Emulation Experts Meeting in The Hague October 06

Chair of Communication Systems - Prof. G. Schneider
Department of Applied Sciences
University of Freiburg
Randolph Welte and Dirk von Suchodoletz
Longtime preservation of digital objects
Project at University of Freiburg (ALUF)

- Project started at around 2004 as for PhD thesis of Randolph Welte and Dirk von Suchodoletz
- Focus is on a broad range of digital objects (DO)
  - “classical” objects like static documents and pictures of a broad range of formats
  - digitized or digital audio and video material
  - several multimedia content, like educational CDROMs, encyclopedias, ...
  - exclusively digital objects like applications for generating other digital objects, scientific use, leisure, ....
  - “digital environment” like operating systems, (graphical) user interfaces
Longtime preservation of digital objects
Different kinds of DO

- Digital world becomes a part of mankind's history, too. So we would like to preserve and migrate more than just static material.
- There are other options for preserving static objects like documents and audio/video – migration, but that strategy could not be trivially applied to the other digital material.
- Emulation is not a completely new concept in computer world:
  - Jeff Rothenberg introduced it for quite a while into the discussion of longtime preservation.
  - Emulation concepts are known since mainframe era.
Longtime preservation of digital objects
Early available types of emulation/emulators

- The original idea for the longtime preservation project at ALUF came from the availability of home computer games on modern IA32 architecture
- A wide range of different emulators or a wide range of disappeared platforms were made available by group of enthusiastic people who would like to preserve the systems and especially the software
- On the other side of the spectrum the company VMware offered at that point its third version of the virtual workstation
- That “workstation” was able to run a complete different operating system ontop of another one
Longtime preservation of digital objects
Experiments with emulators

- Started our experiments with several home computer emulators
- The mess/mame project is bundling the efforts into one component based system
  - multi emulator super system implements a range of different CPUs, IO systems and thus is able to emulate a wide range of former popular home computer systems
  - generally interesting concept to handle different architectures
Longtime preservation of digital objects
Evaluations of emulators

- Looked at other kinds of emulators too
  - Vendors often provide themselves development platforms: copilot/ pose
  - Helps developers to design software without actual availability of hardware
- Tried out Apple system emulation with PearPC
- There is a range of very special architecture emulators available (educational purposes), but did not look deeper into that field
Longtime preservation of digital objects
Evaluation of emulators

- Another field of research was focused on workstation/mainframe emulators
  - systems used in companies and research organizations
    - we had a look at the “Herkules” project – emulator for the IBM mainframe S370 series
    - too complex, too small knowledge on our side
    - at least generated ideas of needed meta data
Longtime preservation of digital objects
Evaluation of emulators

- For most users Intel architecture (IA32) emulation is vital
  - most of DO generated and executed on that platform
- Looked at the simple DOSBOX tool
  - emulates a typical DOS environment without MS- or DR DOS system files needed
  - Available only for LINUX OS, usable for some applications only, mostly games
Longtime preservation of digital objects
VMware – first relevant X86 emulator

- VMware “Workstation” was the first relevant X86 emulator
  - executable on Windows and Linux operating systems
  - defining a complete virtual environment of a PC
  - implementing its own BIOS, IO subsystem, graphics card
  - modeled network and audio interface after some real hardware like the AMD Pcmnet32 netcard and soundblaster SB16 audio card, later the es1371 chip
  - offered IDE and SCSI bus emulation for attachment of real or virtual drives
  - implements floppy disk interface and interfaces to serial, parallel and later on USB devices
Longtime preservation of digital objects
VMware – 3.0 and 3.2

- Started with VMware 3.0 version and then moved to 3.2
Longtime preservation of digital objects
VMware – 3.0 and 3.2

- First major problem occurred – virtual disk image format changed
  - VMware offered a conversion operation implemented directly into the tool
- Other problems occurred with the adaptation of the tool to the Linux environment – if the kernel changed you had to recompile the modules
- Next issue is the non-existent layer of CPU abstraction – VMware only runs on IA32 and later on IA64 CPUs
Longtime preservation of digital objects
VMware – 3.2 / emulator-in-emulator tests

- Did emulation in emulation testing
Longtime preservation of digital objects
VMware – 3.2 / emulator-in-emulator tests

- Run home computer emulator within a Windows98 which itself run in VMware ontop of Linux kernel 2.4
  - demonstrated possibility of interleaving emulators
  - performance was OK
  - user interface was usable: Screen output of the home computer emulator could use the VMware output, keyboard input was passed correctly to the running game

- Next idea was to create a package: Put set of emulators in a whole into a VMware image
  - keeps setup intact
  - allows sending around complete, running package to other users
Longtime preservation of digital objects
VMware – 3.2 / emulator-in-emulator tests

