



## Hochverfügbarkeit mit Linux



Dr. Sebastian Hausmann  
Technical Presales Consultant  
HP  
Hamburg

## Beispiele für Bedrohungen







4/9/2003
Hochverfügbarkeit mit Linux
page 3

## Was kostet Ausfallzeit?

Industry	Business Operations	Average Cost per Hour Downtime
<i>Financial</i>	<i>Brokerage operations</i>	<b>\$6.45M</b>
<i>Financial</i>	<i>Credit card</i>	<b>\$2.6M</b>
<i>Media</i>	<i>Pay-per-view</i>	<b>\$150K</b>
<i>Retail</i>	<i>Home catalog sales</i>	<b>\$90K</b>
<i>Transportation</i>	<i>Airline reservations</i>	<b>\$89K</b>
<i>Media</i>	<i>Telesales</i>	<b>\$69K</b>
<i>Healthcare</i>	<i>Patient record</i>	<b>Loss of life</b>

Source: Strategic Research ([www.sresearch.com](http://www.sresearch.com))

**Beispiel:** Kosten: 1000 pro Stunde  
 Zuverlässigkeit: 99,5 % → **Kosten pro Jahr: 43800**

4/9/2003
Hochverfügbarkeit mit Linux
page 4

## Gründe für Business Continuity



- Haftung (nach KonTraG 1998, AktG, GmbHG)  
ggfs. pers. Haftung
- Basel II (ab 2006)  
Differenzierter, ausgefeilter Risikoansatz führt zur Belohnung in Form von niedriger EK-Zuweisung bei Kreditvergabe
- Versicherungsindustrie (Prämien & Versicherbarkeit)
- Kundenanforderungen (z.B. Ausschreibungen)

4/9/2003

Hochverfügbarkeit mit Linux

page 5

## HP Business Continuity Prozess Modell



4/9/2003

Hochverfügbarkeit mit Linux

page 6

## Klassifizierung der Verfügbarkeit

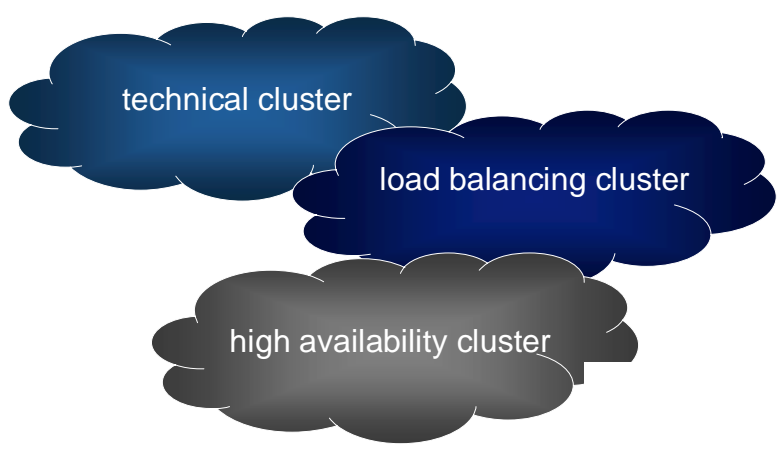
Availability Level	Mögliche Ausfälle	Beispiele
<b>AL4: Fault Tolerant</b>	<b>Keine Unterbrechung bei Fehlern und Reparaturen</b>	<b>Assured Availability HP / Marathon</b>
<b>AL3: Fault Resilient</b>	<b>Verlust von Transaktionen möglich</b>	<b>Oracle Fail Safe</b>
<b>AL2: High Availability</b>	<b>Re-Logon, Neustart, Performanceverlust</b>	<b>HP ServiceGuard MS Cluster</b>
<b>AL1: Data Availability</b>	<b>Shutdown, System nicht verfügbar, Daten</b>	<b>RAID, Datenspiegelung</b>
<b>AL0: Conventional</b>	<b>Server stoppt, Datenverlust, Dateien zerstört</b>	<b>Jeder Server</b>

*Source: Harvard Research Group*

**AL+: Desastertoleranz**      **Wiederherstellung nach Ausfall eines kompletten Rechenzentrums**      **Cluster-Extension**

