

Computing @ GSI

- LINUX @ GSI, overview
- Migration of services
- growing of the LINUX Batch farm
- Migration Windows Desktops -> XP
- security: new firewall configuration
 - statefull routing
 - GSI “free” WLAN
 - DMZ

LINUX @ GSI

- central services (webserver, mailserver,
- desktop
- scientific computing
 - number crunching (batch farm)
 - physics analysis
 - file servers for experiment data

some numbers:

- > 400 LINUX computers
- about 10 file servers
- about 10 group servers
- about 10 compute servers

Migration of services

- migration from AIX/Windows -> LINUX
 - IMAP server
 - old solution: exchange 5.0
 - new solution: OCS?, CYRUS? ,Exchange2003?
 - mailservr:
 - old solution: sendmail on AIX
 - new solution: postfix on LINUX
 - + Spam Filter + Virusfilter
- dhcp
 - testing (automatic configuration based on ORACLE DB, migration from Windows to LINUX soon)
- dns
 - testing, migration from AIX to LINUX soon

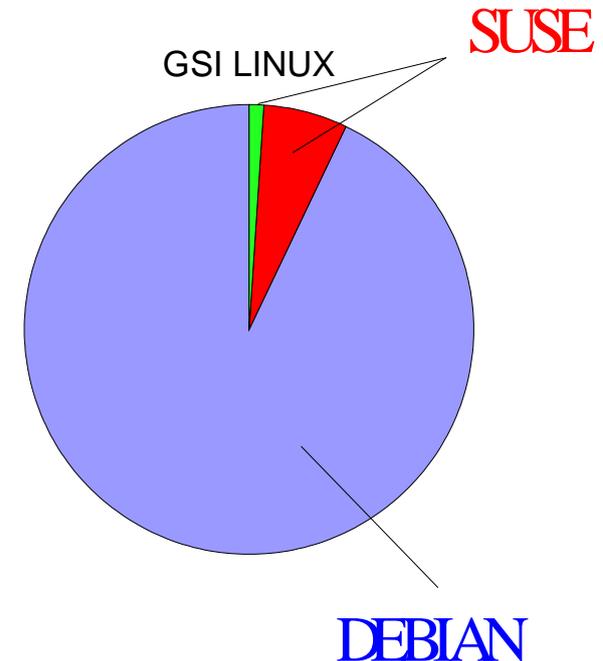
LINUX at GSI:

LINUX cluster: DEBIAN

- stability
 - easy upgrade
- **Laptops** (“stand alone”): **SUSE**
 - hardware detection
 - easy administration for users
 - ... the newest gadgets inside.... ;-)

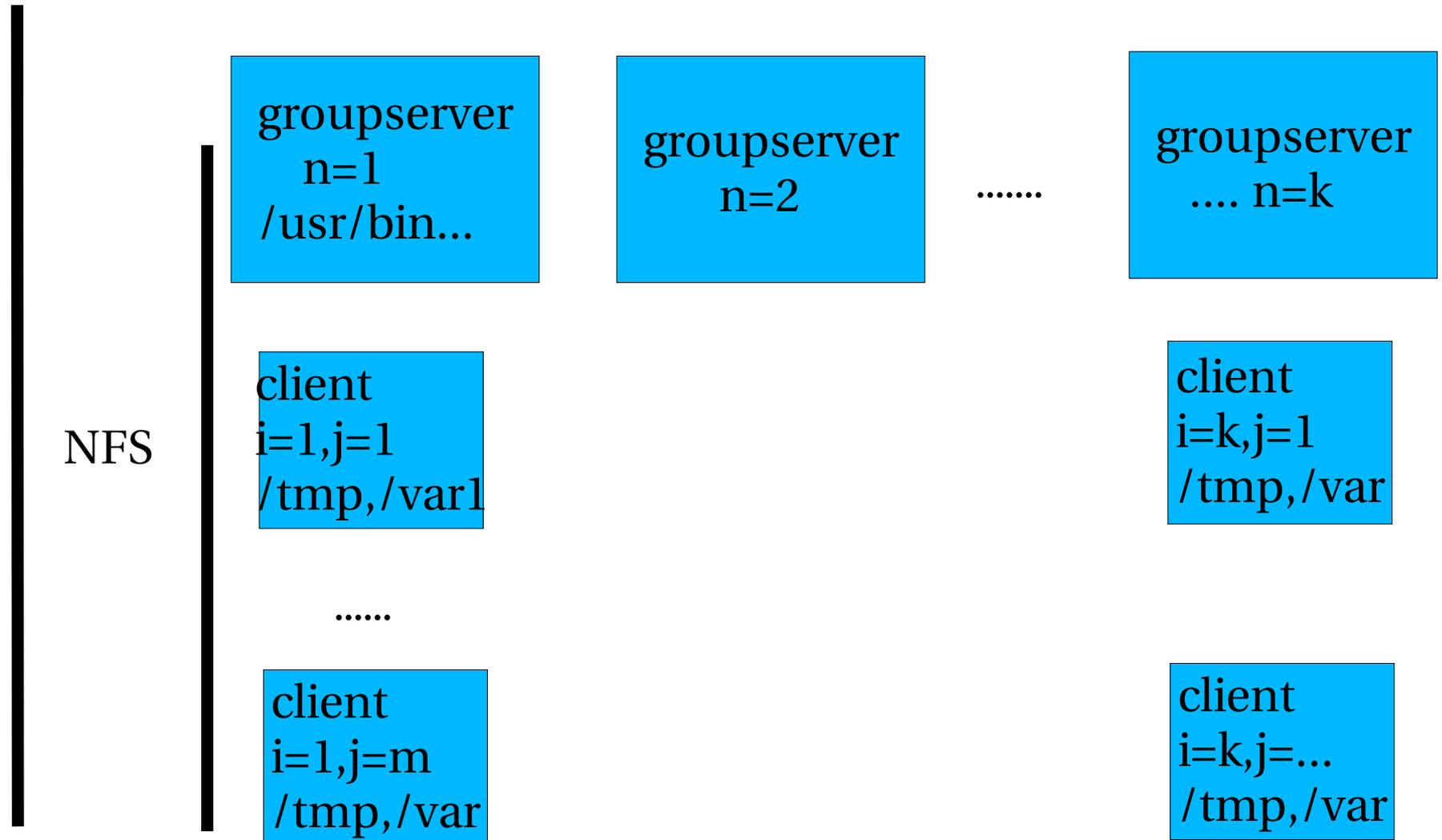
servers for commercial software:

- **SUSE** enterprise server
(required from ORACLE, TSM server)



fileserver
/usr/local,
/data/.....

The GSI LINUX Farm - organisation



LINUX Batch Farm

- 224 CPU's (in 112 smp computers)
 - 96 cpu PIII, 600 MHz, ATX midi tower
 - 128 cpu Xeon 2.4 Ghz, 19"

future: about 100 more CPU's in "pizza box"
and replacement of the 600 MHz boxes (space!)
by 3 Ghz boxes



Host: LX8001
Alias:
Internet: 140.181.96.217
Hardware:
Name: LX8001
Path:
Title Name:

Host: LX8002
Alias:
Internet: 140.181.
Hardware: 00-00-00-00-00-00
Name: LX8002
Path:
Title Name:

Host: LX8003
Alias:
Internet: 140.181.96.775
Hardware:
Name: LX8003
Path:
Title Name:

Host: LX8004
Alias:
Internet: 140.181.96.295
Hardware:
Name: LX8004
Path:
Title Name:

Host: LX8005
Alias:
Internet: 140.181.
Hardware: 00-00-00-00-00-00
Name: LX8005
Path:
Title Name:

Host: LX8006
Alias:
Internet: 140.181.
Hardware: 00-00-00-00-00-00
Name: LX8006
Path:
Title Name:

Ultraview Pro
www.supermicro.com

Windows @ GSI

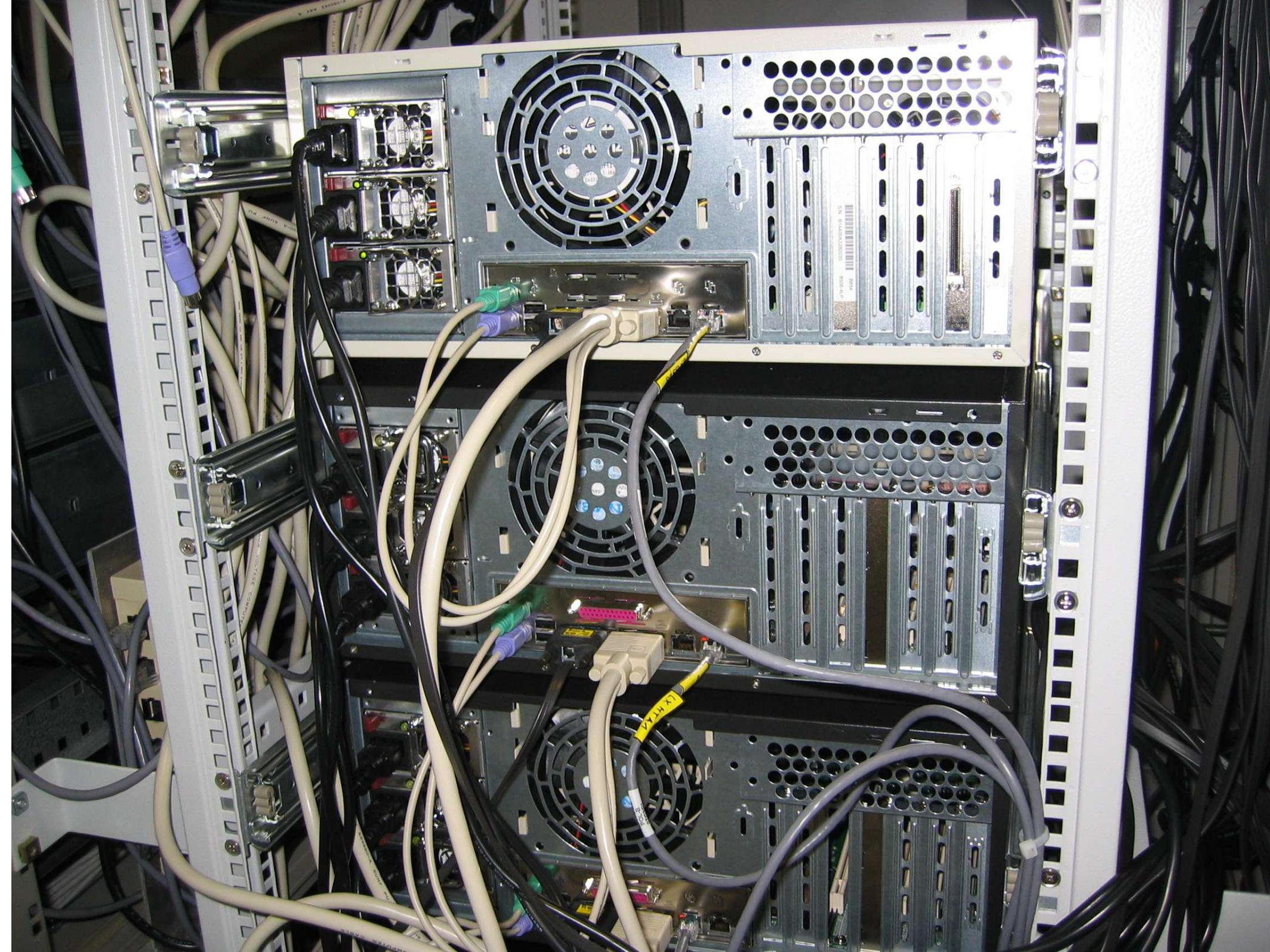
- active directory for the new XP clients
- 2 domain controllers
- 1 print server (Win2K cluster with 2 nodes)
- 2 file servers
- 2 application servers
- 2 web servers
- windows2003 servers
- virus filter McAfee
- new Windows desktops: Windows XP
- Exchange IMAP server + Groupware