- “Execution environment” with UI of home computer emulators
Longtime preservation of digital objects
VMware – 3.2 / emulator-in-emulator results

- Hoped to put the needed knowledge to run the home computer emulators and load the several file types into one single container and provide a convenient user interface
  - our approach was much to complicated and too specific for just one type of emulation
  - put all available application and game ROMs into that image, but how to insert user provided stuff
  - the ROMs were only accessible within running VMware/Win98 but not directly from outside
Longtime preservation of digital objects
VMware – Version 4 and 5 series

- Moved on with the upcoming versions of VMware Workstation, migrated our images
  - type of container format changed again
  - new type of problem: Virtual hardware changed significantly
    - from ISA components to PCI
    - new IO system – from P1/TX chipset to P2/BX
    - most OS required a new set of drivers
  - not for all X86 a set of specific drivers is available – e.g. starts with Windows95 for the guest side
  - could be complicated to run a rather old virtual workstation version on recent host OS
Longtime preservation of digital objects
VMware – alternatives

- VMware product range was extended and pricing changed significantly
  - free player was introduced with version 5.0/5.5
  - VMware server is freely available
    - interesting to host more than one guest running at the same time
    - guest instances are accessible via VNC like interface over the net
- Other players arrived in the market: Microsoft/Connectix VirtualPC, Parallels Workstation
- But all of them commercial products and not Open Source
Longtime preservation of digital objects
VMware – alternatives

● In the meantime (project start was more than two years ago) other interesting projects in the X86 emulation field emerged/matured
  – Bochs/Plex86 is a rather old project aiming on being an alternative to VMware and alike, we will hear more on that later today
  – Qemu is a vital Open Source project using a component approach, so more than just X86 emulation could be done with it
● Especially interesting with Open Source emulation is, that you could integrate these projects in your efforts and pass back work you have done in improving, testing, ...
● So e.g. ISA bus should be available for “eternity”
Longtime preservation of digital objects
VMware and other emulators – IO issues

- During all these experiments the data IO issues arose
  - how to pass data from host environment into the guest (emulated environment) and vice versa
- For not too old systems TCP/IP networking is an option
  - longtime stable standard, uncritical if only internal software network links are used
  - but you cannot rely on availability (driver issues, ...)
- Floppy disk images are very common, bidirectional
- Same applies to ISO image format, but is only one way
Longtime preservation of digital objects
First conclusion

● From all these experiments we got the idea, that meta information is most vital
  – how to install and run a certain emulator
  – how to install / run any software/DO within it
  – which additional software components are needed to run a DO successfully, e.g.
    ● CDROM driver for DOS
    ● Network and audio drivers
● Tools to make floppy and ISO images
● Tools to access the container formats of emulators
Longtime preservation of digital objects  
Project involvement

- Thus our project focuses on to providing a toolbox for emulation environments
  - emulators
  - operating systems
  - applications
  - all additional software/data

- Later on we were brought in contact to the efforts of the KB/NL using emulation and finally became part of PLANETS
Longtime preservation of digital objects
OAIS and dissemination

- In our part we want to focus on the dissemination side of the OAIS model – help the user to access his data
- OAIS (Open Archival Information System) produces a digital object of any type but is not very specific on the type of it
Longtime preservation of digital objects
OAIS and dissemination

- Toolbox should find/provide the user with a “view-path” to an object
  1. get an object from the OIAS archive
     Unpack it and get the metadata which should contain at least the file format and hopefully the generating applications.

  2. if no application was named or if the object consists of a generic application format, like e.g. **png**, **jpeg** or **pdf**, then we should lookup for the file format, followed by getting the applications / environments to view, access, manipulate and execute
Longtime preservation of digital objects
Toolbox and view path

3. the application will require some OS’s to run.
4. the OS requires some emulators
5. emulators require some host system/architecture to be executed on (might be emulated itself)
Longtime preservation of digital objects
Emulation and long-term prospects

- Thus we hope to have for our DO
  - one „view path“ (the steps –all or subsets- are named above)
  - multiple „view pathes“.
  - no „view path“ in the worst case

- To ensure the future usage of the digital objects in our toolbox we could use a „Shell-in-Shell“ method as discussed earlier
Longtime preservation of digital objects
Toolbox and end user services

• Most of the work we have done is not trivially replicatable for the average user
• Thus our aim is to provide sample reference workstations with emulators and “caches” of images already installed for “simplicistic” usage
• Provide a web interface
Longtime preservation of digital objects

End of ALUF presentation

- Thank you for your patience :-(