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:

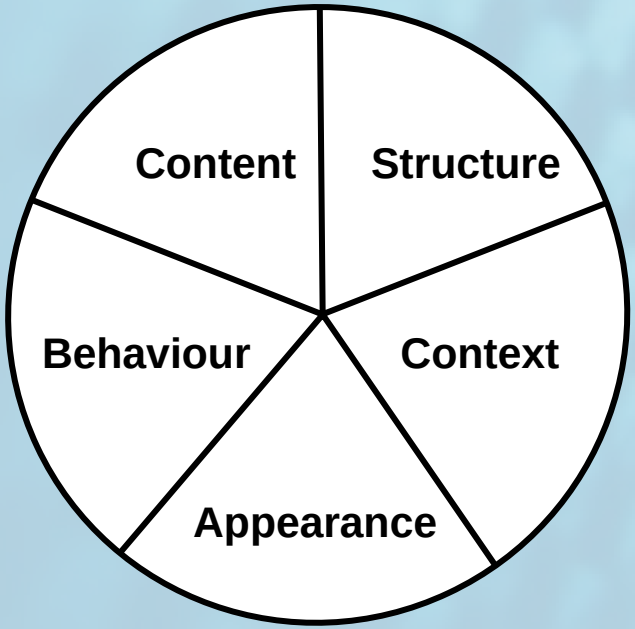
http://www.ecoaction.com.au/res/Image/junkedcomputers.gif

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

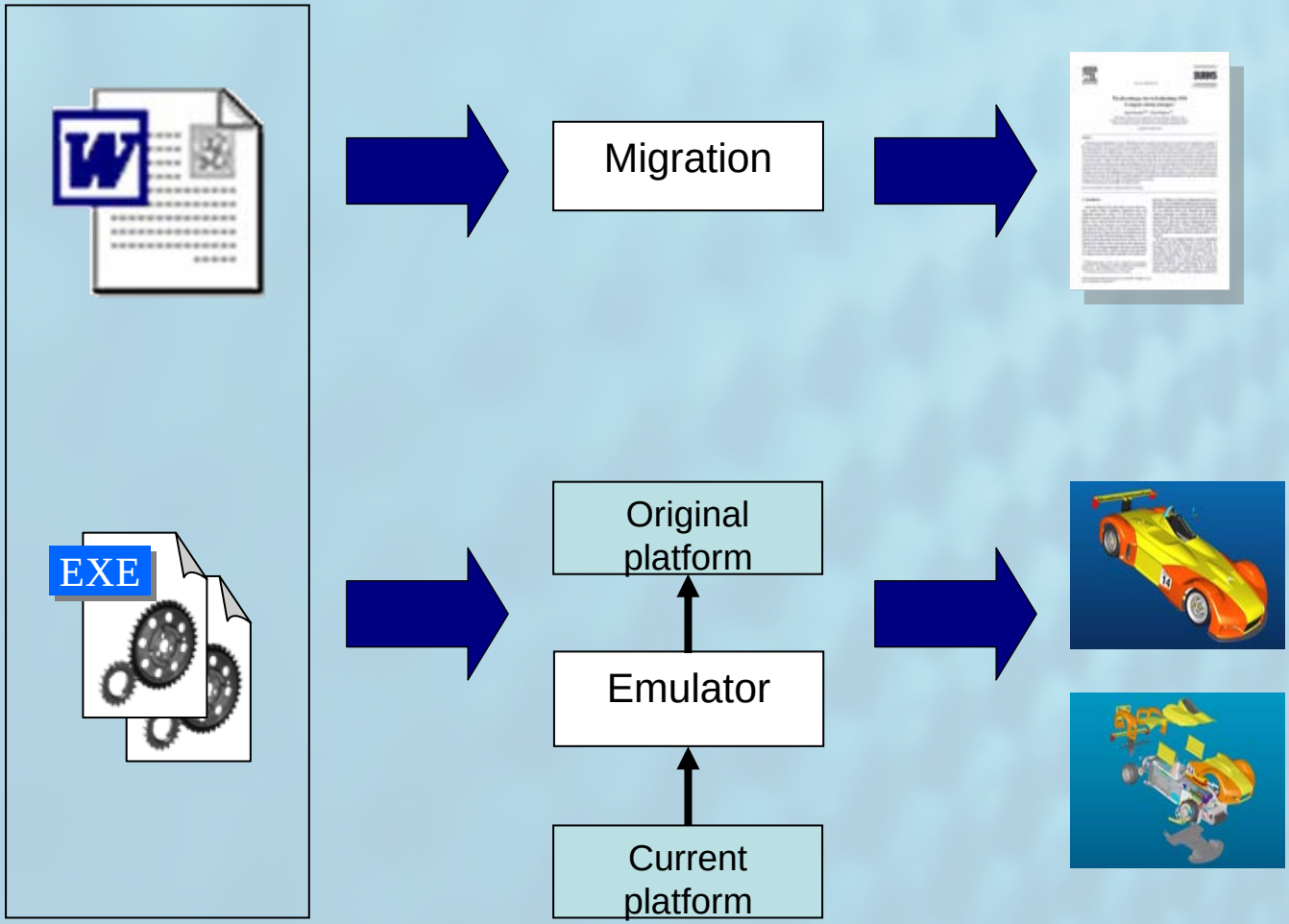
Digital preservation action!

