PARA-VIRTUALIZATION WITH XEN

Ivan Klimek
Ivan.Klimek@cnl.tuke.sk

INTRO

- What is Virtualization ?
- Virtualization approaches
- Xen / para-virtualization
- General usage scenarios
- Advancced usage scenarios
- XEN vs Vserver
- XEN performance isolation

What it is all about ?

Virtualization is a framework or methodology of dividing the resources of a computer into multiple execution environments.
Done by applying one or more concepts or technologies such as:
- hardware and software partitioning,
- time-sharing,
- partial or complete machine simulation,
- emulation,
and many others.

“Virtualization is an abstraction layer that decouples the physical hardware from the operating system to deliver greater IT resource utilization and flexibility.” – www.vmware.com

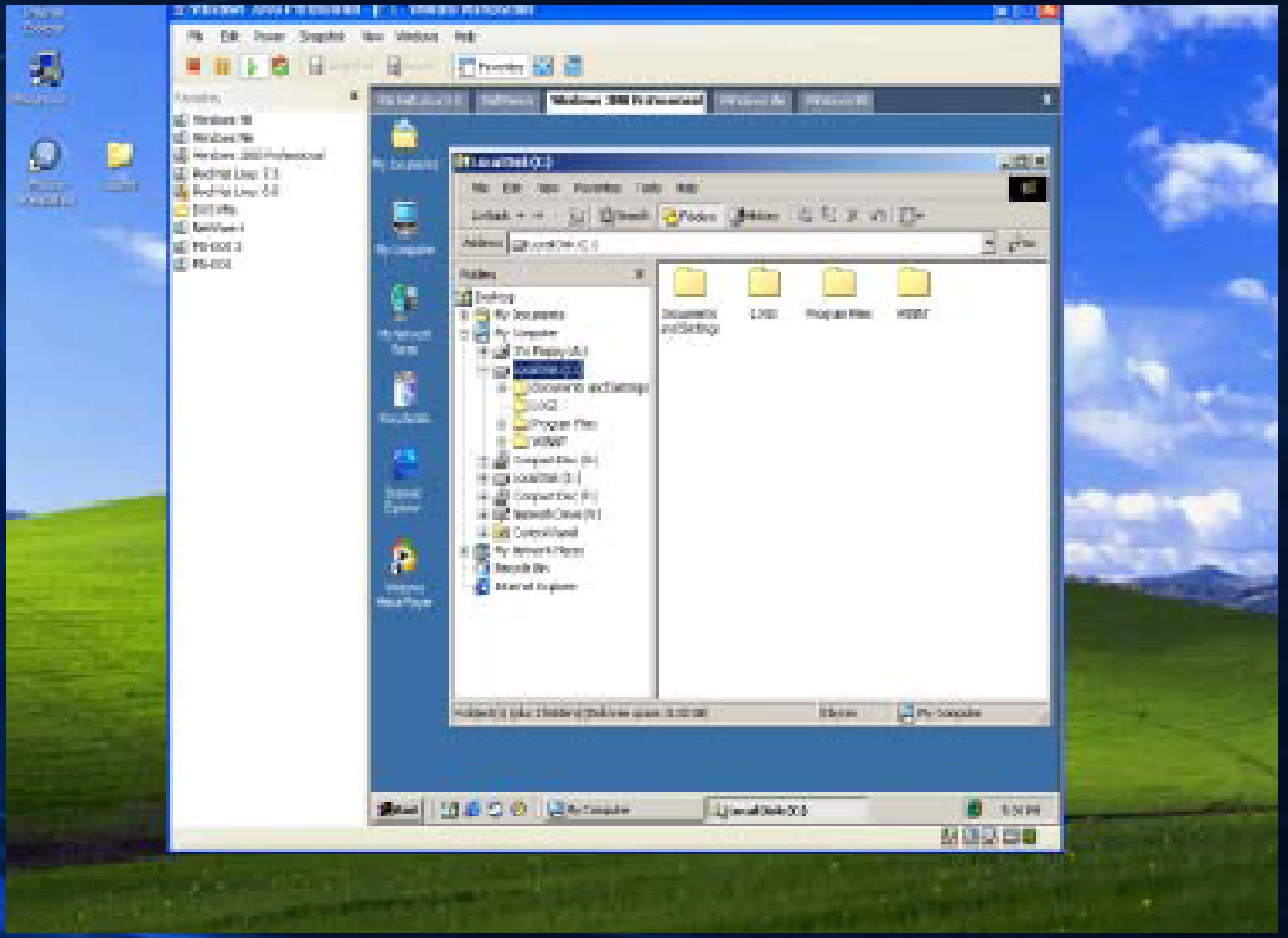
Virtualization approaches

Single OS image: Virtuozo, Vservers, Zones
Group user processes into resource containers
But it's hard to retrofit isolation to conventional OSes

