



Hochverfügbarkeit  
mit Linux

**Dr. Sebastian Hausmann**  
Technical Consultant, HP



© 2004 Hewlett-Packard Development Company, L.P.  
The information contained herein is subject to change without notice

The slide features a blue background on the left and a dark blue background on the right. The HP logo is in the top left. The title 'Hochverfügbarkeit mit Linux' is in white text. The speaker's name and title are in black text. Tux the penguin is in the bottom center. The HP logo with a plus sign is in the bottom right. A copyright notice is at the bottom left.



## Agenda:

Warum Hochverfügbarkeit?

HP ServiceGuard für Linux

Steeleye Lifekeeper





## Beispiele für Bedrohungen



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Was kostet Ausfallzeit?

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

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

**Beispiel:** Kosten: 1000 € pro Stunde  
 Ausfallzeit: Ø 12 Std. (15 Min. – 72 Std.) pro Jahr → **Kosten pro Jahr: Ø 12000 €**  
 (250 – 72000 €)  
 Quelle: Handelskammer Hamburg

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Weitere Gründe für Hochverfügbarkeit

- Haftung (nach KonTraG 1998, AktG, GmbHG)  
ggfs. persönliche 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)

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## HP Business Continuity Prozessmodell

Analyse	Design	Implementierung	Betrieb
Einführungsworkshop Prüfung bestehender Maßnahmen Risikoanalyse Business Impact Analyse Infrastrukturanalyse	Lösungserarbeitung Entscheidungsvorlage	Realisierung der techn. Absicherung Notfallrahmenplanung Alarmplanung Wiederanlaufplanung Datensicherungskonzept Dokumentationserstellung	Übungen Schulung Pflegeservice
 Anforderungskatalog	 Absicherungskonzept	 Notfallhandbuch	

Projekt-Management

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Klassifizierung der Verfügbarkeit

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

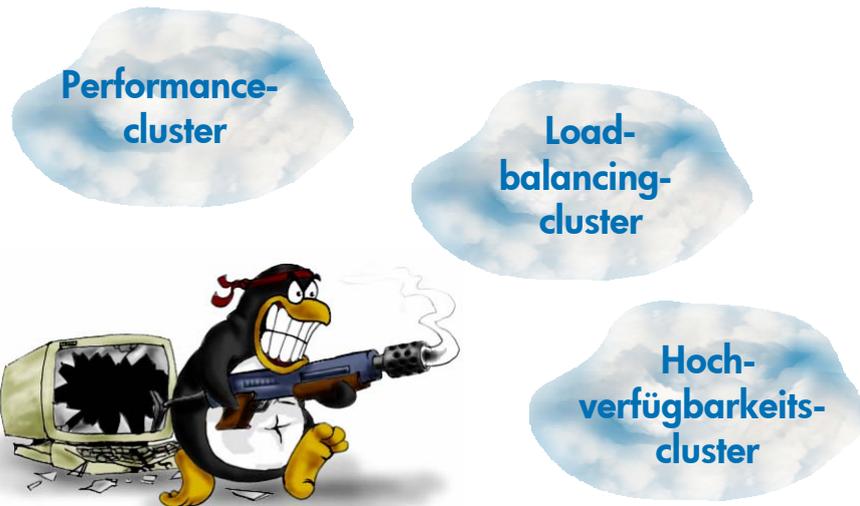
*Source: Harvard Research Group*

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

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Linux und Cluster



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004

... what's going on outside ?



**HA HighAvailability**  
High Availability Linux Project

see <http://linux-ha.org/>

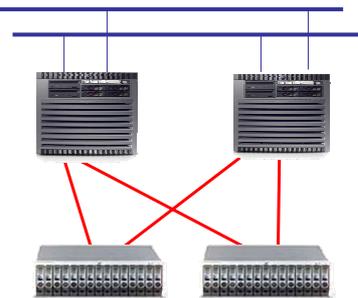
Hochverfügbarkeit mit Linux, DECUS-Symposium 2004





## HP ServiceGuard

- Nicht fehlertolerant, kein Loadbalancing
- Überwacht Hardware und Software
- Erfordert redundante Hardware (no SPOF)
- Läuft auf Standard-Hardware und -Betriebssystem
- Integration von Applikationen ohne Modifikationen
- Mehr als 80.000 Lizenzen verkauft (HP-UX & Linux)
- Unterstützt lokale, Campus-, Metropolitan- und kontinentale Konfigurationen

