

# Server Consolidation with Xen Farming

Ulrich Schwardmann

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Am Fassberg, 37077 Göttingen  
ulrich.schwardmann@gwdg.de

Linux Kongress 2008, 9.10.2008

# Content

Xen Farming

Ulrich  
Schwardmann

- 1 Server Consolidation
- 2 Concept
- 3 RO/RW file systems
- 4 Boot Process
- 5 Updates
- 6 Consistency
- 7 Orchestration
- 8 Security
- 9 Further Plans

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Server Consolidation

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- should be more than just virtualization,
- it should be an advantage for the administrator
- separation of operating system and application administration
- tradeoff between centralization and flexibility in a xen farm

# The Xen Farming Concept

## The Xen Servers Viewpoint

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

From the servers point of view, a xen farm is

- just a couple of xen clients.
- The clients reside in a special network structure
- The clients external network connections is bridged by the xen server.
- The clients use an internal network:
  - to mount a special NFS share
  - and for administration tasks, executed from the xen server:

# The Xen Farming Concept

## Overview and network structure

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

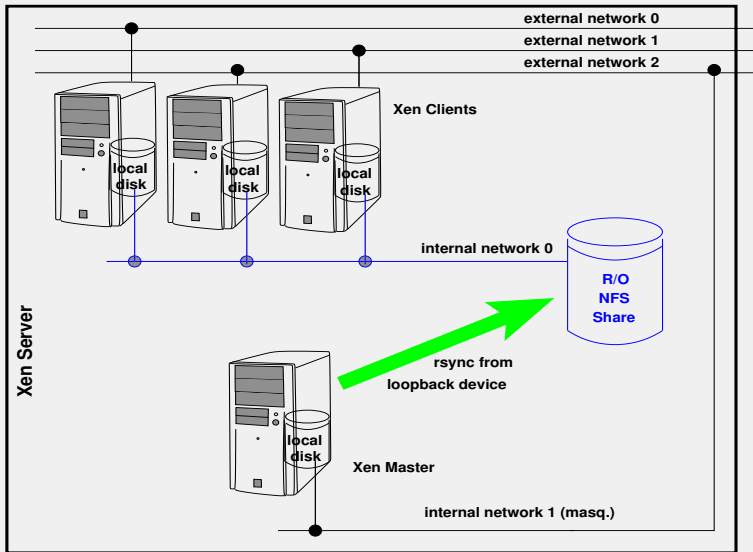
Updates

Consistency

Orchestration

Security

Further Plans



# The Clients Viewpoint

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- The xen clients are the main worker in this concept
- A client is a usual xen machine, but without write access to mayor parts of its installation.
- It imports these parts instead by mounting a readonly file system via its internal network connections.
- During updates of this readonly part the changes from outside have to be transparent to the operating system inside the client.
- This necessitates the use of some shared/network file system.

# The Masters Viewpoint

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- The xen master is the origin of the whole installation and all updates in this installation.
- It has to be alive only during its installation and update processes,
- otherwise only its file system has to be seen by the clients.
- One way could have been to export the masters root file system to the clients directly,
- for safety and security reasons update and file service was completely separated,
- therefore the master is started only for installation and updates.

