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

DAAD Summer School: Aspects of Large Scale High Speed Computing

17th March 2011

Dr. Dirk von Suchodoletz

Faculty of Engineering, University Freiburg

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

Last Lecture / Recap

- Motivation, concept and ideas of Stateless Booting
- Client and server sides in network booting
- Network planning and network boot protocols
- Client side root filesystem, options and challenges for read-write configuration and runtime data
- Overview on system monitoring challenges and solutions

Plan for this Lecture

- Introduction to the topic: Motivation, why virtualize?
- Administrative and economic advantages
- Talk of history and main ideas
- Distinguish full, hardware assisted, para virtualization and tools using it, partitioning, emulation
- Practical application: Running Windows without pain
- Flexible Lecture Pools, server consolidation
- Experiences and other usage scenarios
- Digital preservation

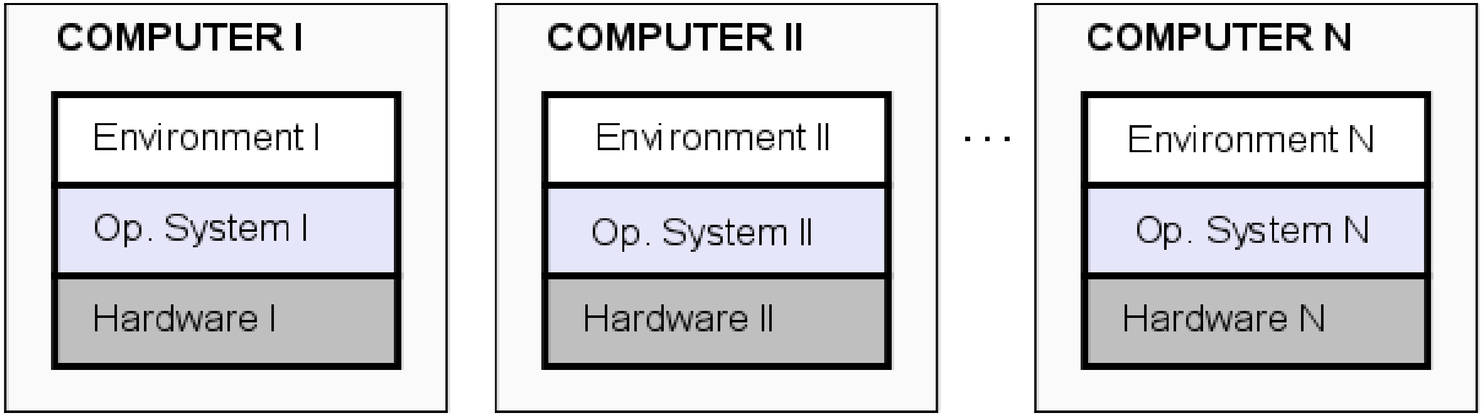
Structure: Motivation

Why Virtualize?

Advantages of Cloud Computing

Motivation to Virtualize

- Economic and administrative driving forces
  - Consolidation of server installations
    - Many companies and organizations want to reduce the number of servers installed, less different hardware to manage
  - Traditional paradigm: One service one server



Motivation to Virtualize

- Economic and administrative driving forces
  - Avoiding conflicts (software versions, conflicting applications)
  - Energy and resource savings
    - Most server hardware is under-utilized and typically runs on low load (DHCP, DNS, MS AD, ...)
    - Servers capacity calculated for peak usage
    - Installed number of servers much greater than compute power needed
    - Concerns with rack + floor space and energy consumption (systems plus cooling)

Motivation to Virtualize

- Economic and administrative driving forces (cont.)
  - Protection
    - Commercial virtualization often coupled with (low-cost) SAN solutions offering better reliability than average server disk installations
    - Cost and complexity of implementing disaster recovery is reduced
    - Image deployment and creation of snapshots on the host is easier than running backup software within each machine
    - Protection against malware

Motivation to Virtualize

- Economic and administrative driving forces (cont.)
  - Deployment
    - Most small and midsize organization and businesses have limited administrative resources
    - Virtualization for them provides less effort and greater speed
  - Stateful cloning/live migration
  - Protecting long running jobs from system changes
    - Snapshot or suspend them for planned maintenance
    - Park long running jobs suspended, when urgent job is to be run on the cluster

Motivation to Virtualize

Albert-Ludwigs-Universität Freiburg

- Economic and administrative driving forces (cont.)
  - Reproducing situations
    - Start several times exactly from same system status
    - Freeze systems states to be inspected by others
    - Malware research, ...