4/9/2003 Hochverfügbarkeit mit Linux page 7

## Linux-Cluster – nichts Neues...

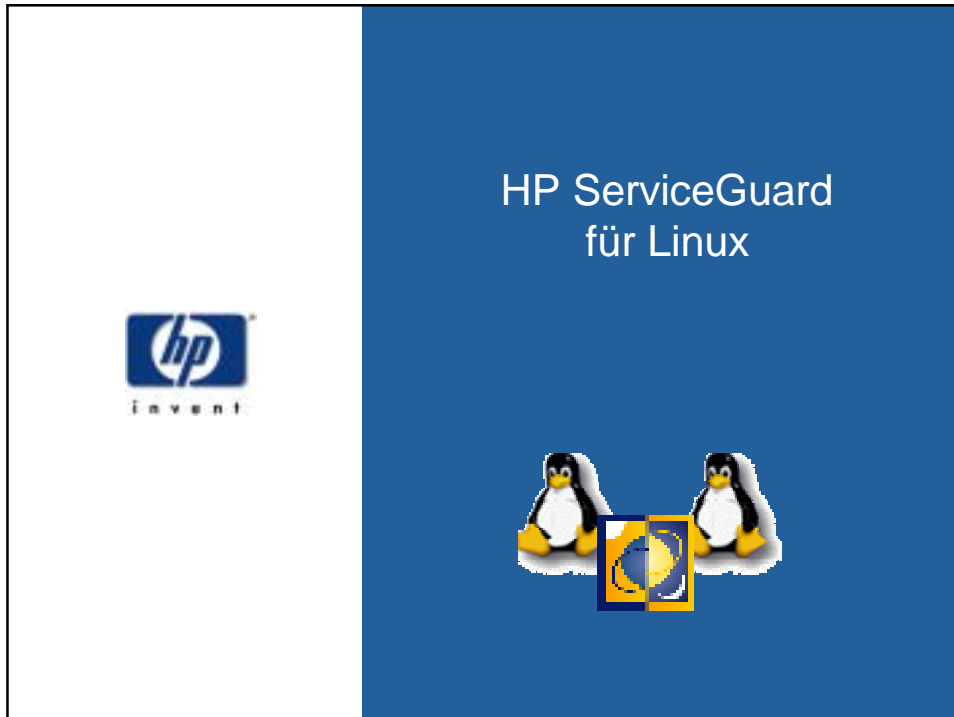


technical cluster



load balancing cluster

high availability cluster


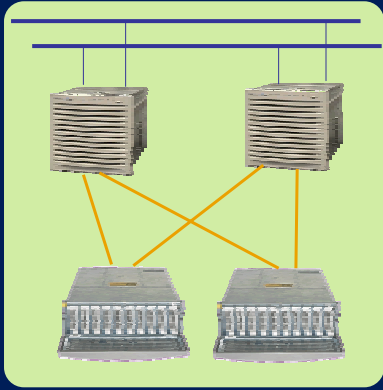
4/9/2003 Hochverfügbarkeit mit Linux page 8



HP ServiceGuard  
für Linux

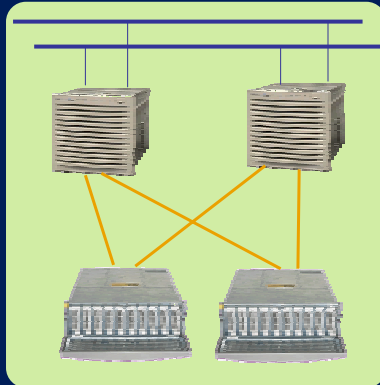
## HP ServiceGuard für Linux Überblick

- Überwachung wichtiger Hardware- und Software-Komponenten
- HP ServiceGuard ermöglicht eine redundante Auslegung der Hardware (**SPOFless**)
- erfordert keine spezielle Hardware, aber...
- es gibt zertifizierte Konfigurationen
- HP ServiceGuard implementiert keine Fehlertoleranz!
- läuft auf Standard-Versionen von HP-UX/Linux
- Integration von Anwendungen ohne Modifikation
- bislang weit über **70.000** verkaufte Lizenzen
- Cluster bis zu einer Größe von **16** Knoten
- Lösungen für lokale, Campus-, Metro- und Continental-Cluster

4/9/2003 Hochverfügbarkeit mit Linux page 10

## MC/ServiceGuard für Linux Überblick



HP ServiceGuard ist in der Lage den Ausfall der folgenden Komponenten zu erkennen und entsprechend zu reagieren:

- System Prozessoren
- System Hauptspeicher
- LAN Medium und LAN Adapter
- Stromversorgung
- Plattenlaufwerke, Volume Gruppen, Logical Volumes, Dateisysteme
- System-Prozesse
- Applikations-Prozesse

Die Fehlererkennung und die Umschaltung der Applikation(en) auf einen gesunden Node geschieht nach definierten Regeln und komplett automatisch. HP ServiceGuard ist in der Lage bei Ausfall einer Komponente je Typs die Verfügbarkeit der Applikationen zu gewährleisten.