Linux based HA server

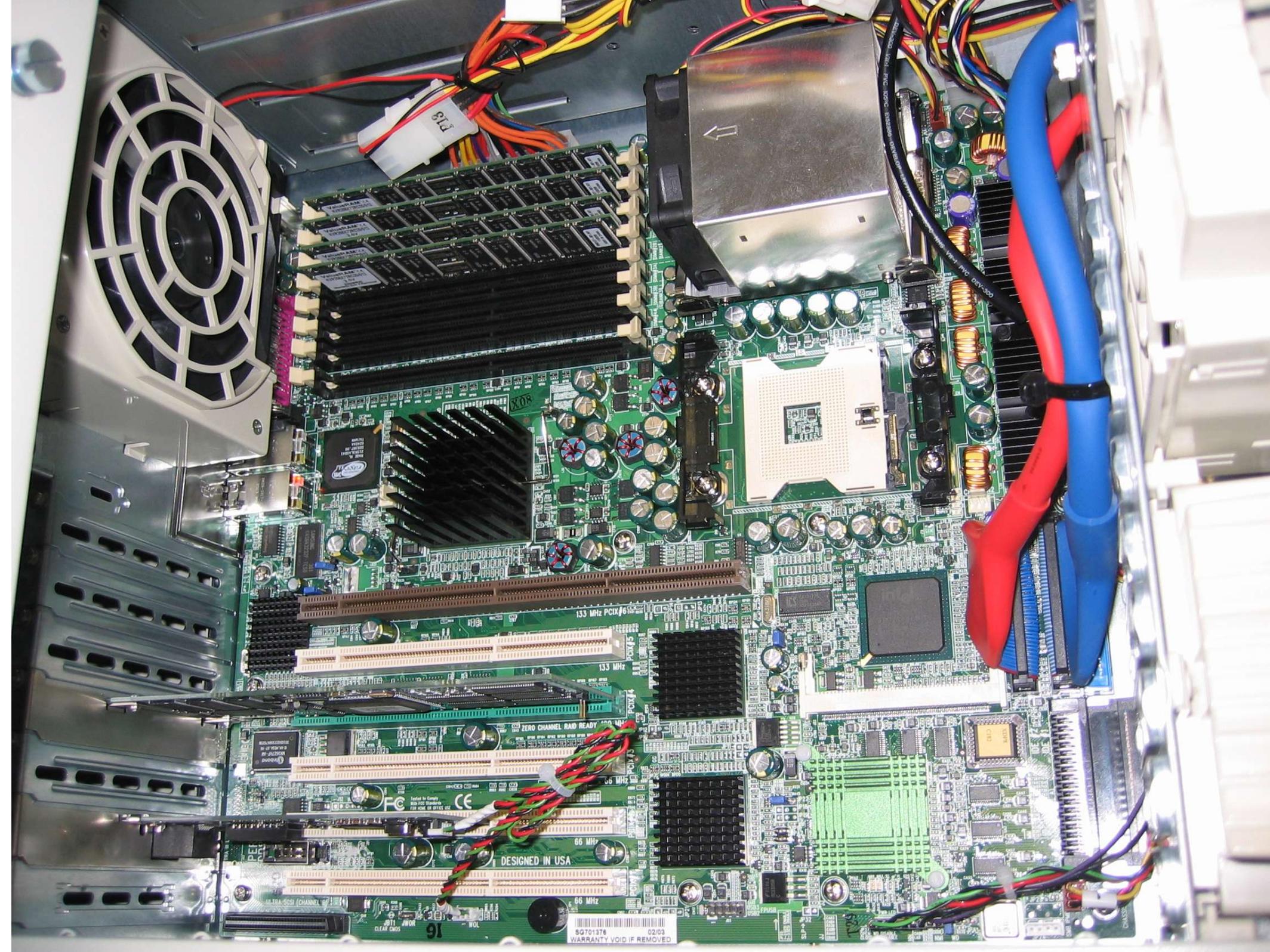
=> Talk from K.Miers

- high availability system for mission critical services
- special hardware with redundant power supplies
- system on hardware RAID I
- important resources on mirrored RAID









