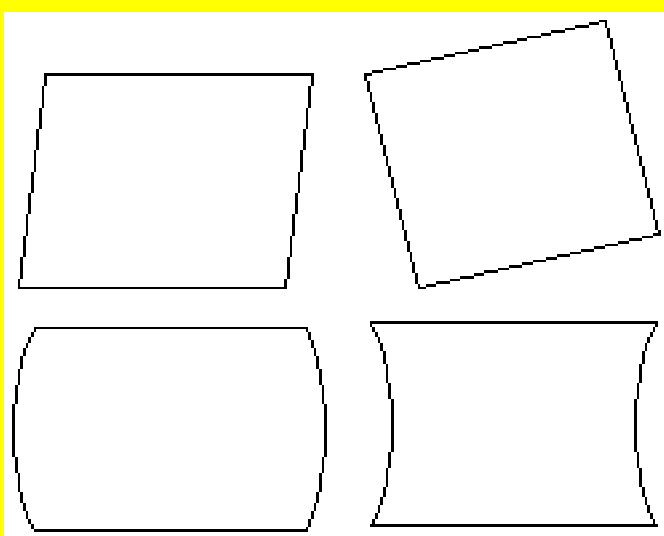
All modern screens are of the *multisync* type. This means, that the screen adjusts itself to the signals received. The individual model has a minimum and maximum horizontal scan frequency. As long as the signals are received within that spectrum, it adjusts itself to the signals.

When the screen receives signals at any given frequency, these signals must be adjusted to fill the screen 100%. That is done through the *digital controller* found in modern screens. Older screens would show a clear black border surrounding the image, whenever the resolution was changed to, lets say 800 x 600 and that is very irritating.

To enable adjustment to maximum screen utilization, the screen must have digital controls electronics. These adjustments are made on the screen control panel. We are talking about:

- Horizontal and vertical size, to have the image fill the maximum usable screen area.
- Horizontal and vertical positioning, to center the image.
- Compensation for trapezoid and pin cushioning.
- Colors and light intensity.

The adjustments can look like this:

Adjustments	**Symbols**
Horizontal and vertical position	
Horizontal and vertical size	
Trapezoid and pin cushioning	

Often screens are preset to a choice of different possible adjustments. In these preset conditions, the image will immediately appear perfect.

However, when you set up a monitor to work under non preset conditions you have to adjust the image yourself. Once that is done the monitor will remember your settings.