4/9/2003

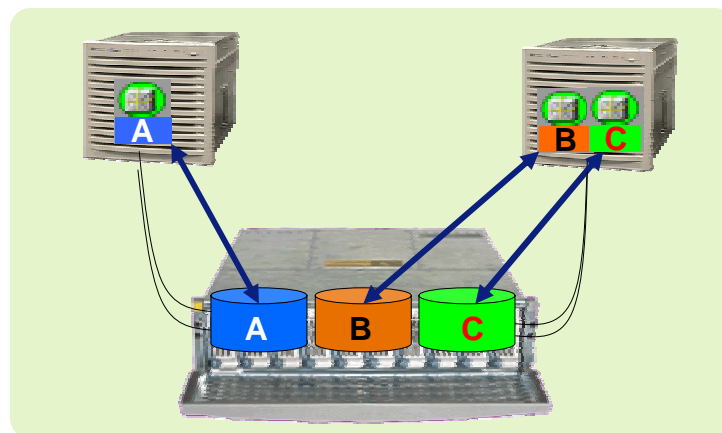
Hochverfügbarkeit mit Linux

page 11

## Plattenkonfiguration



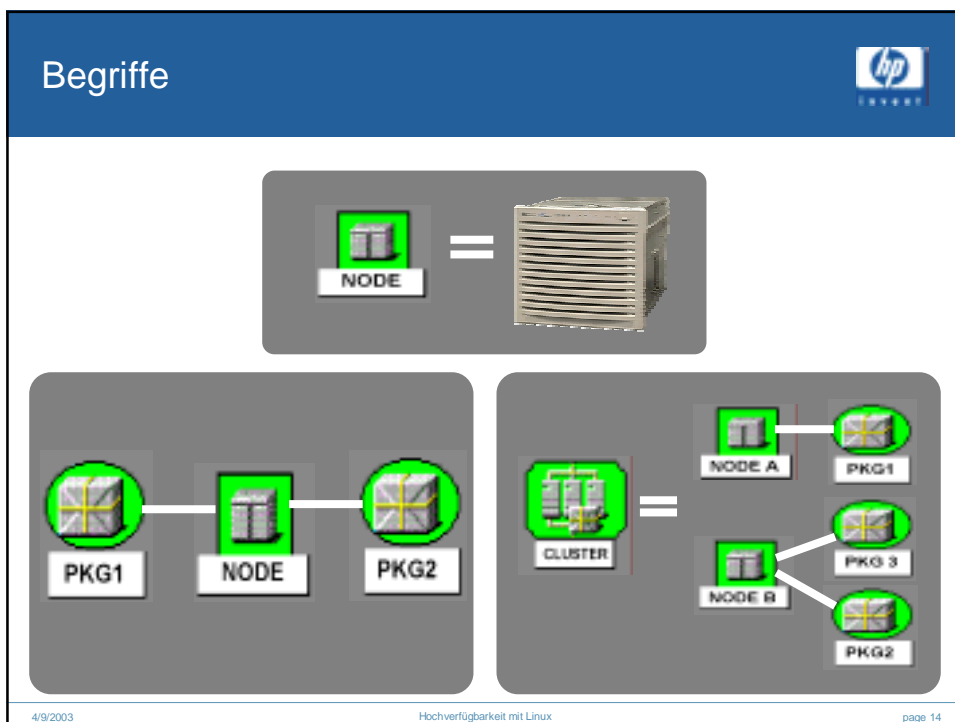
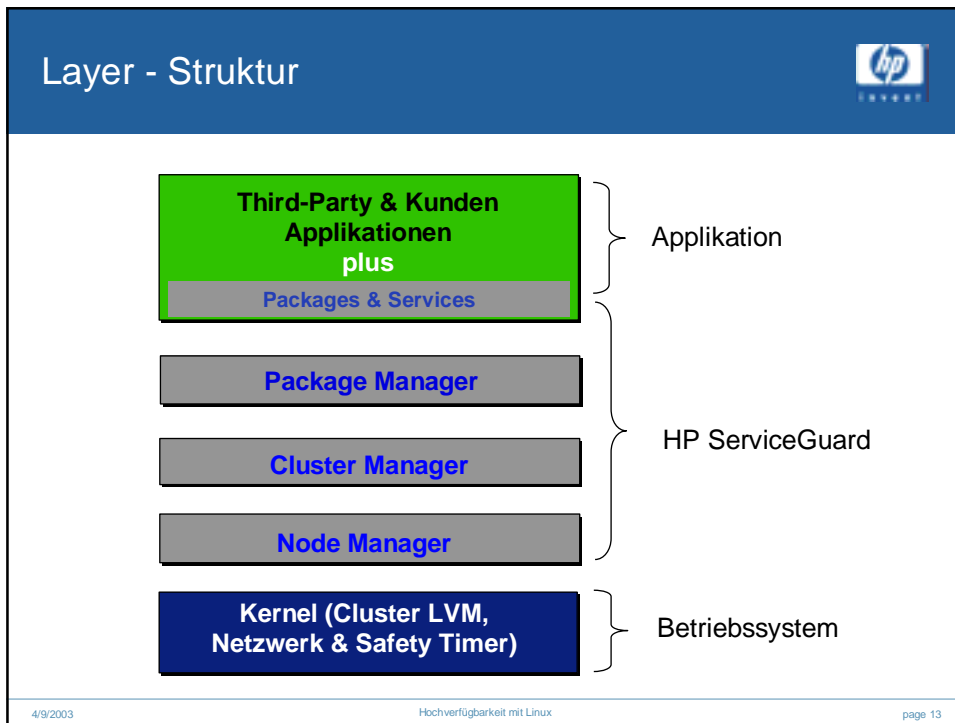
Die **Nodes** eines **Clusters** "sharen" gemeinsame Plattenlaufwerke über Multi-Initiator SCSI-Busse oder über FC-basierende SANs




4/9/2003


Hochverfügbarkeit mit Linux

page 12



## Paket-Definition





Package

**IP Address :**  
15.128.129.130


**Disk space :**  
 Logical Volume    /dev/vg01/lvol1  
                       /dev/vg02/lvol1  
                       /dev/vg02/lvol2

**Processes (über Services):**  
 Application Process 1  
 Application Process 2  
 RDBMS  
 Middleware Process

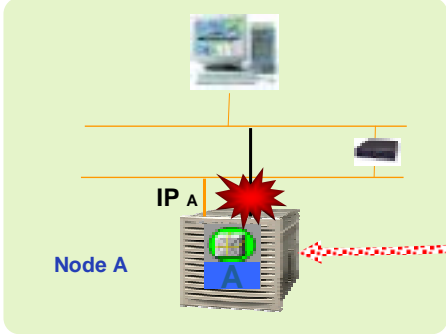
**Start & Stop SCRIPTS**

4/9/2003
Hochverfügbarkeit mit Linux
page 15

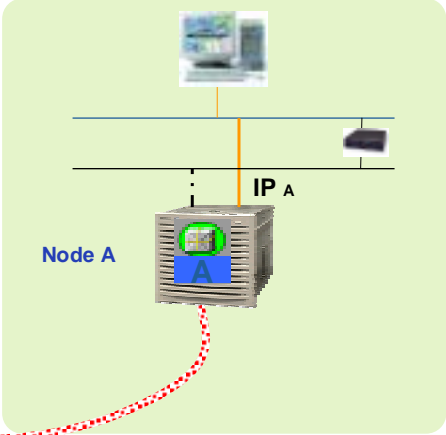
## Lokaler LAN-Ausfall



Transparentes, schnelles **local LAN Failover**. Der TCP/IP Stack wird auf den Standby-Adapter umgeschaltet. Die Clients müssen **kein Reconnect** durchführen.



Node A



Node A


**LAN - Interfaces:**

- 10Base-T
- 100Base-T
- Gigabit Ethernet
- FDDI

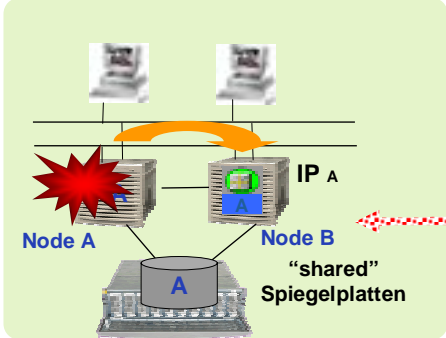
4/9/2003
Hochverfügbarkeit mit Linux
page 16

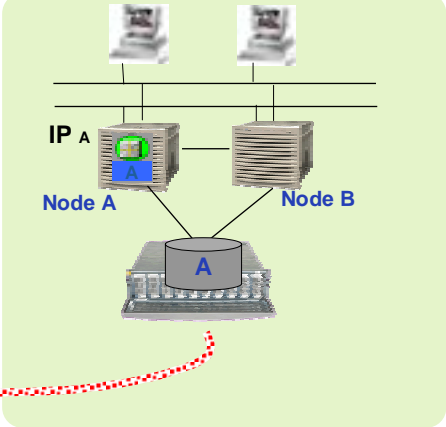


## Paket-Umschaltung


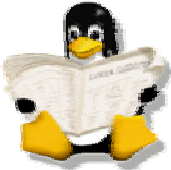


HP ServiceGuard verwendet "relocatable" IP-Adressen und Hostnames. Diese 'virtuelle Hosts' wandern mit dem **Package** auf den anderen **Node**.




