

Emulation for digital preservation in practice: the results

Jeffrey van der Hoeven
Koninklijke Bibliotheek (KB)
National Library of the Netherlands

Co-writers: Bram Lohman and Remco Verdegem

iPres 2007
Beijing, China
October, 2007



planets

digital
preservation
research
and
technology

The cause of all trouble...



Refs:

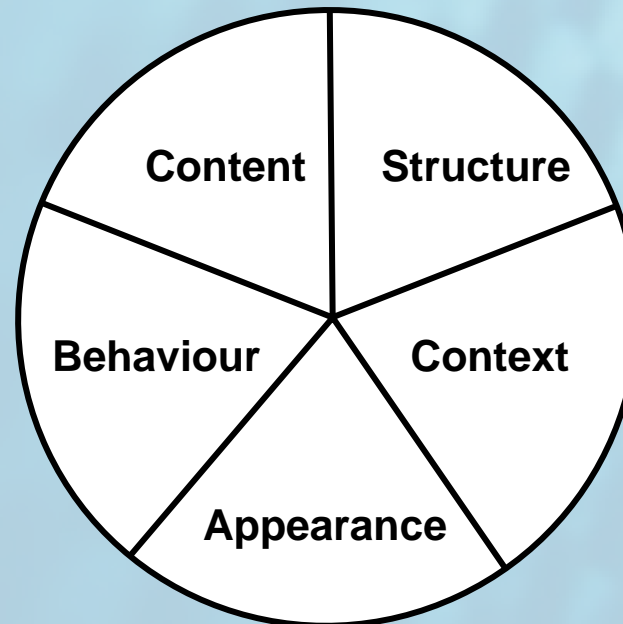
<http://www.ecoaction.com.au/res/Image/junkedcomputers.gif>

<http://www.sinometrecycling.com/Commun/Goods/scrap%20computer%20hard%20disk.JPG>

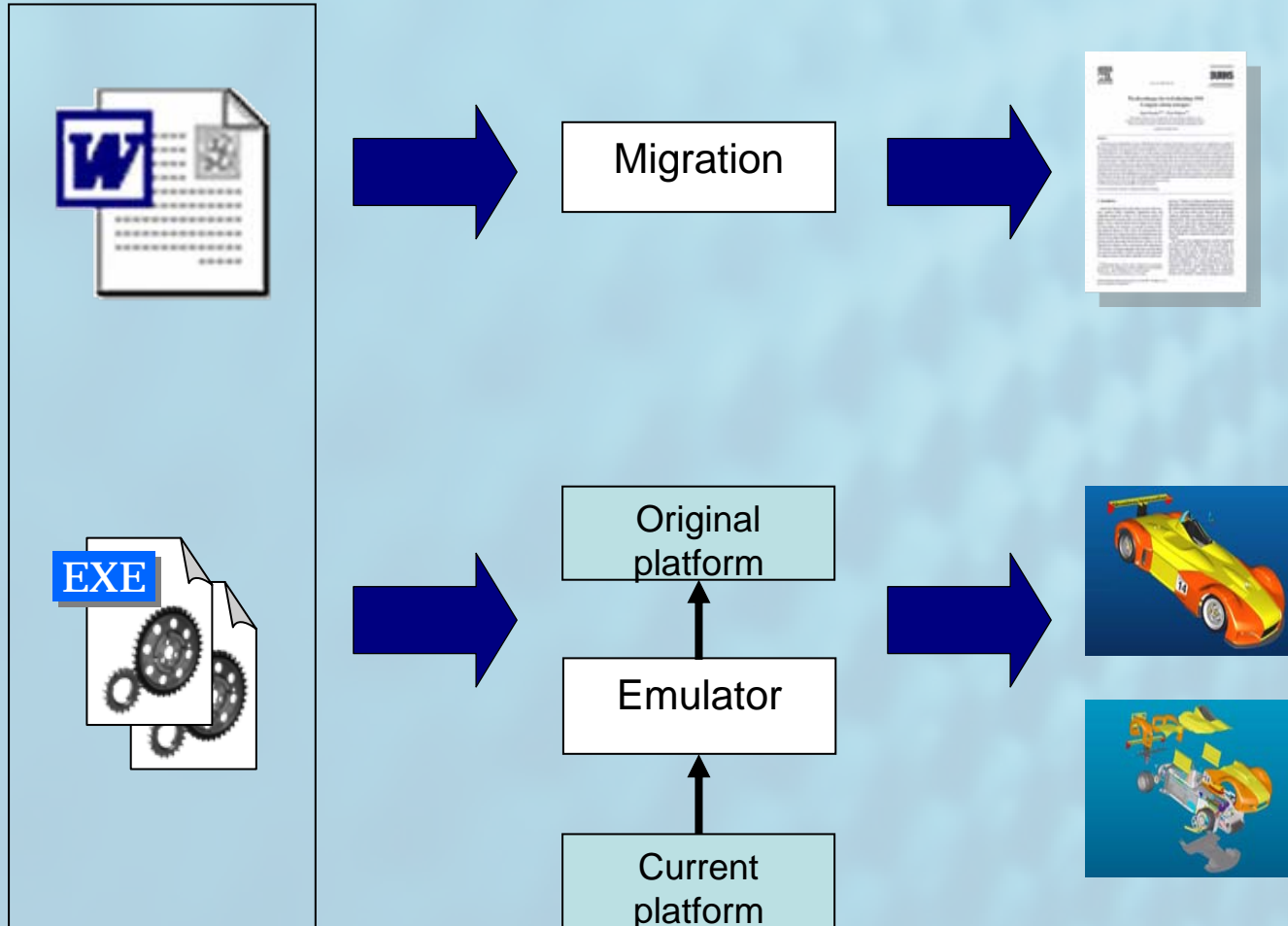
Digital preservation action!

