# Machine Virtualization for Better Hardware Utilization and Efficient Resource Management (second part)

*DAAD Summer School: Aspects of Large Scale High Speed Computing*

*17th March 2011*

Dr. Dirk von Suchodoletz

Faculty of Engineering, University Freiburg

UNI FREIBURG

# Structure: Desktop Virtualization



UNI FREIBURG

Practical Application

Running Windows Painlessly

Offer Flexible Lecture Pools

# Virtualization in Freiburg



UNI FREIBURG

- Using since 2003/4 in production scenarios
- Nevertheless, not all of Live migration, Dynamic resizing, Snapshotting, Isolation, Provisioning, available in all software products, especially not the free/Open Source variants
- Main application in the beginning: Desktop virtualization
  - Running other X86 operating systems on top of Linux Desktop, especially Windows
  - Extending this by offering a session chooser to host wide variety of different courses



- Challenges
  - Lecture pools have to be very robust
  - Wide variety of lectures run: Starting with simple How-to-Use-Office up to complex statistical, geographical software packages
  - One, unified installation is impossible
    - Software conflicts
    - Conflicting requirements of lecturers and users
  - Multi-boot on same disk not an option
    - Difficult to handle, how to make updates if courses are run on other system, ...

# Flexible Desktop Systems





- Number of lecture pools in the computer center and university library (~80 machines)

# Flexible Desktop Systems





- Windows is not easy deployable in a stateless manner like discussed for Linux in first lecture (no idea why)
  - BartPE & Co. not official solutions and pretty restricted in a number of aspects
- Traditional solution: Use several, removable IDE disks with Windows variants on it
  - Reinstall disks if a new software, course setup is required
  - Exchange disks between courses
- Tedious task, not very flexible, introducing lots of errors

# Flexible Desktop Systems



UNI FREIBURG

- Completely new approach: Use Stateless Linux system and virtualization
- Main goal: Easy provisioning, malware protection, isolation, software conflict avoidance
- Started with VMware Player
  - Free, but proprietary
  - Still in use
  - Looking into alternatives like QEMU/KVM and VirtualBox
- Requirements like USB access not as convenient as with VMware Player yet

