ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

550 Jahre
Albert-Ludwigs-
Universität Freiburg
1457 – 2007

5th Black Forest Grid Workshop
23rd April 2009

Stateless Compute Cluster

Fast Deployment and Switching of Cluster Computing Nodes for easier Administration and better Fulfilment of Different Demands

Dirk von Suchodoletz
Chair of Communication Systems
Prof. Gerhard Schneider

Institute for Computer Science /
Computer Center of the University of Freiburg

Stateless Compute Cluster

- Structure of my presentation
  - Motivation of this presentation
  - Stateless operation of computers
  - Principles of operation
  - Boot procedures and menus

Speaker / Project

- Dirk von Suchodoletz - dsuchod@uni-freiburg.de, Assistant of the CommSys professorship
- Project coordinator of a stateless Linux project used in pool systems to run WinXP in VM-Player
- Large computer pool environments up to 400+ machines, experiences in diskless Linux deployment since mid of 1990th
- Clusters have quite similar requirements, idea of extending the project

Motivation of this Talk

- Grid and cluster should be optimal/optimized to reduce administration efforts
- Coming from Linux desktop pool operation
  - Deploying same principles of pool OS distribution into the cluster domain
  - Using idle times of desktops/pools for cluster computing – they offer reasonable compute power, think of pools during night, weekends and holidays
  - No additional rack space, cooling, power, …

Motivation of this Talk

- Typical administration tasks
  - OS roll-out to large number of nodes
  - Permanent updates of these machines
  - Fast inclusion of newly installed nodes
- Tests and experiments
  - Test new OS, versions
  - Test specific environments
  - If OS directly installed difficult to handle (partition the disk, more installation time and efforts, ...)

Stateless Computer Operation

- Idea: 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 (just rebooted into other system)

Requirements: High-capacity Networks

- In most institutions: well expanded backbone with permanent increase of bandwidth and reduction of latency, while highly available
- Powerful Ethernets with up to 1GBit/s to the personal desktop
- Decreasing differences to avarage cluster nodes of the ethernet connected type
- Plain sailing: clients root filesystem transported via network