Three basic criteria for defining action:

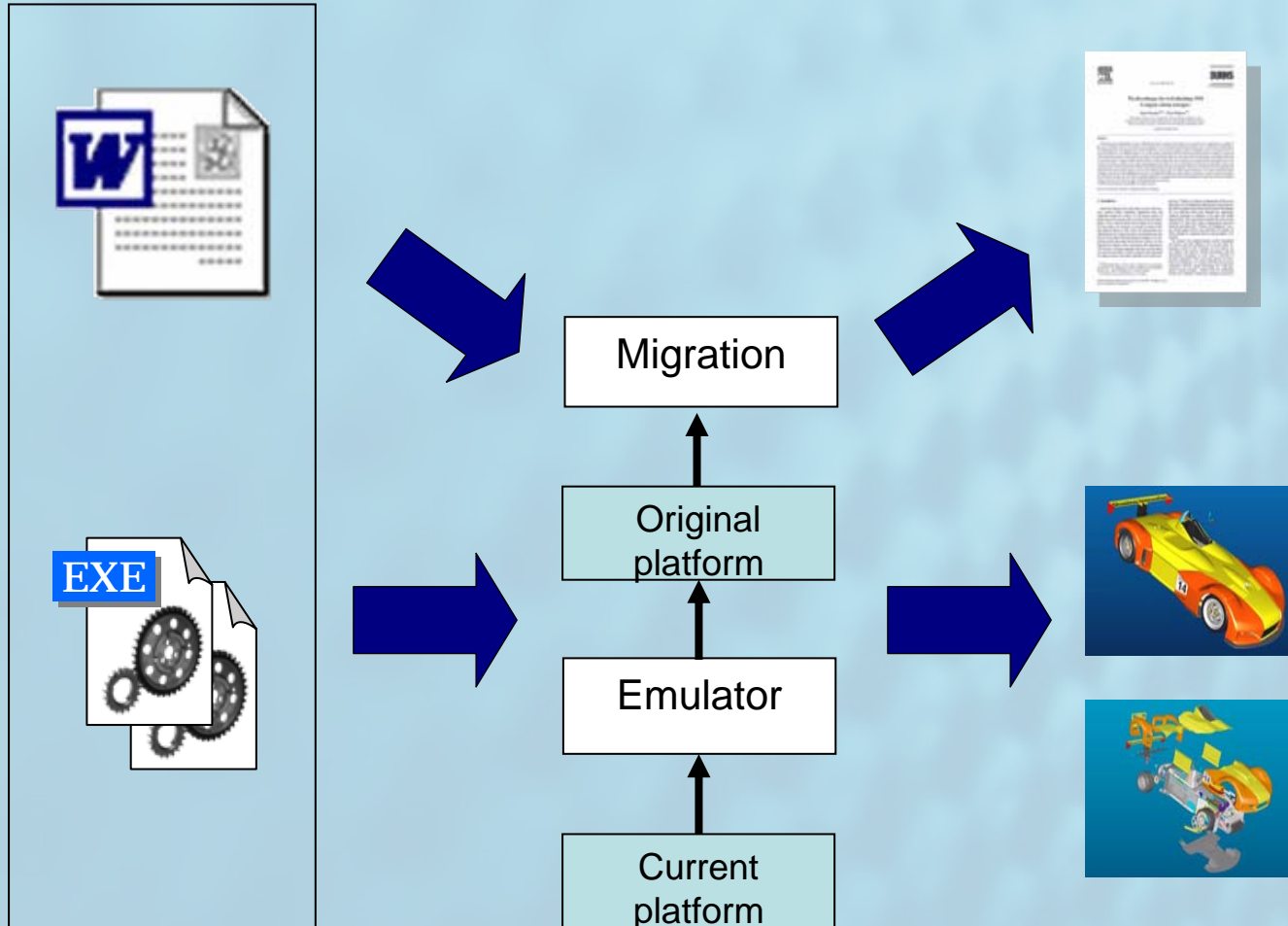
- Policy of the institution
- (Future) user requirements
- Kind of digital object:



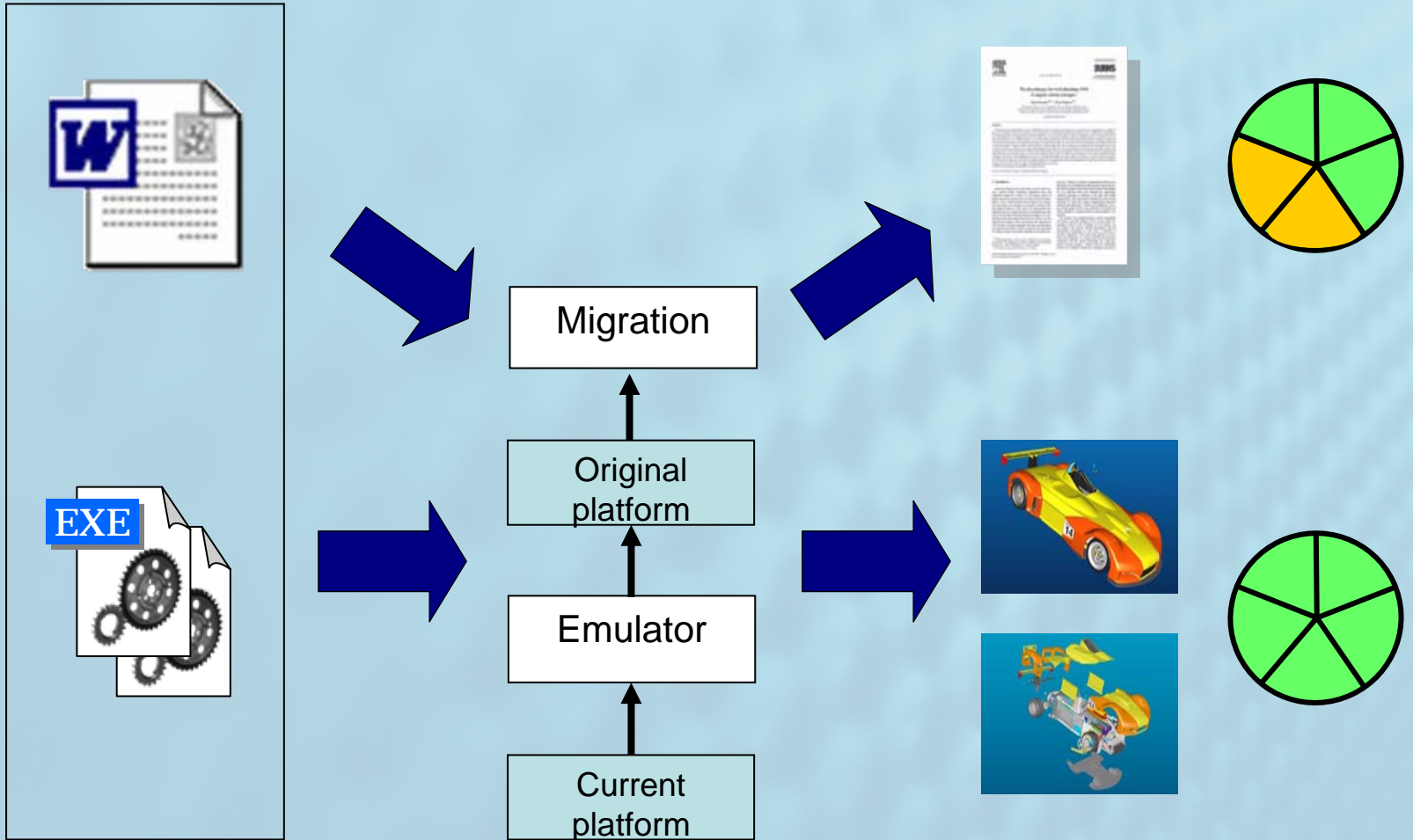
Migration & emulation



Migration & emulation



Migration & emulation



Emulation

= adapt the computer environment to render the digital object authentically.

Pros

- Rendering of original computer environment
- No changes to digital object
- Proven technology

Cons


- Complexity
- Initial costs (effort)
- Isolated process limiting information reuse
- Knowledge of original environment required

Never applied to an operational digital archiving environment.

- ❑ Held on 20 October 2006
- ❑ Focused on emulation for DP
- ❑ Attended by selected group of experts in the field of digital preservation, emulation and IT.
- ❑ States that:
 - ❑ “Emulation is a vital piece of the puzzle for retaining long-term access to the wide range of digital objects.”
 - ❑ “...important steps to be taken to make emulation appropriate for DP.”

