

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:

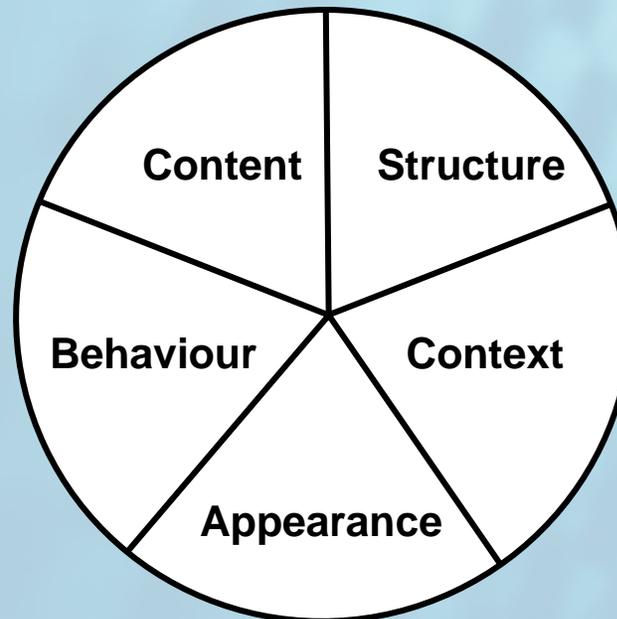
<http://www.ecoaction.com.au/res/Image/junkedcomputers.gif>

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

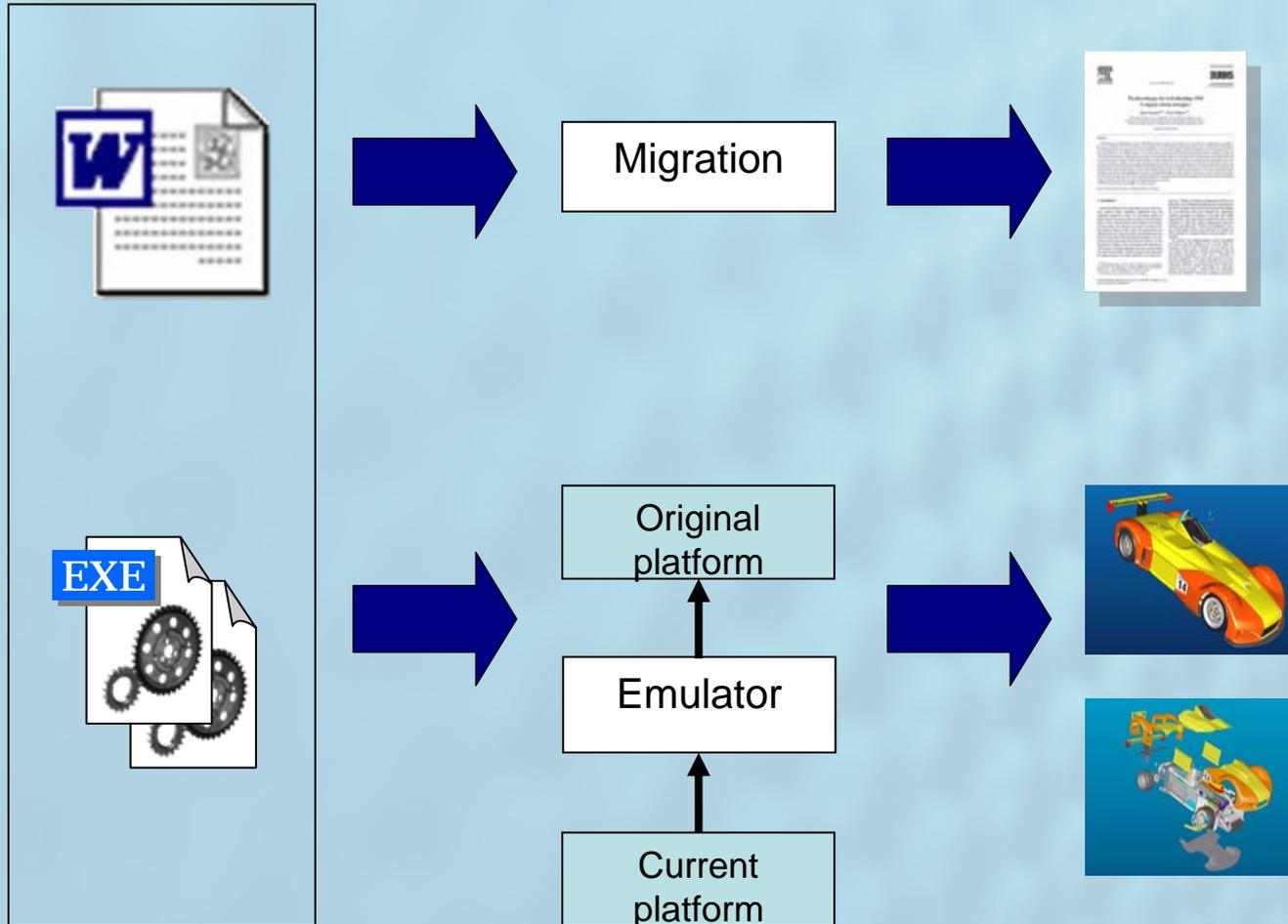
Digital preservation action!