# Stateless Windows



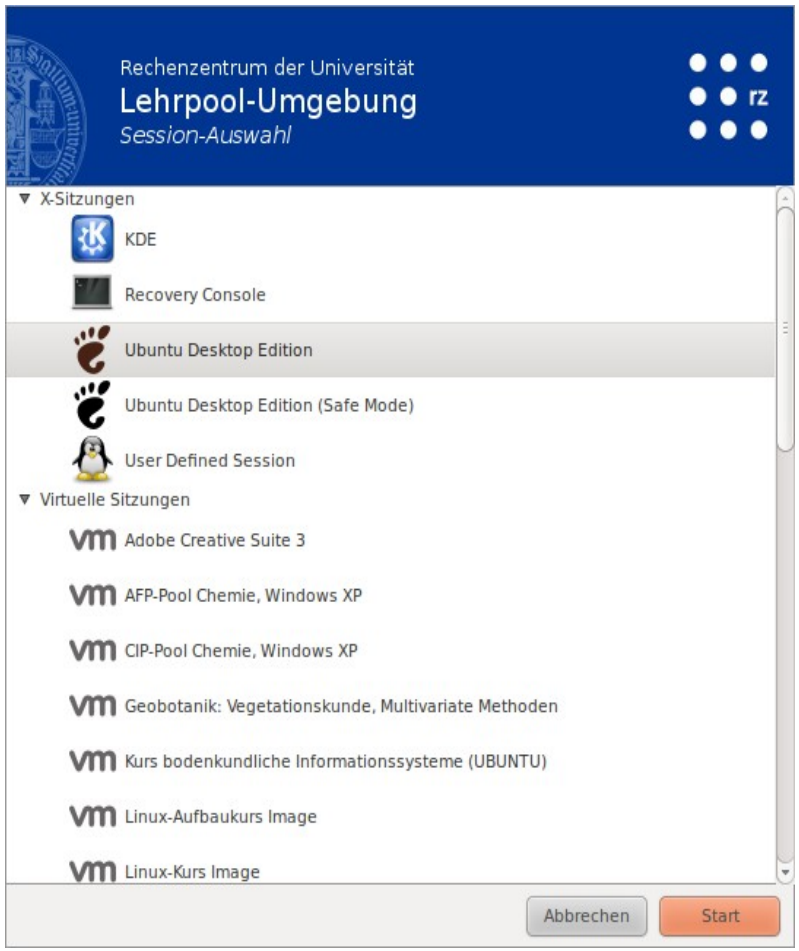


- Virtual machine images provided via read-only NFS from a central fileserver
- Mounted by every client
- Virtual machine runs in non-persistent mode
- All run-time data is stored locally on the client and discarded after the session
- Virtual machine configured automatically by scripts

# Windows Sessions



- User logs into Linux environment
- Presented with a session selection tool programmed for that purpose
- All active lecture pool virtual machines listed after standard Linux desktop sessions
- Additional information on contained applications is provided



# Windows Configuration





- Each virtual machine is configured by the lecturer on his/her machine
- Provide a prepared image with:
  - Most (networking) services switched off
  - No indexing
  - No permanently running virus, malware scanners
  - No Windows swap file
  - Share mounting (home directory and other file shares) and printer configuration utility installed

# Windows Configuration





- Images then extended by the lecturers of each course for their needs
- Simple to handle as Windows is used in standard desktop mode
- No special requirements on network based installations or restrictions
- Distribution of workload within institution
  - Pool administration is separated from lecture preparations
  - Asynchronous preparation: Completely independent on courses running on the hardware
  - No silly software installation demands

# Conclusions



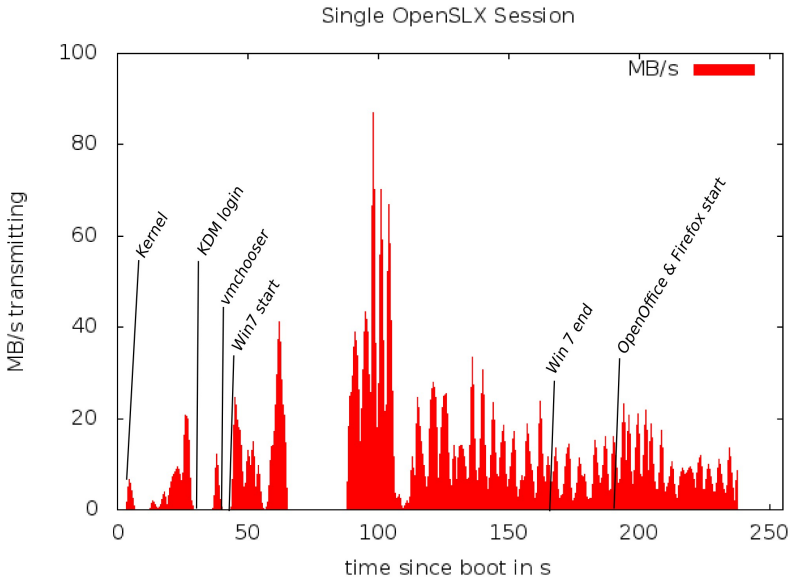


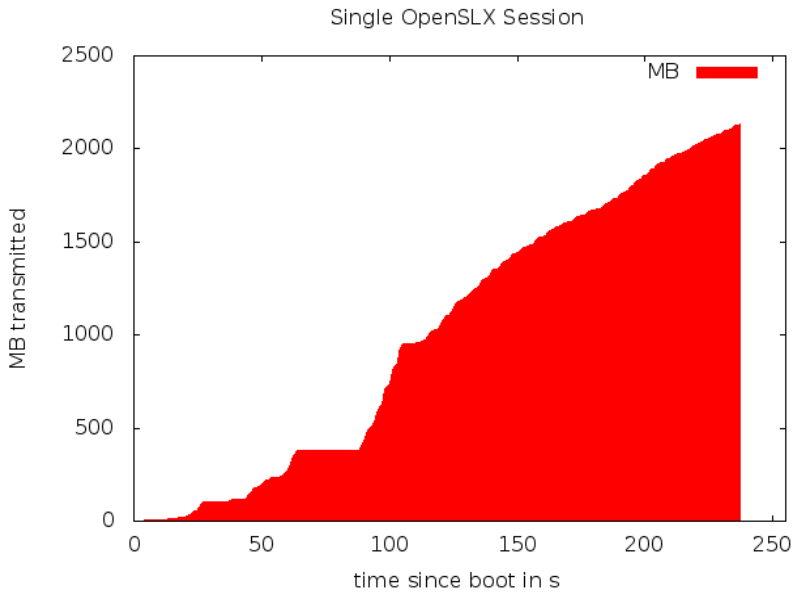
- After a phase of trial&error – significantly reduced efforts to run a Windows lecture pool
- Result: Stateless Windows, like stateless Linux
- Other courses possible too: Regular Linux networking with root access for participants in virtual machines
- Courses could be switched with virtually no delay
  - One course ending, students leaving, logging off
  - Next course starting few minutes later
- Several hundred courses run, often over 20 different system images offered