Full virtualization: VMware, VirtualPC, QEMU
Run multiple unmodified guest OSes
Hard to efficiently virtualize x86

Para-virtualization: Xen
Run modified or new OSes on a specialized guest architecture

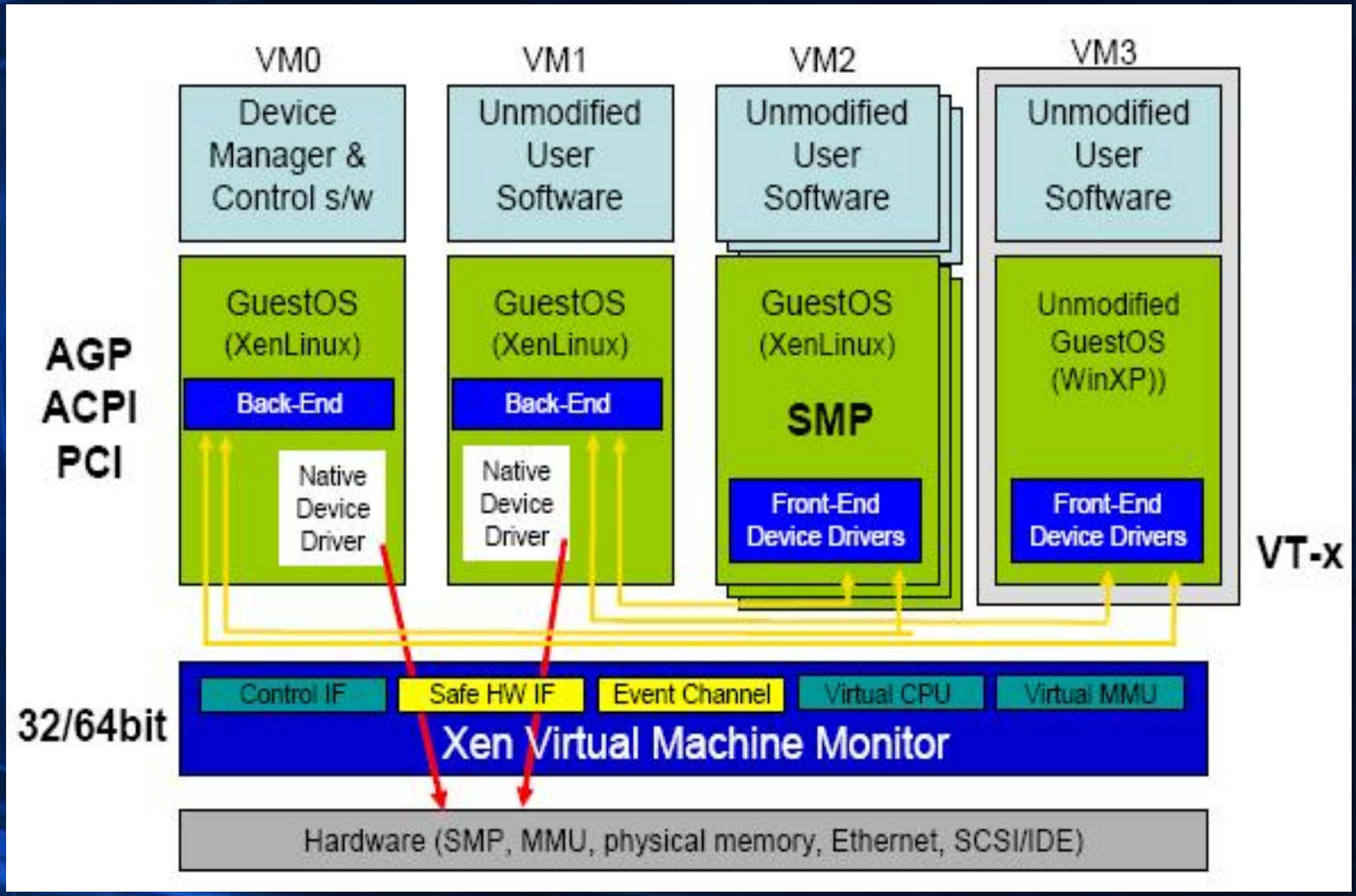
Full Virtualization



Xen & Paravirtualization

Support full-featured multi-user multi-application OSes
– Contrast with Denali approach: thin OSes for lightweight services
OSes are ported to a new 'x86-xeno' architecture
– Similar to x86, but call to Xen for privileged operations
– Porting requires access to source code
The need to modify kernel to understand virtualised environment