Three basic criteria for defining action:

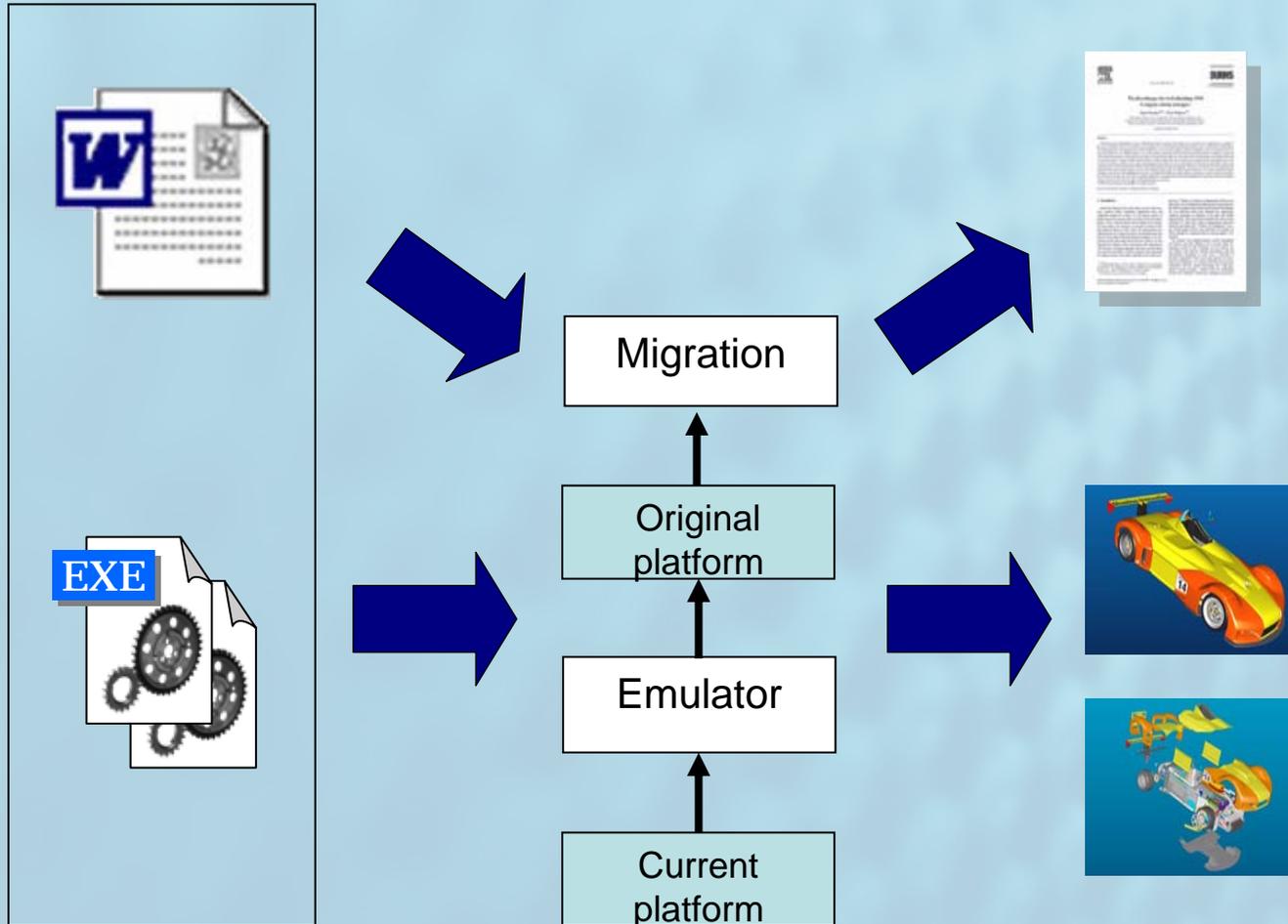
- Policy of the institution
- (Future) user requirements
- Kind of digital object:



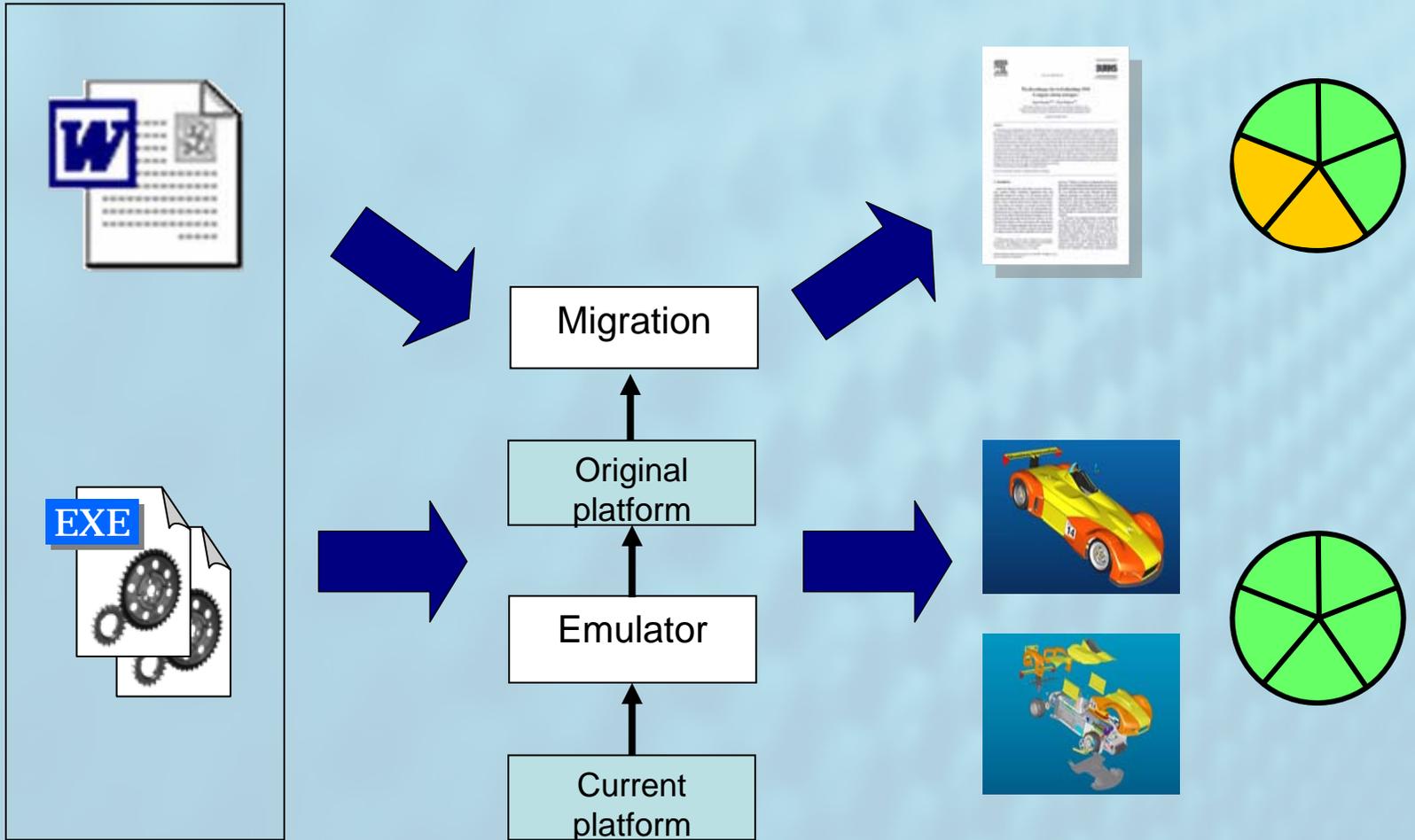
Migration & emulation



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

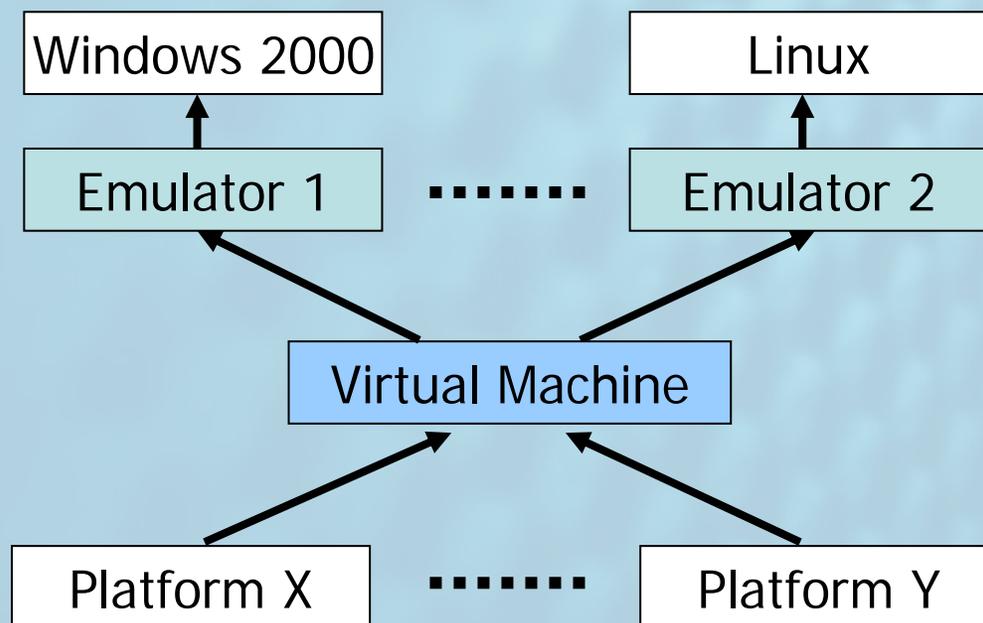
Project outline

- 
- A vertical black arrow pointing downwards, indicating the chronological order of the project milestones.
- 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?