# Conclusions



- Insignificantly longer setup time as two operating systems are started
- Higher network peek loads – remember the graphs of last lecture (2GByte + of network traffic until Windows 7 session was prepared)



# Conclusions





- All sessions
  - Look same on every desktop running same virtual machine
  - Stateless – if rebooted, clean slate (no chance for malware, easy reset after misconfiguration)
- Disadvantage – no traditional Windows profiles, persistent desktop decorations, …
  - Often not really a problem (or even an advantage: Just think of avoiding to copy a desktop profile containing an ISO image of 4.7GByte of size)

# Conclusions





- After a phase of trial&error – significantly reduced efforts to run lecture pools
- VMplayer is tricky to run in optimal memory configuration, especially if using no disk at all
- Optimizations: Using RamZSwap available with newer kernels and overbooking of TempFS
- Windows in Bridged-Network mode produced new problem: Same SSID/machine name of every running instance
- Windows 7 annoys with CPU “driver” reconfiguration
- Ethernet VLAN problems in earlier VMware versions

# Optimizations





- Use “Shared Folders” instead of CIFS mounts
- Idea: Use suspend, resume to speed up system start
  - Every session has to be started nevertheless (from same image), why not producing a snapshot from common part and resume from there
  - Problem: Different base hardware prevents this as CPU is virtualized only (different types of CPU, AMD, Intel, (no) virtualization extensions)
- Use Linux KVM for better memory utilization
- Allow easier configuration of multiple virtual machines on a single desktop machine

# Structure: Virtualization Applications

UNI FREIBURG

Service Consolidation

Experiences of Usability

General Challenges

# Server Consolidation





- Traditional service consolidation with virtual machines
- Used VMware Server to test and prepare a wide range of different Linux systems
  - Preparation images to be cloned for the stateless Linux pool
  - Testings of configurations, software installations
  - More than 30 different systems installed
  - Typically less then five run in parallel, rest suspended

# VMware Server



UNI FREIBURG

- VMware Server on our SUN V20z Dual-Opteron 6GByte machine of 2005
  - Started with version 1, had to switch to version 2 because of stopped kernel support for newer Linux versions
  - Okay performance for testing stuff
  - Not much fun with server operation on it
  - Mounting problems with the server console to access the systems directly (KVM style)
  - Number of problems with the web interface
- Stopped using it last year