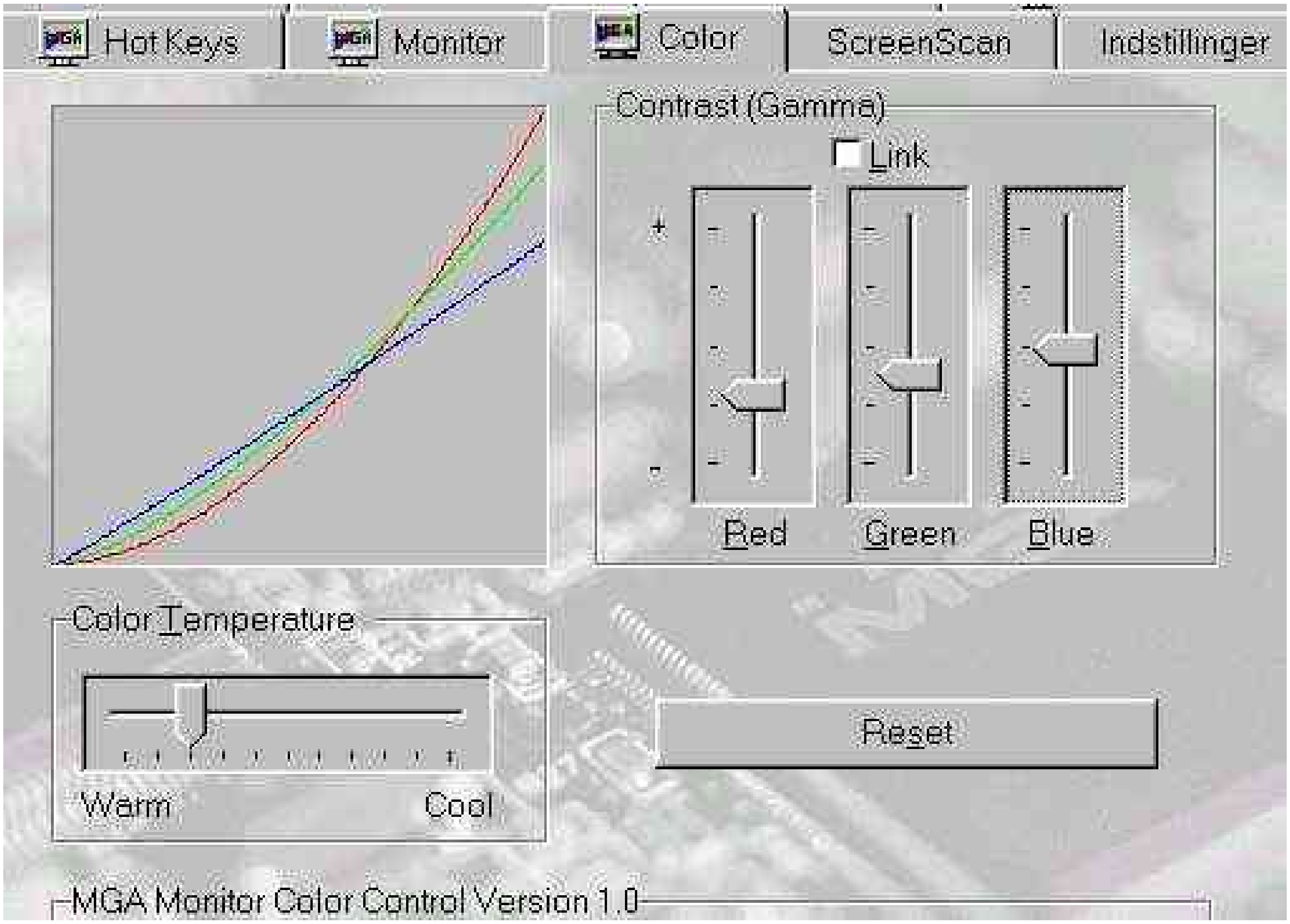
There are no international standards for the design of these digital controllers. They are quite different from monitor to monitor and not all easy to work with. However, working with adjustments is a minor problem, relative to other monitor qualities.

Color adjustments

Top

The screen can show the colors in different heat ranges. The better screens with digital controllers usually have at least two temperature ranges to choose from. I prefer 6500 degrees. 9300 is somewhat colder.

Similarly, some video cards can adjust the screen color temperature like Matrox here:



You should try the different color temperatures. They have a significant effect on the image appearance.

Aperture grill pitch

Top

Often you see the term dot pitch or *aperture grill* pitch. It is measured in millimeters. The numbers indicate the average distance between individual screen dots. The smaller the better. That provides a finer grain screen. For large CRT monitors (21"), the dot pitch can be 0.31 mm or 0.28 mm.

Otherwise, a dot pitch of 0.28 mm or 0.25 mm is considered sufficiently good for ordinary 15" and 17" screens. A few monitors offer 0.22 mm dot pitch.

Screen savers

Top

Early monitors had low quality phosphor coatings. That could cause a screen image to "burn-in" if left unattended. You could clearly see that in work places, where the PC was used for only one program. That program image remained clearly on the screen, after the PC was shut down.

That led to screen savers. In my recollection, Norton's Commander was one of the first of this kind. After a selected number of minutes without activity, the screen switches to moving stars, as if you were flying through space. This prevents the regular image from burning in.

CRT monitors have improved a lot since then - the screen image will not "burn in" in a modern CRT. At the same time, screen savers have developed into an art form of their own. Windows is born with a number of choices in screen savers. Also, many programs include a screen saver or two as an extra *feature* . Some provide a series of images, such as "celebrity cars," showing movie celebrities with their fancy cars.

Use the screen savers. They can spice up day-to-day work. And please always use a screen saver with your TFT monitor!