Three basic criteria for defining action:

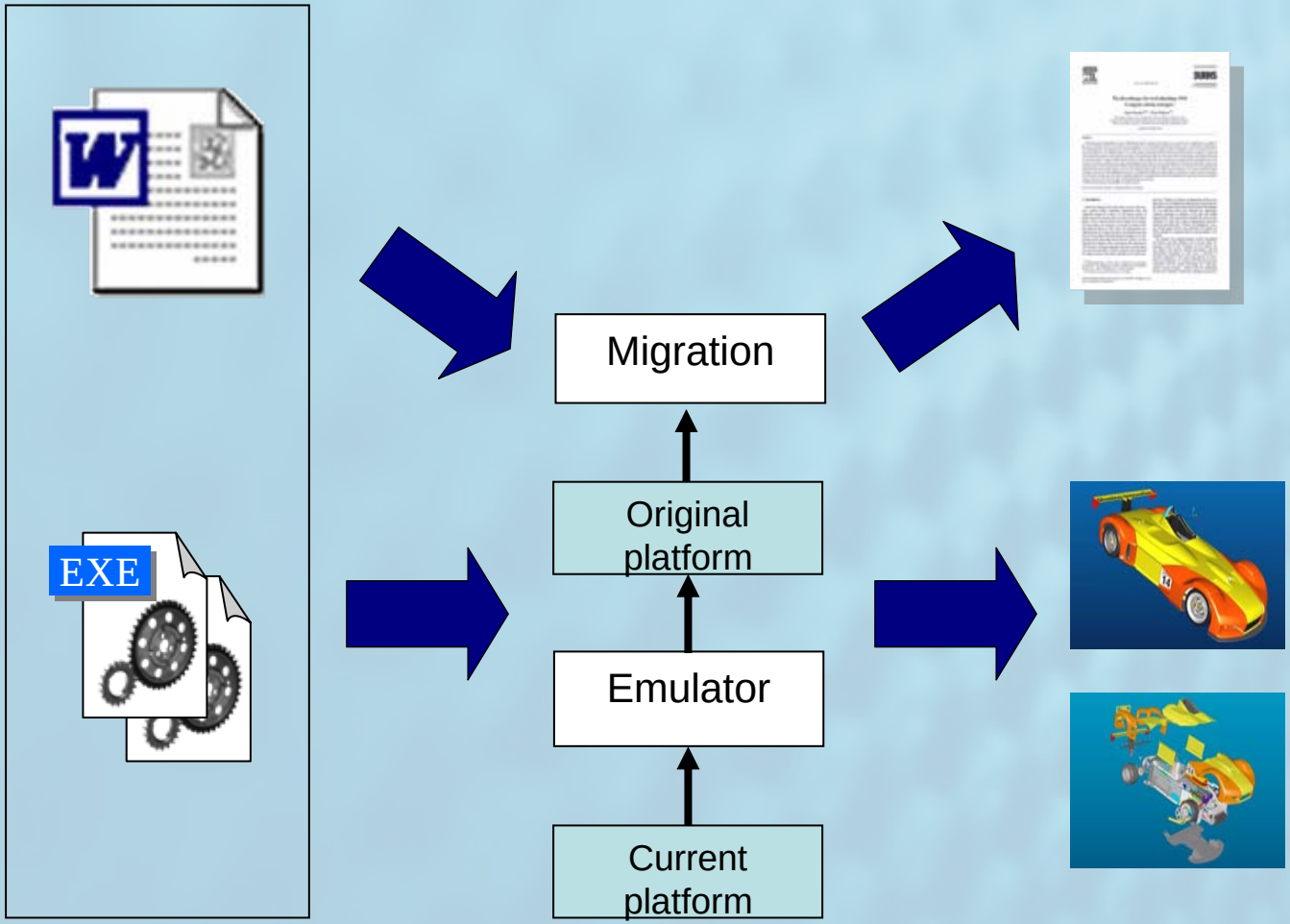
- Policy of the institution
- (Future) user requirements
- Kind of digital object:



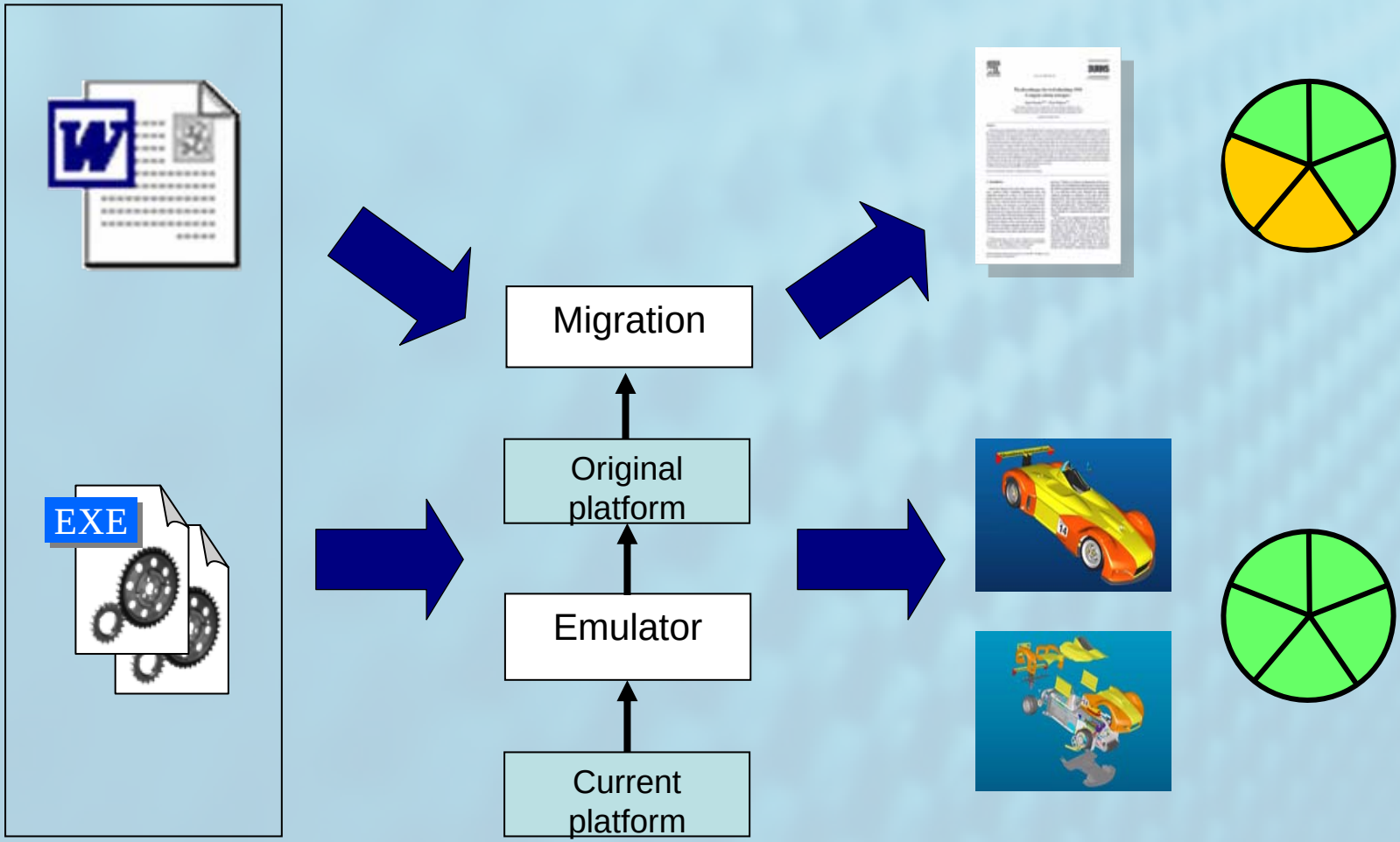
Migration & emulation



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.

Emulation Expert Meeting 2006

- 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

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.

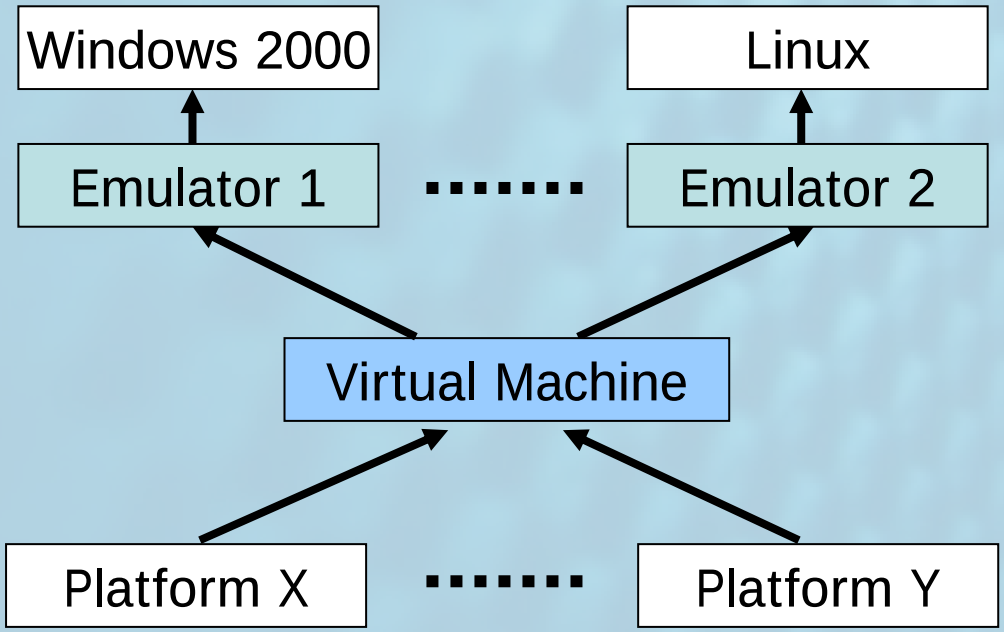
2006 Tessella leads development.
Jeff Rothenberg supports project.