# XEN on Linux

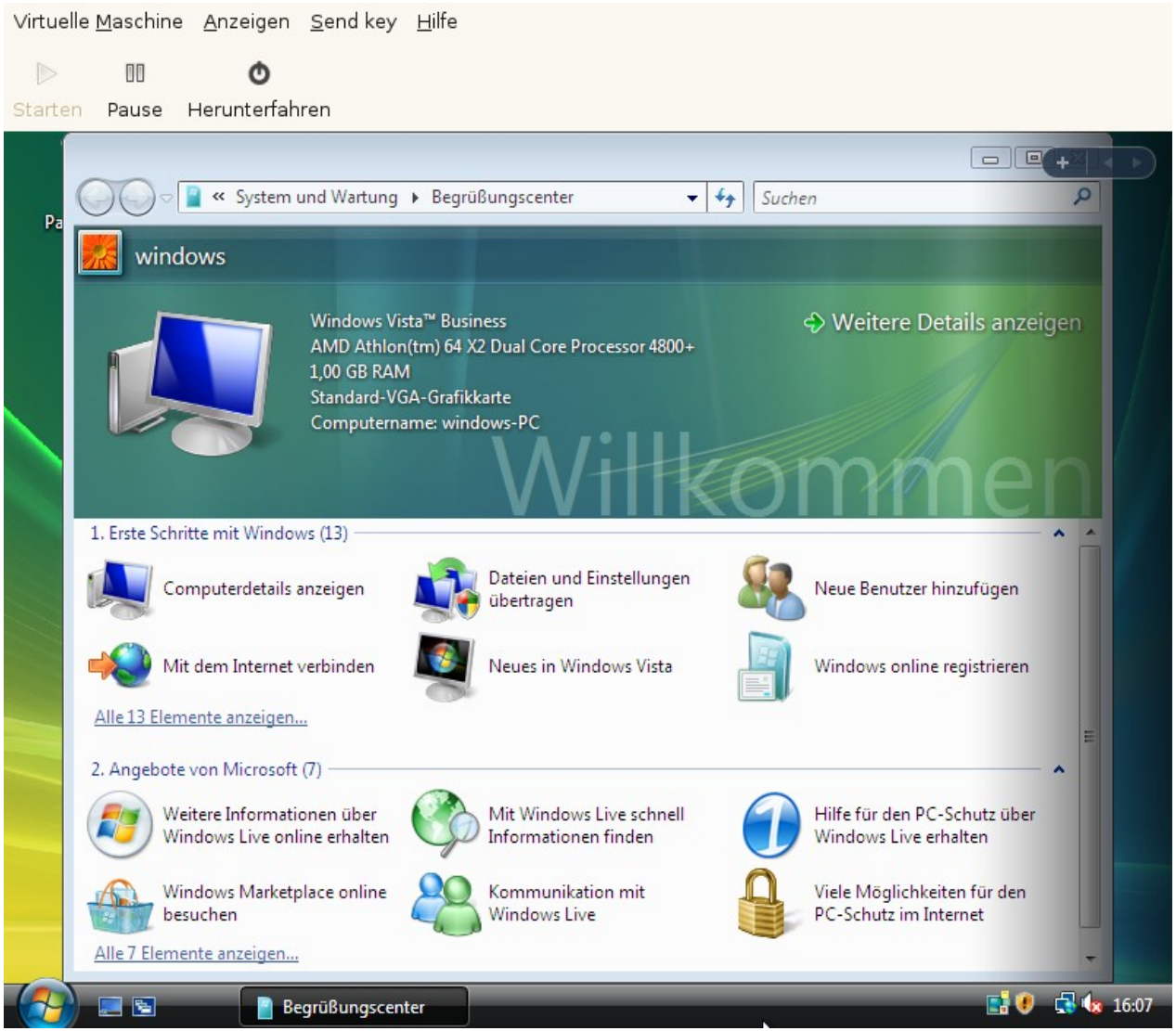




- After disappointing performance of VMware Server
- Number of tests with XEN
  - Unclear project status, no official Linux kernel inclusion
  - More complex to setup
  - Missing Linux kernel support for a while
  - Test systems were running OK
  - Big issue: Proper console access to manage machines

# XEN on Linux






- On CPUs with virtualization extension, Windows installable
- Not as comfortable as VirtualBox, VMware on Linux

# VMware ESX





- Unsolved console issues let us switch to VMware ESX run by another team in the computer center
- Number of machines on professional hardware
- Comparably expensive, license fees, hardware, SAN, ...
- One server machine for testing
  - Root filesystem exports for Linux Stateless clients
  - Good performance, near to real hardware, but restricted regarding memory and disk space

# VMware ESX





- Couple of test machines from the old VMware Server
  - Rather easy system conversion within the VMware product portfolio
  - Good performance
  - Disadvantage: Only the Windows based management console is really usable
- Proper failover strategy with multiple servers
  - No system outage yet
  - Seamless extension and maintenance of the system

# VMware ESX





- Advantages of proper backup and snapshotting not used
  - Nevertheless – no significant failures yet
- Generally: Consolidation of hardware, reductions in resource consumption (machines, energy, space)
- Took a while until access privileges and network assignments were properly configured
- Typical cloud user experience
  - Most mature management interface
  - Offering all necessary virtualization advantages

# Rabbit Virtual Machines





- Virtualization a nice concept, but
  - A new virtual machine is just a mouse click away!
  - But also a virtual machine needs to be managed and maintained (think of added complexity for system monitoring)
  - Define strict rules for the provisioning of a new virtual machine
  - Strict cost models protect from uncontrolled requests for new VMs
  - Continuously control the utilization of VMs (workload, period of utilization,….)