Environmental standards

Top

Screen radiation is a pollutant. There is no concrete evidence that screen radiation can cause illness. However, artificially generated radiation must be unwelcome in our environment. Consequently, industry standards have been developed for acceptable radiation levels.

Since the early nineties, the Swedish MPR-2 standard established limits for monitor electrostatic radiation.

Since then came the stricter TCO-92. It limits the permitted amount of low level radiation and establishes standards for electrical and fire safety. Usually TCO means *Total Cost of*

Ownership but here it refers to the Swedish Confederation of Professional Employees (Tjänstmännens Central Organisation). They defined strong standards for emissions.

Finally, we have TCO-95, which is the strictest standard. Similar to TCO-92, it also includes regulations on ergonomics (including refresh rates), maximum energy consumption, environmentally friendly production and recycling facilities. The best screens comply with this standard. Screens adhering to the TCO standards are more expensive. Obviously since they are better screens.

The flat TFT screens do not emit *any* radiation at all and they consume considerably less energy than the radiating screens. This is another indication that TFT may be the standard screen of the future.

The VESA DPMS system is an energy saving technology, which includes both screen and video card. A modern 17" screen consumes about 100 watt in normal use. With DPMS the screen switches to two energy saving modes. First, power consumption drops to 25 watts and finally again drops to 8 watt.

- Next page
- Previous page

To learn more

Top

Read about sound cards in Module 7c.

Read about digital sound and music in Module 7d .

[Main page] [Contact] [Karbo's Dictionary] [The Software Guides]

KarbosGuide.com. Module 7a.6

The flat panel monitors

The contents:

- Introduction to flat panel monitors
- Eye Ergonomics
- The Digital Interface
- Next Generation Monitors?

- Next page
- Previous page

The Digital Flat Panel Monitors

Top

The big, heavy traditional CRT monitors will eventually be phased out. To day we see them replaced by the flat andLCD (Liquid Crystal Display) monitors, also known from labtops.

It may be a few years before this technology will be dominating, but it is bound to happen. The the flat panel monitors are excellent, and they are available; the prices have gone down. Today a 17.3" LCD costs as much as a 21" CRT monitor did 4 years ago.

The LCD screen is flat, since it contains no cathode ray tube (CRT). Instead the screen image is generated on a flat plastic disk, where millions of transistors create the pixels.

Here you see a Siemens Nixdorf 3501T. It was my first TFT monitor (from 1997), and it produces a sharp high resolution image - better than any other I had ever seen:

Eye Ergonomics

Top

The digital flat panel monitors are also called "soft" screens, since their images seems to have a "softer" quality than those from traditional CRT monitors. The image does not flicker thus causing less eye strain.

People, like myself, who have become accustomed to these soft images will not return to the traditional monitors. I cannot express this with enough emphasis: The flat display is the best monitor available. It is so good to your eyes!

Modern research has shown that a steadily illuminated screen image is a very important element in a good work environment. The eye responds to all light impressions, and the brain interprets all light impressions continually. When a mediocre monitor flickers, the brain will continually receive superfluous light impressions "noise" to sort out. Thus the brain works permanent overtime interpreting the screen flicker. No wonder that people get tired from watching their monitors.

At the same time the LCD screen is by far the most environmentally safe product. These flat screens emit *zero* radiation, and they consume significantly less power than the traditional monitors. Another reason to expect LCD screens to become the monitors of the future.

No refresh rate

A big advantage in the LCD screen is that it does not flicker. Traditional CRT monitors flicker all the time which is not ideal. Of course the best CRT monitors have a high refresh rate (85 Hz or more), which provides a very stable image with no noticeable flicker. But the LCD screen does not flicker at all (when digitally connected). They have a refresh rate of 0 Hz!

Please notice that looking at LCD displays, you may read information like:

- Pixel Frequency 65MHz
- Horizontal 30 ~ 50KHz
- Vertical: 55 ~ 70Hz

This indicates that there *is* a refresh rate. There is, but it is only working when the screen image is *changing.* So if you move a window across the screen, the changes will be updated with a refresh rate of 60 Hz or what ever you choose.