Full statement can be downloaded from www.kb.nl

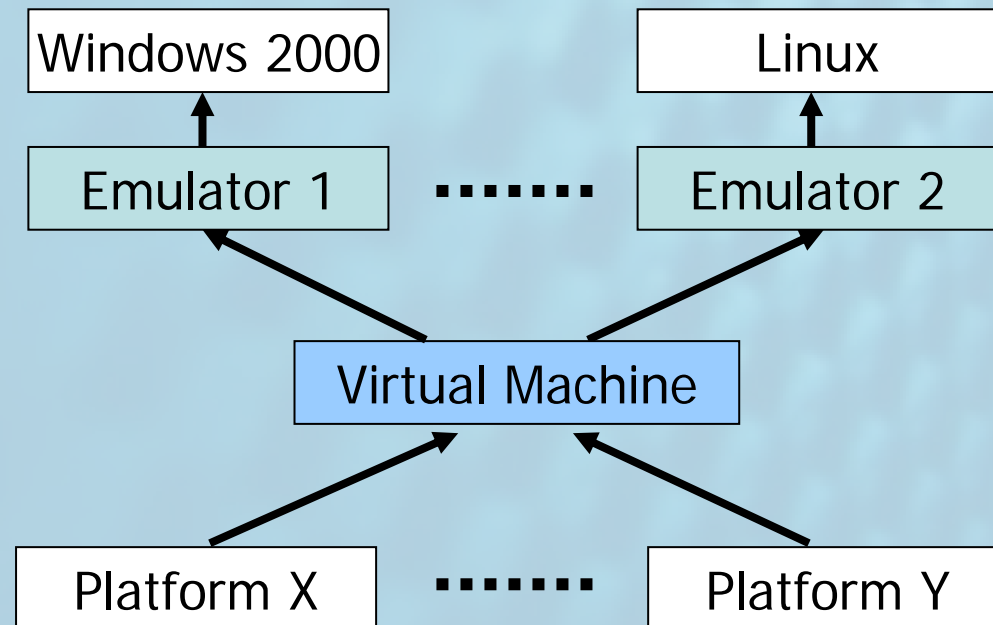
Project outline

- 
- A thick black vertical arrow pointing downwards, indicating the chronological order of the project events.
- 2004 KB preliminary study: feasibility emulation
- 2005 New DP-proof design: modular emulation.
KB and Nationaal Archief start joint project.
Goal : build and test modular emulator
Scope: PDF, databases, multimedia apps.
- 2006 Tessella leads development.
Jeff Rothenberg supports project.
- 2007 First release of modular emulator.
On July 1st, Dioscuri becomes part of Planets.

What is modular emulation?

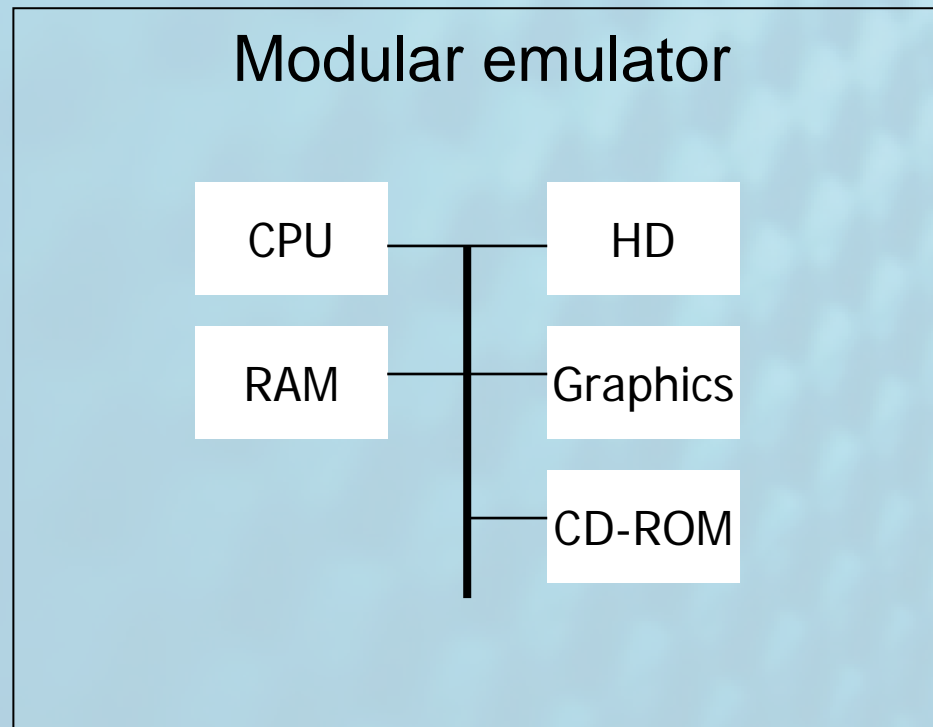
Two key features: durability and modularity

- ❑ Durable -> emulator has to endure time. This can be done by making the emulator portable to a wide range of computer platforms using a Virtual Machine (VM).



What is modular emulation?

- ❑ Modular -> emulator consists of modules. Each module emulates the functionality of a hardware component. This way, the modular emulator can be configured much like a real computer.