Xen architecture

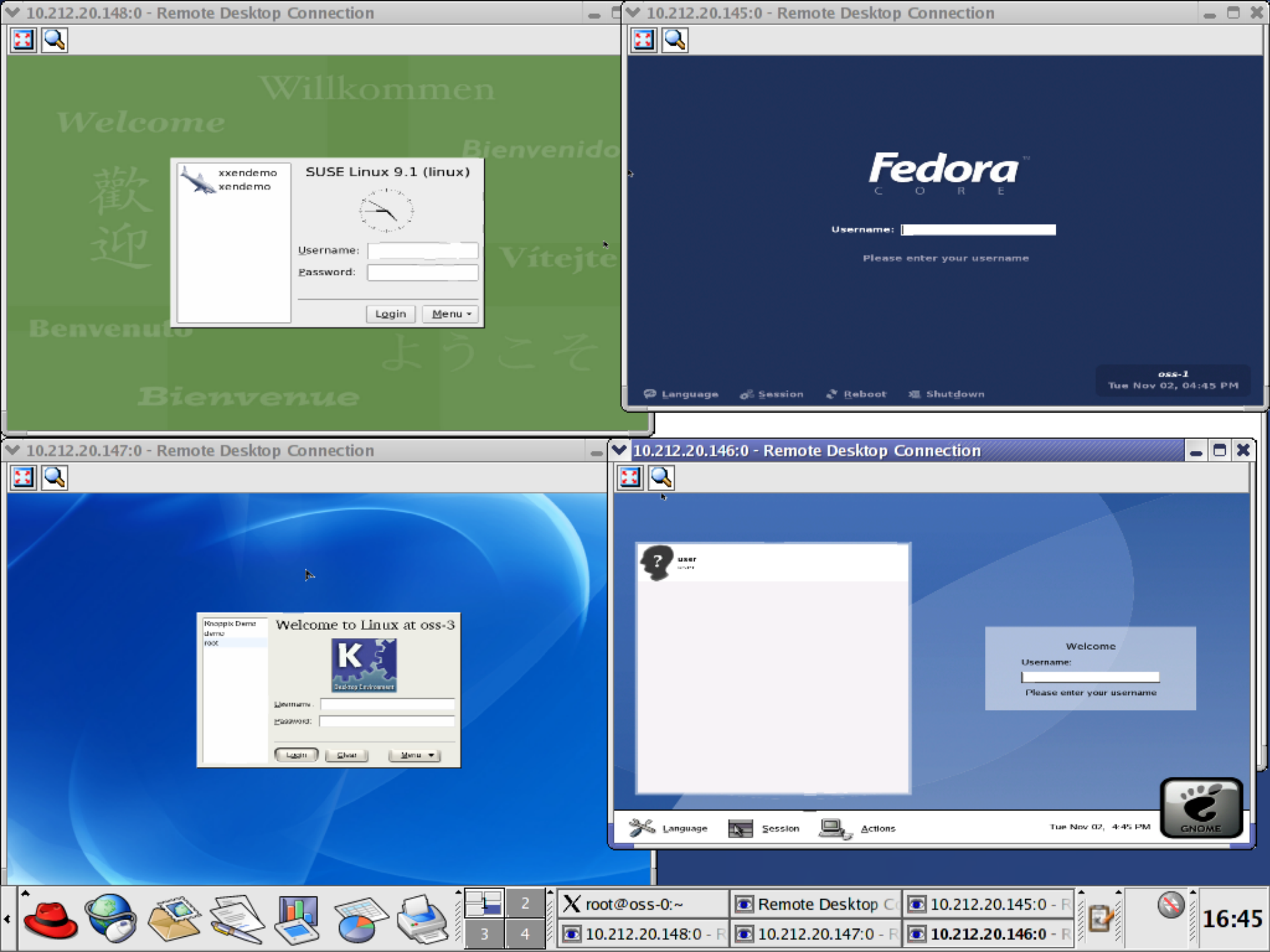


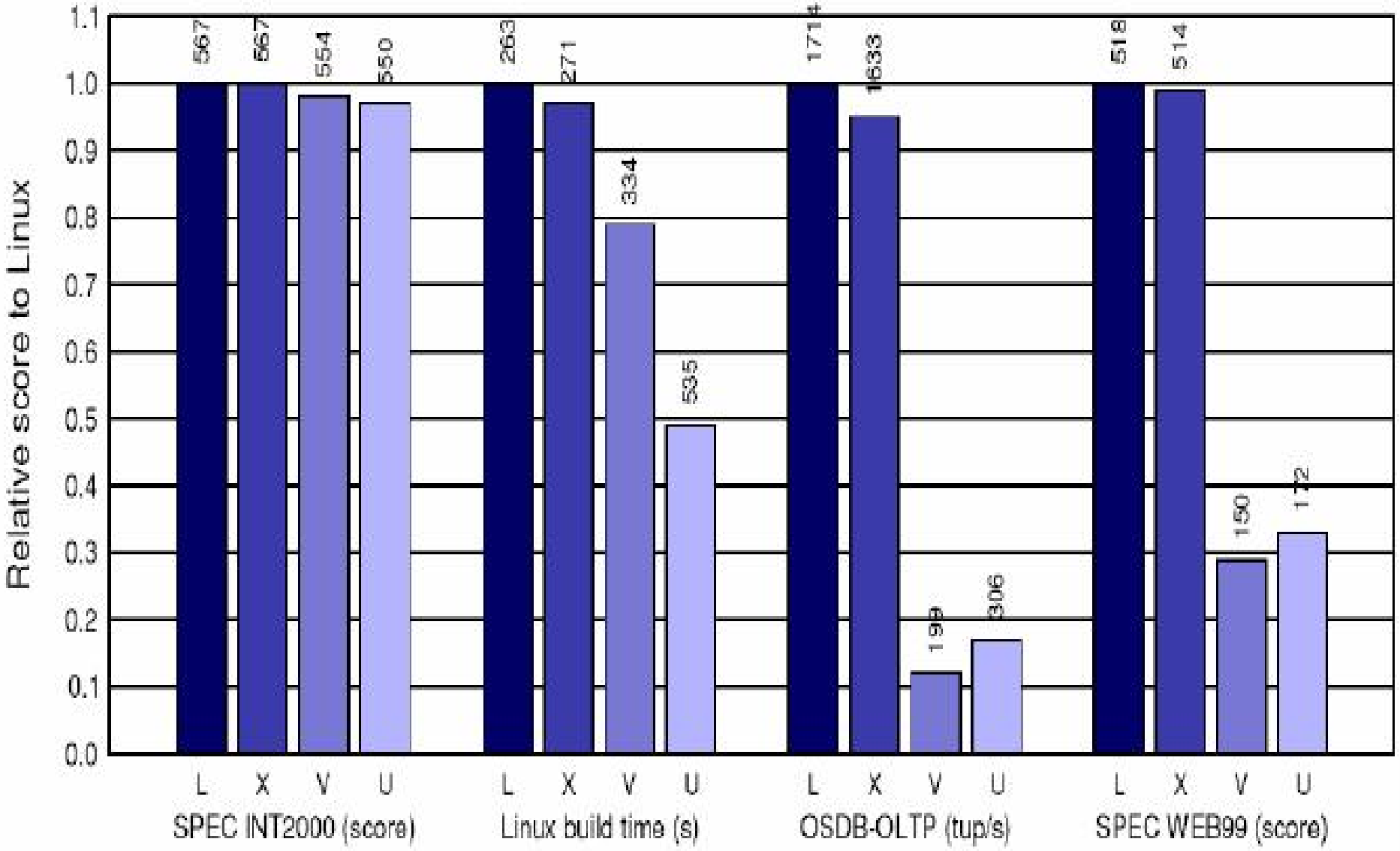
Xen architecture

- Xen runs one dom0 with privileged kernel which is the host system
- Under that are the domU kernels - the virtual machines
- Complete separation - each virtual machine can run its own kernel version (compiled with xen + virtual device drivers)
- Memory, CPU time, HDD, network addresses can be assigned per machine.
- Virtual machine migration