2007 First release of modular emulator.
On July 1st, Dioscuri becomes part of Planets.

What is modular emulation?

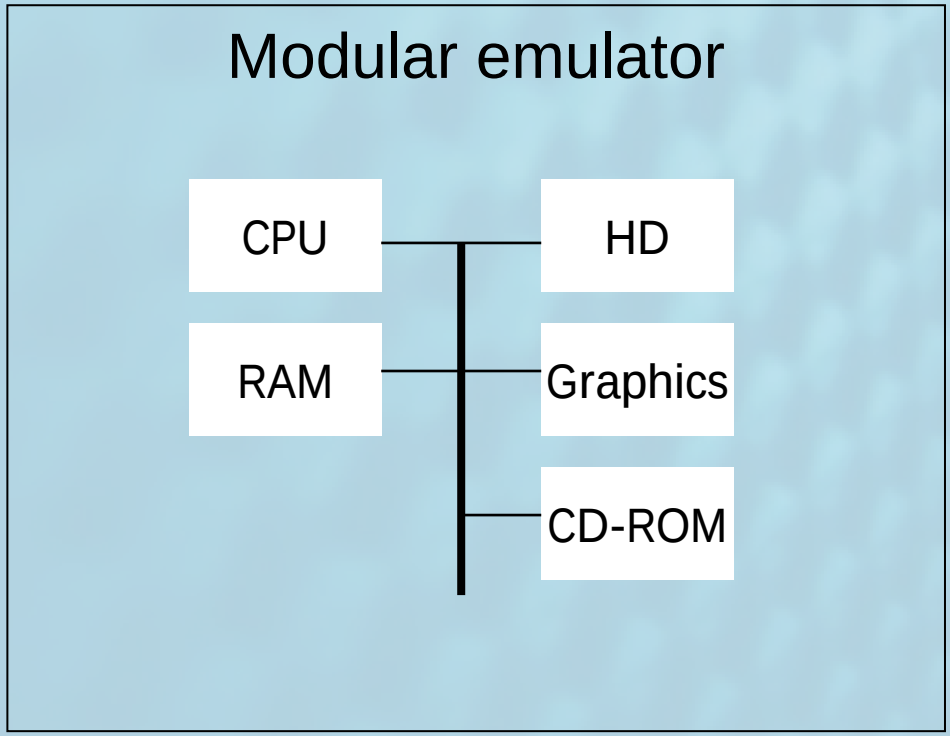
Two key features: durability and modularity

- 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).