Richard

Bill

Jeffrey

Jeff

Hilde

Bram

Remco



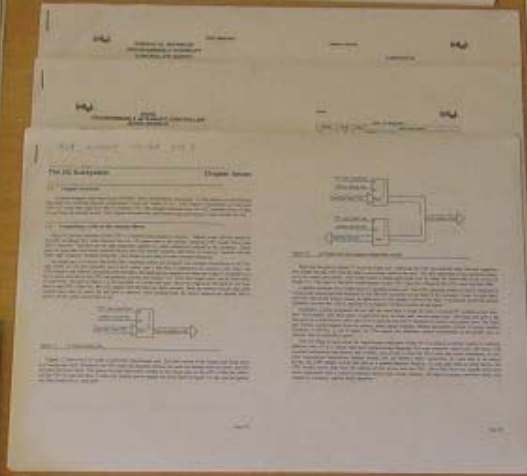
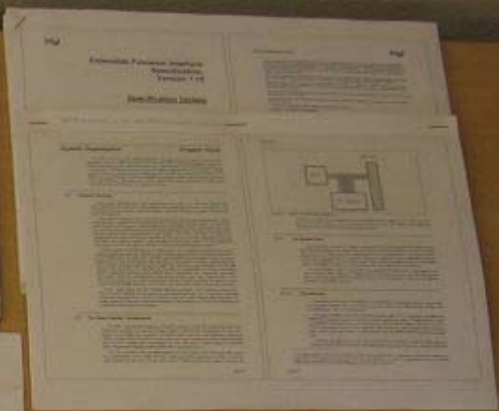


Embedded BIOS 4.1

The BIOS (Basic Input/Output System) is a software program that is stored in a non-volatile memory (ROM) on the motherboard. It is responsible for initializing the hardware components of the computer and providing a platform for the operating system to run.

Key features of the Embedded BIOS 4.1 include:

- Support for up to 640K of random access memory (RAM).
- Support for up to 1024K of hard disk storage.
- Support for up to 16 floppy disk drives.
- Support for up to 16 serial ports.
- Support for up to 16 parallel ports.



Intel 80486DX40 Pinout

Pin	Function
1	\overline{CS}
2	\overline{MEMEN}
3	\overline{MEMRST}
4	\overline{MEMRST}
5	\overline{MEMRST}
6	\overline{MEMRST}
7	\overline{MEMRST}
8	\overline{MEMRST}
9	\overline{MEMRST}
10	\overline{MEMRST}
11	\overline{MEMRST}
12	\overline{MEMRST}
13	\overline{MEMRST}
14	\overline{MEMRST}
15	\overline{MEMRST}
16	\overline{MEMRST}
17	\overline{MEMRST}
18	\overline{MEMRST}
19	\overline{MEMRST}
20	\overline{MEMRST}
21	\overline{MEMRST}
22	\overline{MEMRST}
23	\overline{MEMRST}
24	\overline{MEMRST}
25	\overline{MEMRST}
26	\overline{MEMRST}
27	\overline{MEMRST}
28	\overline{MEMRST}
29	\overline{MEMRST}
30	\overline{MEMRST}
31	\overline{MEMRST}
32	\overline{MEMRST}

Intel 845G/845GL/845GV Chipset

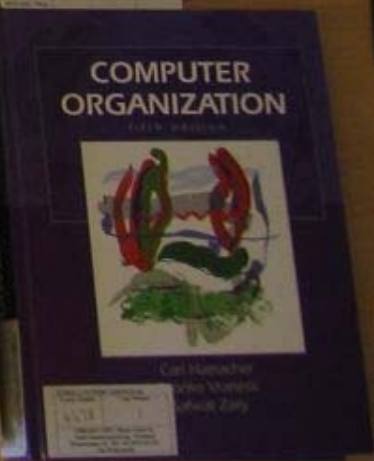
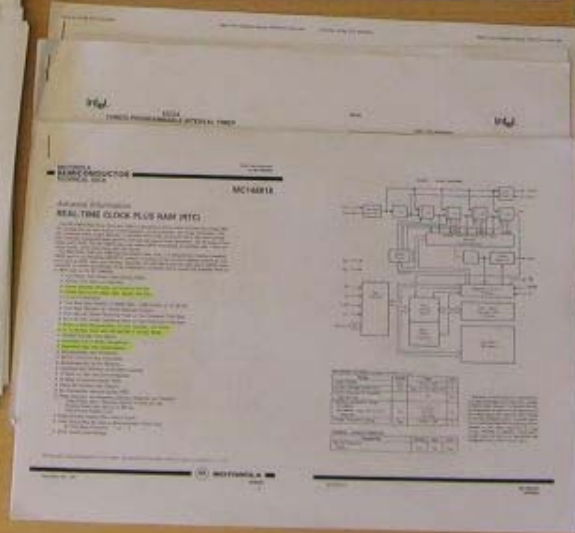
Intel 845G/845GL/845GV Chipset