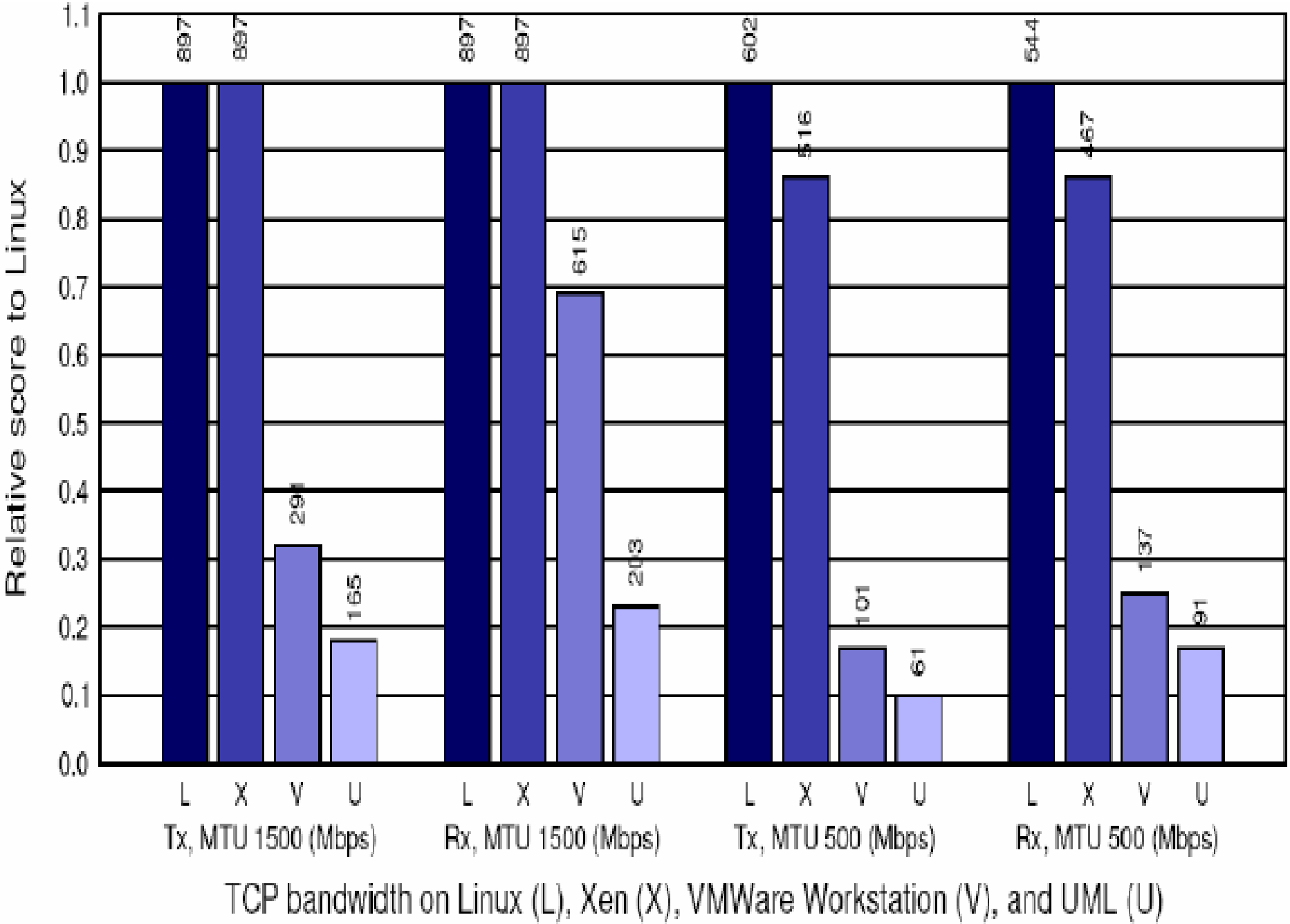
XEN features

- Minimal overhead – 2~5%
- Near native performance
- No downtime
- Systems are files, can be easily backed up and restored
- Live relocation / migration in ms
- Domains are completly independent
- PERFORMANCE ISOLATION (CPU time, HDD, RAM, network)
- Virtual networking / clustering possibilities
- Open-source

