ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

550 Jahre
Albert-Ludwigs-
Universität Freiburg
1457 – 2007

Workshop: Virtualized IT infrastructures and their management

23./24. October 2008 Garching/München

Broadening the Battle-zone

Some thoughts on the extension of grid/cluster computing into the desktop and pool PC domain

Michael Janczyk
Dirk von Suchodoletz
Chair of Communication Systems
Prof. Gerhard Schneider

Computer Center of the University of Freiburg

Broadening the Battle-zone

- Structure of this presentation
  - Short overview on the chair for communication systems and the affiliation with the computer center
  - Foundation for new ideas and approaches
    - I. Motivation of this presentation
    - II. Current status
    - III. Experiments Matrix
    - IV. Virtualization for clusters
    - V. Preliminary results
    - VI. Conclusion
  - Discussion

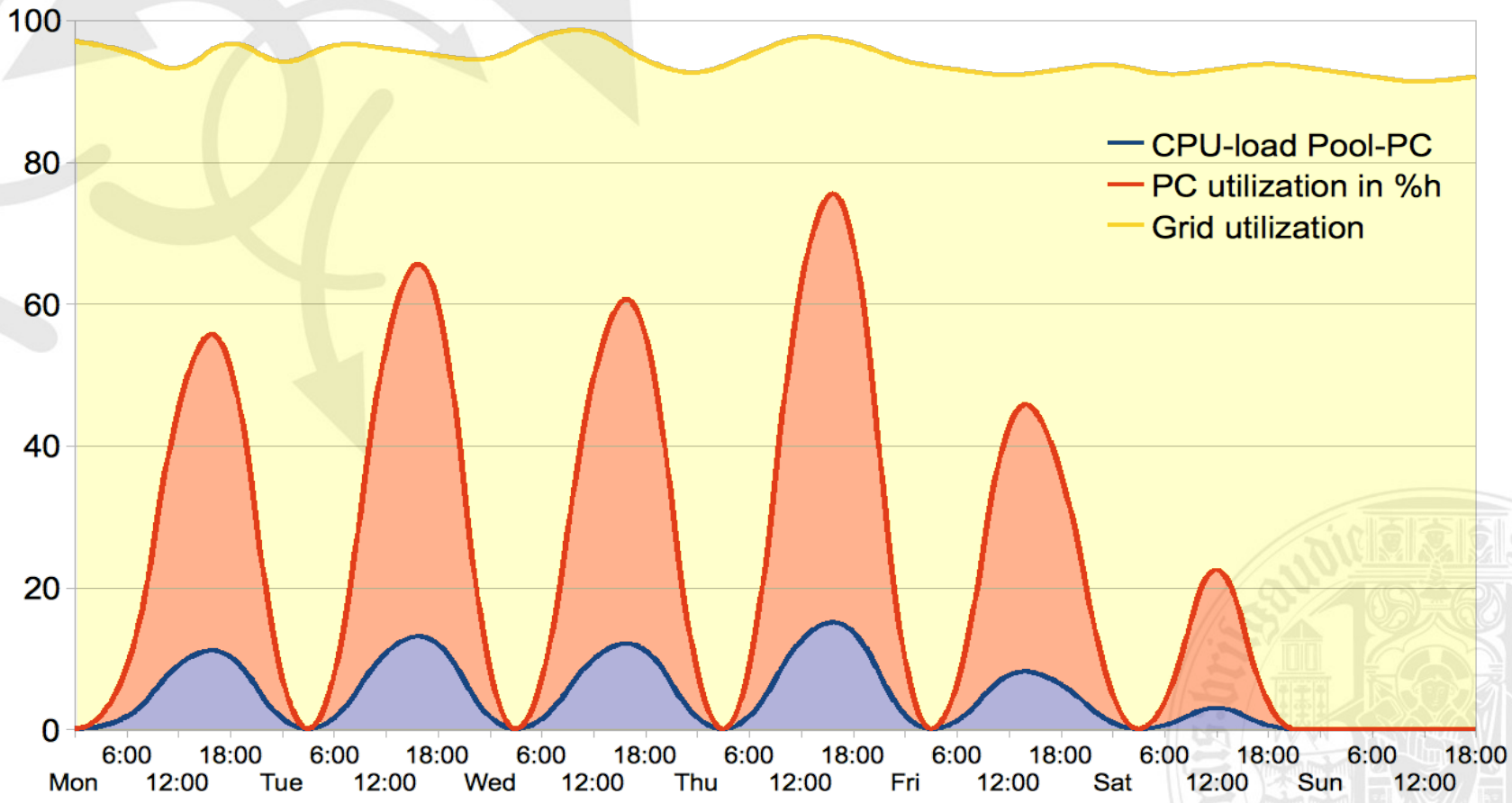
Chair of Communication Systems

- Prof. Gerhard Schneider is head of Computer Center too, thus the connection
- Focus on practical issues of computer operation
- Research in the fields of long-term preservation, location based services/location awareness, identity and site management, …
- Research group: Dirk von Suchodoletz, Michael Janczyk

Service Unit Computer Center