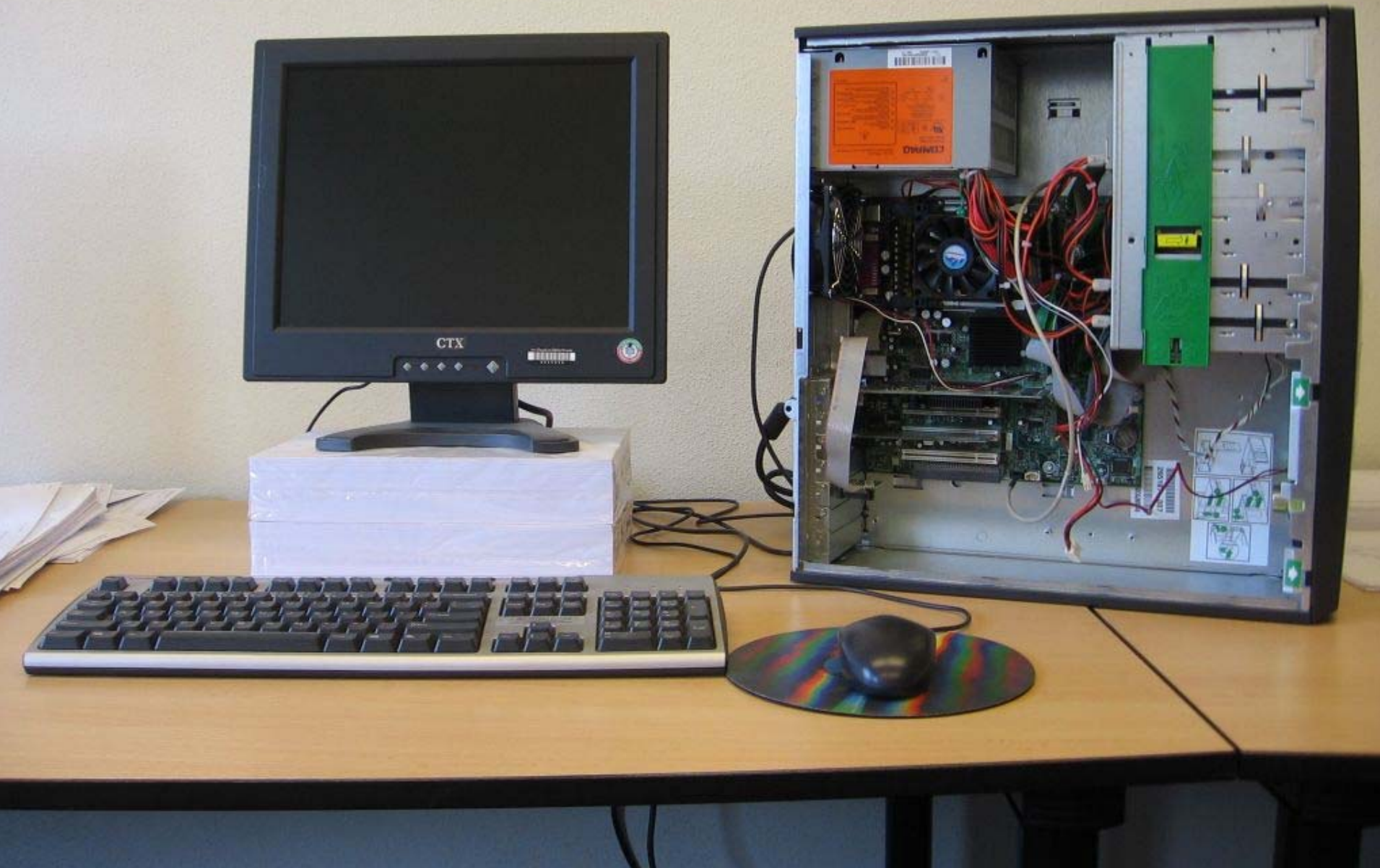
This document provides the specifications for the Intel 845G/845GL/845GV Chipset. It includes a block diagram showing the chipset's connection to the system bus and various peripheral devices. The chipset is designed to manage data flow between the system bus and various peripheral devices, including graphics, audio, and network controllers.

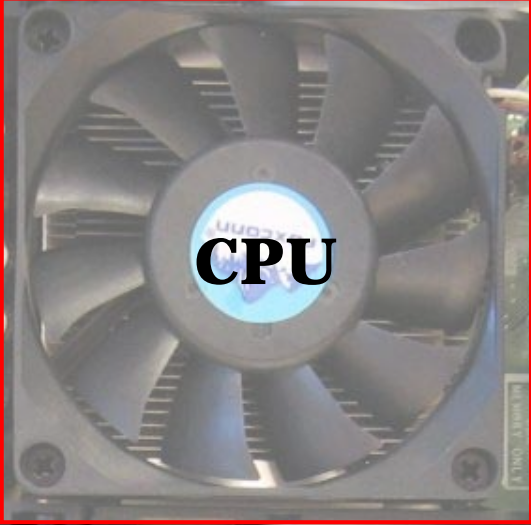
VESA®

VESA®

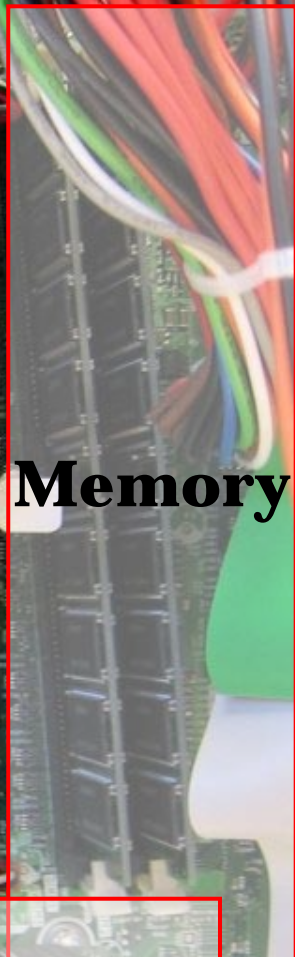
This document provides the specifications for the VESA® standard. It includes a block diagram showing the standard's connection to the system bus and various peripheral devices. The standard is designed to manage data flow between the system bus and various peripheral devices, including graphics, audio, and network controllers.