Motivation to Virtualize

- Economic and administrative driving forces (cont.)
  - Agility
    - Virtualization and cloud computing helps organizations and businesses adapt server resources to address changes in workload demands
    - Makes it easier to bring up new services
    - Changing capacity requirements (different demand in peak/off-peak business)
    - Changing capacity offering/availability
  - Metering of job/machine resource consumption
  - Resource consumption enforcement

Motivation to Move into Clouds

- Economic and administrative driving forces (cont.)
  - Freedom of choice
    - Virtualization allows organizations and enterprises purchases based on competitive pricing without worrying about the overhead of supporting multiple hardware vendors
  - Planned maintenance – easier to replace hardware parts if machine is not directly attached to it
  - Emulating an environment for legacy software

Motivation to Move into Clouds

- Virtualization is prerequisite to cloud operation, as it allows/provides
  - Live migration
  - Dynamic resizing
  - Taking snapshots
  - Isolation
  - Provisioning
- Clouds seen as a chance to enable solutions that otherwise would be difficult and expensive
- IaaS, SaaS, infrastructure/software as a service, ...

Structure: X86 Virtualization

History of Virtualization

Full, Hardware Assisted, Para Virt.

Partitioning, Emulation

Examples

History of Resource Virtualization

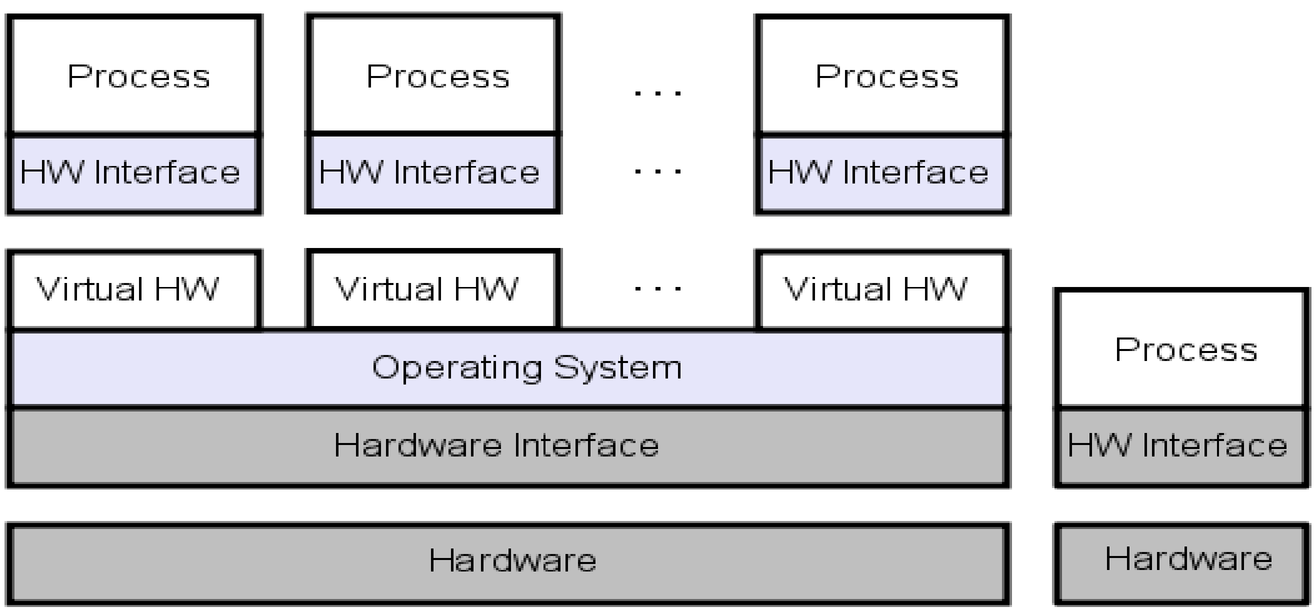
- Idea not really new concept but just repetition of / driven by the old dream: Better utilization of hardware
- Concepts already known
  - Introduction of operating systems helped with resource sharing for applications
  - Introduction of schedulers to manage limited CPU resources and run processes in "parallel"
  - Virtual address space to extend expensive physically available RAM