- Computer center manages IT infrastructure
  - Runs a number of Grid/Cluster systems
  - Offers different types of computer pools
  - Supervises some more pools in different faculties
  - Administrates the network infrastructure, hands out IPs for the LAN, DNS, DHCP
- Thus interested in
  - Solutions in effective resource utilization
  - Discussion of new approaches for consulting, resource planning and purchasing

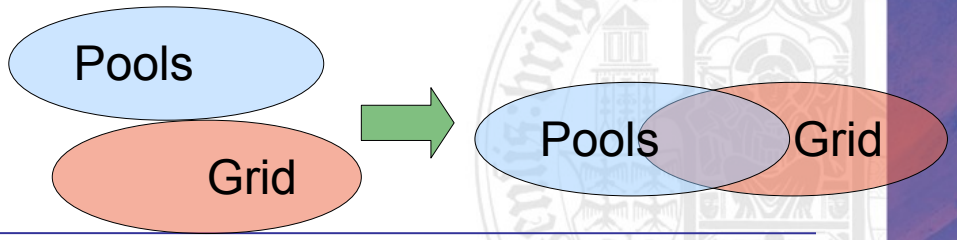
Utilization of Ordinary Pool PCs



I. Motivation

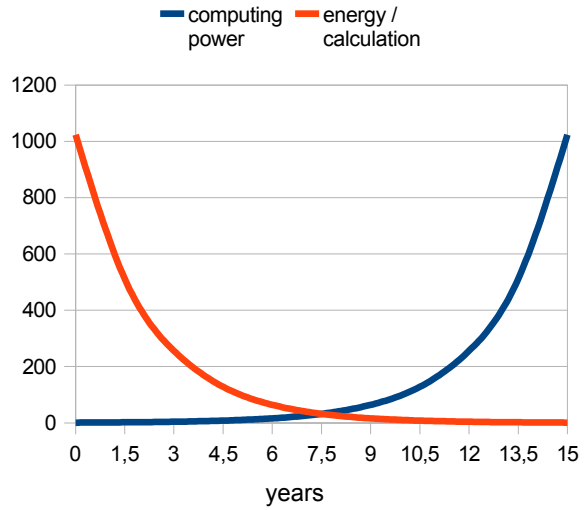
- Large amount of PC available
  - Pools:
    - CPU most of the time idle using electricity
    - Or switched off
  - Both states undesired

- Need of computational power and time
- Separate infrastructure
- Merging resources



Motivation II

- Lifetime of a PC about 2-4 yrs. (Warranty 3 yrs.)
- Computing power rises
- Energy consumption remains quite stable
- Regular replacements
- Green-IT:
  - No wasting of resources
  - Less hardware / duplicates
  - Better energy to performance ratio



II. State of Present Computer Technology

- Desktop- and pool PCs offer reasonable compute power and based on the identical architecture like today's cluster/grid nodes
- Powerful ethernets with up to 1GBit/s to the personal desktop