# Rabbit Virtual Machines





- An unused VM wastes resources (e.g. memory)
- The lifecycle management of VMs needs new operational processes
- Lifecycle manager software could help to manage and monitor virtual machine parks
- Server virtualization alone is not the magic bullet
- It should be one building block of a holistic infrastructure optimization concept
- It should come along with other initiatives like I/O- and storage virtualization
- A common management of physical and virtual entities could significantly reduce complexity

# Structure: Uses of Virtualization



UNI FREIBURG

Network Experiments

CPU Cycle Harvesting

Digital Preservation Aware Systems

# Networking Experiments





- Virtualization software like VMware Server, ESX or XEN not only allow for sophisticated hardware configuration but offer different types of network connections
  - Network bridges to the physical Ethernet
  - Routed networking and completely virtual networking
  - Multiple virtual (Ethernet) switches configurable
  - New nodes easy to deploy
- Thus possible to produce complex but completely virtual network setups



- Benefits of Virtual Networks
  - Create networks with different characteristics
  - Adapt to service demands
  - Optimized topology
  - Adjustable link properties (e.g. bandwidth)
- Dynamic reconfiguration
  - Within hours without leaving your work desk and fiddling with cables and network hardware
  - Network can adapt to changing business rules

# Networking Experiments



UNI FREIBURG

- Encapsulation: Different networks don't interfere with each other
  - Use different techniques in parallel, e.g. Ipv4/IPv6
  - Allow for smooth transitions
- Add/test new functionality (new IP based protocols) without disturbing legacy network