Operating System Resource Sharing

Albert-Ludwigs-Universität Freiburg

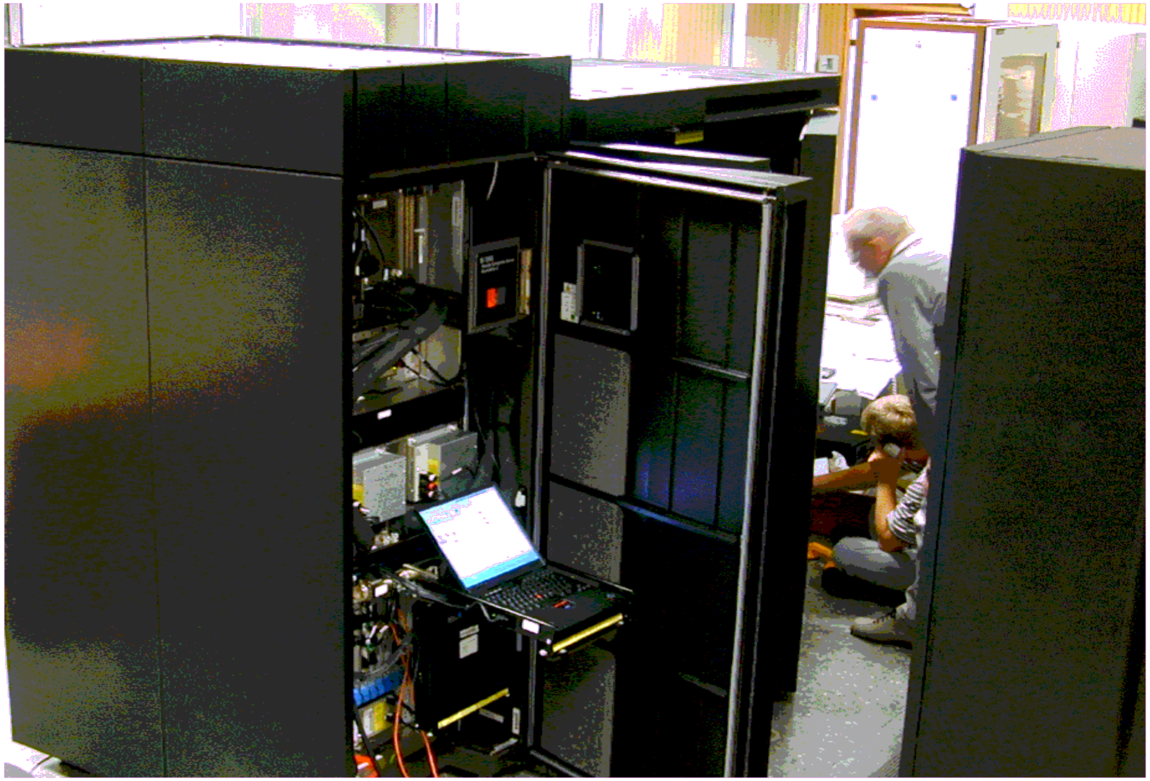
UNI FREIBURG

- One process does not reserves the complete hardware
  - Processes get virtual hardware assigned and get isolated from each other
  - CPU might run other process when first one is waiting for IO (e.g. on periphery)



History of Virtualization

- Idea pretty old – used by IBM zSeries main-frames since 30 years
- Logical Partitioning (LPAR)
- Complete separation of I/O from processing
- Dynamic sharing of network and I/O adapters