Stateless Computer Operation

- Disk-based installations increase the maintenance efforts
- Dramatic decreased administration because of centralization
  - attendence of central servers instead of decentral nodes
  - new clients are simple to add
  - easy replacement of failing machines
  - rather different operating systems and/or operations could be run on just same machine

III. Experiments Matrix

- Different terms must be satisfied:
  - Major requirement: Desktop user shouldn't notice any performance degradation
  - Desktop / cluster user shouldn't interfere with other environment
  - Environments must be separated strictly (no reading/writing of data, …)
  - Easy deployment (small administrative overhead)

III. Virtualization

- For separation of desktop/cluster computing environment
- Easy to separate environments
- Secure
  - Overhead of virtualization techniques is to be estimated
  - What is necessary to prior desktop processes (no mouse lags, programs start in acceptable time…)

Linux and Virtualization Options

- Linux as Open Source OS
- Platform for most grids / clusters
- Degree of para- / virtualization
- Full virtualization:
  - binary translation: VMware Server
  - hardware-assisted virtualization: KVM
- Paravirtualization: Xen
- Operating system-level virtualization: Linux VServer

Full Virtualization I

- Binary translation
  - instructions are analyzed / translated on the fly
- VMware Server
  - chosen because of the maturity of product
  - Experience gained at computer center
- Easy to
  - Install,
  - Configure,
  - Test on stateless clients
- For first test, does not allow cluster

Full Virtualization II

- Hardware-assisted virtualization
  - uses virtualization extension of the CPU
- Kernel-based Virtual Machine (KVM) chosen
- Integrated in linux kernel since 2.6.20
- Only kernel mode part
- User mode: Modified QEMU
- Qemu easy to use
  - configuration of the VM through command line

Paravirtualization

- Patch of kernel necessary
- Xen hypervisor loads in ring 0
  - Host kernel runs on ring 1
- Xen best known for paravirtualization
- Most complex solution