118

133 MHz PCI-X #6

133 MHz PCI-X #4

66 MHz PCI-X #3

66 MHz PCI-X #2

66 MHz PCI-X #1

SG701376 02/03
WARRANTY VOID IF REMOVED

DESIGNED IN USA

FC

CE

ZERO CHANNEL RAID READY

1.5 TB SAS CHANNEL #1

1.5 TB SAS CHANNEL #2

1.5 TB SAS CHANNEL #3

1.5 TB SAS CHANNEL #4

1.5 TB SAS CHANNEL #5

1.5 TB SAS CHANNEL #6

1.5 TB SAS CHANNEL #7

1.5 TB SAS CHANNEL #8

1.5 TB SAS CHANNEL #9

1.5 TB SAS CHANNEL #10

1.5 TB SAS CHANNEL #11

1.5 TB SAS CHANNEL #12

1.5 TB SAS CHANNEL #13

1.5 TB SAS CHANNEL #14

1.5 TB SAS CHANNEL #15

1.5 TB SAS CHANNEL #16

1.5 TB SAS CHANNEL #17

1.5 TB SAS CHANNEL #18

1.5 TB SAS CHANNEL #19

1.5 TB SAS CHANNEL #20

1.5 TB SAS CHANNEL #21

1.5 TB SAS CHANNEL #22

1.5 TB SAS CHANNEL #23

1.5 TB SAS CHANNEL #24

1.5 TB SAS CHANNEL #25

1.5 TB SAS CHANNEL #26

1.5 TB SAS CHANNEL #27

1.5 TB SAS CHANNEL #28

1.5 TB SAS CHANNEL #29

1.5 TB SAS CHANNEL #30

1.5 TB SAS CHANNEL #31

1.5 TB SAS CHANNEL #32

1.5 TB SAS CHANNEL #33

1.5 TB SAS CHANNEL #34

1.5 TB SAS CHANNEL #35

1.5 TB SAS CHANNEL #36

1.5 TB SAS CHANNEL #37

1.5 TB SAS CHANNEL #38

1.5 TB SAS CHANNEL #39

1.5 TB SAS CHANNEL #40



Experiment Data on IDE-SCSI RAID_s