# Files System: writeable and readonly parts

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

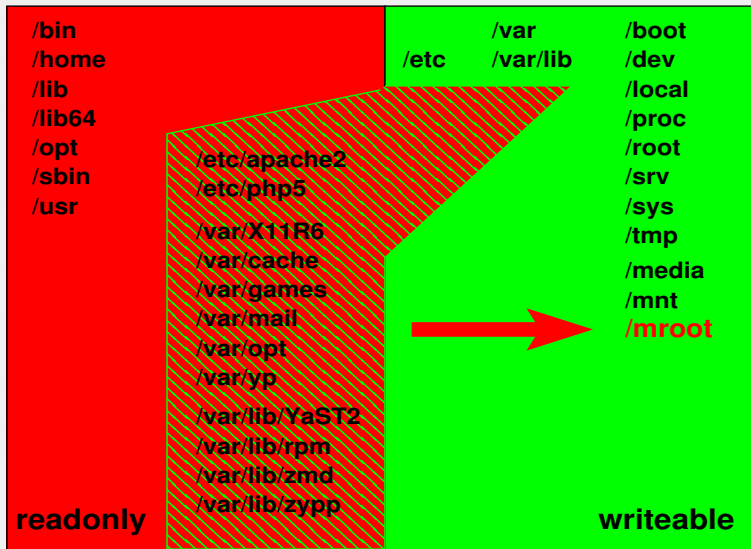
Further Plans

- Executables, libraries, static data etc. can be hold in a readonly region.
- Files, that have to be writeable:
  - all files that are changed by the operating system during lifetime
  - all configuration files must be configurable by the administrator
  - home directories
  - temporary directories
  - mount points
  - directories for application data
- typical: 4.5 GB in the readonly area and about 150 MB in the writeable area



# Files System: writeable and readonly parts

Xen Farming

Ulrich  
SchwardmannServer  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Boot Process

## Concept of Diskless Clients (DXS)

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- DXS have no disk at all, they get their files from an NFS server.
- DXS are developed to administrate a greater amount of computers, mainly used as X server (D.v.Suchodoletz)
- advantage: the processes use the local hardware
- DXS get the complete network configuration, kernel, initial ram disk and file system via TFTP boot from the server.
- for xen farming only parts of the file system must be imported
- problem: how gets the kernel access to the remote nfs share during its boot process.

# Boot Process

applied to Xen Farming Clients

Xen Farming

Ulrich  
Schwardmann

- Client has special boot option for internal network config.
- This is read by the init process of the initial ram disk
- It contains the ip-address, the NFS server, the gateway and the broadcast setting of this client.
- The init process can set up a network interface according to these settings.
- Init starts the portmapper, loads the appropriate drivers for the nfs mount and mounts the readonly file system from the NFS server.
- The mount point of the readonly file system has to be moved to the correct place before:
- initrd-init starts '/sbin/init', out of the NFS share.
- Everything continues as with a usual OS init
- A debug function can be implemented into the initrd, that starts a shell before starting '/sbin/init'

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Boot Process

The relevant lines in init of initrd

Xen Farming

Ulrich  
Schwardmann

```
mount -o remount,$fsoptions $rootdev /root
/bin/mount --move /mroot /root/mroot
/bin/mount --move /dev /root/dev
cd /root
umount /proc
umount /sys
# Export root fs information
ROOTFS_BLKDEV="$rootdev"
export ROOTFS_BLKDEV
exec /bin/run-init -c ./dev/console /root \\\
$init $init_args $runlevel
```

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Updates

## of the Master and the NFS Share

Xen Farming

Ulrich  
Schwardmann

- An update of the xen farm has its origin in an update of the master.
- After starting the master a usual update process is performed.
- In the case of a kernel update, the old `'/lib/modules'` files are saved (see below).
- The master is shutted down afterwards.
- The virtual disk of the xen master is loopback mounted.
- The xen server synchronizes the nfs share from there.
- This updates all files in the readonly part of the clients file systems.
- The properties of NFS ensures a transparent change of files for the clients processes.

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Updates

of initial ram disk and kernel modules

Xen Farming

Ulrich  
Schwardmann

initial ramdisk:

- Each new kernel comes with a new initrd
- In this environment we need to modify the init file inside the new initial ram disk.
- Therefore the new initrd is unpacked,
- the init file is modified (exchanged),
- the initrd is packed again
- and copied to the correct place in the NFS share.

kernel modules:

- the kernel modules of the old kernel version, that still runs on the clients, is still available
- therefore a reboot, which is in the application administrators responsibility, can be postponed.