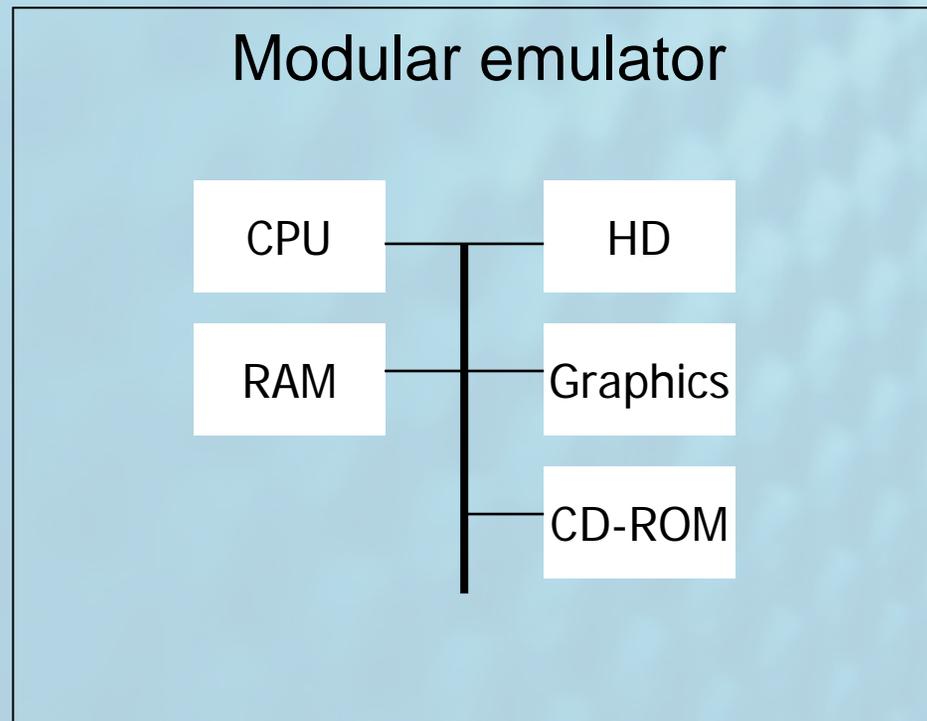
Two key features: durability and modularity

- ❑ 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

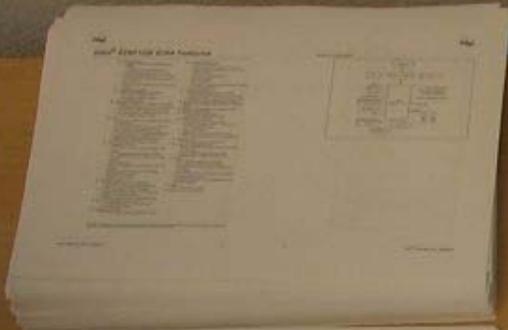
Jeff

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.

External Floppy Controller Specifications Version 1.0

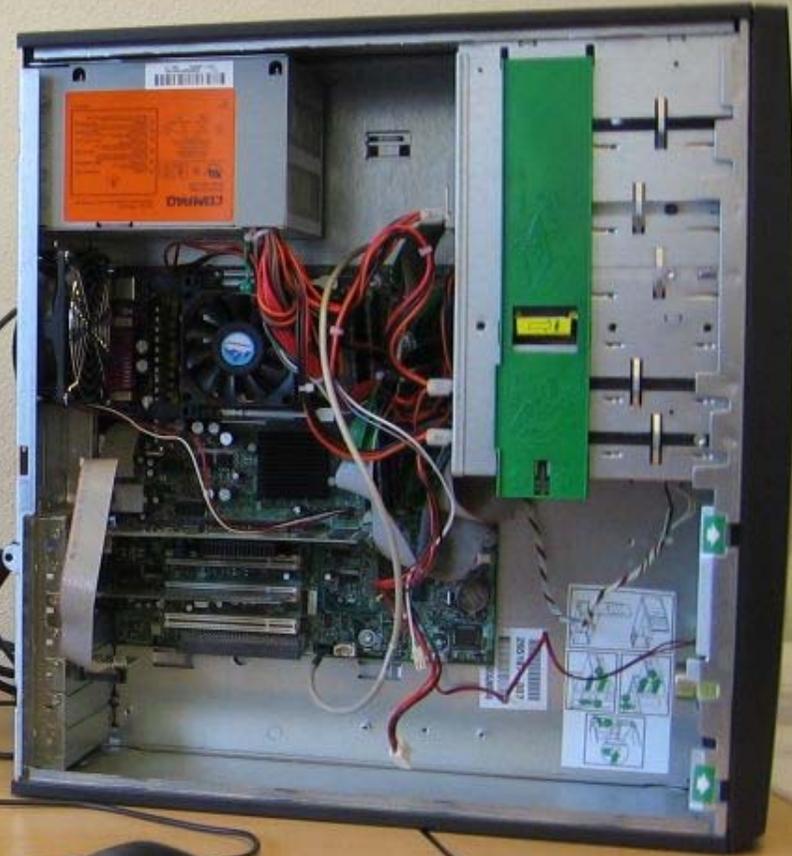
This document provides the specifications for the external floppy controller. It includes information about the controller's capabilities, supported floppy disk formats, and the required hardware connections.