Old Idea for New Architecture

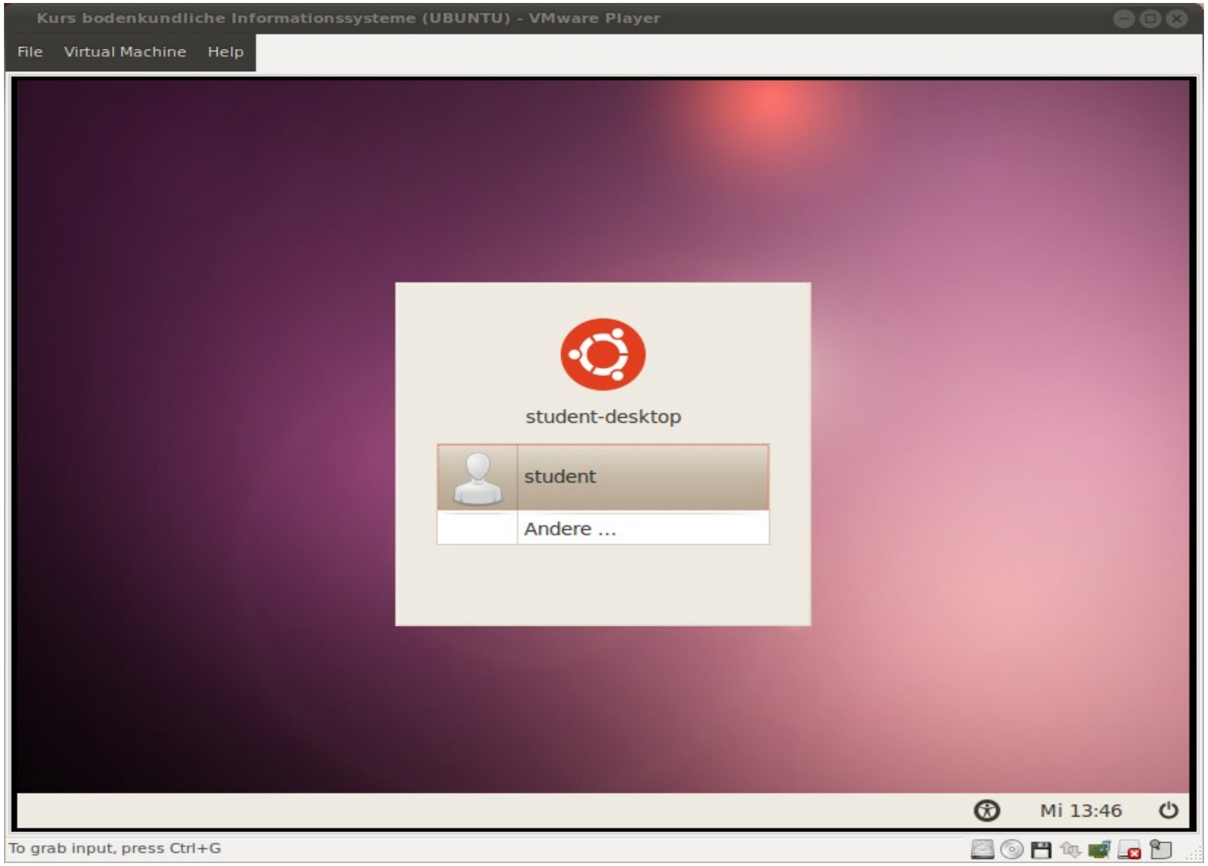
- Introduced stuff like Parallel Sysplex : Virtualization across servers via Coupling Facility
- DB2: Parallel database with shared data access and unlimited scalability
- z/VM: 2nd level of system virtualization
  - Virtualization of processor, memory, network, I/O, hardware emulation, …
- Founders of VMware members of the z/VM group from IBM
  - In theory X86 was not virtualizable
  - VMware Workstation (end of 1990ies) was the first X86 virtual machine

Virtualization Concepts

- Virtualization can be realized on different layers: "Real" virtualization or para virtualization on layer 1
  - Partitioning of the real hardware into several virtual hardware sets
  - Virtual hardware is pretty much of the same type the real hardware
  - Several (unmodified) operating systems run on the same physical hardware without interfering each other
  - Operating systems have not to be of the same type but need to be executable on the real hardware

Full Virtualization

- Using binary translation, e.g. VMware product suite (VMware Player Linux on Linux)