CPU



Memory



Graphics card



**PIC
RTC**



BIOS



ModuleProcessor

Module
Memory

ModuleVideo

Module
PIC
RTC

Module
Clock
CMOS

Results

- ❑ Dioscuri – modular emulator for digital preservation
- ❑ Current version: 0.2.0 (Beta)
- ❑ Programmed in Java using JVM
- ❑ Capable of:
 - ❑ Running MS-DOS, FreeDOS, Linux 16-bit (ELKS)
 - ❑ Norton Commander 3.0, WordPerfect 5.1, DrawPerfect 1.1, many games like PC-versions of PacMan, Tetris, Chess, Ironman and many more. Even DOS-based webbrowser Arachne works!
 - ❑ XML-based module configuration
 - ❑ Text extraction from emulated environment into the clipboard of host computer.
 - ❑ Running on many platforms like Intel Windows, PowerPC Mac, Sun Sparc Solaris.

menuPanel

Left		Files	Commands	Options	Right	
		Name	Name	Name	Name	
1	√ Brief		123view	exe	qbasic	exe
	Full		Io	sys	qbasic	hlp
M	Info		Msdos	sys	read	me
a	Tree		attrib	exe	sys	com
a	On/Off	Ctrl-F1	autoexec	bat	undelete	exe
c	-----		chkdsk	exe	unformat	com
c	√ Name		command	com	xcopy	exe
c	eXtension		config	sys		
d	tiMe		dbview	exe		
e	SiZe		edit	com		
e	Unsorted		edit	hlp		
f	-----		fdisk	exe		
f	Re-read		format	com		
h	Drive...	Alt-F1				
l						
mem	exe					
nc	exe					
ncsmall	exe					
Io.sys		33430	11-11-91	5:00a		

ca D:\EMULAT~1\WORKSP~3\emu\images\floppy\nc\NC.EX

Left		Files	Commands	Options
		Date	Time	
	Brief	-11-06	18:33	
1	Full	-10-06	9:45	
	Info	-11-06	15:29	
	Tree	-02-07	13:04	
	On/Off	-03-06	10:53	Ctrl-F1

1	Name	-12-05	15:14	
	eXtension	-08-06	17:11	
	tiMe	-12-05	11:59	
	SiZe	-12-05	20:20	
	Unsorted	-07-06	15:21	

	Re-read	-10-06	17:06	
	Drive...	-02-07	11:00	

A:\>

1Help 2User 3View 4Edit 5Ccpu 6R

Things to improve

- Performance must be increased (limit of 100x slower)
- Data extraction and insertion
- More modules:
 - improved CPU (32-bit)
 - Mouse
 - Sound
 - Network
 - ...
- Module library
- Replacing JVM by more universal virtual machine

Planets**KB**

2007

Improve Dioscuri
(32-bit, mouse, etc.)

Test elaboration with
Dioscuri in reading
rooms

2008

Experiments and
module library

First case: combining
emulation with web
archiving

2009

Integration with
interoperability
framework (IF)