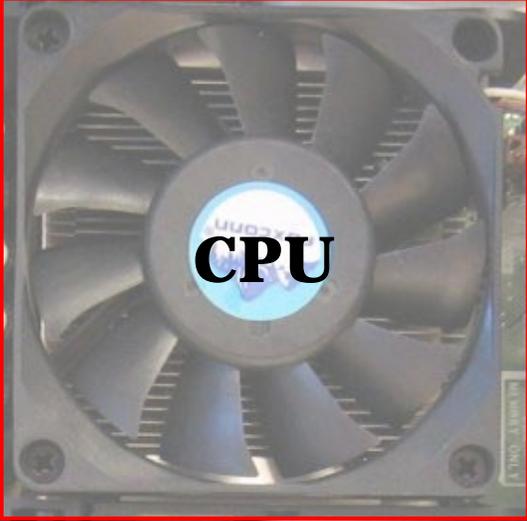


Intel 80486 Architecture

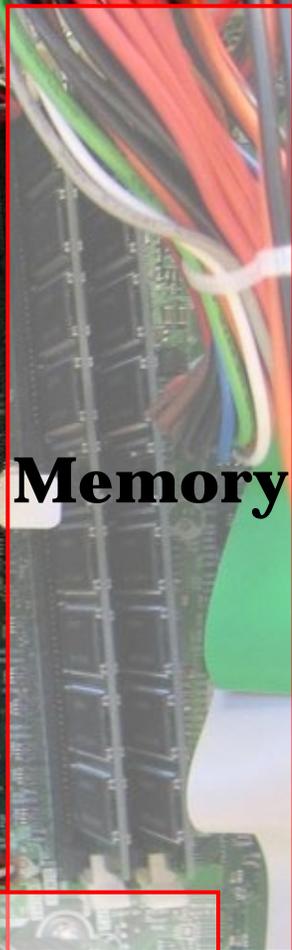
This document describes the architecture of the Intel 80486 processor. It details the internal components, including the cache, bus, and execution units, and explains how they interact to perform various tasks.