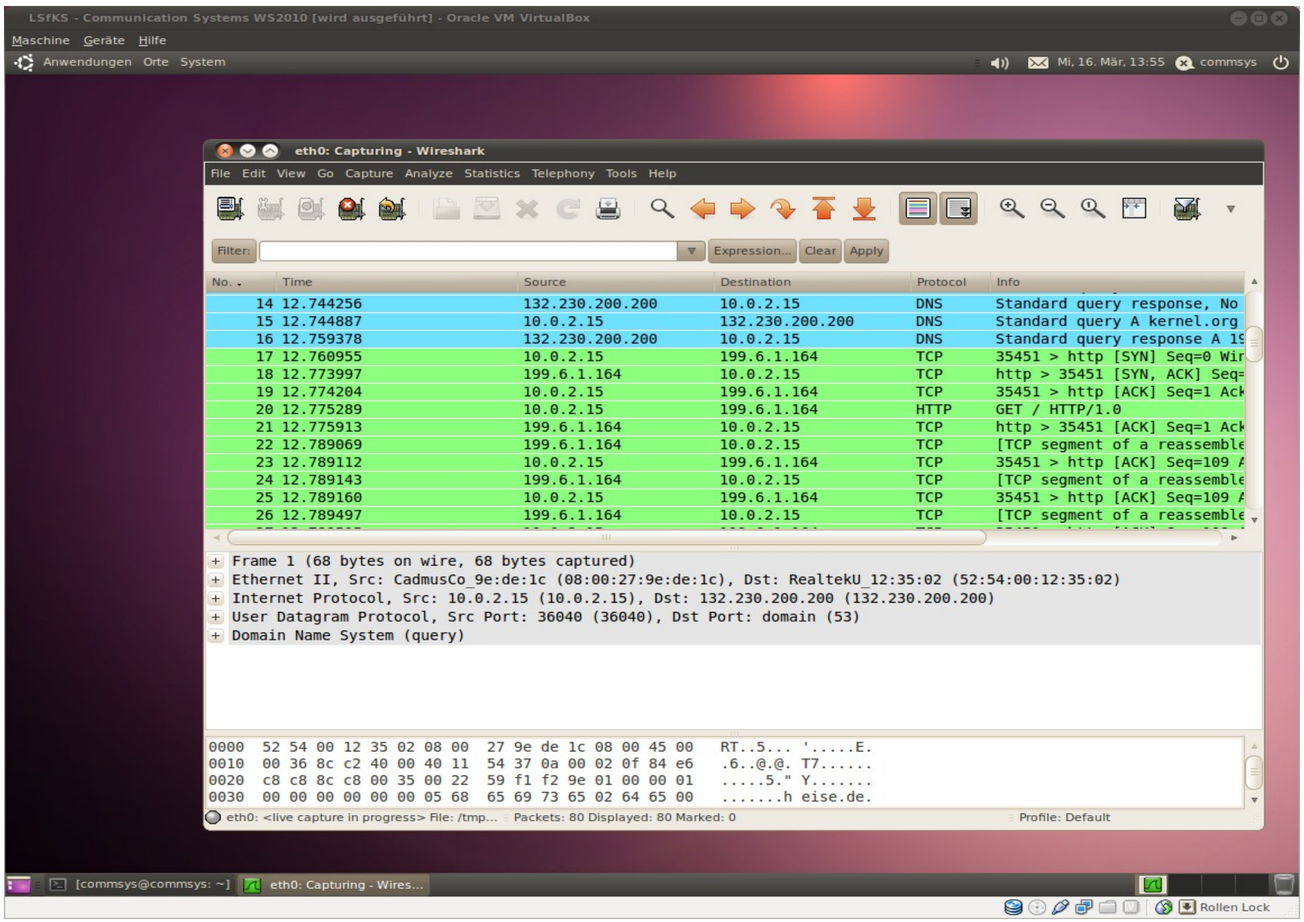


Full Virtualization

- Advantage: Every standard X86 CPU could be used
- Disadvantage: Privileged operations have to be captured and to be translated
  - More complex
  - Instructions are analyzed / translated on the fly
- Thus more popular and used in other commercial products like Microsoft Virtual PC, Parallels, ...

Full Virtualization

- Popular commercial, Open Source solution is VirtualBox (Innotek, Sun Microsystems, Oracle)