To many users this does not really matter; using Office programs, most of the time the screen image does not change, hence it does not flicker. Obviously it is a problem if you expect to use your flat panel monitor to show full motion videos or games.

The digital interface

Top

The most important thing about the flat panel monitor is that it is connected to a *digital* graphics port. Unfortunately, this is not always the case.

Back in 1997, when I got my first flat panel monitor (the Siemens 3501T mentioned above) it was only available with a (total proprietary) digital graphics adapter. This was a 1st generation flat panel monitor. Later the manufactures found out to add an analog port in the displays. This way people could buy a flat panel monitor and reuse their exixting graphics adapter. It is a marketing stunt, which should not be followed!

The only way to benefit from a flat panel monitor is to feed it digitally.

Our latest monitor

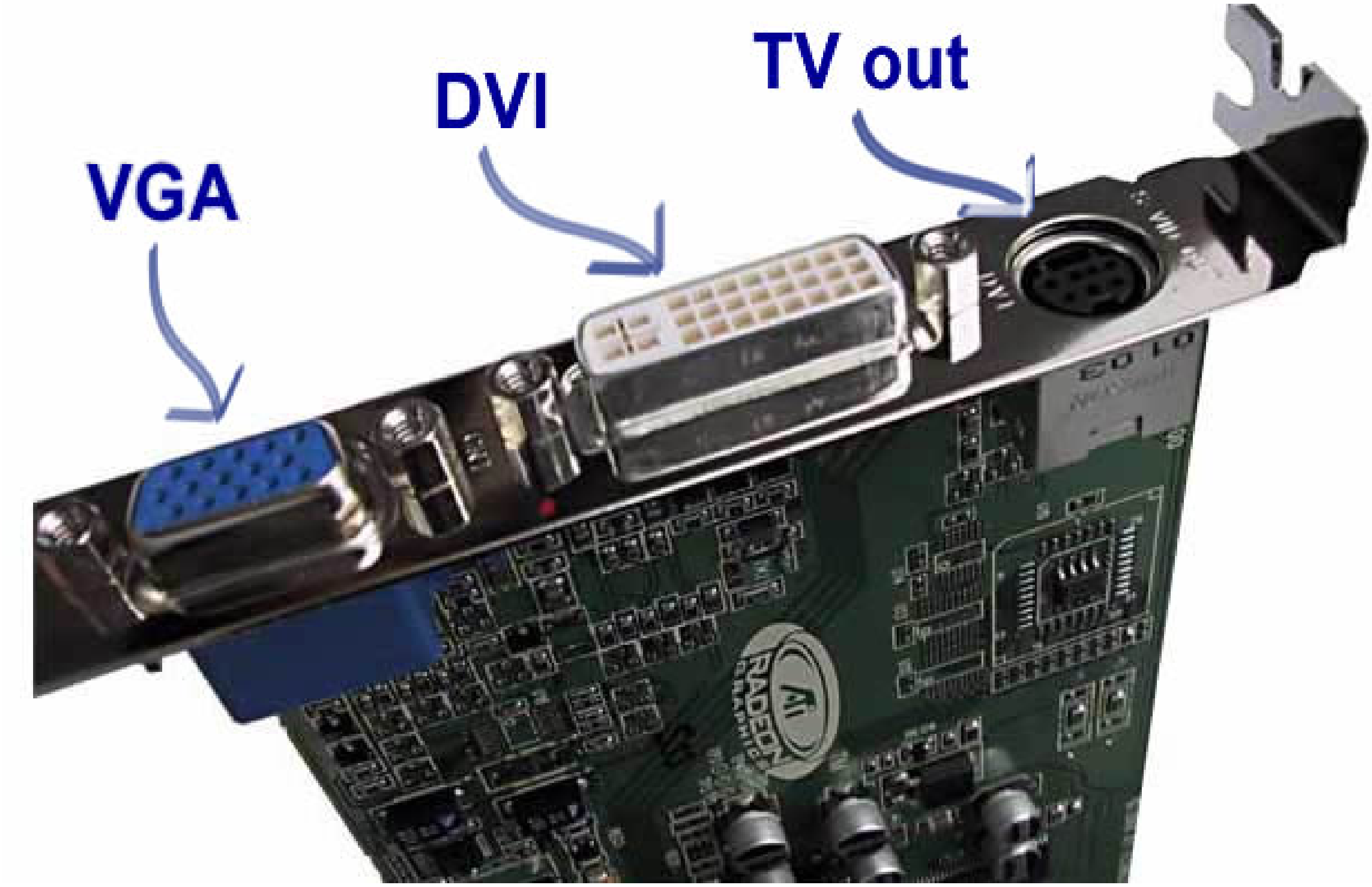
In spring 2001 I bought a new flat panel monitor; it is a Dell model 1701FP. A nice 17.3" monitor, which I paid around $900 for (the price have decreased later). The monitor holds both an analog (VGA-) port and a digital port (DVI).