Pin	Signal	Direction
1	AD[15]	Out
2	AD[16]	Out
3	AD[17]	Out
4	AD[18]	Out
5	AD[19]	Out
6	AD[20]	Out
7	AD[21]	Out
8	AD[22]	Out
9	AD[23]	Out
10	AD[24]	Out
11	AD[25]	Out
12	AD[26]	Out
13	AD[27]	Out
14	AD[28]	Out
15	AD[29]	Out
16	AD[30]	Out
17	AD[31]	Out
18	AD[32]	Out
19	AD[33]	Out
20	AD[34]	Out
21	AD[35]	Out
22	AD[36]	Out
23	AD[37]	Out
24	AD[38]	Out
25	AD[39]	Out
26	AD[40]	Out
27	AD[41]	Out
28	AD[42]	Out
29	AD[43]	Out
30	AD[44]	Out
31	AD[45]	Out
32	AD[46]	Out
33	AD[47]	Out
34	AD[48]	Out
35	AD[49]	Out
36	AD[50]	Out
37	AD[51]	Out
38	AD[52]	Out
39	AD[53]	Out
40	AD[54]	Out
41	AD[55]	Out
42	AD[56]	Out
43	AD[57]	Out
44	AD[58]	Out
45	AD[59]	Out
46	AD[60]	Out
47	AD[61]	Out
48	AD[62]	Out
49	AD[63]	Out
50	AD[64]	Out
51	AD[65]	Out
52	AD[66]	Out
53	AD[67]	Out
54	AD[68]	Out
55	AD[69]	Out
56	AD[70]	Out
57	AD[71]	Out
58	AD[72]	Out
59	AD[73]	Out
60	AD[74]	Out
61	AD[75]	Out
62	AD[76]	Out
63	AD[77]	Out
64	AD[78]	Out
65	AD[79]	Out
66	AD[80]	Out
67	AD[81]	Out
68	AD[82]	Out
69	AD[83]	Out
70	AD[84]	Out
71	AD[85]	Out
72	AD[86]	Out
73	AD[87]	Out
74	AD[88]	Out
75	AD[89]	Out
76	AD[90]	Out
77	AD[91]	Out
78	AD[92]	Out
79	AD[93]	Out
80	AD[94]	Out
81	AD[95]	Out
82	AD[96]	Out
83	AD[97]	Out
84	AD[98]	Out
85	AD[99]	Out
86	AD[100]	Out
87	AD[101]	Out
88	AD[102]	Out
89	AD[103]	Out
90	AD[104]	Out
91	AD[105]	Out
92	AD[106]	Out
93	AD[107]	Out
94	AD[108]	Out
95	AD[109]	Out
96	AD[110]	Out
97	AD[111]	Out
98	AD[112]	Out
99	AD[113]	Out
100	AD[114]	Out
101	AD[115]	Out
102	AD[116]	Out
103	AD[117]	Out
104	AD[118]	Out
105	AD[119]	Out
106	AD[120]	Out
107	AD[121]	Out
108	AD[122]	Out
109	AD[123]	Out
110	AD[124]	Out
111	AD[125]	Out
112	AD[126]	Out
113	AD[127]	Out
114	AD[128]	Out
115	AD[129]	Out
116	AD[130]	Out
117	AD[131]	Out
118	AD[132]	Out
119	AD[133]	Out
120	AD[134]	Out
121	AD[135]	Out
122	AD[136]	Out
123	AD[137]	Out
124	AD[138]	Out
125	AD[139]	Out
126	AD[140]	Out
127	AD[141]	Out
128	AD[142]	Out
129	AD[143]	Out
130	AD[144]	Out
131	AD[145]	Out
132	AD[146]	Out
133	AD[147]	Out
134	AD[148]	Out
135	AD[149]	Out
136	AD[150]	Out
137	AD[151]	Out
138	AD[152]	Out
139	AD[153]	Out
140	AD[154]	Out
141	AD[155]	Out
142	AD[156]	Out
143	AD[157]	Out
144	AD[158]	Out
145	AD[159]	Out
146	AD[160]	Out
147	AD[161]	Out
148	AD[162]	Out
149	AD[163]	Out
150	AD[164]	Out
151	AD[165]	Out
152	AD[166]	Out
153	AD[167]	Out
154	AD[168]	Out
155	AD[169]	Out
156	AD[170]	Out
157	AD[171]	Out
158	AD[172]	Out
159	AD[173]	Out
160	AD[174]	Out
161	AD[175]	Out
162	AD[176]	Out
163	AD[177]	Out
164	AD[178]	Out
165	AD[179]	Out
166	AD[180]	Out
167	AD[181]	Out
168	AD[182]	Out
169	AD[183]	Out
170	AD[184]	Out
171	AD[185]	Out
172	AD[186]	Out
173	AD[187]	Out
174	AD[188]	Out
175	AD[189]	Out
176	AD[190]	Out
177	AD[191]	Out
178	AD[192]	Out
179	AD[193]	Out
180	AD[194]	Out
181	AD[195]	Out
182	AD[196]	Out
183	AD[197]	Out
184	AD[198]	Out
185	AD[199]	Out
186	AD[200]	Out
187	AD[201]	Out
188	AD[202]	Out
189	AD[203]	Out
190	AD[204]	Out
191	AD[205]	Out
192	AD[206]	Out
193	AD[207]	Out
194	AD[208]	Out
195	AD[209]	Out
196	AD[210]	Out
197	AD[211]	Out
198	AD[212]	Out
199	AD[213]	Out
200	AD[214]	Out
201	AD[215]	Out
202	AD[216]	Out
203	AD[217]	Out
204	AD[218]	Out
205	AD[219]	Out
206	AD[220]	Out
207	AD[221]	Out
208	AD[222]	Out
209	AD[223]	Out
210	AD[224]	Out
211	AD[225]	Out
212	AD[226]	Out
213	AD[227]	Out
214	AD[228]	Out
215	AD[229]	Out
216	AD[230]	Out
217	AD[231]	Out
218	AD[232]	Out
219	AD[233]	Out
220	AD[234]	Out
221	AD[235]	Out
222	AD[236]	Out