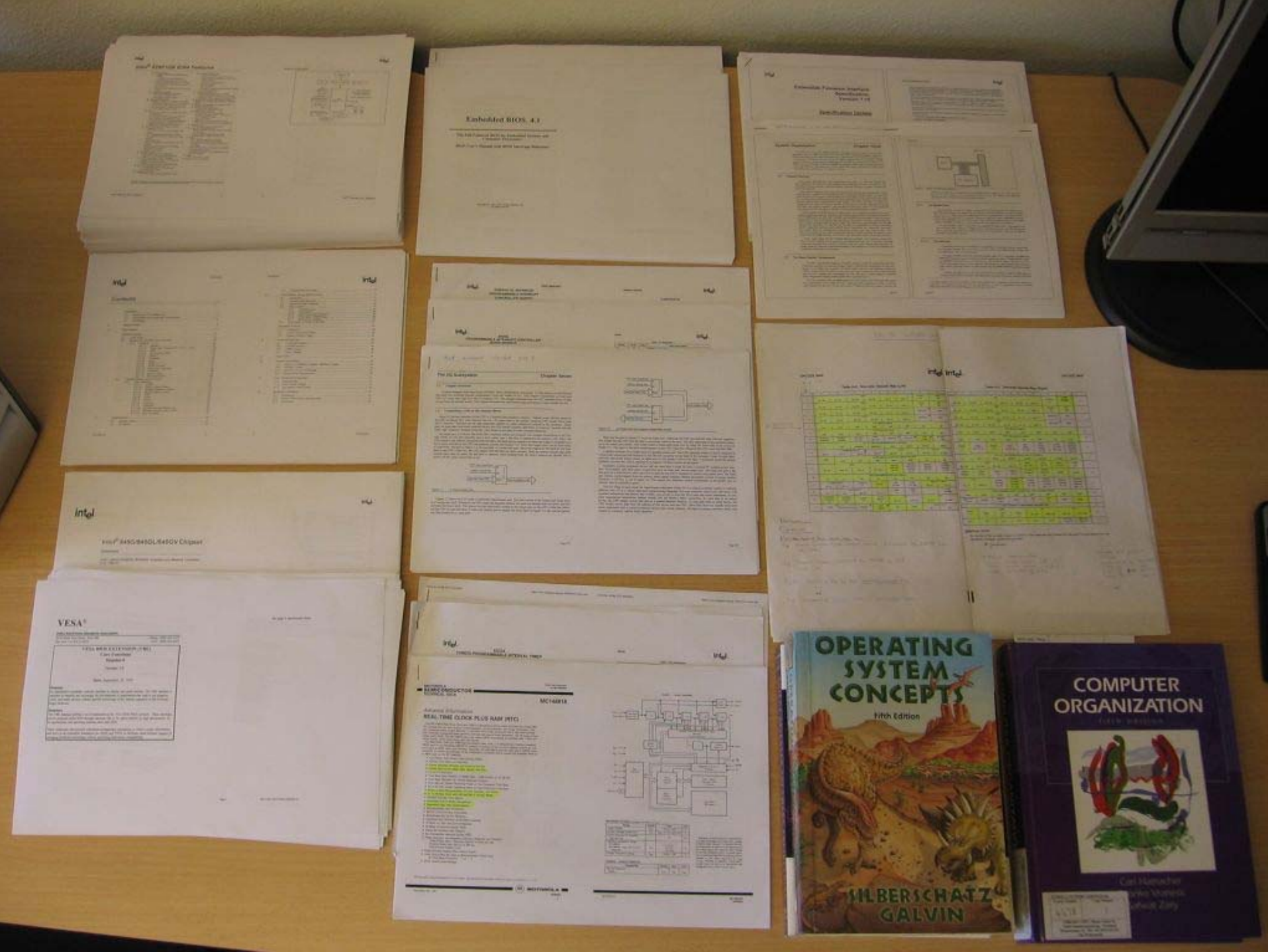


What is modular emulation?

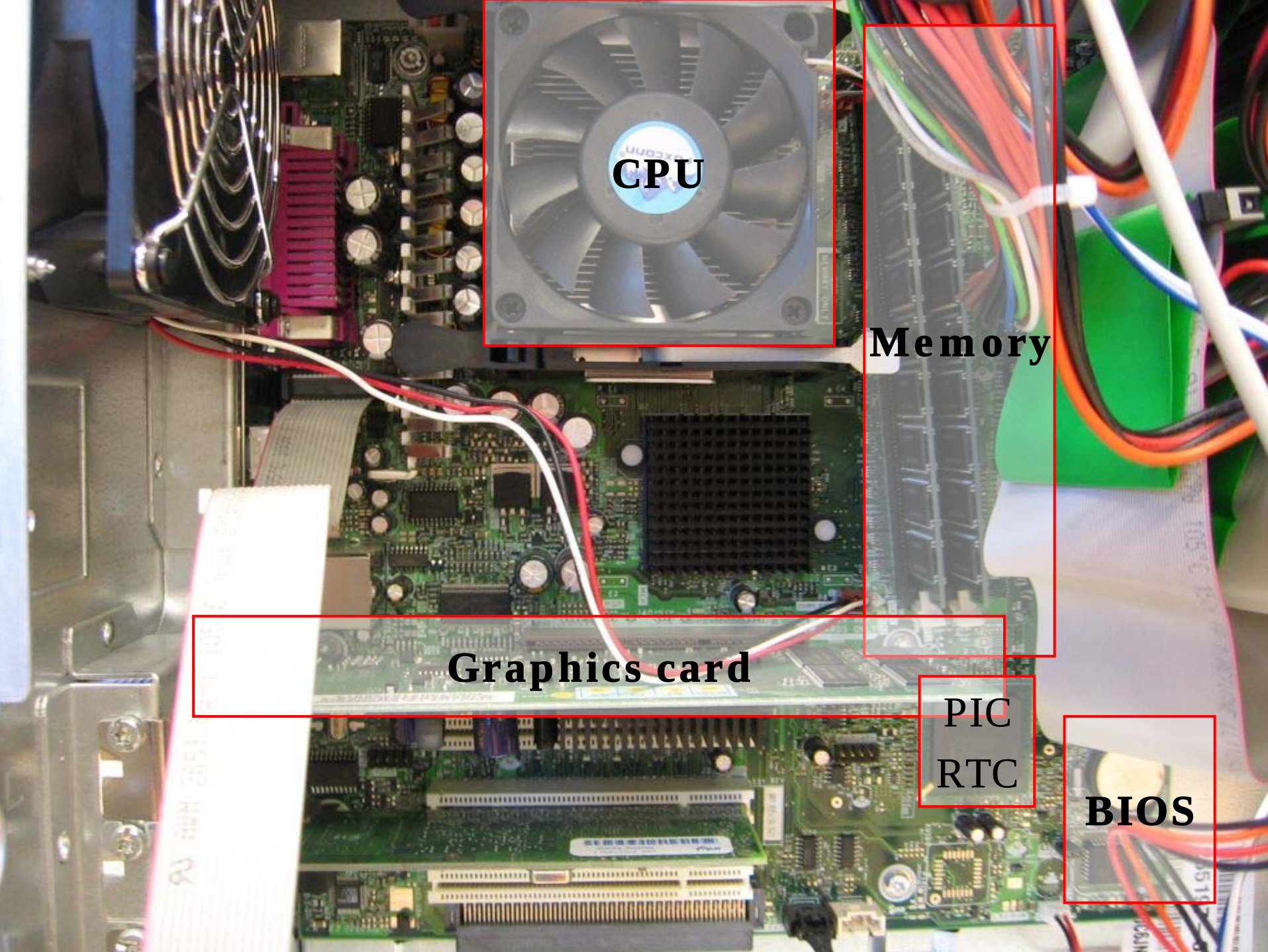
- Modular -> emulator consists of modules. Each module emulates the functionality of a hardware component. This way, the modular emulator can be configured much like a real computer.