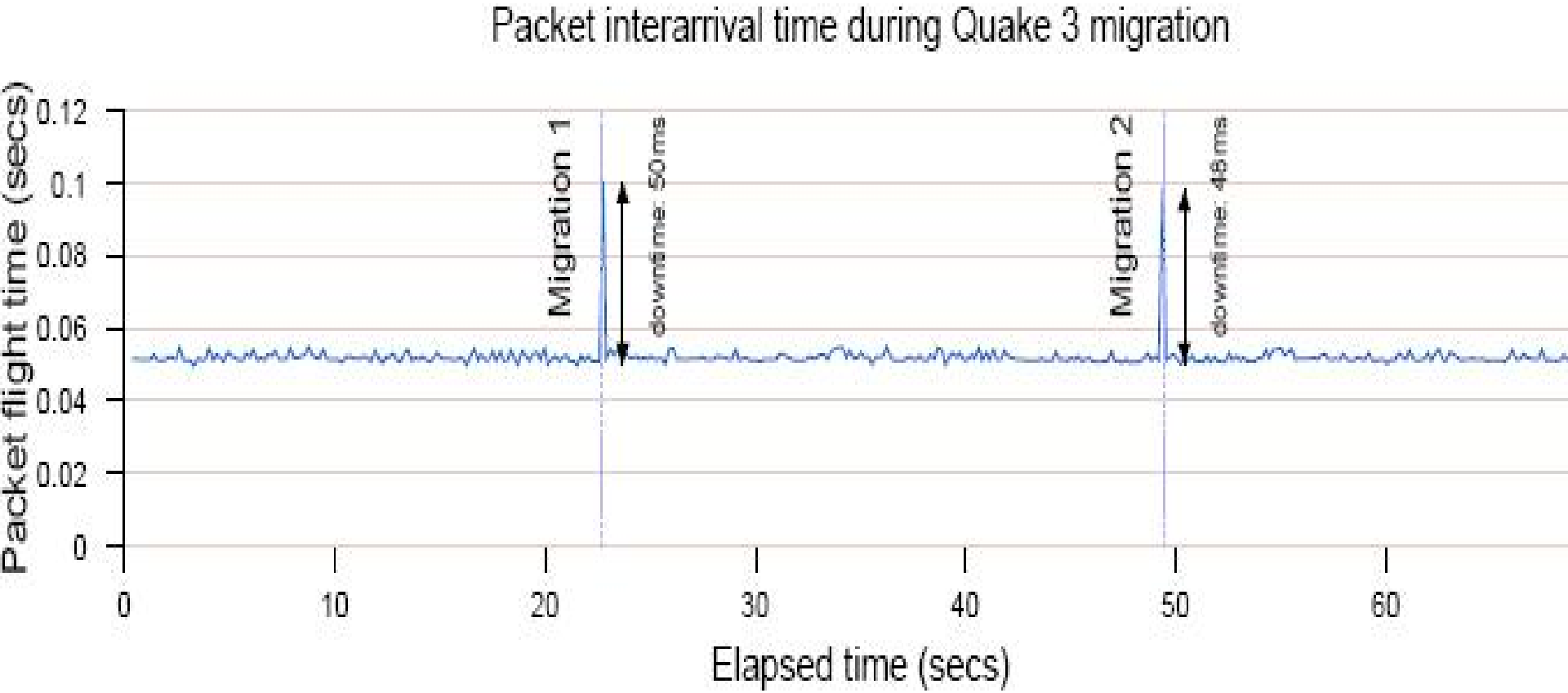

10.212.20.148:0 - Remote Desktop Connection
Willkommen
Welcome
Bienvenido
歡迎
Vítejte
Benvenuto
ようこそ
Bienvenue
xxendemo
xendemo
SUSE Linux 9.1 (linux)
Username:
Password:
Login
Menu
10.212.20.145:0 - Remote Desktop Connection
Fedora
CORE
Username:
Please enter your username
Language
Session
Reboot
Shutdown
oss-1
Tue Nov 02, 04:45 PM
10.212.20.147:0 - Remote Desktop Connection
Welcome to Linux at oss-3
10.212.20.146:0 - Remote Desktop Connection
user
Welcome
Username:
Please enter your username
Language
Session
Actions
Tue Nov 02, 4:45 PM
GNOME
root@oss-0:~
Remote Desktop C
10.212.20.145:0 - R
10.212.20.148:0 - R
10.212.20.147:0 - R
10.212.20.146:0 - R
16:45



Benchmark suite running on Linux (L), Xen (X), VMware Workstation (V), and UML (U)


Relative score to Linux
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
897
897
897
897
602
544
516
467
615
291
203
165
137
101
91
61
L X V U
Tx, MTU 1500 (Mbps)
L X V U
Rx, MTU 1500 (Mbps)
L X V U
Tx, MTU 500 (Mbps)
L X V U
Rx, MTU 500 (Mbps)
TCP bandwidth on Linux (L), Xen (X), VMWare Workstation (V), and UML (U)

Quake 3 Server relocation



General usage scenarios

- Server Consolidation.
Move multiple servers onto a single physical host with performance and fault isolation provided at the virtual machine boundaries.
- Hardware Independence.
Allow legacy applications and operating systems to exploit new hardware.
- Multiple OS configurations.
Run multiple operating systems simultaneously, for development or testing purposes.
- Kernel Development.
Test and debug kernel modifications in a sand-boxed virtual machine, no need for a separate test machine.

General usage scenarios

- Cluster Computing

Management at VM granularity provides more flexibility than separately managing each physical host, but better control and isolation than
single-system image solutions, particularly by using live migration for load balancing.

- Hardware support for custom OSes

Allow development of new OSes while benefiting from the wide-ranging hardware support of existing OSes such as Linux.