223	AD[237]	Out
224	AD[238]	Out
225	AD[239]	Out
226	AD[240]	Out
227	AD[241]	Out
228	AD[242]	Out
229	AD[243]	Out
230	AD[244]	Out
231	AD[245]	Out
232	AD[246]	Out
233	AD[247]	Out
234	AD[248]	Out
235	AD[249]	Out
236	AD[250]	Out
237	AD[251]	Out
238	AD[252]	Out
239	AD[253]	Out
240	AD[254]	Out
241	AD[255]	Out
242	AD[256]	Out
243	AD[257]	Out
244	AD[258]	Out
245	AD[259]	Out
246	AD[260]	Out
247	AD[261]	Out
248	AD[262]	Out
249	AD[263]	Out
250	AD[264]	Out
251	AD[265]	Out
252	AD[266]	Out
253	AD[267]	Out
254	AD[268]	Out
255	AD[269]	Out
256	AD[270]	Out
257	AD[271]	Out
258	AD[272]	Out
259	AD[273]	Out
260	AD[274]	Out
261	AD[275]	Out
262	AD[276]	Out
263	AD[277]	Out
264	AD[278]	Out
265	AD[279]	Out
266	AD[280]	Out
267	AD[281]	Out
268	AD[282]	Out
269	AD[283]	Out
270	AD[284]	Out
271	AD[285]	Out
272	AD[286]	Out
273	AD[287]	Out
274	AD[288]	Out
275	AD[289]	Out
276	AD[290]	Out
277	AD[291]	Out
278	AD[292]	Out
279	AD[293]	Out
280	AD[294]	Out
281	AD[295]	Out
282	AD[296]	Out
283	AD[297]	Out
284	AD[298]	Out
285	AD[299]	Out
286	AD[300]	Out
287	AD[301]	Out
288	AD[302]	Out
289	AD[303]	Out
290	AD[304]	Out
291	AD[305]	Out
292	AD[306]	Out
293	AD[307]	Out
294	AD[308]	Out
295	AD[309]	Out
296	AD[310]	Out
297	AD[311]	Out
298	AD[312]	Out
299	AD[313]	Out
300	AD[314]	Out
301	AD[315]	Out
302	AD[316]	Out
303	AD[317]	Out
304	AD[318]	Out
305	AD[319]	Out
306	AD[320]	Out
307	AD[321]	Out
308	AD[322]	Out
309	AD[323]	Out
310	AD[324]	Out
311	AD[325]	Out
312	AD[326]	Out
313	AD[327]	Out
314	AD[328]	Out
315	AD[329]	Out
316	AD[330]	Out
317	AD[331]	Out
318	AD[332]	Out
319	AD[333]	Out
320	AD[334]	Out
321	AD[335]	Out
322	AD[336]	Out
323	AD[337]	Out
324	AD[338]	Out
325	AD[339]	Out
326	AD[340]	Out
327	AD[341]	Out
328	AD[342]	Out
329	AD[343]	Out
330	AD[344]	Out
331	AD[345]	Out
332	AD[346]	Out
333	AD[347]	Out
334	AD[348]	Out
335	AD[349]	Out
336	AD[350]	Out
337	AD[351]	Out
338	AD[352]	Out
339	AD[353]	Out
340	AD[354]	Out
341	AD[355]	Out
342	AD[356]	Out
343	AD[357]	Out
344	AD[358]	Out
345	AD[359]	Out
346	AD[360]	Out
347	AD[361]	Out
348	AD[362]	Out
349	AD[363]	Out
350	AD[364]	Out
351	AD[365]	Out
352	AD[366]	Out
353	AD[367]	Out
354	AD[368]	Out
355	AD[369]	Out
356	AD[370]	Out
357	AD[371]	Out
358	AD[372]	Out
359	AD[373]	Out
360	AD[374]	Out
361	AD[375]	Out
362	AD[376]	Out
363	AD[377]	Out
364	AD[378]	Out
365	AD[379]	Out
366	AD[380]	Out
367	AD[381]	Out
368	AD[382]	Out
369	AD[383]	Out
370	AD[384]	Out
371	AD[385]	Out
372	AD[386]	Out
373	AD[387]	Out
374	AD[388]	Out
375	AD[389]	Out
376	AD[390]	Out
377	AD[391]	Out
378	AD[392]	Out
379	AD[393]	Out
380	AD[394]	Out
381	AD[395]	Out
382	AD[396]	Out
383	AD[397]	Out
384	AD[398]	Out
385	AD[399]	Out
386	AD[400]	Out
387	AD[401]	Out
388	AD[402]	Out
389	AD[403]	Out
390	AD[404]	Out
391	AD[405]	Out
392	AD[406]	Out
393	AD[407]	Out
394	AD[408]	Out
395	AD[409]	Out
396	AD[410]	Out
397	AD[411]	Out
398	AD[412]	Out
399	AD[413]	Out
400	AD[414]	Out
401	AD[415]	Out
402	AD[416]	Out
403	AD[417]	Out
404	AD[418]	Out
405	AD[419]	Out
406	AD[420]	Out
407	AD[421]	Out
408	AD[422]	Out
409	AD[423]	Out
410	AD[424]	Out
411	AD[425]	Out
412	AD[426]	Out
413	AD[427]	Out
414	AD[428]	Out
415	AD[429]	Out
416	AD[430]	Out
417	AD[431]	Out
418	AD[432]	Out
419	AD[433]	Out
420	AD[434]	Out
421	AD[435]	Out
422	AD[436]	Out
423	AD[437]	Out
424	AD[438]	Out
425	AD[439]	Out
426	AD[440]	Out
427	AD[441]	Out
428	AD[442]	Out
429	AD[443]	Out
430	AD[444]	Out
431	AD[445]	Out
432	AD[446]	Out
433	AD	