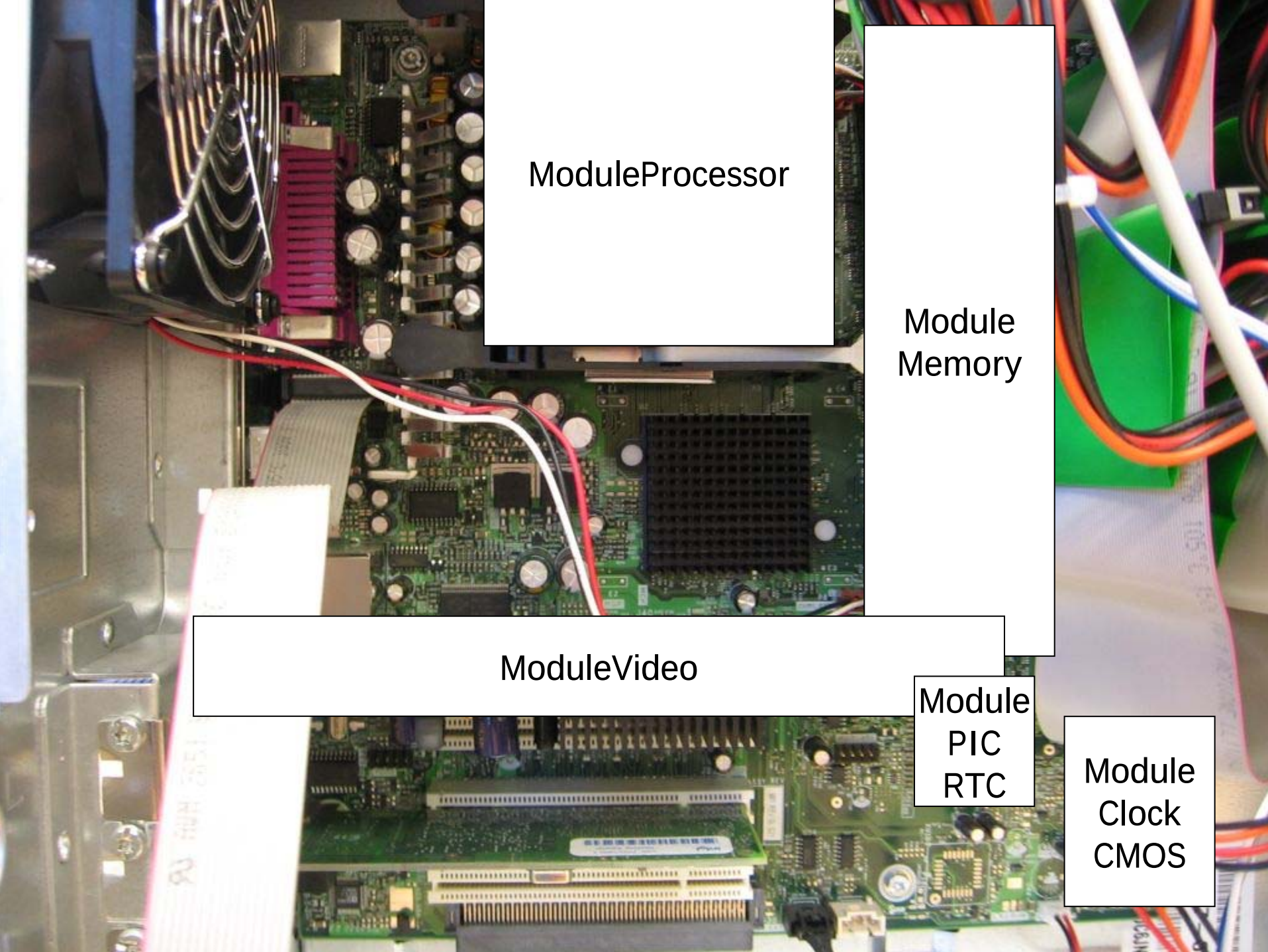



Richard
Bill
Jeffrey
Jeff
Hilde
Bram
Remco


Embedded BIOS 4.1
VESA
REAL TIME CLOCK PLUS RAM (RTC)
OPERATING SYSTEM CONCEPTS
Fifth Edition
SILBERSCHATZ
GALVIN
COMPUTER ORGANIZATION