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

# Updates

of the xen clients writeable part

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- The update is performed on the clients itself by synchronizing with the NFS share.
- In principle all the updates on the master should be available on the clients too.
- Therefore there has to be some synchronizing mechanism for all files in the writeable part
- But there are changes on files done by the application administrators, that should not be overwritten by this procedure
- Usually update mechanisms have policies, that do not overwrite critical configuration files.
- But this does not work here, because the update is done on the master, where no changes in the configuration is performed.
- Necessary: a special update policy for this part of the file system

# Consistency after Update

## general aspects

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- An installation or update of the system introduces new files and replaces old files.
- a conflict can occur there, where a new file of an update or installation has the same name of a file, that is introduced or changed by an administrator.
- Since all files in the readonly part of the file system can't be changed by the administrator, the change of these files is controlled by the usual update mechanisms.
- Therefore only those files, that are in the writeable file system, can have a conflict.



# Consistency after Update

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- these conflictable files are therefore exactly those files,
  - that are in the writeable part
  - that differ from the corresponding files before update
  - or do not exist in the masters file system before update
  - and that are written by the installation or update
- one has to compare therefore all files in the writeable part of the client with the corresponding files on the master file system
- all files that differ and all additional files in the writeable part of the client indicate a conflict
- these files are left untouched in our model, the new version is copied as corresponding file with extension '.xennew' beside these files.
- we decided to use MD5-sum as an indicator for chngement.

# The Orchestration of the Xen Farm

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- done by a couple of simple shell scripts, and
- a configuration skript, that holds all variables used.
- These scripts are used to:
  - create and update the master,
  - create and update the clients,
  - or create and update the nfs share,
  - get access to the clients from the xen server,
  - find out there properties
  - communicate to the clients administrators

# Security

Xen Farming

Ulrich  
SchwardmannServer  
Consolidation

Concept

RO/RW file  
systems

Boot Process

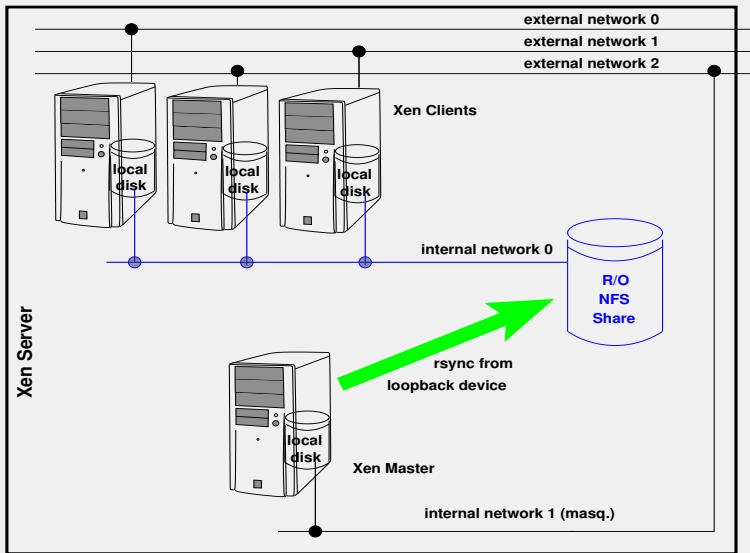
Updates

Consistency

Orchestration

Security

Further Plans



# Further Plans

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

- update of writeable fs-part is not really consistent at the moment
- some scripts are still missing
- proceed from proof of concept stage to an open project
- include other distributions (than SLES10)
- integration into virt-manager would be nice

# Thanks

Xen Farming

Ulrich  
Schwardmann

Server  
Consolidation

Concept

RO/RW file  
systems

Boot Process

Updates

Consistency

Orchestration

Security

Further Plans

to the other people working in this project:

**Tim Ehlers, GWDG**

**Christian Boehme, GWDG**

and

# Thanks for Your Attention !!

# Questions ??