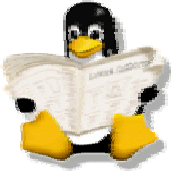


4/9/2003
Hochverfügbarkeit mit Linux
page 17


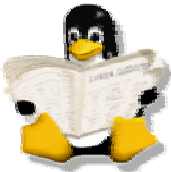
## Design-Ziele

- ein Sourcecode für alle HP SG-Varianten
- langfristig Funktionalität wie HP-UX-Version
- Standard HP-Softwaresupport
- Keine Kernel-Anpassungen  
oder  
Anpassungen unter GPL
- Nutzung vorhandener Open-Source-Software:
  - o Linux LVM
  - o Linux Software RAID
  - o ReiserFS
  - ...


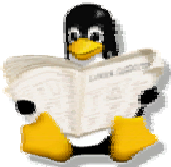
## Eigenschaften


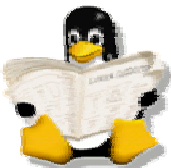
- IA-32 Architektur
- 2 SCSI-Nodes oder bis zu 8 FC-Nodes
- Active / Active
- maximal 150 Packages im Cluster
- maximal 900 Services je Package
- maximal 7 Heartbeat LANs
- "shared" FC/SCSI Plattenlaufwerke
  
- Red Hat Professional 7.3 (SCSI)
- Red Hat Advanced Server 2.1 (FC)
  
- CLI und SGMgr Unterstützung
- Online Rekonfiguration von Nodes, Packages, VGs
- Quorum Server Support für 100 Nodes / 50 Cluster
- Software RAID Support
- Smart Array Support
- HP AutoPath Support


## Applikations-Integrationen

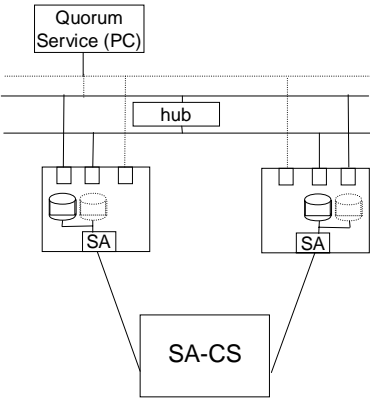
- drei Arten von Anwendungs-Integrationen:
  - offizielle "Produkte"
  - getestete Script-Lösungen
  - Whitepapers
  
- Voraussetzungen (Auszug):
  - Reboot/Powerfail-Resistenz
  - automatische Start-/Stop-Prozeduren
  - keine Abhängigkeiten zu CPU-IDs / MAC-Adressen
  - NFS-Locks vermeiden
  - kein Binding an den Host-Namen
  - feste TCP-Ports und DNS verwenden
  - möglichst keine Daten im Root-Bereich
  - keine lokale Peripherie verwenden
  - Clients brauchen ein Reconnect-Verfahren

  	<h2>unterstützte Hardware</h2> <p><b>Server:</b></p> <ul style="list-style-type: none"> <li>• HP ProLiant DL360 G3</li> <li>• HP ProLiant DL380 G2</li> <li>• HP ProLiant DL380 G3</li> <li>• HP ProLiant DL580 G2</li> </ul> <p><b>Netzwerkkarten:</b></p> <ul style="list-style-type: none"> <li>• integrierter Netzwerkport</li> <li>• NC3134 Fast Ethernet NIC 64 PCI Dual Port 10/100</li> <li>• NC6136 Gigabit SX Server Adapter, PCI 64-bit/66-MHz</li> <li>• NC7770 PCI-X Gigabit Server Adapter</li> </ul> <p><b>Anschluss – Mass-Storage:</b></p> <ul style="list-style-type: none"> <li>• Smart Array 5i and 532</li> <li>• Emulex LP952, LP982, LP8000, LP9000, LP9002</li> <li>• Brocade 2400, 2800, 3240, 3250, 3800, 12000</li> <li>• McData FC64M</li> </ul> <p><b>Mass-Storage:</b></p> <ul style="list-style-type: none"> <li>• Smart Array Cluster Storage</li> <li>• VA7100, VA7400, VA7410</li> <li>• XP 48, 128, 512, 1024</li> </ul> <p style="text-align: right;"><i>Config-Guide beachten!</i></p>
--	---

  	<h2>Lieferumfang</h2> <p>HP Serviceguard for Linux ProLiant Cluster with crossover cable (2-node license) 305199-B21</p> <p>HP Serviceguard for Linux (single license) 307754-B21</p> <p>Two CDs in the product:</p> <ol style="list-style-type: none"> <li>1. Serviceguard Distributed Components CD Serviceguard Manager Software (HP-UX, Windows and Linux) A.03.00 Quorum Server (HP-UX and Linux) A.02.00 and related documentation</li> <li>2. Serviceguard for Linux CD Serviceguard for Linux software A.11.14 Serviceguard for Linux patches (Bonding for Advanced Server 2.1 only) Drivers (Emulex, LVM 1.0, memory allocation library, SG/Mgr API) and related manuals</li> </ol>
--	--