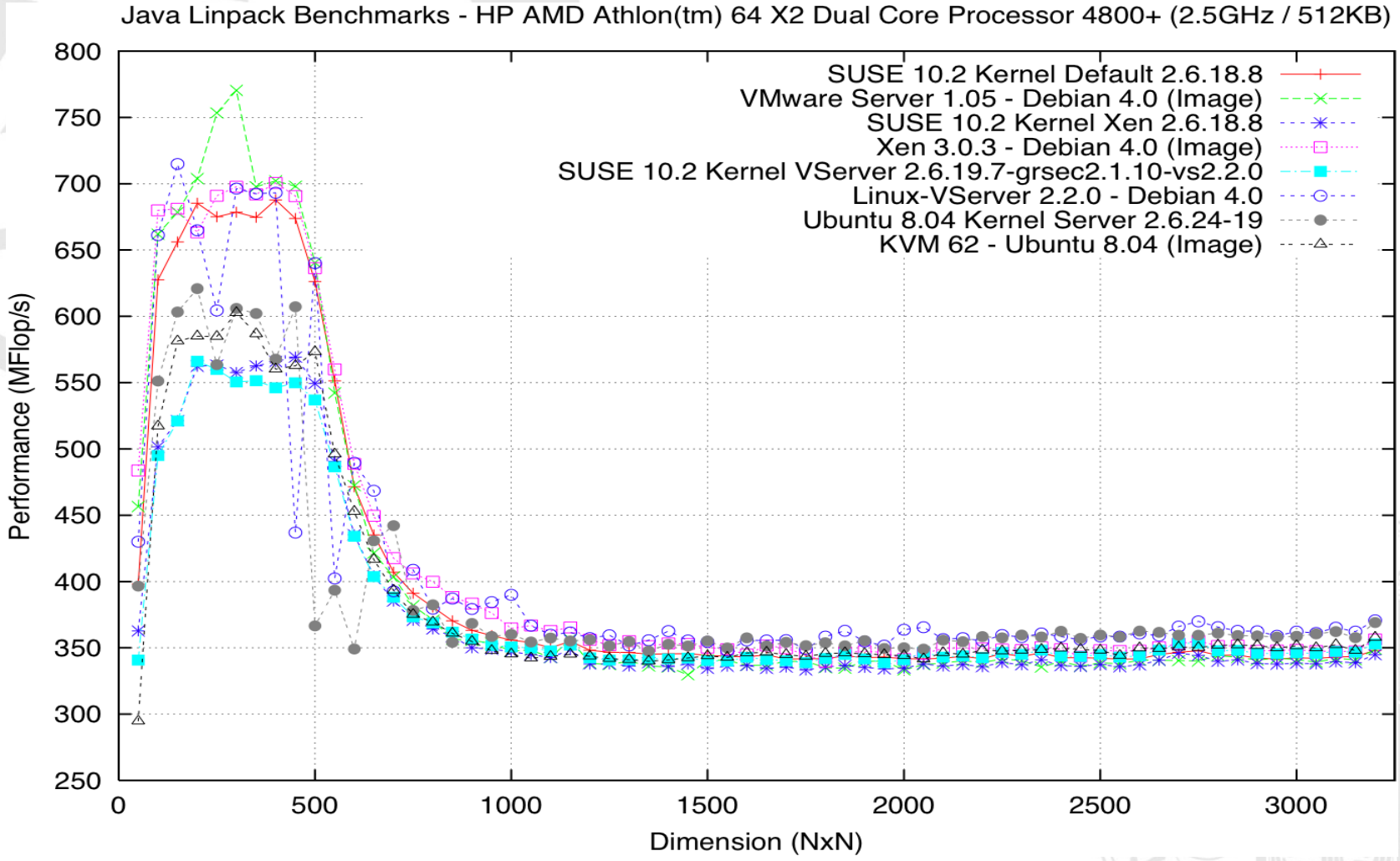
OS-level Virtualization

- Patch of kernel necessary
- Technique also called jails
  - For each system one container
  - Only user space part
- One kernel for all systems, less overhead
- Linux-VServer well known and for free