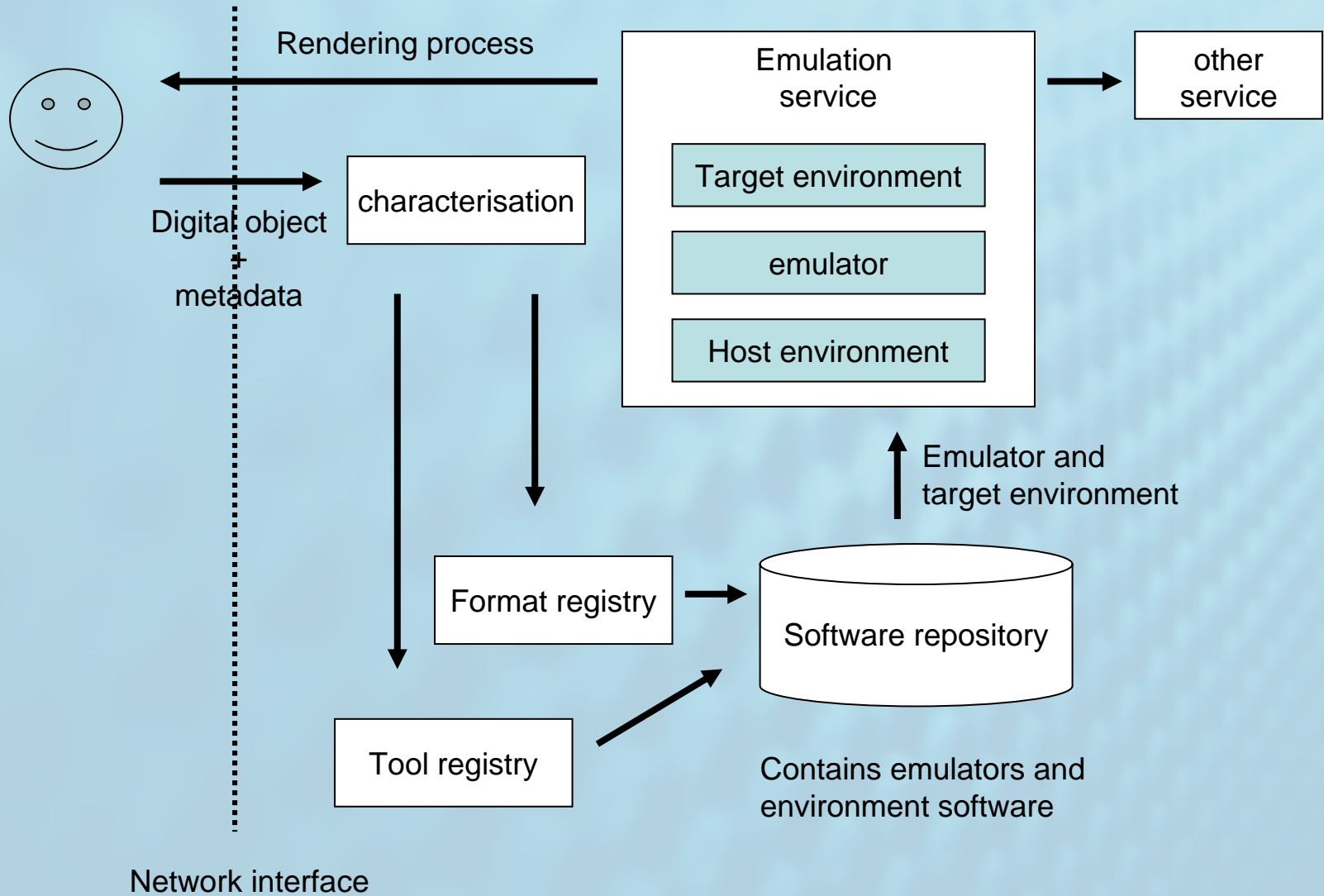
Integration with
e-Depot process flow

2010



Emulation service

Future situation



Things to note

- Software repository
- Disk image preparation
 - Pre-installed disk images
 - Created on demand
- Old documentation (manuals, tutorials, tips & tricks)
- Data insertion and extraction
 - Transfer text
 - Transfer images
 - Transfer files
 - Transfer (user-defined) stream

Dioscure - the modular emulator for digital preservation - Mozilla Firefox

Bestand Bewerken Beeld Geschiedenis Bladwijzers Extra Help

<http://dioscure.sourceforge.net/> Google

KB nationaal archief

DIOSCURE

Dioscure - the modular emulator

Dioscure is an x86 computer hardware emulator written in Java. It is designed by the digital preservation community to ensure documents and programs from the past can still be accessed in the future.

The Dioscure emulator has two key features: it is durable and flexible. Because it is implemented in Java, it can be ported to any computer platform which supports the Java Virtual Machine (JVM), without any extra effort. This reduces the risk that emulation will fail to work on a single architecture in the future, as it will continue to work on another architecture.

Dioscure is flexible because it is completely component-based. Each hardware component is emulated by a software surrogate called a module. Combining several modules allows the user to configure any computer system, as long as these modules are compatible. New or upgraded modules can be added to the software library, giving the emulator the capability to run these.

Dioscure is the best choice to retain access to your old documents, games and other applications!

Latest news

3 September 2007
 Dioscure version 0.2.0 is out now! This release is capable of running various versions of MS-DOS, FreeDOS 0.9 Beta (Included in package) and ELKS (Embeddable Linux Kernel Subset). The following improvements have been made:

- ◆ Fixed minor bugs in CPU instructions and added some new instructions
- ◆ Fixed keyboard status LEDs on GUI
- ◆ Improved system timer and PIT
- ◆ Improved GUI for image file selection
- ◆ Improved keyboard controller for mouse/keyboard control
- ◆ Added functionality to copy text to clipboard (text mode only)

For a complete overview of all changes in recent versions, please check the changelog. See the disk images download section for various new disk image downloads.

 **Download: Dioscure version 0.2.0**



Dioscure

- [Idea and key features](#)
- [Digital Preservation](#)
- [Screenshots](#)
- [Latest news!](#)

Downloads

- [Latest version](#)
- [All versions / sourcecode](#)
- [Disk images](#)

Documentation

- [User manual](#)
- [Reference docs](#)
- [Javadoc](#)
- [Changelog](#)

Support

- [FAQ](#)
- [Forum](#)

Development

- [Buglist \(tracker\)](#)
- [Feature requests \(tracker\)](#)
- [Roadmap](#)

Contact

- [About project team](#)
- [Join development!](#)
- [Mailinglist](#)

References:

Dioscuri website:

<http://dioscuri.sourceforge.net>

Emulation project KB-NL / Nationaal Archief and

Emulation Expert Meeting (EEM) 2006:

http://www.kb.nl/hrd/dd/dd_projecten/projecten_emulatie-en.html

Contact:

jeffrey.vanderhoeven@kb.nl