## SCSI Configuration : Support only 2-node clusters






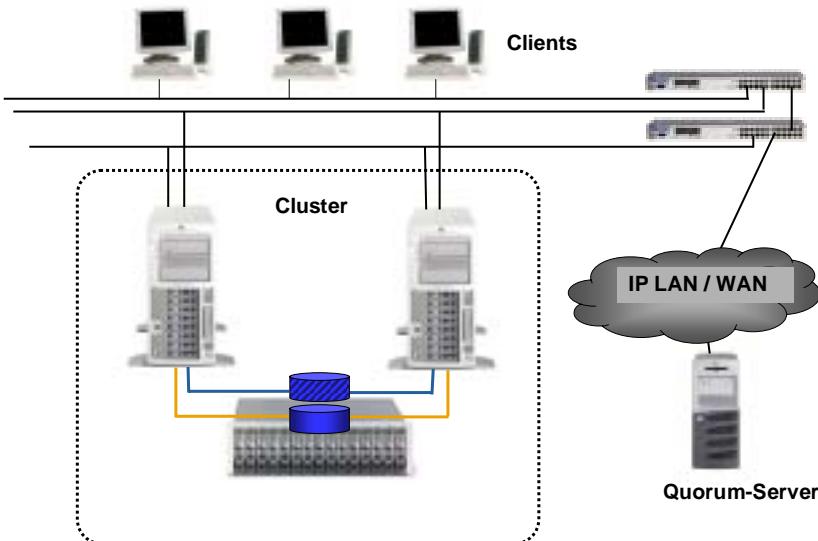
with ProLiant and SA-CS  
on RedHat 7.3 with 2.4.18 kernel

**HW RAID**

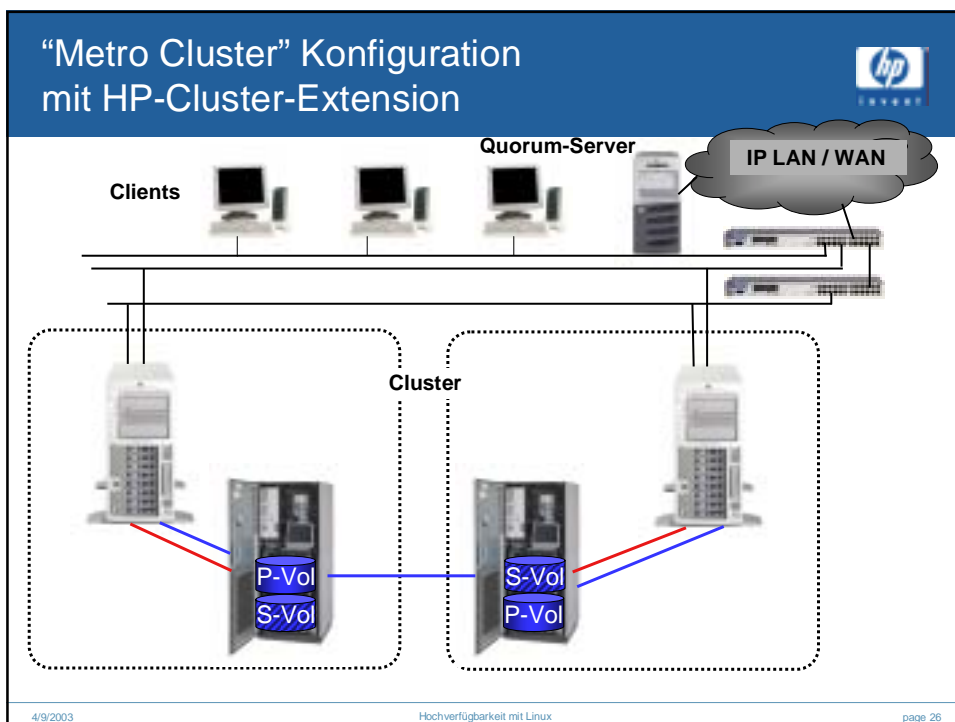
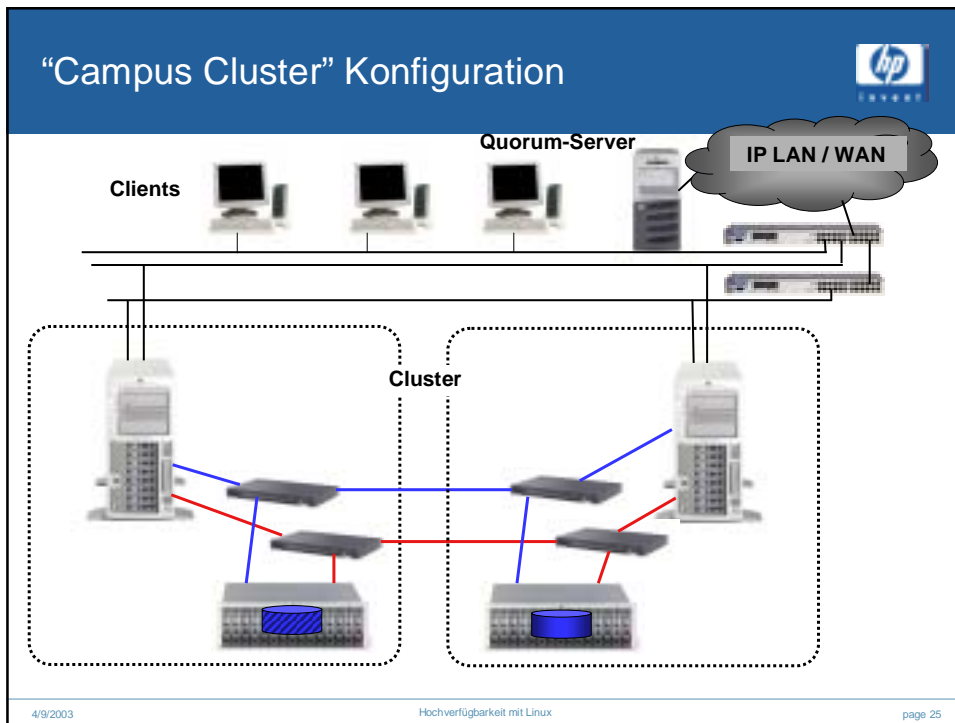
4/9/2003
Hochverfügbarkeit mit Linux
page 23