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- Each domain network interface is connected to a virtual network interface in dom0, by a point to point link (effectively a .virtual crossover cable.).
- Trafic on these virtual interfaces is handled in domain 0 using standard Linux mechanisms, for bridging, routing, rate limiting, etc. Xend calls on two shell scripts to perform initial configuration of the network and configuration of new virtual interfaces.
- By default, these scripts configure a single bridge for all the virtual interfaces. Arbitrary routing / bridging configurations can be configured by customizing the network suport scripts.

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- **XEN WORLDS** is being developed to provide a method for performing assignments and lab work in information assurance, operating systems and networking courses that require root access to the individual machines, or the entire network.
- The project is aimed at creating a versatile "**virtual lab**" where an entire **network of virtual machines**, (a Xen World), can be provided to each student that will allow on-campus and distance education students 24/7 access via SSH, allow students to turn-in a virtual machine or an entire network as the finished product and allow for grading to occur directly on those machines instead of grading a few select artifacts such as configuration files, programs or outputs.

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

Xen Worlds cluster

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- Using XEN it is possible to create a whole virtual network which could include a **firewall, honeypod, IDS** .... all in one PC.
- All of this virtual machines consuming together only a **few percent** of the available CPU time.
- All managed from a **single point**.
- **Automatic** honeypod – IDS - firewall actions.
- Scripts can automatically check the status of the network, pinging from Dom0, if some errors ocurred, close network connectivity, and restart the affected virtual machine.
- **SO MANY POSSIBILITIES** .....
- This is the **future** of protecting stand-alone PCs
- We are working on it allready ....

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- The **XENOPPIX** project is a customized KNOPPIX with XEN included.

- very interesting project, but for my mostly becouse the KNOPPIX **terminal server integration.**

- **Possibility to network boot directly to XEN**

- they are using PXE network boot, becouse XEN will only boot from the GRUB bootloader, and PXE-GRUB is the only version of GRUB supporting PXE.

- but PXE-GRUB supports only one type of network card the eepro100, so this is the WRONG WAY !!!

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- **I managed to boot XEN directly without PXE-GRUB, with a few tricks it is possible to use PXE-LINUX ....**

- not every NIC supports PXE, but there are some other methods too ....

- But why to boot directly to XEN ???

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- Global research projects using volunteers to participate in public grids. For exmp. SETI or looking for a cure for cancer

- Insecure technology, cant be used in goverment sector

- how many PCs does the govermet have ? How many PCs are in all of the schools, generaly in all goverment paid organizations ? Really many.

- Why dont use them ? We have the right technology allready ....

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- Thru booting a whole segment of PCs directly to XEN dom0, which could start another domain with restricted acces to HDD …
- the central point of this topology the machine running termminal server, acting as the gateway for domUs, not for dom0s, ACL blocking traffic to and from dom0s. Couse dom0s have acces to HDDs …
- the central point can be booted from a CD or via internet, (xenoboot)
- XEN virtual network security features can be implemented too

- this way it is possible to build a secure grid of whole LANs, with maximum security ....

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- using clustering and XEN, it is possible to create clusters with great granularity, centrally managed, with central points for VM generation, VM repositorys, services without downtime, with fast deployment ....

- the **DREAM** system

- I am studying this .... my own project called U.S.A. (uber-Server Architecture) :)