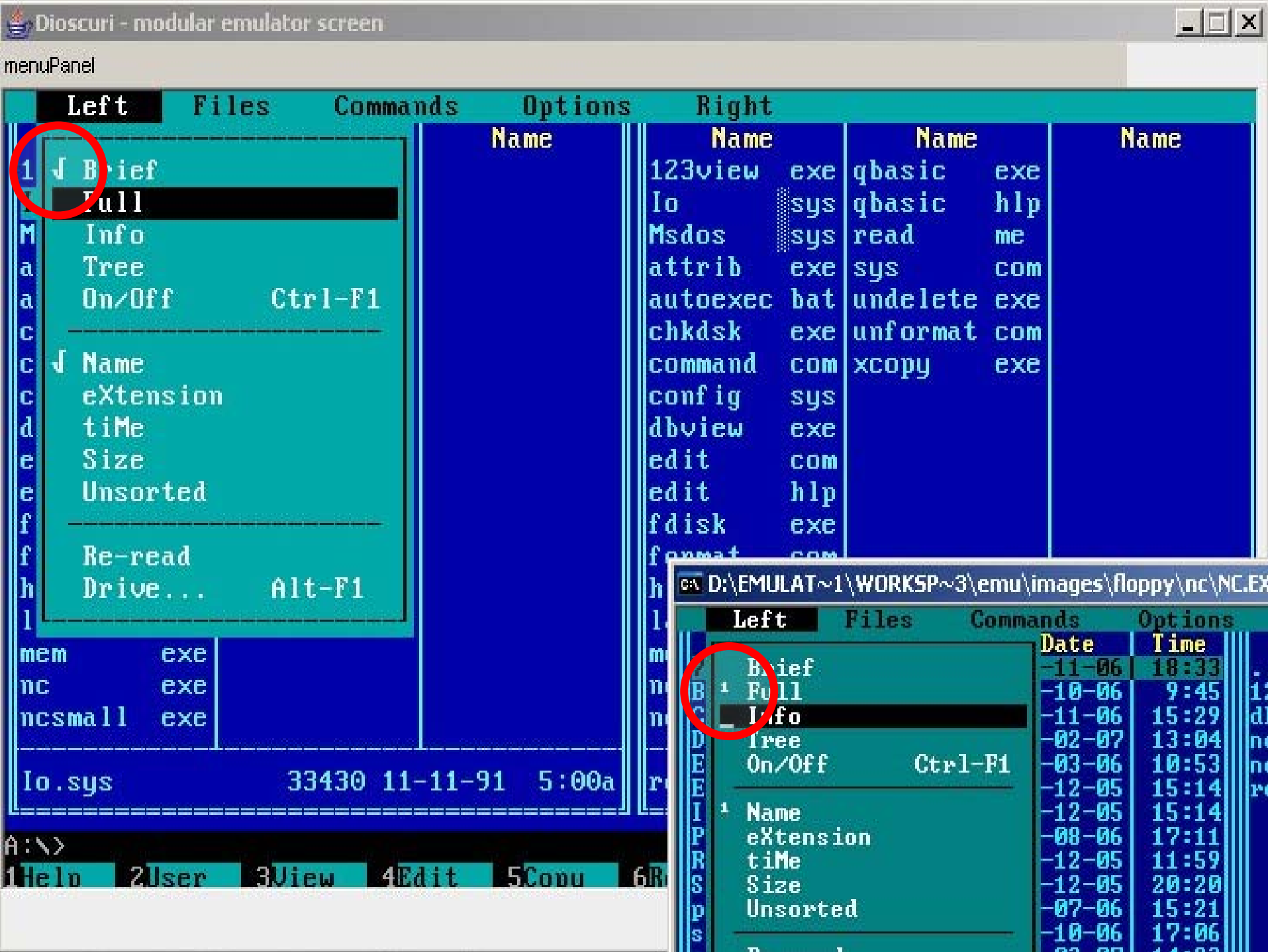

CTX


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.