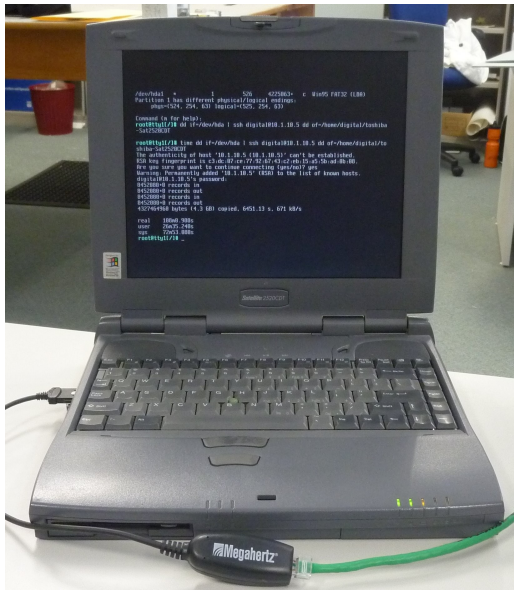
# CPU Cycle Harvesting



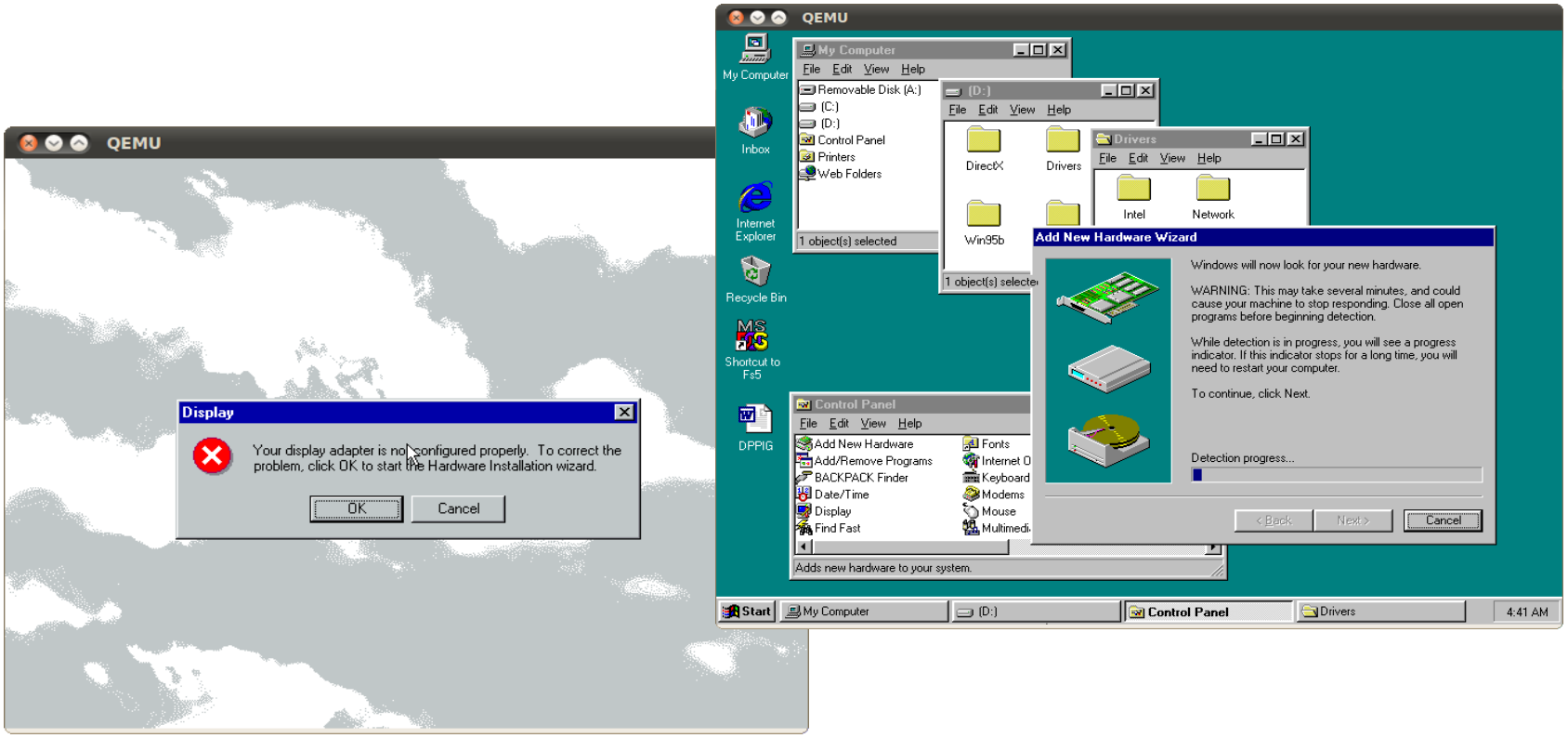


- Idea – more abstract distributed computing than Seti@home
- Original plan: Reboot lecture pool machines into cluster mode
  - Organizational problem – when to do, how to ensure availability of machines for reboot
  - Problem: Long running jobs
- Second approach: Run second virtual machine to "harvest" unused CPU, allow for long running jobs
- Virtualization interesting for cluster computing nevertheless – optimize hardware usage with different job profiles, ...

- Virtualization later on emulation (remember first lecture) could be used to preserve complex digital objects like databases, CMS, …
- Recovering Complete Machines
  - Dumping entire hard disks including partitioning




- … and make them accessible in virtual machines
- Adaptation to new “hardware” environments by installing appropriate drivers



# Built-in Digital Preservation





- Future: Many machines life starts virtual
- More and more complex systems for industry processes, supply chain management, …
- Preserve entire networks of those machines to recover
  - Entity relationships
  - Fullfill legal requirements
- Virtualization tools have to provide stable interfaces
  - Not yet the case, e.g. VMware virtual machine and system images changed significantly over time

# Outlook





- Next lecture, Monday
  - Double lecture starting 2pm in Computer Lab #4
- Starting with the practical part
  - After introduction to system in use in Freiburg
  - Demonstration of a basic network booting Linux system
  - Configuring necessary services, like DHCP, TFTP, looking into PXElinux