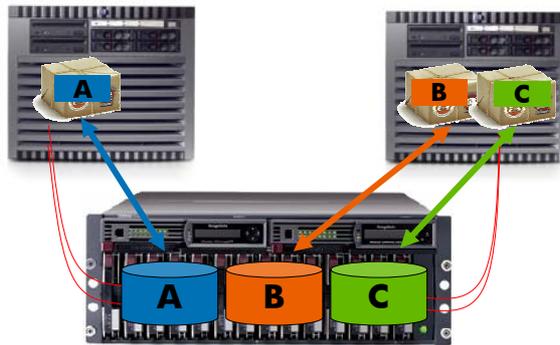


Hochverfügbarkeit mit Linux, DECUS-Symposium 2004

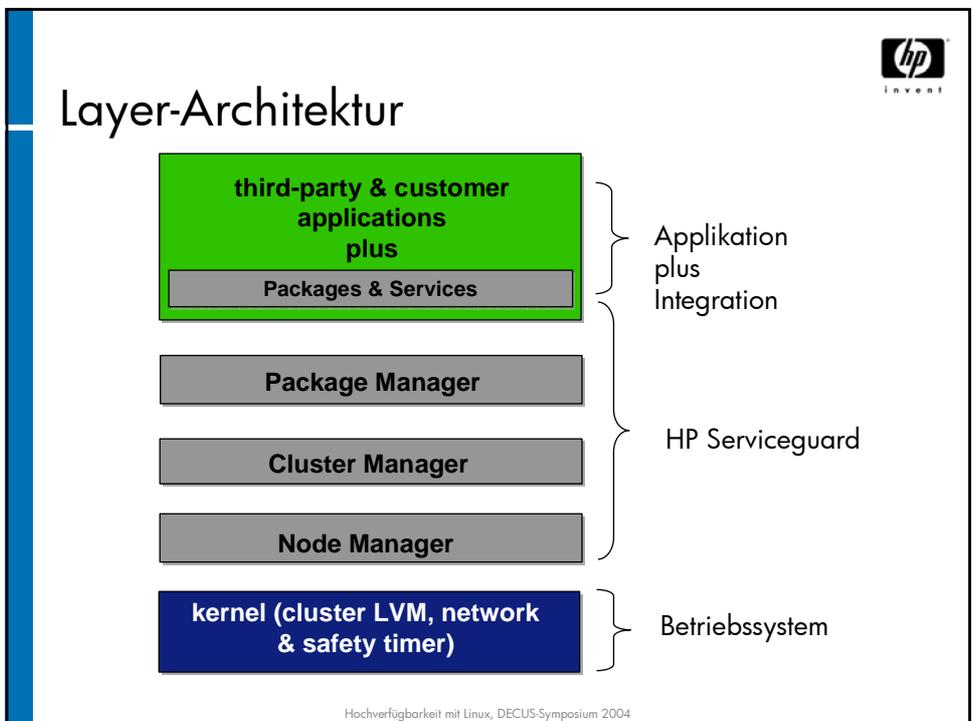
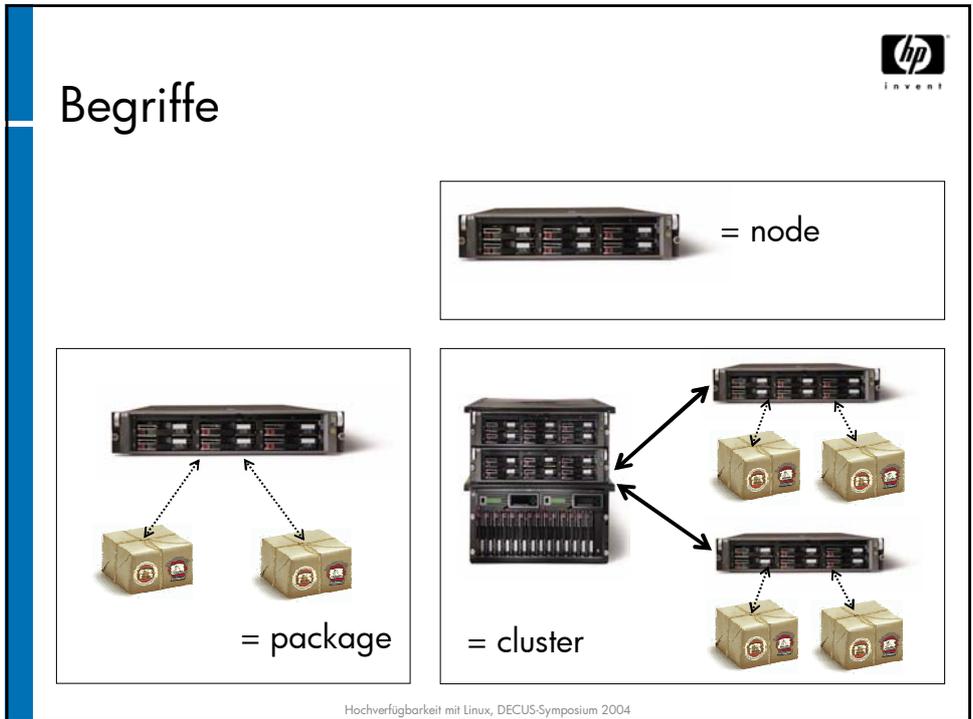