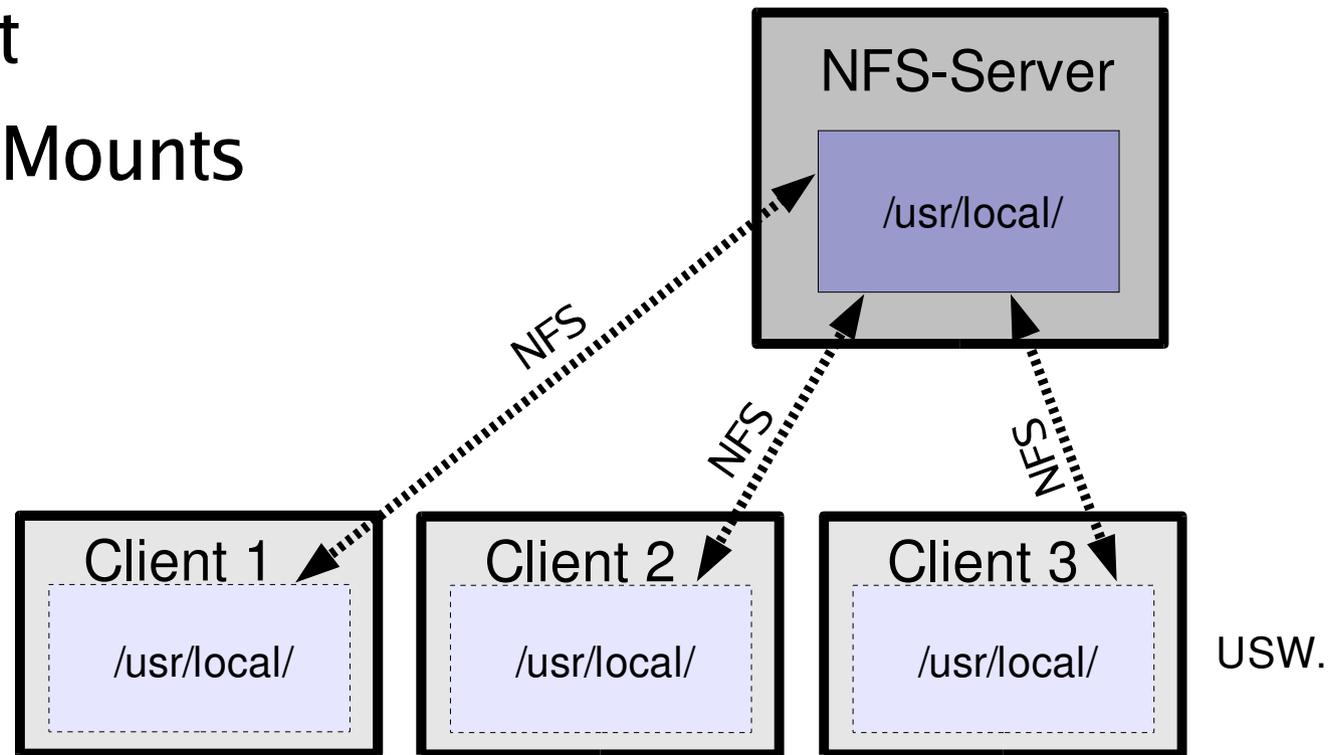
Walter Schoen, GSI

Linux-Ausfallsicherheit für NFS

- Problemstellung
- Lösungsansatz
- Tools
- Implementierung / Testbetrieb

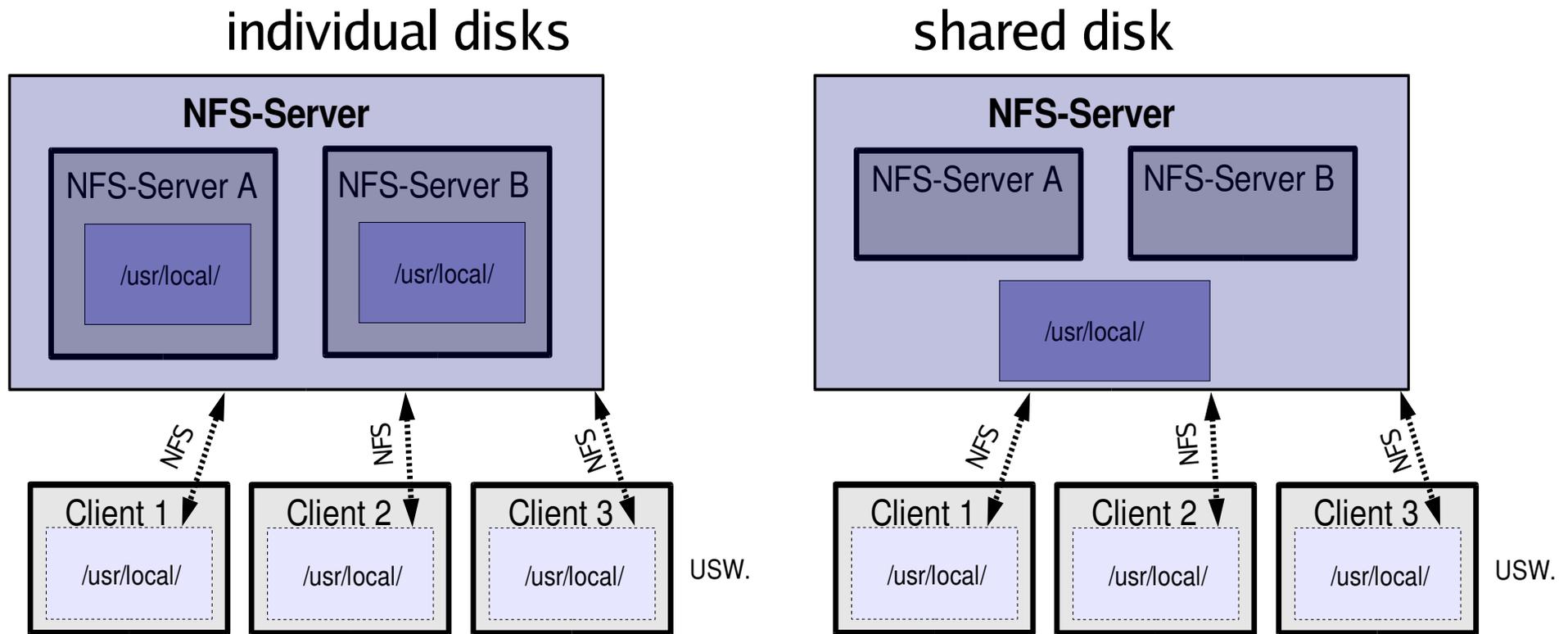
Problemstellung

- Gemeinsames /usr/local auf einem NFS-Server für alle Clients
- Ausfall des NFS-Servers:
 - Arbeit steht
 - Stale NFS Mounts



Lösungsansatz

2 identische Server mit Hot-Standby:



Problem: Synchronisation des Dateisystems

Linux-Tools

heartbeat

- Gegenseitige Überwachung der Server
- Starten der Dienste

drbd

- Synchronisation des Dateisystems (/usr/local)

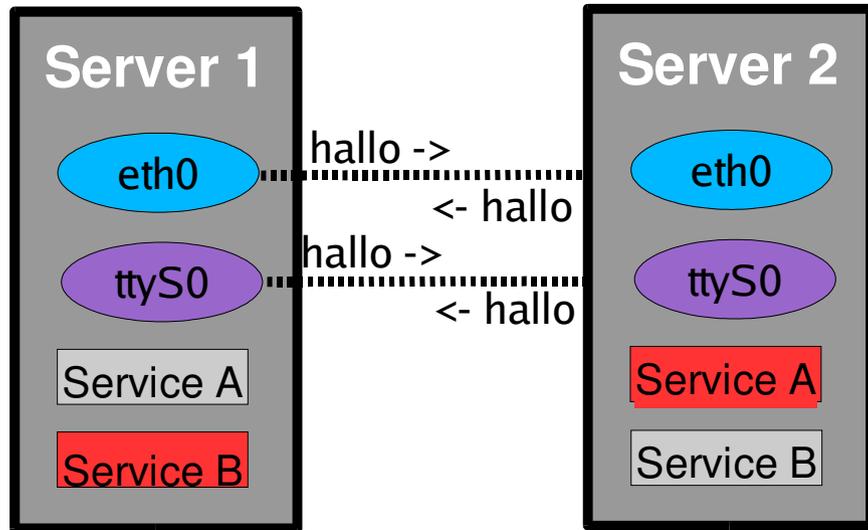
mon

- Systemüberwachung

heartbeat

- Wie weiß der Backup-Server, wenn der Master versagt?
- Beide kommunizieren über Ethernet und die serielle Schnittstelle (beides wegen Redundanz) und tauschen in regelmäßigen Abständen Nachrichten aus (Herzschlag).
- Wenn diese Nachrichten ausfallen, dann ist der Rechner ausgefallen.
- Der Backup-Server startet den NFS-Dienst.

heartbeat



Normalbetrieb:

Server 1 - Master für Service B

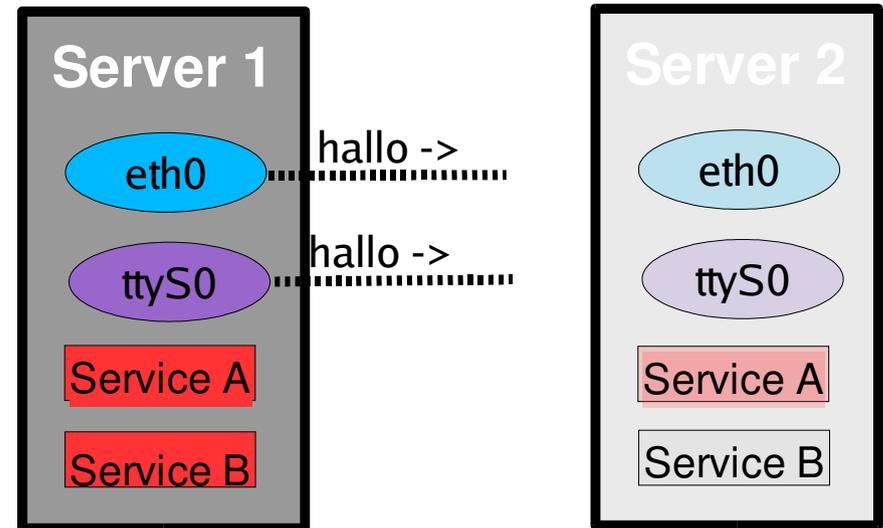
Server 2 - Master für Service A

Fehlerfall:

Server 2 fällt aus

“heartbeat- Antwort bleibt aus

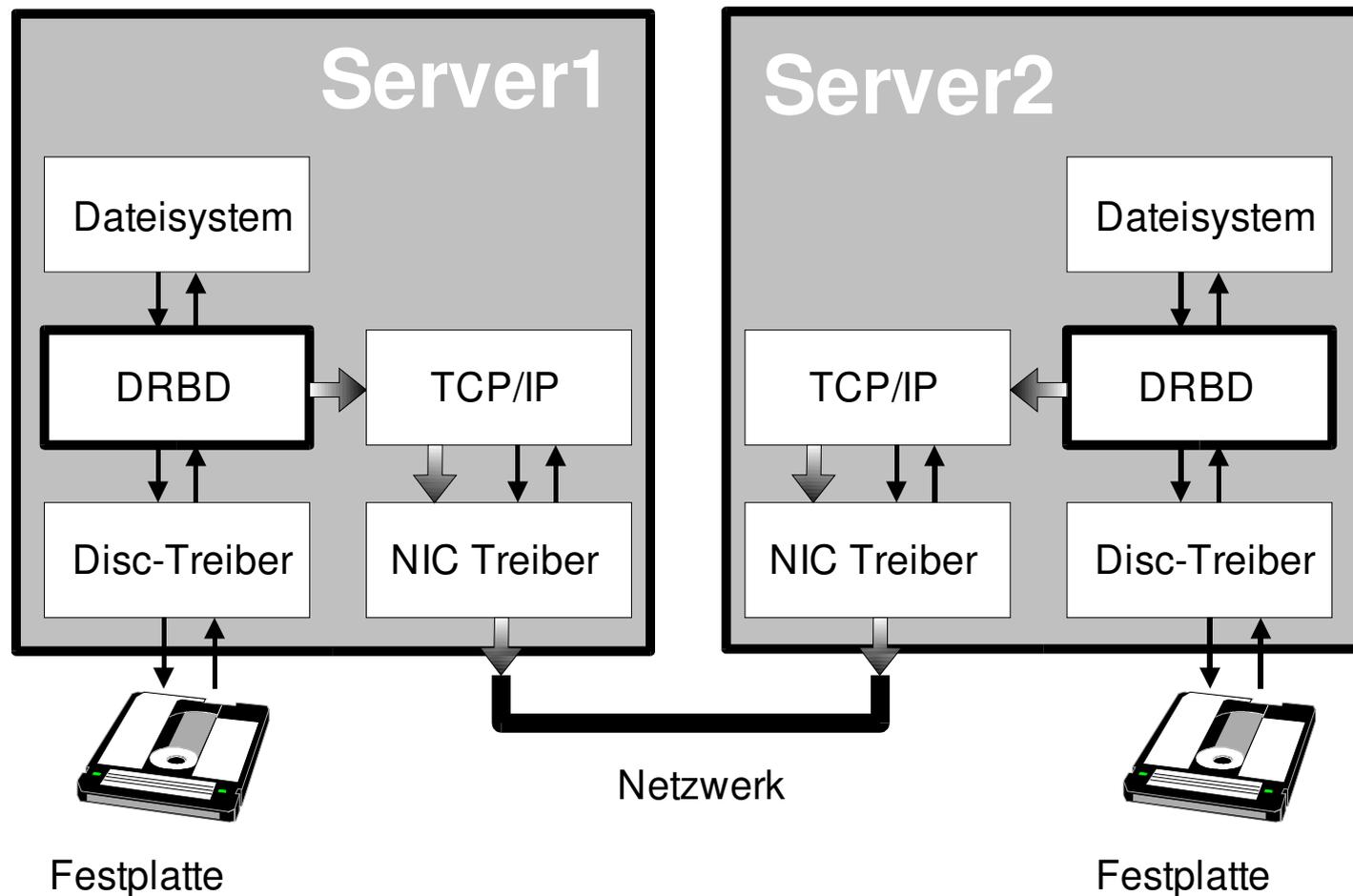
Server 1 übernimmt Service A



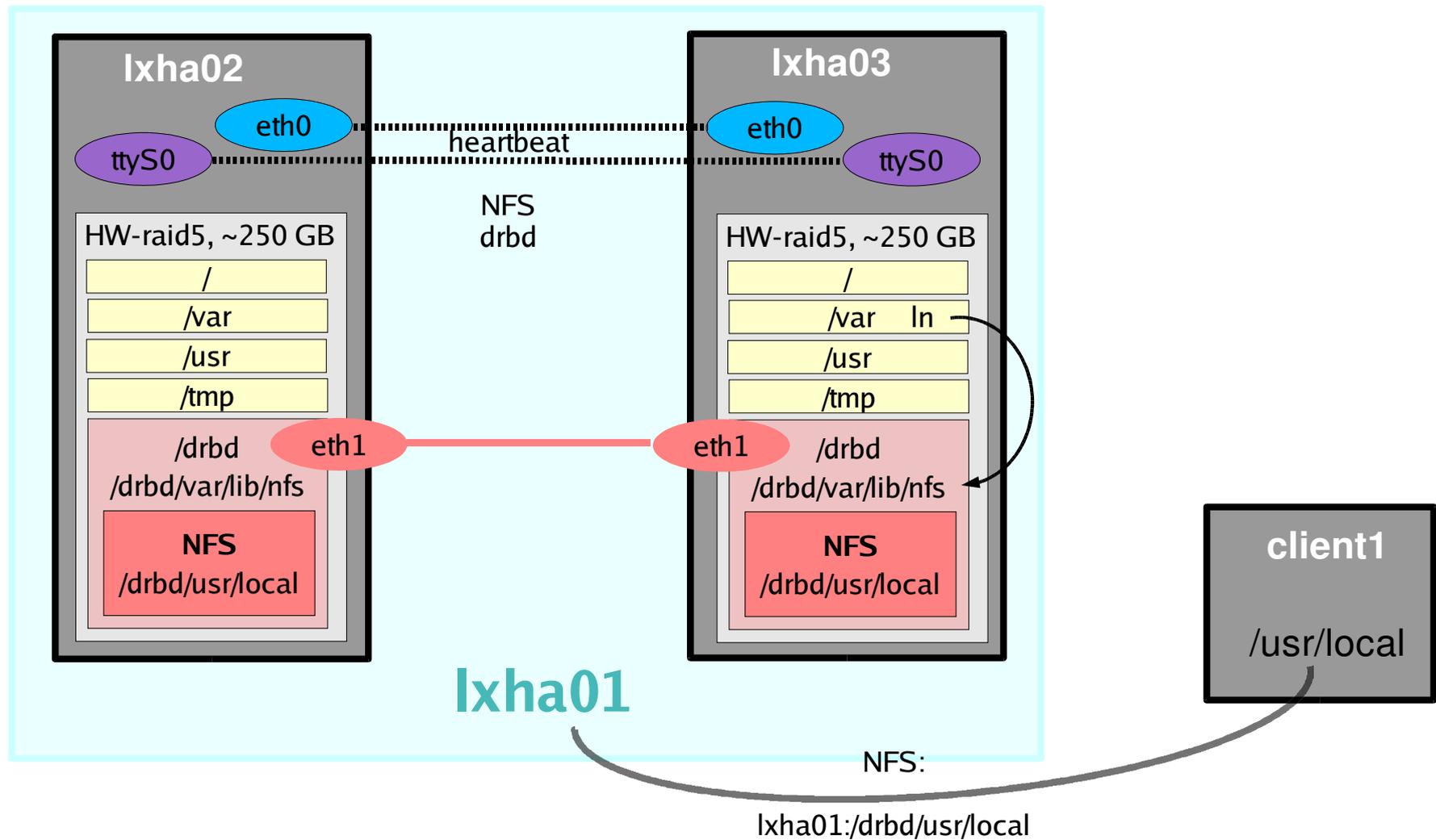
drbd

- Distributed Replicated Block Device
- Kernelpatch für eine Zwischenschicht für Block-Devices
- über diese Schicht Spiegelung von Partitionen über das Netzwerk
- im Prinzip:
 - ➔ ein RAID-1 Verbund über das Netzwerk

Funktionsweise drbd



Implementierung



mon

- Service Monitoring Daemon:
 - Ressourcen-Monitoring
 - Netzwerkverfügbarkeit
 - Serverprobleme
 - Umgebungsvariablen (z. B. Temperatur)
- Überwachung erfolgt mit individuellen Skripten
- Im Fehlerfall wird eine Aktion getriggert (Mail, Restart ...)

Wozu ein Monitoring-Tool für heartbeat?

- Heartbeat prüft, ob der andere Server lebt.
- Es überwacht nicht den Service selbst.
- Mögliche Fehlerfälle:
 - heartbeat fällt aus, der Service nicht
 - beide heartbeat-Verbindungen unterliegen einem Hardwaredefekt
 - Master-Server fällt “teilweise” aus
 - ein Service fällt aus, heartbeat nicht

Monitoring des Systems