CPU



Memory



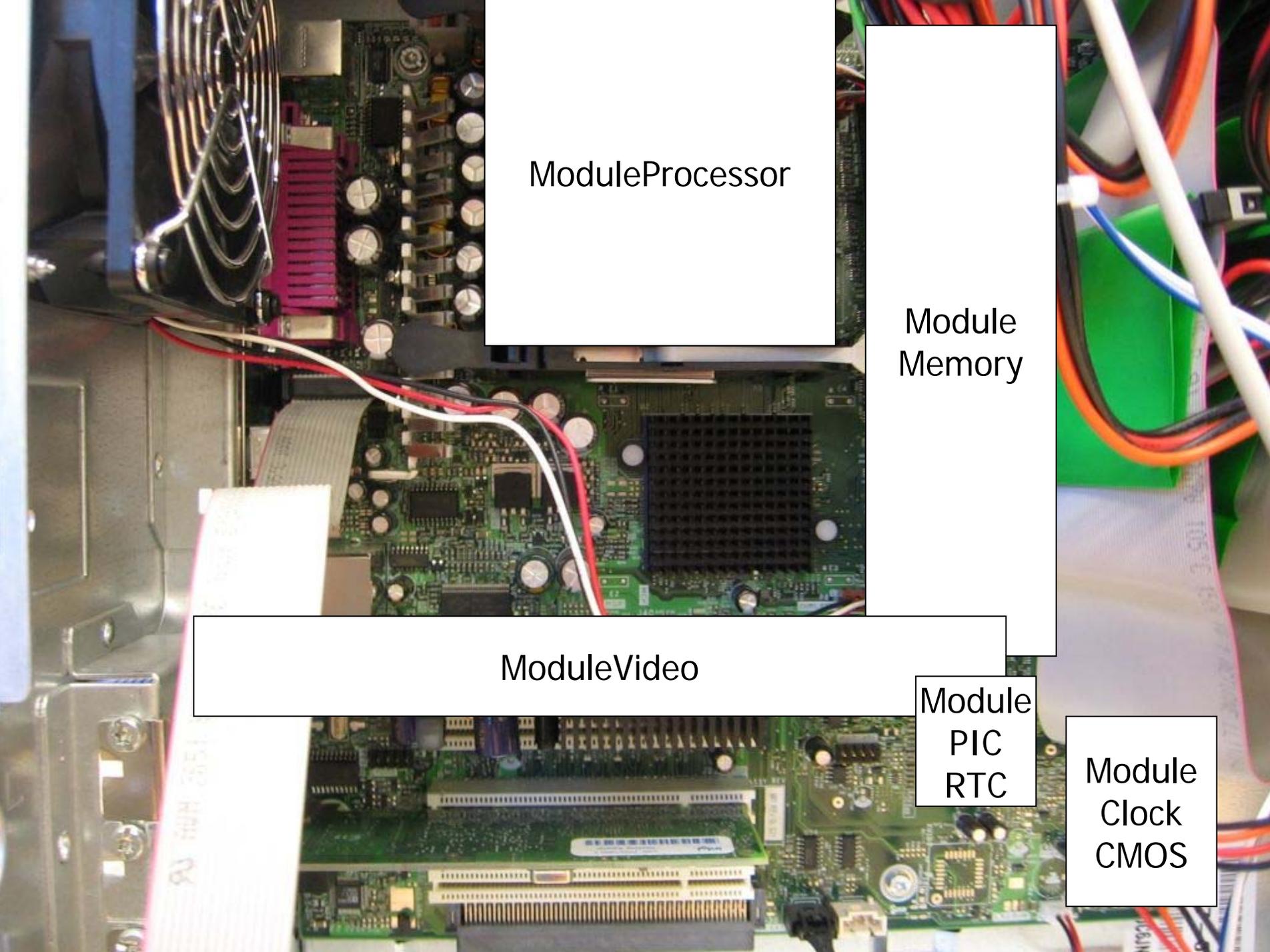
Graphics card



**PIC
RTC**



BIOS



ModuleProcessor

Module
Memory

ModuleVideo

Module
PIC
RTC

Module
Clock
CMOS

- 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
	Full				Io	sys
M	Info				Msdos	sys
a	Tree				attrib	exe
a	On/Off		Ctrl-F1		autoexec	bat
c	<hr/>					
c	√ Name				undelete	exe
c	eXtension				unformat	com
d	tiMe				xcopy	exe
e	SiZe				command	com
e	Unsorted				config	sys
f	<hr/>					
f	Re-read				dbview	exe
h	Drive...		Alt-F1		edit	com
l	<hr/>					
mem	exe				edit	hlp
nc	exe				fdisk	exe
ncsmall	exe				format	com
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
<hr/>				
1	Name	-12-05	15:14	
	eXtension	-12-05	15:14	
	tiMe	-08-06	17:11	
	SiZe	-12-05	11:59	
	Unsorted	-12-05	20:20	
		-07-06	15:21	
		-10-06	17:06	
		-02-07	14:00	

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)

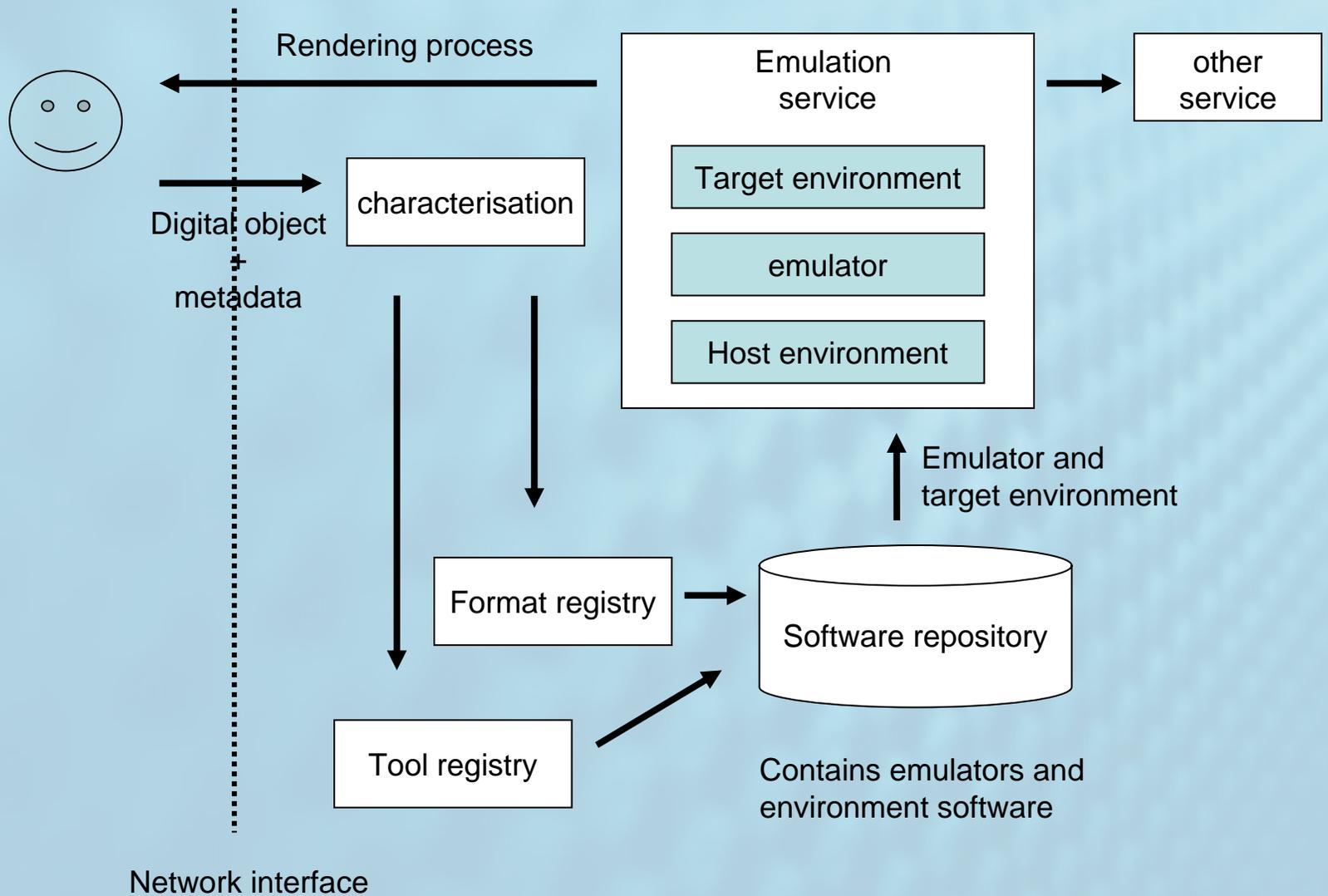
Integration with
e-Depot process flow

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