A 17.3" flat panel monitor has a visible area much bigger than that from a traditional 17" CRT monitor - you can compare it to a 19" CRT. The 17.3 inches is the visible diameter.

I was told, that it should work fine using the analog port with my existing Matrox G400 graphics controller. The manual emphased that one should use the 60Hz mode. I installed the hardware, and it worked fine. Only the image was terrible! It was flickering and very unsharp, kind of "dirty".

The weird thing is, that the Dell manual holds almost nothing on these issues. But through testing we found out that the 75 Hz analog mode was the best possible. But it was not satisfactory, not at all. Having paid quite a lot of money, we decided to go for a digital graphigs controller. We found an ATI Radeon VE, which turned out to be a great card at a reasonable price:

From the box, the product seems to be aimed gamers. To those I am sure, that the RADEON chipset and the 32 MB of DDR RAM is fine. To us, the very important issue was, that there is a DVI connector on the board:

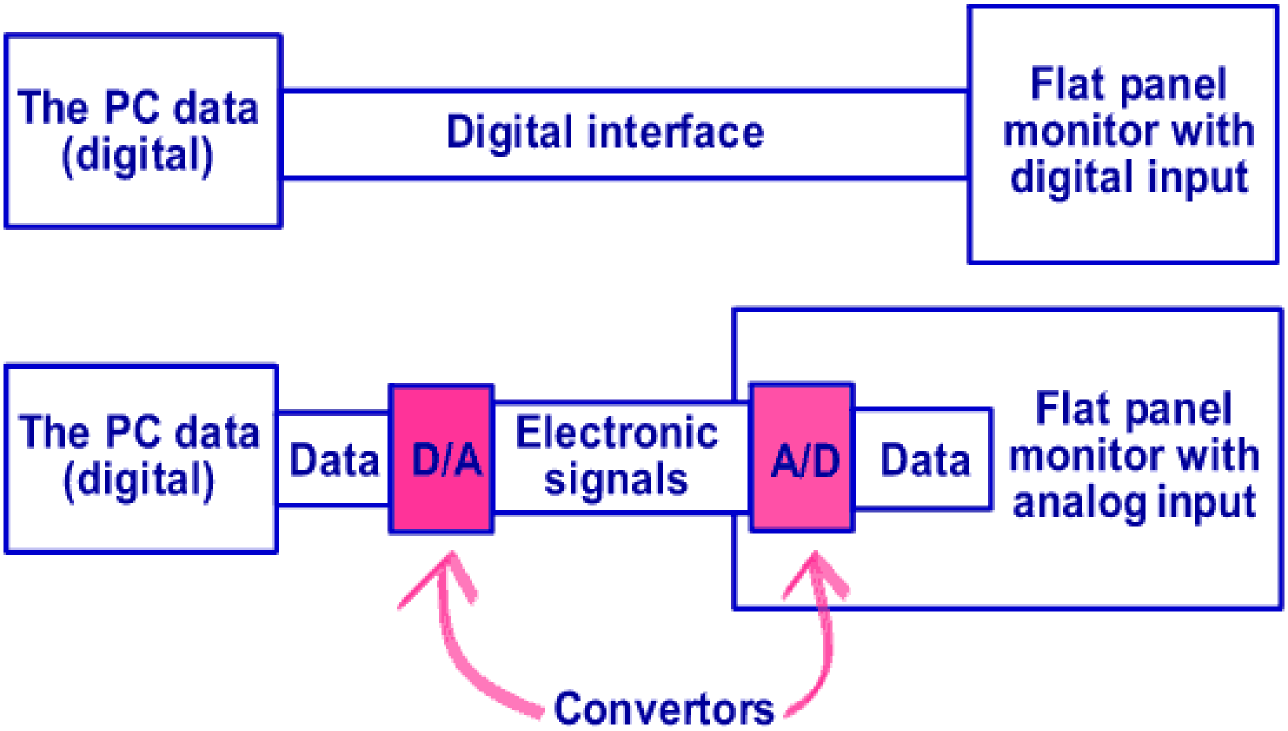


With the DVI connector in use, the DELL flat panel monitor works absolutely perfectly. Thinking about it, it is incredible, that the company does not tell this in the manual. With digital interface the image is extremely sharp and completely flicker free.

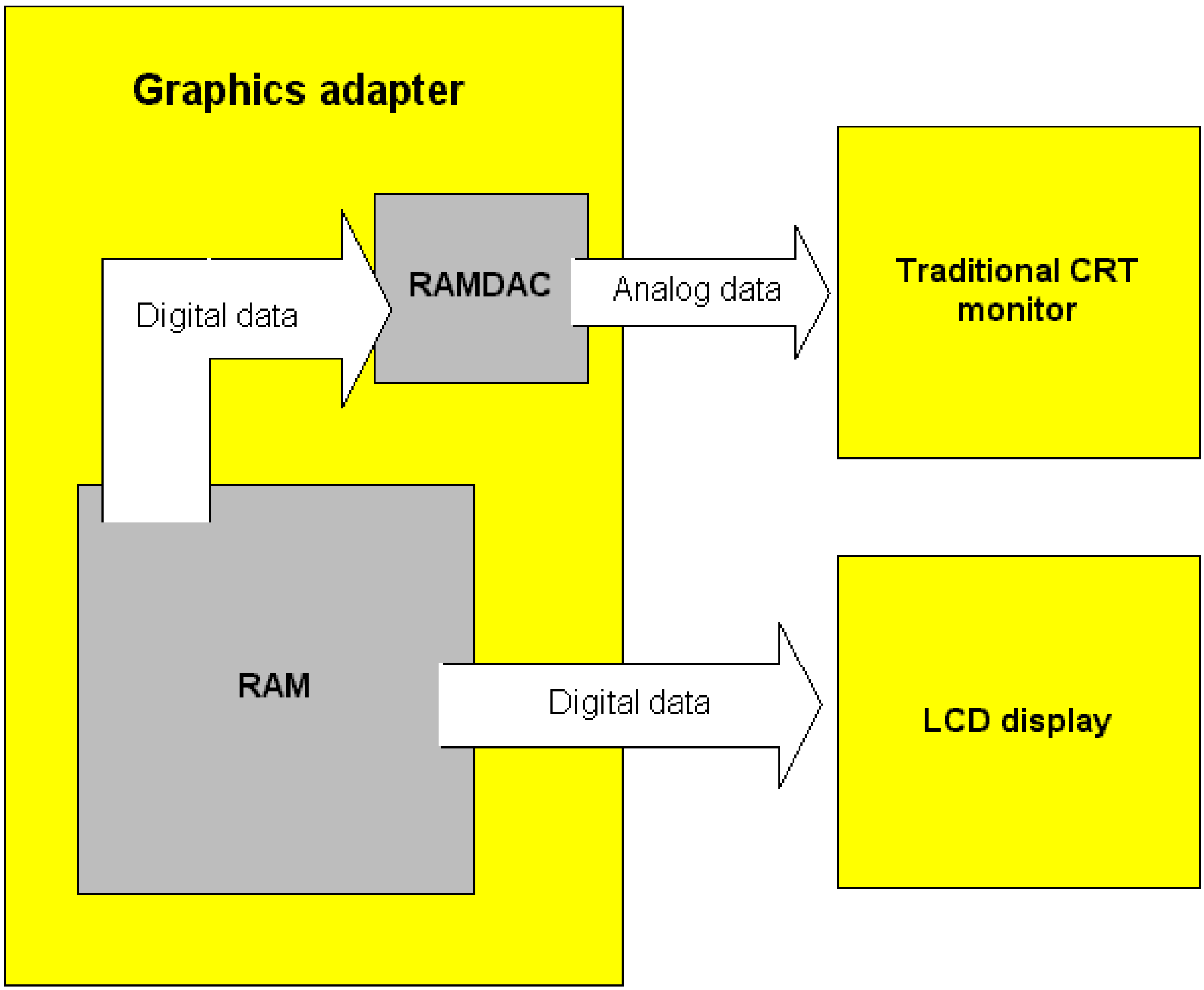
A flat panel monitor is digital by nature. There is no analog electronics included, and that is the big advantage of this technology. Hence, the monitor should not be connected through an analog interface. In fact, using the analog interface, you get to conversions, which both add noise to the final image. First the graphics adapter has to convert the digital data of the PC to analog electronical signals. Then these analog signals have to be converted back til digital information to feed the display.