## „Shared Storage“

Die **Knoten** eines **Clusters** „sharen“ gemeinsame Plattenlaufwerke über Multi-Initiator-SCSI-Busse oder über FC-basierende SANs



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004





## Paket

Ein Paket definiert alle Ressourcen, die zum Betrieb einer Applikation auf einem Cluster-Knoten notwendig sind.



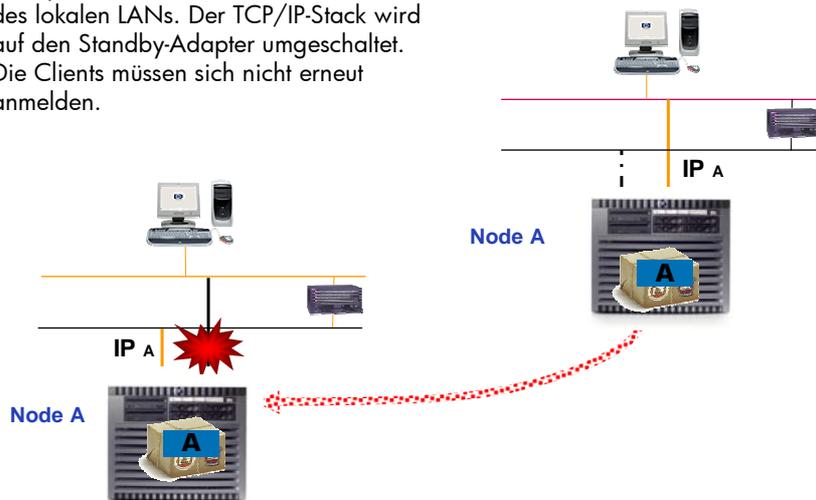
- IP Address :**  
15.128.129.130 on subnet 15.136.129.0
- disk space :**  
VG=vg01 LV=/dev/vg01/lvol1 FS=/appl1  
LV=/dev/vg02/lvol1 FS=/appl2
- processes (via services):**  
application process 1  
application process 2  
RDBMS  
middleware processes
- start & stop scripts**

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Lokaler LAN-Ausfall

Transparentes und schnelles Umschalten des lokalen LANs. Der TCP/IP-Stack wird auf den Standby-Adapter umgeschaltet. Die Clients müssen sich nicht erneut anmelden.