## “Local Cluster” Konfiguration





4/9/2003
Hochverfügbarkeit mit Linux
page 24



## Die Vorteile von HP Cluster Extension



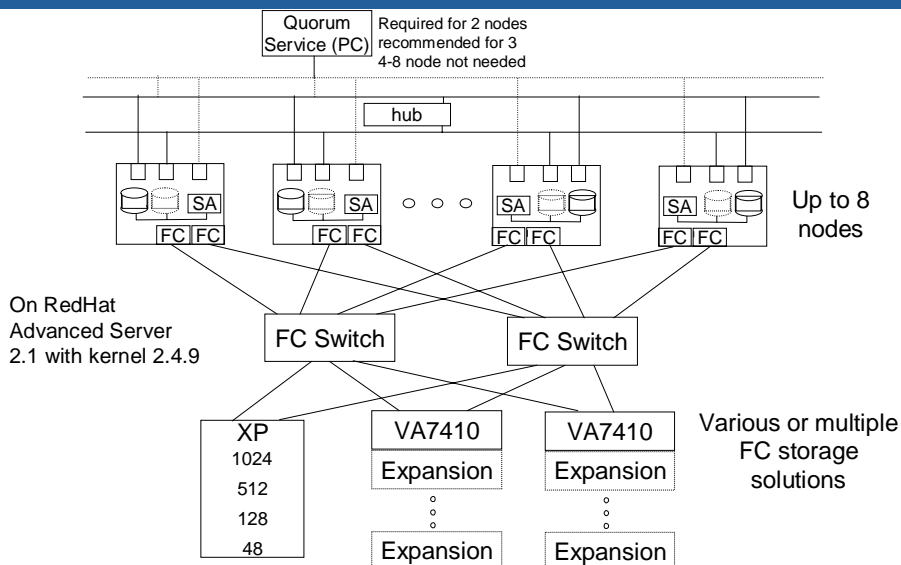
- im Prinzip keine Distanz-Limitationen
- asynchroner Modus für lange Distanzen (keine Signal Latency)
- Load Balancing zwischen den XP Systemen
- bessere Performance als OS Mirroring beim Schreiben
- die Resynchronisation erfolgt immer auf Basis von Deltas (Tracks)
- implementiert ein Quorum für die Verfügbarkeit der Daten
- der Spiegel ist immer "schreibgeschützt"
- man kann einen "schreibbaren" Snapshot der Produktionsdaten erzeugen (z.B. für Wartungsarbeiten)
- eine Technologie für AIX, HP-UX, Windows2000 und Linux Cluster

4/9/2003

Hochverfügbarkeit mit Linux

page 27

## FC Configuration: Support up to 8 nodes



4/9/2003

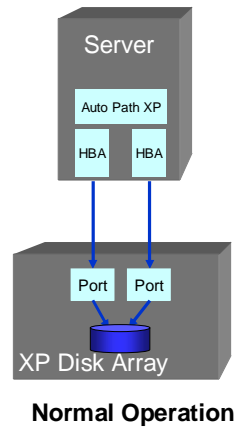
Hochverfügbarkeit mit Linux

page 28

## Autopath



- Was macht HP AutoPath?
  - Stellt Path Failover und Failback (Recovery) zwischen einem Linux Server und HP XP/VA Plattensystemen zur Verfügung
  - Dynamisches Loadbalancing über verschiedene Pfade
- Wie ist dies implementiert?
  - Der HP AutoPath Treiber ist ein “Dynamically Loadable Module” welches den IO über die Pfade steuert
  - Der Treiber stellt die Summe aller Pfade als “virtuelles Block Device” transparent dar (/dev/AP[a-z])



4/9/2003

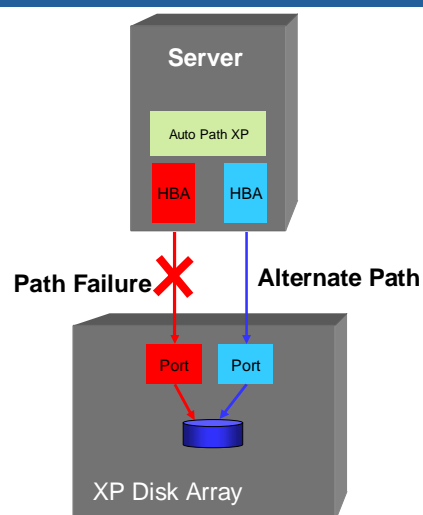
Hochverfügbarkeit mit Linux

page 29

## Autopath



- der Failover ist transparent für die Applikation
- unterstützt Fibre Channel und Fibre Channel SANs
- fällt ein Pfad aus, verwendet AutoPath weiterhin die verbliebenen Pfade und die definierte Policy
- AutoPath “reaktiviert” automatisch sobald der ausgefallene Pfad wieder zur Verfügung steht.



4/9/2003

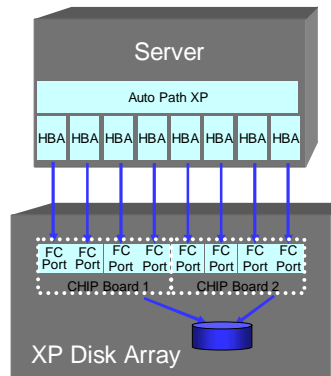
Hochverfügbarkeit mit Linux

page 30

## Load-Balancing



- maximal 8 Pfade zu einer Platte (LUN)
- Dynamisches Load Balancing läuft über alle installierten Adapter (HBAs)
- 4 Load Balancing Policies:
  - "no load balancing"
  - "round robin"
  - "shortest queue first"
  - "shortest response time first"
- unterstützt derzeit alle HP XP und HP VA Disk Arrays
- Load Balancing Support in Cluster-Umgebungen (HP MC/ServiceGuard)



4/9/2003

Hochverfügbarkeit mit Linux

page 31

## User-Interface



### Service Guard Manager:

- Java GUI
- uses cluster object manager, no snmp (as ClusterView)
- currently only for monitoring
- runs on all Java supported platforms (1.3.0 release req.)