Full Virtualization

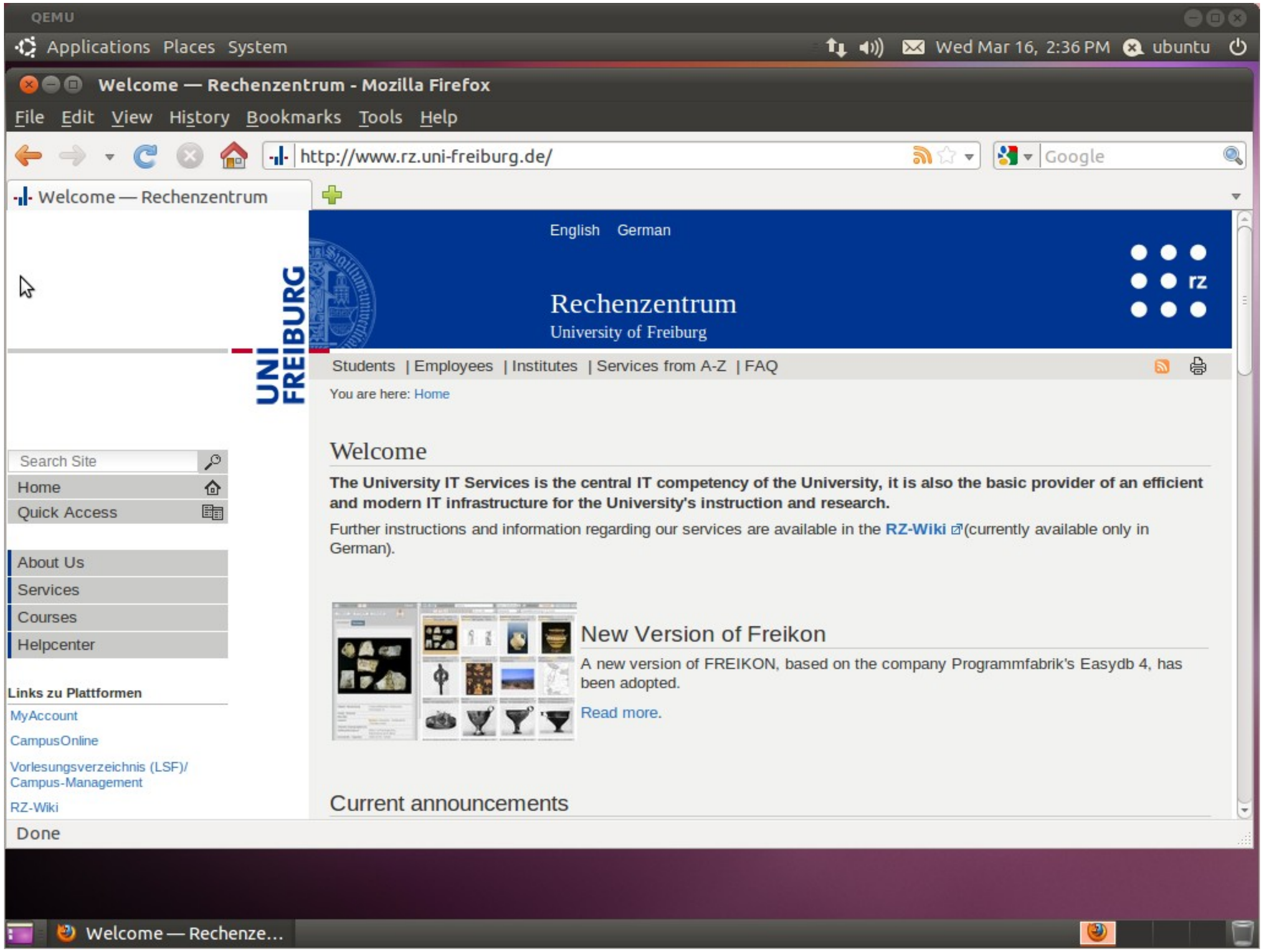
- Hardware-assisted virtualization
  - Newer than binary translation
  - Requires hardware support, CPU has to provide virtualization extension (just check it on your linux machine via the /proc/cpuinfo interface)
  - Intel: vmx, AMD: svm
  - Still not standard in value CPUs
- QEMU/KVM popular choice on Linux, split into
  - Kernel part (KVM, different modules/abstractions for Intel/AMD)
  - Peripheral/user mode part: Modified QEMU

Full Virtualization

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

- QEMU/KVM (Linux on Linux)