IV. Preliminary Results I

- Java Linpack benchmark
- Test on host and guest (VM)
- No user interaction

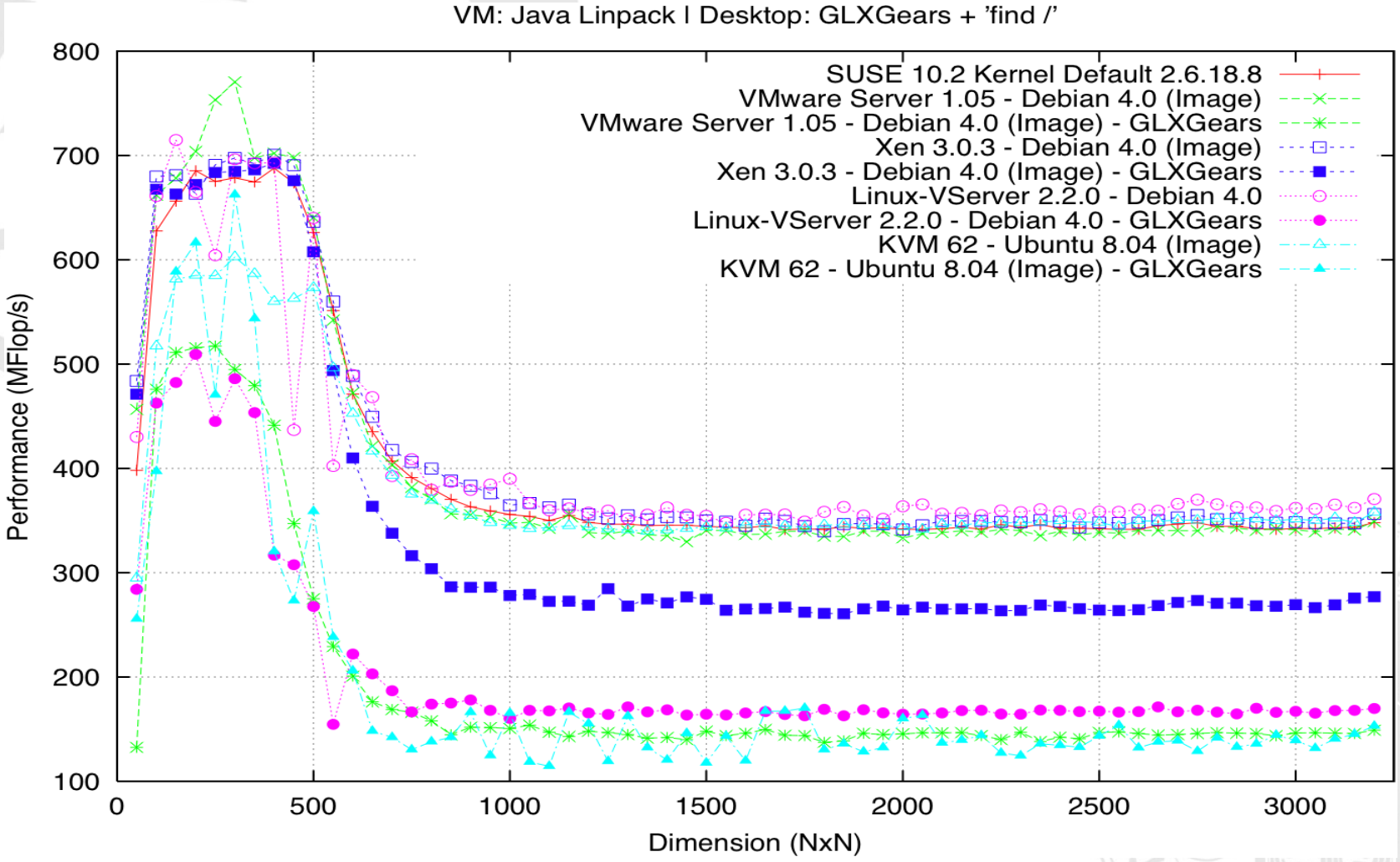
Preliminary Results I



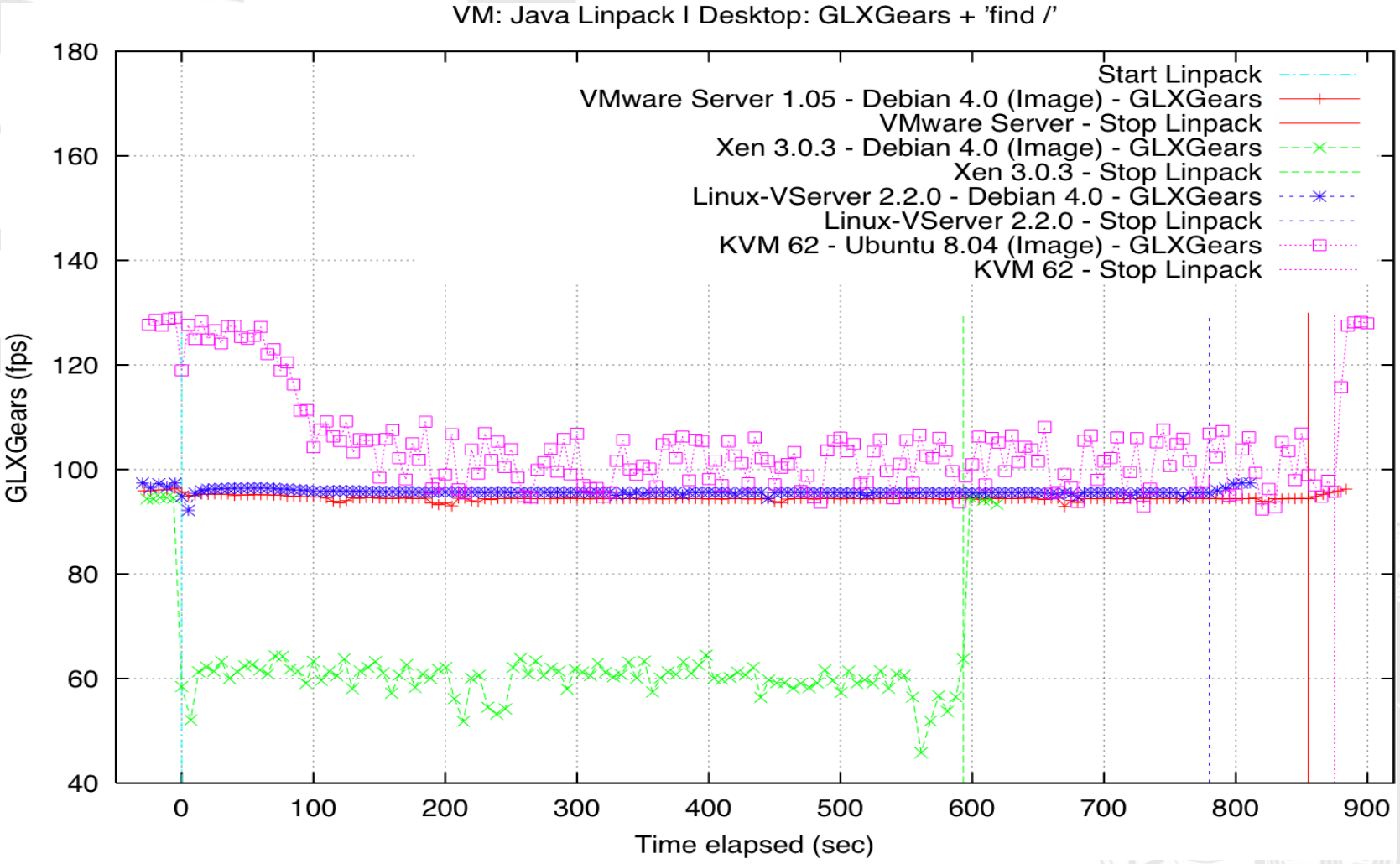
Preliminary Results II

- Java Linpack benchmark
- User interaction simulated by glxgears and find
  - Glxgears simulates graphical interaction
  - Find loop simulates I/O interaction
- Glxgears results of KVM not comparable to the others since other OS and X.Org version was used

Preliminary Results II



Preliminary Results II



Experiment Results

- Best results with Java Linpack were delivered by Linux Vserver (as it could be expected)
- Close behind follow VMware Server and KVM
- But KVM is rather new in the arena
  - Has to be tested in an equivalent environment
  - Shows leaps of performance in synthetic setup
- Xen shows significant decrease the performance of our synthetic setup
  - Fails major requirement: Desktop user shouldn't notice any performance degradation

Conclusion - Deployment

- Requirements for running virtualization in large scale
- VMware Server, KVM and VServer easy to deploy in stateless environments, thus no adding great efforts to add cluster feature to standard desktop
  - Kernel modules
  - Tool setup scripting for automated startup
  - Guest environment operation (how to load guest systems)
- Xen more difficult to install: Client bootup procedures have to be greatly modified

Further Research

- Real-life tests have to be executed with real cluster computing applications
- User interaction should be tested manually as well
  - Use and feel
- In large scale cluster deployment: Cluster scheduler have to handle higher frequency of node status changes
  - Nodes might crash (be rebooted, switched off) more often than classic cluster machines
  - Fluctuant compute power – so more difficult forecasting of job length
  - Scheduling suitable jobs into desktop cluster

Questions!? / Contact Information

Lehrstuhl für Kommunikationssysteme
Rechenzentrum der Universität
Herrmann-Herder-Str. 10
79104 Freiburg

Tel. +49 761 203 4698 / 8058
Fax +49 761 203 4640

mj0@uni-freiburg.de
dsuchod@uni-freiburg.de

www.ks.uni-freiburg.de
portal.uni-freiburg.de/rz