### command line interface:

```
cmapplyconf, cmcheckconf, cmdeleteconf, cmgetconf,
cmhaltcl, cmhaltnode, cmhaltpkg, cmhaltserv,
cmmakepkg, cmmodnet, cmmodpkg, cmquerycl, cmreadlog,
cmreadlog, cmruncl, cmrunnode, cmrunpkg, cmrunserv,
cmscancl, cmviewcl, cmviewconf
```



4/9/2003

Hochverfügbarkeit mit Linux

page 32



## Quorum-Server



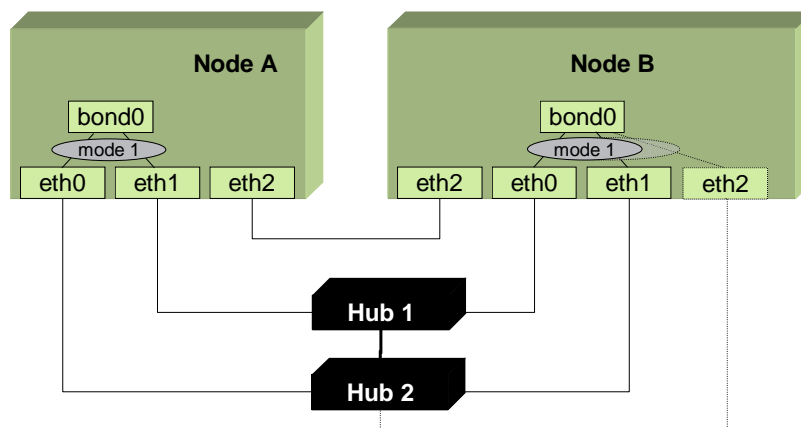
- Beim Start und im Fehlerfall regelt der QS die Initialisierung des Clusters.
- Der QS läuft nicht auf einem Cluster-Knoten.
- QS benutzt TCP/IP und wartet auf Anfragen von ServiceGuard-Knoten.
- Das erste Release unterstützt nur eine IP-Adresse für den QS.
- Das erste Release unterstützt nur einen QS pro Cluster.
- Der QS-Prozess läuft mit hoher Priorität (`nice -20`).
- Auf QS Maschinen können andere Applikationen laufen.
- QS kann als SG-Paket auf einem anderen Cluster laufen.
- Ein QS kann für mehrere Cluster genutzt werden:  
Maximal 5 Cluster und maximal 20 Knoten.
- Der QS wird für die Cluster-Konfiguration benötigt (`cmapplyconf`).
- Der QS wird für die Cluster-Neubildung benötigt, falls verbleibende Knoten keinen Quorum bilden können ( $\leq 50\%$  der Knoten).
- **kein SPOF**

4/9/2003

Hochverfügbarkeit mit Linux

page 33

## Bonding




4/9/2003

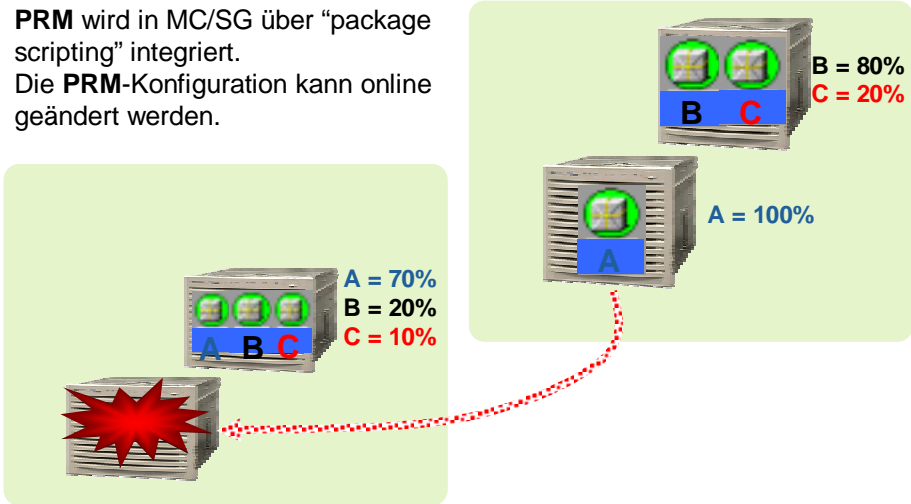
Hochverfügbarkeit mit Linux

page 34

## Process-Resource-Manager



**PRM** wird in MC/SG über "package scripting" integriert.  
Die **PRM**-Konfiguration kann online geändert werden.



4/9/2003 Hochverfügbarkeit mit Linux page 35

## Steeleye Lifekeeper




## Lifekeeper: fast wie Serviceguard



- Active/Active and Active/Passive Configurations up to **16 Nodes**
- N+1 and Cascading Failover
- Local restart avoids failover
- No Interaction needed
- Provides multiple paths to data to insure data availability
- Manual failover for Maintenance
- Replicated Storage clusters (using internal disks)
- Many Application-Recovery-Kits (ARK)
- Open Source Software Development Kit for custom Applications
- Java based GUI for Administration
- Support for Online-extension without Interruption

4/9/2003

Hochverfügbarkeit mit Linux

page 37

## unterstützte Hardware

### Server:

- HP Proliant ML350 G2/G3
- HP Proliant ML370 G2/G3
- HP Proliant DL360 G2/G3
- HP Proliant DL380 G2/G3
- HP Proliant DL530 G1/G2
- HP Proliant DL580 G1/G2

### Netzwerkkarten:

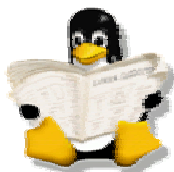
- Standardkarten

### Anschluss – Mass-Storage:

- Smart Array 532 Controller
- Smart Array 5i (Plus) Controller
- Fibre Channel Host Adapter Kit FCA2214