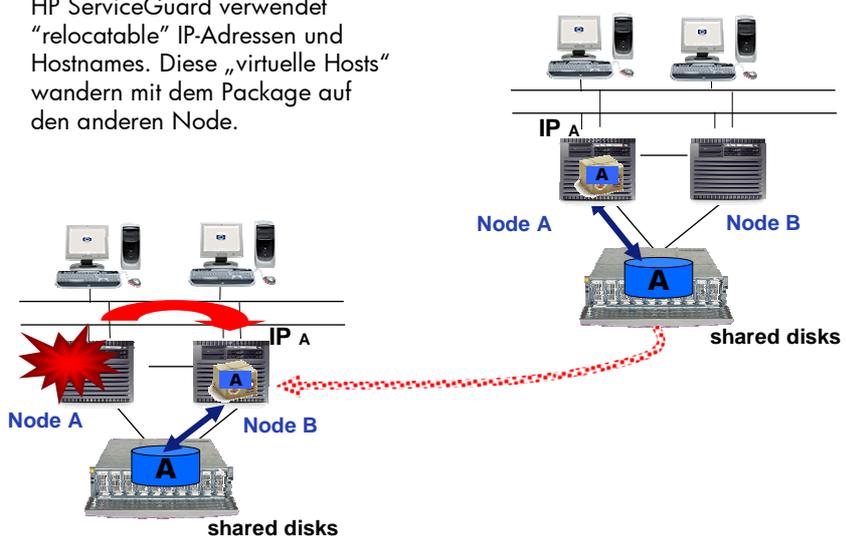


Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Paket-Umschaltung

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



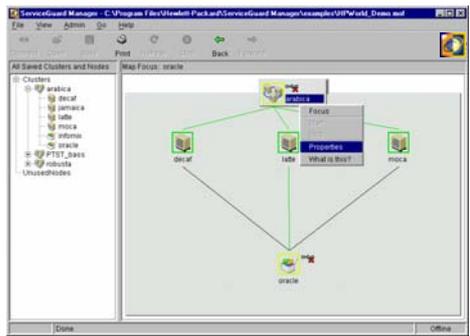
Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Management

### Serviceguard-Manager:

- Java GUI
- uses cluster object manager (no SNMP)
- supports many platforms
- monitoring and **controlling**



### command line interface:

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



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Designziele

- 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:
  - Linux LVM
  - Linux Software RAID
  - ReiserFS
  - ...

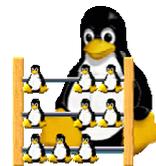


Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Eigenschaften der Version A.11.15

- IA-32- und IA-64-Architektur
- 2 SCSI-Nodes oder bis zu 16 FC-Nodes
- Active / Active
- maximal 150 Packages im Cluster
- maximal 900 Services je Cluster
- maximal 200 virtuelle IP-Adressen
- maximal 7 Heartbeat LANs
- "shared" FC/SCSI Plattenlaufwerke
- SUSE Linux Enterprise 8 (UL 1.0)
- Red Hat Enterprise Linux AS 3 
- CLI und SGMgr Unterstützung
- Online Rekonfiguration von Nodes, Packages, VGs
- Quorum Server Support für 100 Nodes / 50 Cluster



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Unterstützte Distributionen

**A.11.14:**

- RedHat AS 2.1 with kernel 2.4.9-e3/e25/e27 (SCSI & FC)
- Redhat 7.3 with kernel 2.4.18 (SCSI only)
- SuSE ES8 UL 1.0 with kernel 2.4.19 (SCSI & FC)
- Linux LVM for Kernel (patches supplied for RedHat)
- ext2, ext3 Filesystem and ReiserFS

**A.11.15:**

- **RedHat EL AS 3.0 (SCSI & FC) April 2004**
- SuSE ES8 UL 1.0 with kernel 2.4.19 of SP2a (SCSI & FC)
- SuSE ES8 UL 1.0 with kernel 2.4.21 of SP3 (IPF & FC only)
- Linux LVM for kernel
- ext2, ext3 Filesystem and ReiserFS


Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Unterstützte Hardware

<p><b>Server:</b></p> <ul style="list-style-type: none"><li>• HP ProLiant DL360 G3</li><li>• HP ProLiant DL380 G2/G3 (auch Packaged-Cluster)</li><li>• HP ProLiant DL560</li><li>• HP ProLiant DL580 G2</li><li>• HP ProLiant ML 350 G3</li><li>• HP ProLiant ML 370 G3</li><li>• HP ProLiant BL20p</li><li>• HP ProLiant BL40P</li><li>• HP Integrity rx2600</li><li>• HP Integrity rx5670</li></ul>	<p><b>Storage:</b></p> <ul style="list-style-type: none"><li>• XP48, XP128, XP512, XP1024</li><li>• VA7x00</li><li>• EVA3000/5000 (nur mit SecurePath, 2 Knoten)</li><li>• MSA1000 (nur mit SecurePath, 2 Knoten)</li><li>• external SCSI storage (nur ProLiant)</li></ul> <p><b>HBAs:</b></p> <ul style="list-style-type: none"><li>• supported Proliant SCSI HBAs</li><li>• A6826A for IPF</li><li>• FCA2214 for IA32</li><li>• Emulex for existing configs</li></ul>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Config-Guide beachten!**  
<ftp://ftp.compaq.com/pub/solutions/enterprise/ha/linux/svcguard-certmatrix.pdf>