- similar project called **VIRTUAL WORKSPACES**

ADVANCED USAGE SCENARIOS

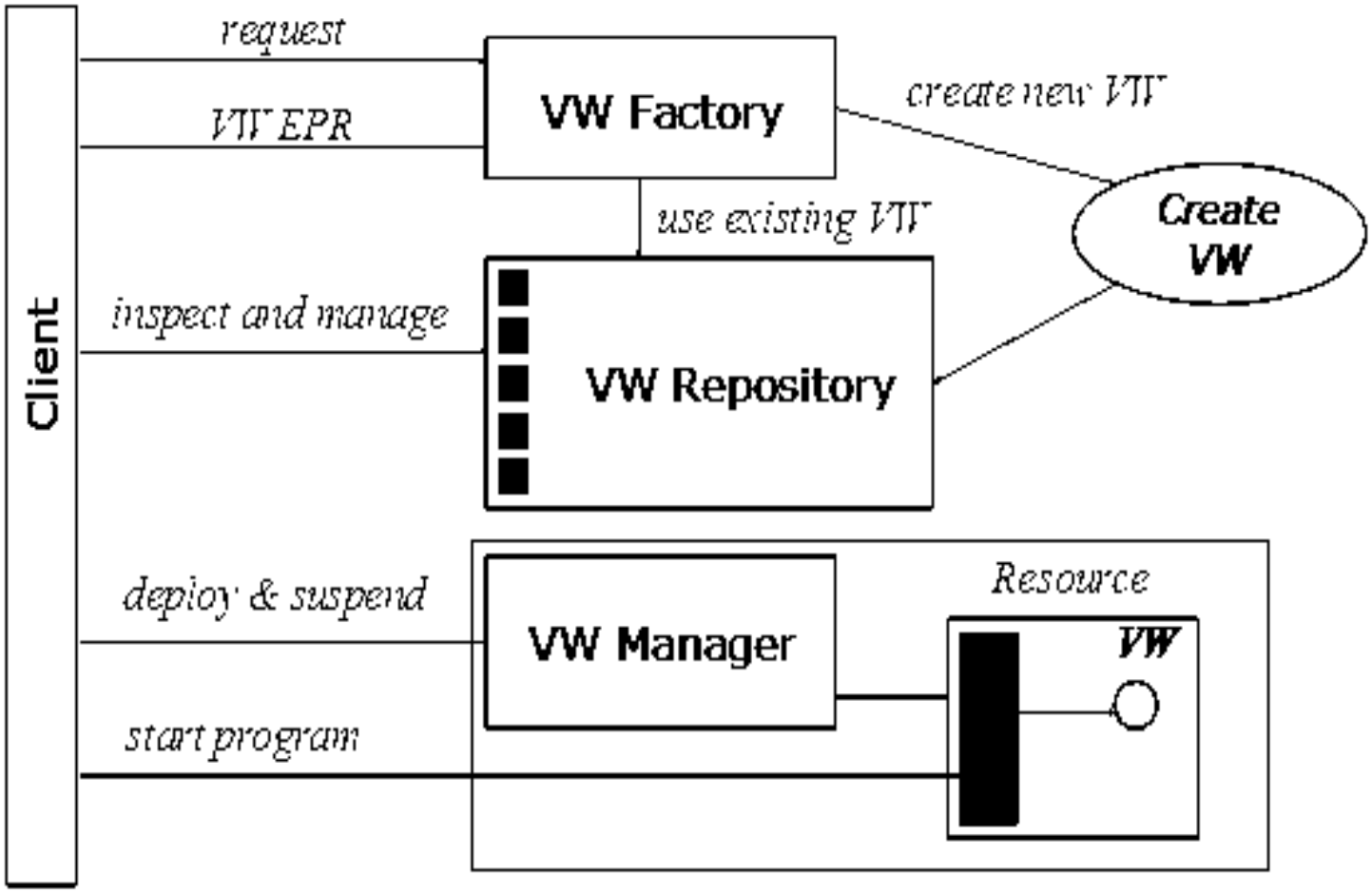
VIRTUAL NETWORKS

What are Virtual Workspaces?

Basic workspace: a Unix account on a remote machine
Software configuration requirements
submit node for a Grid3 cluster
Resource allocation requirements
Use exactly X memory, at least X disk space, Z bandwidth…
Sharing and isolation properties
Unix account, sandbox, various kinds of virtual machines …
And others…
Workspace can be managed and refined
Manage lifetime
A workspace can be deployed on a resource
A workspace can have various implementations

ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS



ADVANCED USAGE SCENARIOS

VIRTUAL NETWORKS

- what if all the home PCs could be replaced by thin clients ?
- separate virtual machines for any user, connection through vnc, streaming video quality, no need for buying new HW ....
- masive server farms would be needed,
- very fast connection links needed
- PC on demmand,
- similar to the process which are all the big companies going through, similar to Citrix, Application on demmand ....
- only a question of time when someone will start with this

Xen vs Vserver

Vserver :

- Improved chroot + kernel modifications, not complete separation between servers
- All machines run the same kernel. Contexts are added to the kernel provide process, disk and memory separation
- All virtual servers use the same network device with different aliases – no virtual nics
- Identical files/trees can be shared between virtual servers (immutable)

Xen vs Vserver

Vserver

+ Minimal performance penalty
+ Virtual machines can share common files -> reduced space requirements

- Not complete separation
- Restricted to one kernel

Xen

+ Complete separation
+ Can mix kernels
+ Virtual machine migration

- VM management on cluster

XEN PERFORMANCE ISOLATION

- managing CPU time using schedulers, Xen includes kernel boot time options for scheduling. Similiar to traditional Linux schedulers that divide CPU time for userland processes, XEN schedules resources between Vms.
(BVT, sEDF scheduler)

- Disk Scheduling with Quality of Service Guarantees
(YFQ)

- you can assign as much RAM as you want to any of the virtual machines, that same applies for disk space

XEN IS THE FUTURE

Thank you !

Ivan.Klimek@cnl.tuke.sk