Using the digital interface, each pixel consists of three transistors, which each is mapped to the corresponding memory cell holding the image info. A purely digital to digital transmission with no electronical noise involved - that is the way to produce a stunning image!

Here you see an illustration of the differences between the two setups:



The A/D (analog to digital, the RAMDAC of the video card) and D/A (digital to analog) conversions only reduce image quality on a flat panal monitor - nothing else! Hence, the digital interface by-pass the RAMDAC of the graphics controller:



Panel Quality

It is difficult to produce a flat panel display without flaws. Most panels sold have a few defect pixels, where one or more transistors are gone.

The bigger the panel gets, the more flaws you find (due to the increasing number of transistors). This helps keeping up the prices - the manufactures have to throw away a large percantage of the production - you cannot repair flawed pixels.

All vendors have some kind of quality policy in this area. Some only accept up to 3 or 5 pixel flaws per panel. Others accept up to 15, if they are not situated in the middle of the display. When you buy a flat panel display you should make sure that you can return it if the number of pixel-flaws is to big. This can be hard to achieve.

In the summer 2001, a German survey showed that some vendors take advantage of the consumers ignorance in this area. More than 30% of the flat panel monitors were of second range quality and should not have been brought to the shops. Obviously it has been tempting to some companies to sell some of the displays, which should have been dismissed.

Screen savers

Using a LCD display, you should remember to install the Windows screen saver. I use "Black screen" after 5 minutes. Like in the "old days", the monitor may get damaged from longer periods of showing the same image.

Next Generation Monitors?

Top

The monitor technology is advancing very rapidly. An interesting development comes from a British invention LEP - *Light Emitting Plastic* . It is an ordinary thin, flexible plastic (polymer), which is sandwiched together with a thin film of indium tin oxide and aluminum. Thin-film transistors control the oxide layer, causing the huge plastic polymer molecules to become light emitting.

These LEP screens should have these advantages:

- They are completely flat and lightweight.
- They consume only small amounts of electric power.
- They do not require background illumination, which the LCD crystals do.
- They emit light, which is visible from all angles of view.

These screens was expected to be available in year 2002/2003, but lately there have been no much indication that they will come. Currently work is being done with prototypes, which have a resolution of 200 dpi. That corresponds to a resolution of 2200 X 1600 pixels in a 15" screen. So maybe we can look forward to an extremely high screen resolution.

I would like to fantasize about future Coca-Cola bottles with a built-in video display in the plastic bottle! By the way, these polymer plastic materials are finding their way into other parts of the data processing technology. Work is being done on developing different storage media, hard disks in terabytes size and RAM modules based on polymers. These "organic" storage media should also be significantly cheaper to produce that the traditional products.

See Cambridge Display Technology web .

- Next page
- Previous page

To learn more

Top

Read about sound cards in Module 7c.

Read about digital sound and music in Module 7d .

Karbo's Dictionary. Please choose a letter:

A B C D E F G H I J K L M N O P Q R S T U V W X - Å

Click Here