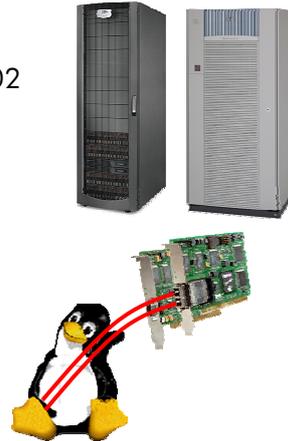


Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## multipathing support

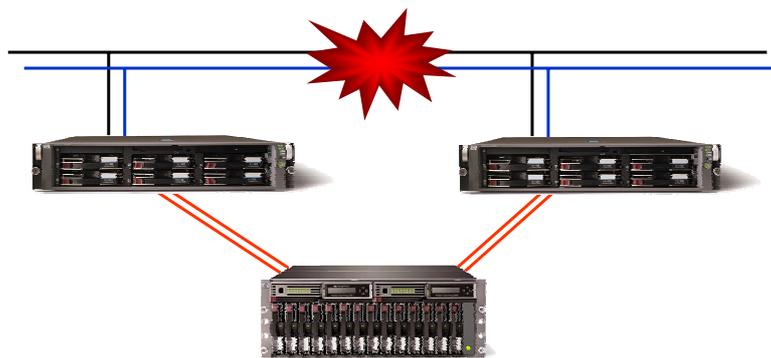
- hp Autopath for Linux for kernel 2.4.2-2 only (A.11.13)
- raidtools multipathing is supported for
  - VA
  - XP
- support for SecurePath 3.0A with A.11.14.02
  - EVA3000 (2 nodes only)
  - EVA5000 (2 nodes only)
  - MSA (2 nodes only)
- new Qlogic multipath driver for
  - XP disk arrays
  - VA disk arrays



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## „Split Brain“-Situation



Die Clusterservices können nicht mehr miteinander kommunizieren  
beide Knoten versuchen, alle Cluster-Pakete zu übernehmen

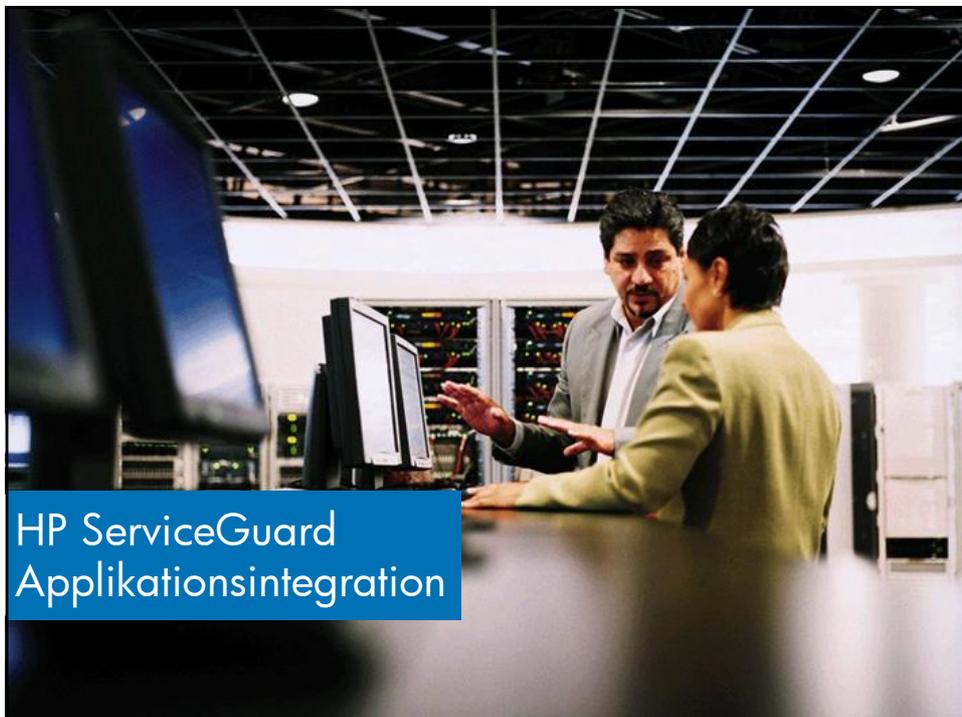
Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Split-Brain → 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 mit Port 1238 auf Anfragen von ServiceGuard-Knoten.
- Nur eine IP-Adresse für den QS möglich.
- Nur ein QS pro Cluster möglich.
- Auf QS Maschinen können andere Applikationen laufen.
- QS kann als ServiceGuard-Paket auf einem anderen Cluster laufen.
- Ein QS kann für mehrere Cluster genutzt werden:  
Maximal 50 Cluster und maximal 100 Knoten.
- Der QS wird für die Cluster-Konfiguration benötigt (cmapplyconf).
- Der QS wird für die Cluster-Neubildung benötigt, falls verbleibende Knoten kein Quorum bilden können ( $\leq 50\%$  der Knoten).
- Also ist ein QS notwendig für einen 2-Knoten-Cluster und optional für 3 bis 16 Knoten.
- Ab Version **A.11.15** gibt es auch eine „Cluster-Lock-LUN“ für Cluster mit **2 bis 4** Knoten.
- Die Cluster-Lock-LUN ist eine „shared LUN“ mit einer festen Partitionsgröße von 100kB.
- **kein SPOF**