### Mass-Storage:

- RA4100
- MSA1000
- MA8000



*Quickspecs beachten!*

## Application Recovery Kits



Linux	NT & Windows 2000
<ul style="list-style-type: none"> <li>• Apache Web Server (SSL)</li> <li>• Sendmail</li> <li>• Oracle 8i,9i</li> <li>• Informix</li> <li>• MySQL</li> <li>• DB2 (WE/EE/EEE)</li> <li>• Print Services</li> <li>• IP Recovery</li> <li>• NFS Server</li> <li>• NAS Recovery</li> <li>• SAP R/3</li> <li>• Samba</li> </ul>	<ul style="list-style-type: none"> <li>• Microsoft IIS</li> <li>• Microsoft Exchange</li> <li>• Oracle 8i,9i</li> <li>• Informix</li> <li>• Microsoft SQL Server</li> <li>• Print Services</li> <li>• IP Recovery</li> </ul>

4/9/2003 • LVM Hochverfügbarkeit mit Linux page 39

## Zusammenfassung





## HP ServiceGuard und Steeleye Lifekeeper

### LifeKeeper recommended when:

- customer already has LifeKeeper installed
- customer requires broad hardware support
- SuSE or Red Hat is preferred distribution
- ra4100, msa1000, ma8000, or non-hp (e.g. emc) storage required
- mixed Windows/Linux environment with common HA look and feel

### ServiceGuard recommended when:

- customer already has ServiceGuard for hpux installed
- customer requires single vendor hardware/software solution
- Red Hat is preferred distribution
- VA and XP storage will be deployed in the Linux environment
- mixed HPUX/Linux environment with common HA look and feel

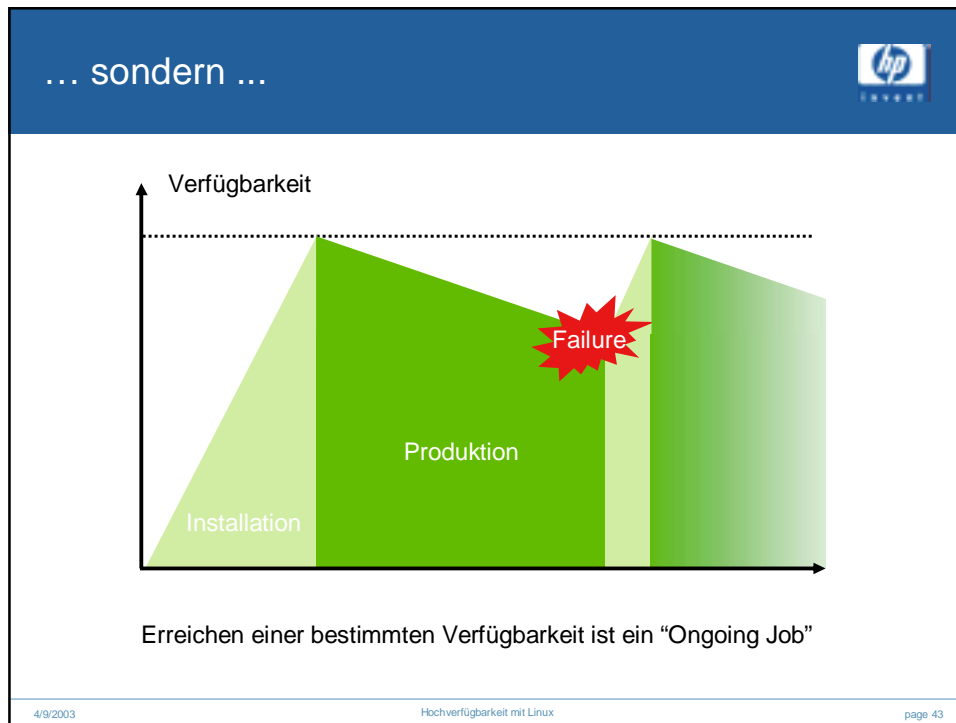
4/9/2003
Hochverfügbarkeit mit Linux
page 41

## Verfügbarkeit ist nicht nur Technologie

Quelle: GartnerGroup  
October 1999

Verfügbarkeit	Ausfallzeit in Std/Jahr
99%	87,6
99,9%	8,76
99,95%	4,38
99,99%	0,876
<b>99,999%</b>	<b>0,0876 (= ca 5 Minuten)</b>

4/9/2003
Hochverfügbarkeit mit Linux
page 42



... braucht noch Dinge "on top"

- **Training, Consulting und Support**
  - HP Education Services
  - HP Consulting
  - HP Support Services (bis zum Mission Critical Support)
- **konsequentes System Management**
  - HP Service Control Manager
  - HP Linux Multi-System Management
- **Datensicherungsverfahren**
  - HP OmniBack
- **Überwachung der Infrastruktur**
  - HP OpenView Operations
  - HP TopTools für Linux

4/9/2003 Hochverfügbarkeit mit Linux page 44

## Service and Support



H4810T:

MC/ServiceGuard Implementation Service (for Linux)

H4400A/H4402A/H4404A Option 6A8:

s/w updates with Next Day/Same Day/24x365 h/w support

H4401A Option 6A8:

8x5 phone-in support plus software updates

H4405A Option 6A8:

Personalized Support Services (PSS)

24x365 phone-in support plus software updates

H4385A Option 6A8:

Critical Support Services (CSS)

24x365 phone-in support plus software updates

4/9/2003

Hochverfügbarkeit mit Linux

page 45

## ... und noch ein paar URLs



**Linux bei HP:**

<http://www.hp.com/go/linux>

**MC/ServiceGuard:**

<http://h18000.www1.hp.com/solutions/enterprise/highavailability/linux/serviceguard>

**Hochverfügbarkeit:**

<http://www.hp.com/products1/unix/highavailability>

**Process-Resource-Manager:**

[http://resourcemanagement.unixsolutions.hp.com/WaRM/prm\\_linux/index.html](http://resourcemanagement.unixsolutions.hp.com/WaRM/prm_linux/index.html)

4/9/2003

Hochverfügbarkeit mit Linux

page 46