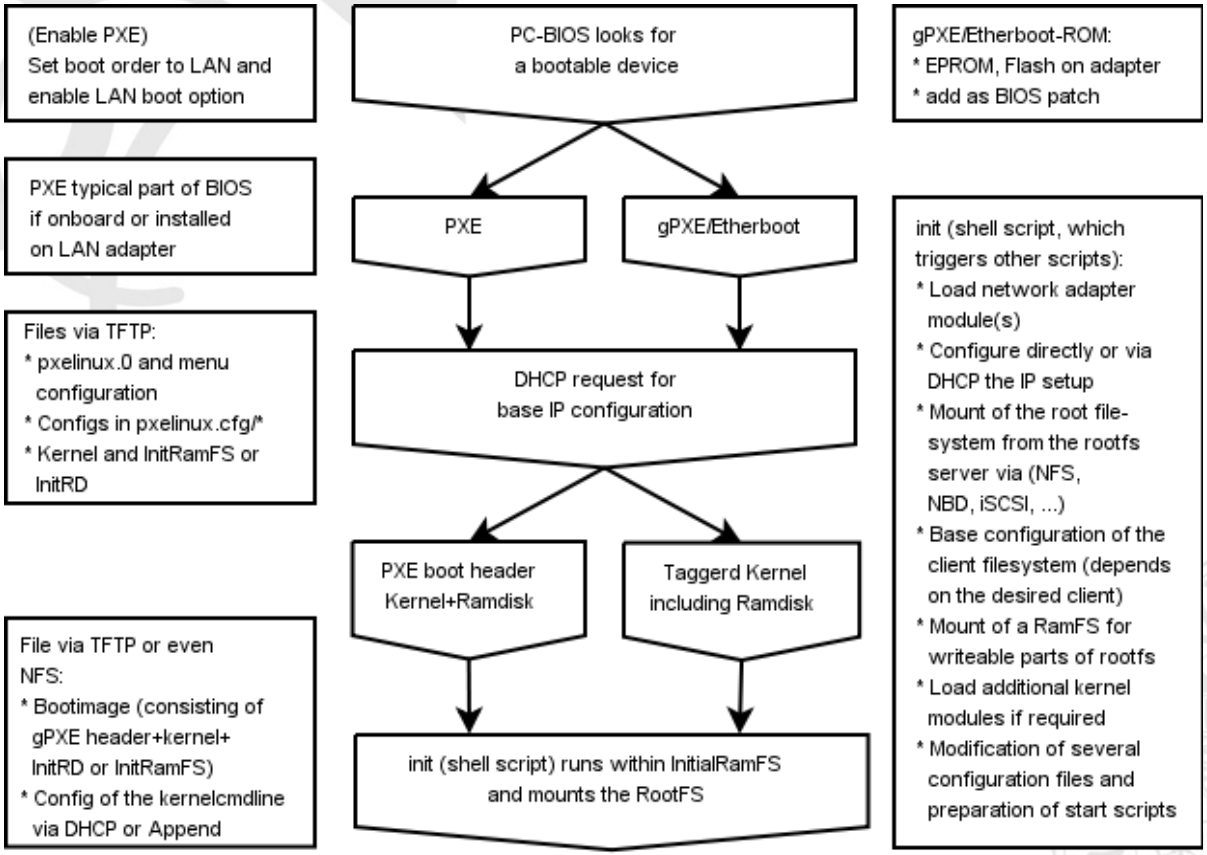
Stateless Linux I

- Different projects on the net and research at the professorship
  - Run already a number of maschines
  - Computer pool operation at the CC and other faculties
  - By now: general offering of remote boot services for third party in the university
- To achieve this objective
  - Offering of a simple standardized base environment extendable by third party

Stateless Linux II

- Client side
  - Does not need any installation to hard disk, but could use any disk for local scratch
  - All clients share the same root filesystem, which is stored on one/some servers
  - Clients are configured automatically during bootup
  - Bootup speed is despite network transfer often better then in disk based installations
  - Present system offers despite shared rootfile-system per client configuration