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004





## Applikations-Integrationen

- drei Arten von Applikations-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



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Applikationsintegration

- **Produkte:**
  - HP ServiceGuard Extension for SAP for Linux
  - HP ServiceGuard for Linux Oracle database toolkit
- **Toolkits:**
  - Apache
  - MySQL
  - NFS
  - Samba
  - SendMail

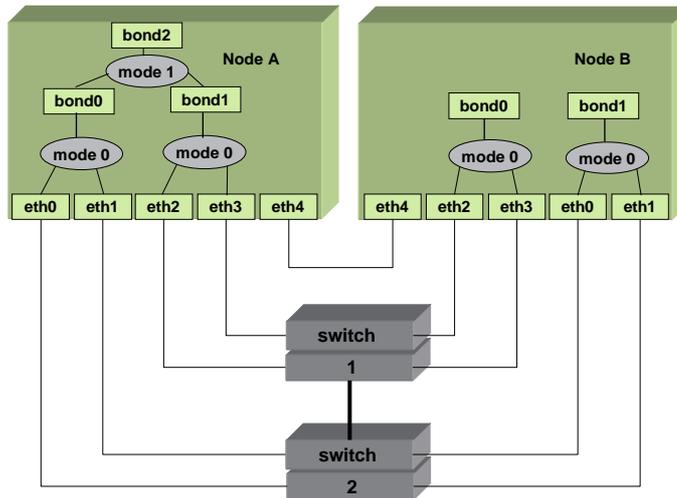
kostenlos runterzuladen:  
<http://www.software.hp.com>



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



# „Bonding Driver“



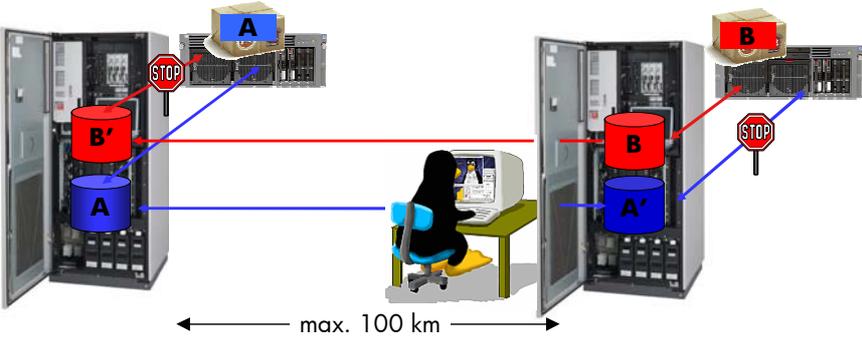
Hochverfügbarkeit mit Linux, DECUS-Symposium 2004





## Continuous Access XP

- Unterstützt synchronen und asynchronen CA-Betrieb
- Aktiv-Aktiv-Konfiguration
- Datenspiegelung durch Festplattensystem transparent für Cluster-Package-Manager