Dioscuri - modular emulator screen
menuPanel
Left Files Commands Options Right
Brief
Full
Info
Tree
On/Off Ctrl-F1
Name
eXtension
tiMe
Size
Unsorted
Re-read
Drive... Alt-F1
Name
123view exe
lo sys
Msdos sys
attrib exe
autoexec bat
chkdsk exe
command com
config sys
dbview exe
edit com
edit hlp
fdisk exe
mem exe
nc exe
ncsmall exe
qbasic exe
qbasic hlp
read me
sys com
undelete exe
unformat com
xcopy exe
lo.sys 33430 11-11-91 5:00a
A:\>
1Help 2User 3View 4Edit 5Copy
D:\EMULAT~1\WORKSP~3\emu\images\floppy\nc\NC.EX
Left Files Commands Options
Brief
Full
Info
Tree
On/Off Ctrl-F1
Name
eXtension
tiMe
Size
Unsorted
Date Time
-11-06 18:33
-10-06 9:45
-11-06 15:29
-02-07 13:04
-03-06 10:53
-12-05 15:14
-12-05 15:14
-08-06 17:11
-12-05 11:59
-12-05 20:20
-07-06 15:21
-10-06 17:06

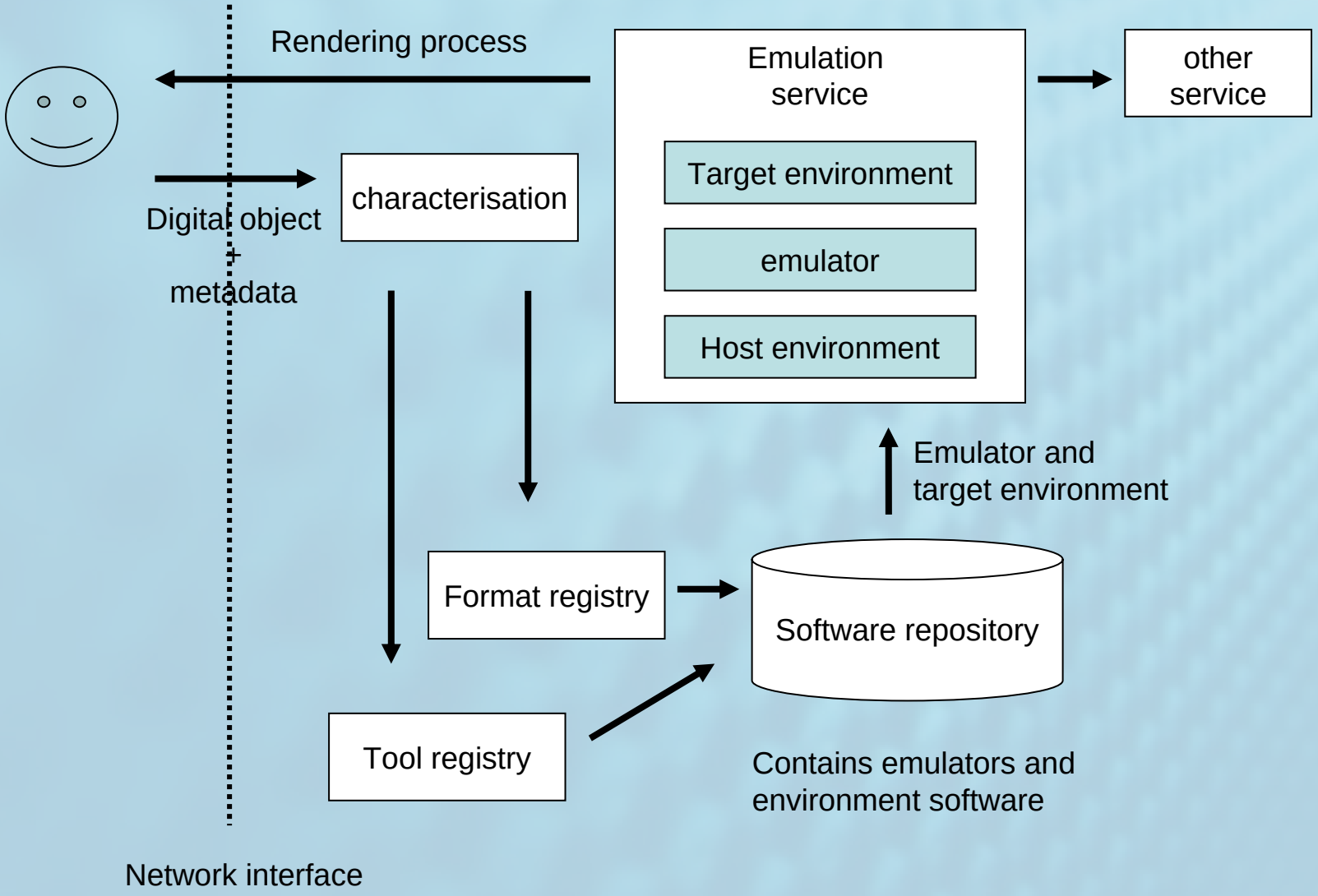
Things to improve

- 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

Next steps

	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

http://dioscuri.sourceforge.net

Questions?

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