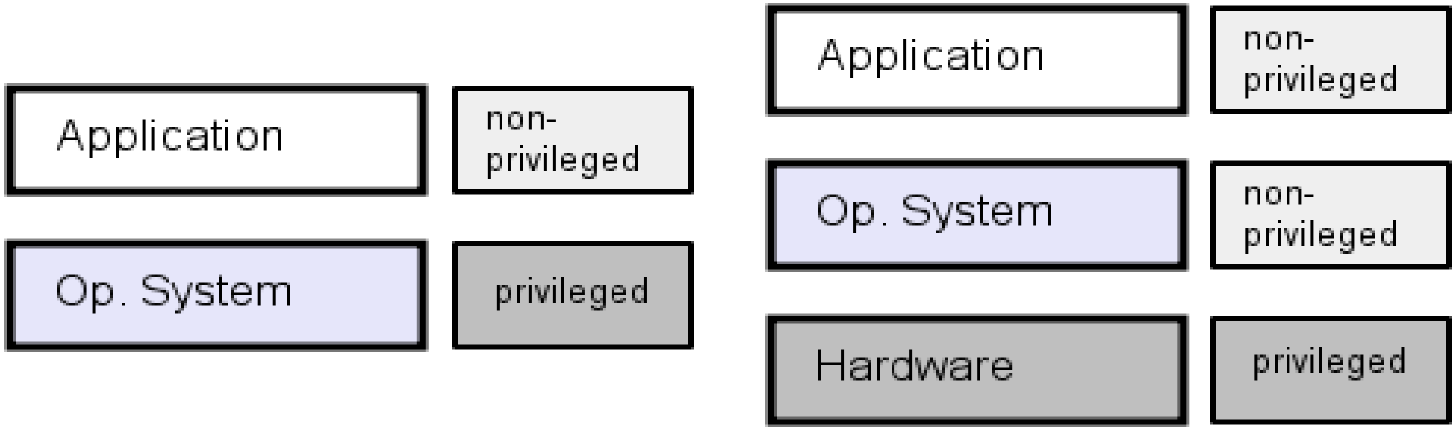


Para Virtualization

- Paravirtualization supports high performance
- Using modified operating systems kernel
  - Guest OS is aware of the virtualization
  - Guest OS performs hypercalls instead of system calls
- XEN – the standard example around for a while
  - Allows exclusive hardware allocation to virtual machines, e.g. network interfaces

Real vs. Para Virtualization

- Kernel is judge on resource assignment – kernel mode and user mode
- Privileged / non-privileged access



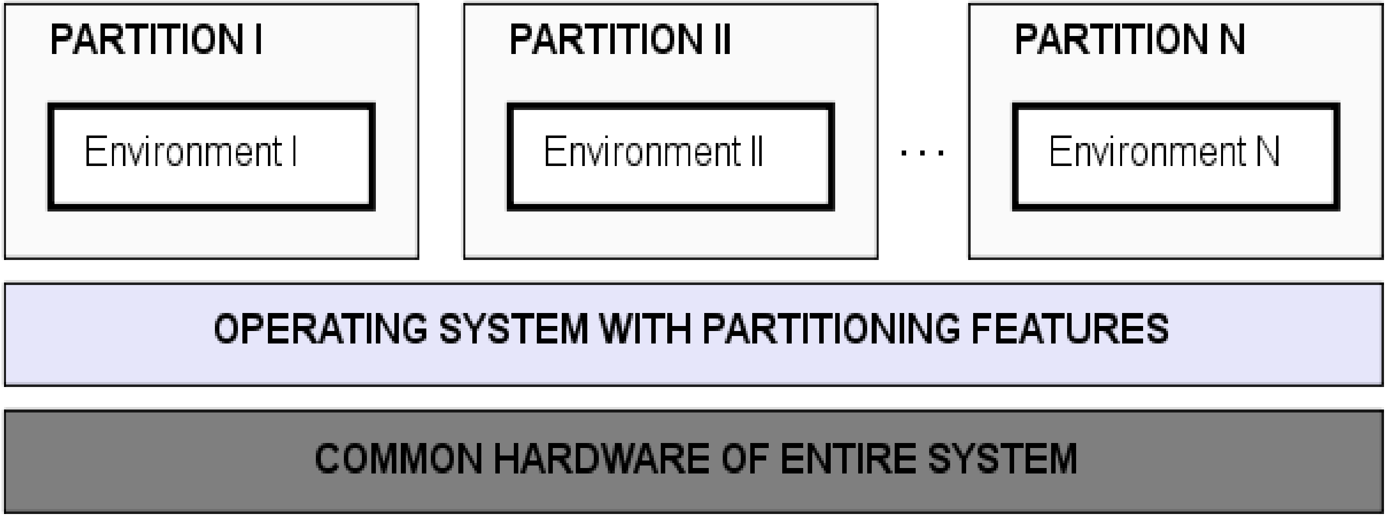
Partitioning

- Virtualization could be realized on a higher layer – instead of running completely different operating systems, do partitioning within running OS
  - Different partitions run on the same kernel, no different operating systems possible
  - Creation of different environments with its own users, processes and filesystems
  - Kernel has to isolate partitions from each other
  - Less overhead than full virtualization

Partitioning

Albert-Ludwigs-Universität Freiburg

- Examples: Linux Virtuozzo, Linux OpenVZ



Break

- Short break, then continue with
  - ...
  - Desktop virtualization
  - Other usage scenarios
    - Networking experiments
    - CPU “cycle stealing”
  - Digital preservation scenarios