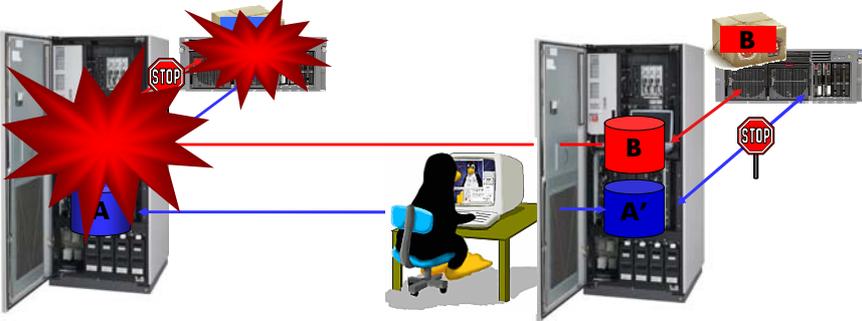
← max. 100 km

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Continuous Access XP

- Selbst nach Zerstörung eines Rechenzentrums laufen die Applikationen weiter.



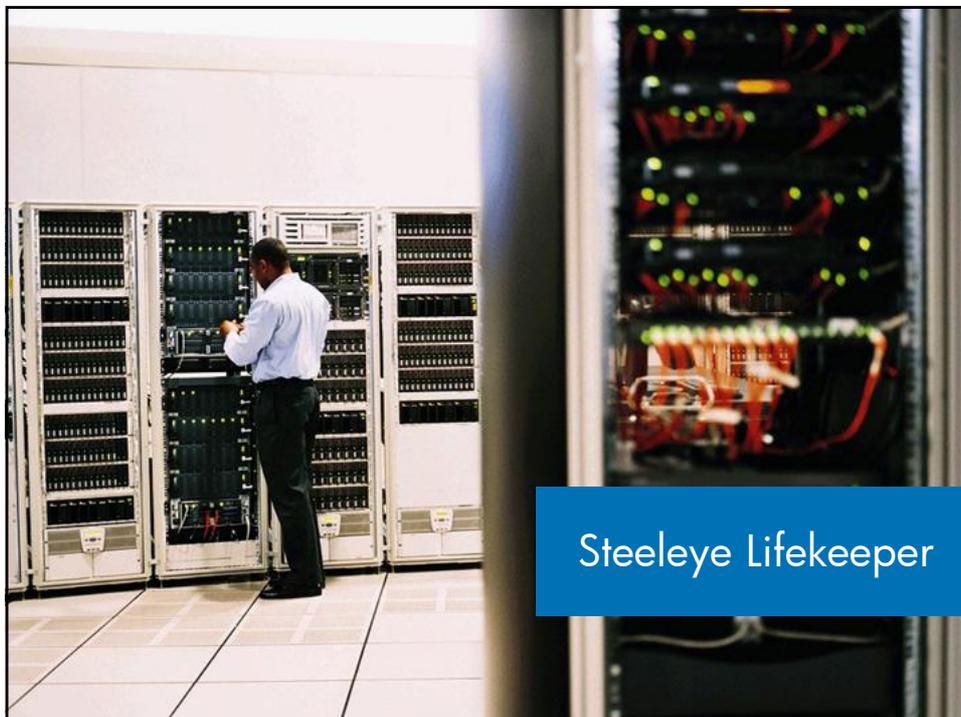
Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## 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

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004





## 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



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## 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
- HP Proliant BL20p / BL40p

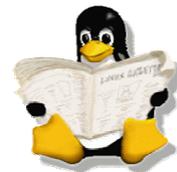
### Anschluss – Mass-Storage:

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

### Mass-Storage:

- RA4100
- MSA1000
- MA8000

Quickspecs beachten!



Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



## Application Recovery Kits

<p><b>Linux</b></p> <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><li>• LVM</li></ul>	<p><b>NT &amp; Windows 2000</b></p> <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>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004



# Zusammenfassung



**LifeKeeper™**  
for Linux®



## HP ServiceGuard und Steeleye Lifekeeper



### LifeKeeper falls:

- LifeKeeper schon im Einsatz ist.
- ra4100, ma8000 oder Nicht-HP-Storage (z.B. EMC) zum Einsatz kommen soll.
- gemischte Windows/Linux-Umgebung mit gemeinsamen "HA look and feel".

### ServiceGuard falls:

- ServiceGuard schon im Einsatz ist (auch für HP-UX).
- eine Lösung aus einer Hand gewünscht ist.
- die Festplattensysteme VA oder XP zum Einsatz kommen sollen.
- gemischte HP-UX/Linux-Umgebung mit gemeinsamen "HA look and feel".

Hochverfügbarkeit mit Linux, DECUS-Symposium 2004