Stateless Linux III – Bootup procedure



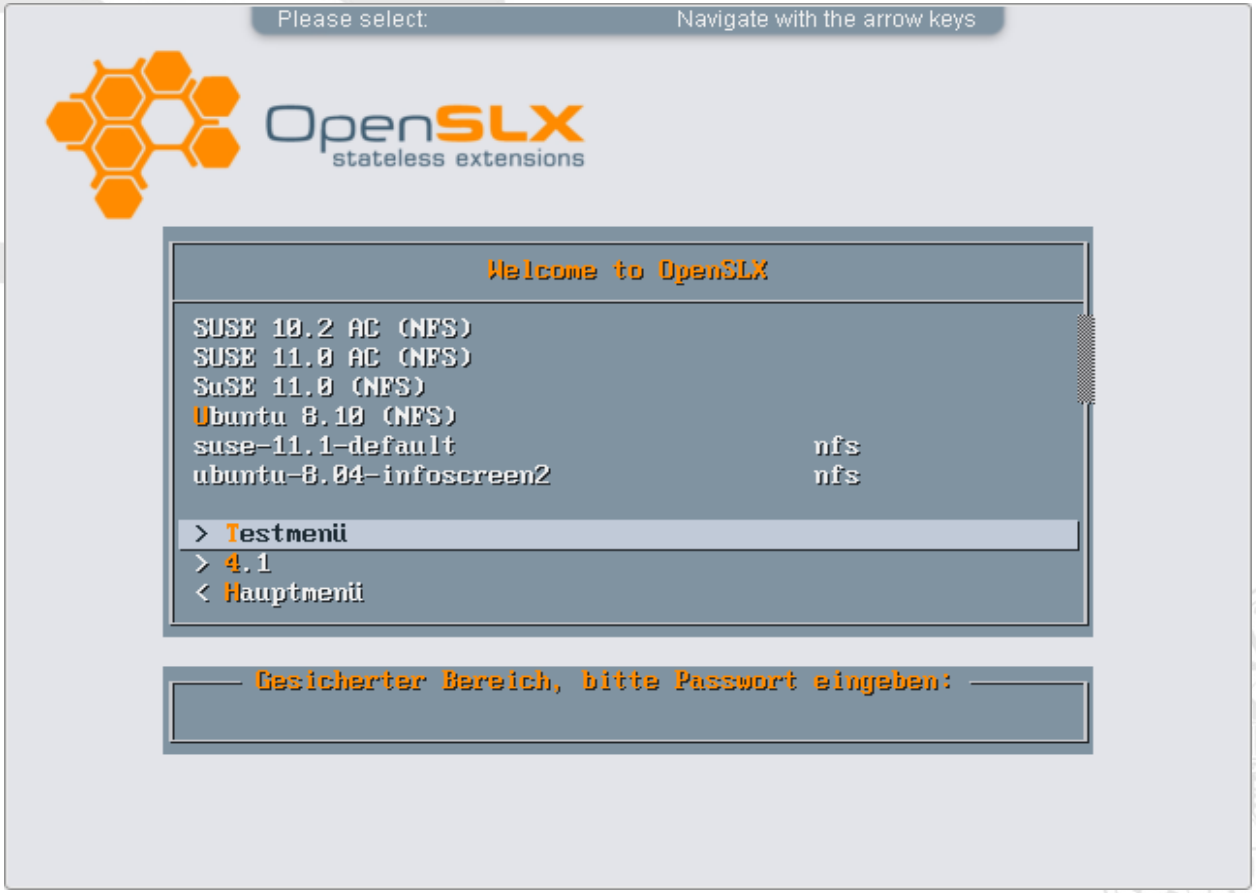
Stateless Linux IV

- Servers side
  - One server is able to host several different rootfilesystems and large number of clients
  - Avarage hardware requirements, lots of RAM, fast disks, broad network connection to the backbone improve performance
  - Using redundant servers and failover it is easy to have simple maintenance
  - Using standard Internet protocols, like DHCP, TFTP and NFS

Technological Base: PC und PXE

- Typical desktop PC, cluster node offers the capability of PXE booting
- Integrated into centrally managed DHCP, depends simply on additional TFTP
- Lots of free and commercial boot products which could even be chained
  - Offering the option of sophisticated boot menus (perfect for flexible test environments)
  - Using PXE/SysLinux here

PXE Boot Menus



Software Base: OpenSLX Project

- Software suite to setup average Linux distributions for network/stateless booting
- Two major parts
  - Server and system setup done by admin
    - Installation of Linux distribution(s)
    - Configuring of additional componets (plugin system) and clients (DB)
    - Exporting the systems via NFS/NBD ...
    - Creating boot configuration and PXE menus

Software Base: OpenSLX Project

- Client setup
  - Scripts running on the client to configure hardware and services (on every client during every boot)
  - Translucent filesystems help to simplify even more ...
- General idea: using as much of the distribution code and packages
- Minimizing the adaptations to diskless operation

Software Base: OpenSLX Project

- Project provides a set of commandline tools for each admin task
- ... and provides setup scripts for a range of distribution like Ubuntu, SuSE, Debian, RedHat/Scientific Linux
- Well tested concept for lecture pool setups for over ten years ...

Testing needed!

- If any value seen in this appoach – more real life experience required
- Basic Integration of Scientific Linux done, but
  - Which additional components required
  - Concept usable for cluster setups
  - Mixed mode tests (automatic re-booting of pool machines into different modes)

Conclusion

- Advantages
  - Ultra-fast deployment of cluster nodes
  - High flexibility and fast switching of operating systems, versions and dynamic reconfiguration
  - Easy advance testing
- Possible disadvantages
  - Network/server risk (of failing)
  - Network load and response time of root filesystem servers
  - Differences of disk-based/less operation

Questions!? / Discussion!

Thank you very much for your patience and please do not hesitate to ask questions, give comments, ... !

Chair of Communication Systems
Computer Center of Freiburg University
Herrmann-Herder-Str. 10
79104 Freiburg

Tel. +49 761 203 4652 / 8058
Fax +49 761 203 4640

dsuchod@uni-freiburg.de

www.ks.uni-freiburg.de/projekte/ldc / www.openslx.org