



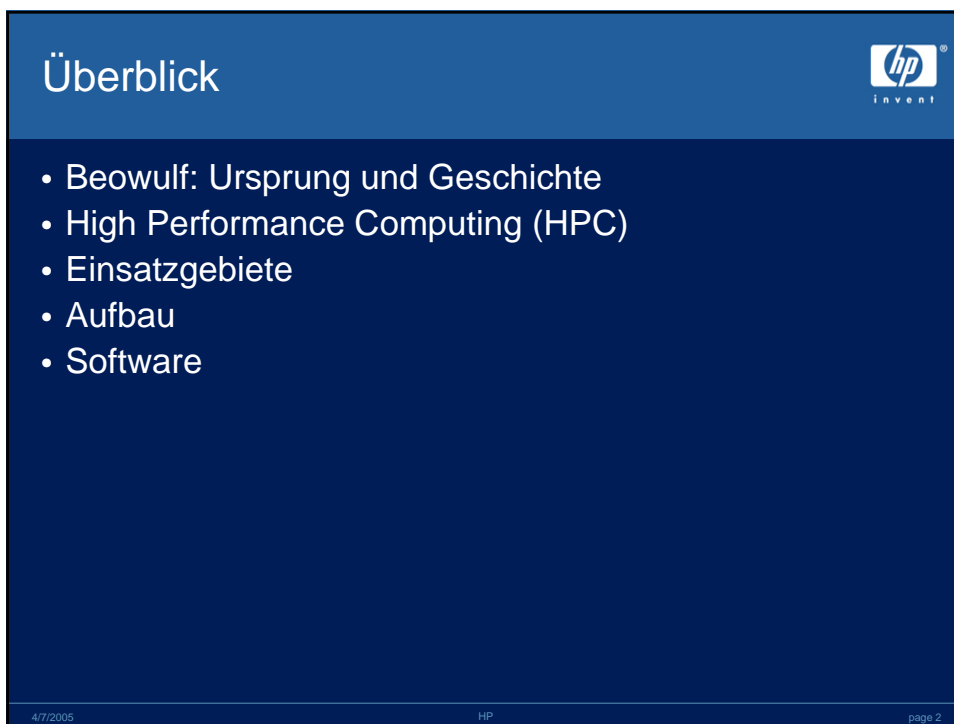
The slide features a white vertical bar on the left side containing the HP logo and the word 'invent'. The main area is a solid blue background. The title 'HP Beowulf Cluster' is positioned at the top right. Below it, the name 'Hanul Sieger, HP' is centered. At the bottom left, there is a large white graphic consisting of a plus sign followed by the lowercase letters 'hp'. In the bottom right corner, there is a small copyright notice: '© 2003 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.'

 HP Beowulf Cluster


Hanul Sieger, HP



© 2003 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice.



The slide has a dark blue background. The title 'Überblick' is in the top left corner. The HP logo and 'invent' are in the top right corner. A bulleted list is in the center. At the bottom, there is a footer with the date '4/7/2005', the text 'HP', and 'page 2'.

Überblick 

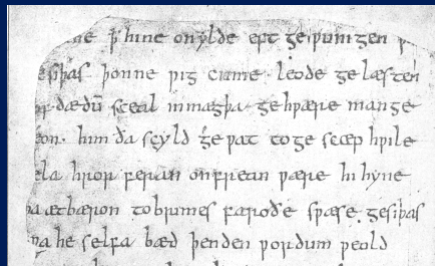
- Beowulf: Ursprung und Geschichte
- High Performance Computing (HPC)
- Einsatzgebiete
- Aufbau
- Software

4/7/2005 HP page 2

Woher kommt "Beowulf"?



- "Beowulf" ist das früheste, erhaltene Epos in englischer Sprache. Die Geschichte handelt von einem Helden Beowulf, der, stark und mutig, das Monster Grendel besiegt.



Was sind Beowulf Cluster?



- "Beowulf" ist ein allgemeiner Name für skalierbare Cluster, mit folgenden Eigenschaften:
- Standard Hardware
 - Eigenes (privates) Systemnetzwerk
 - Open-Source Software und -Infrastruktur
- Anlaufstelle im Internet: www.beowulf.org

Kurze Geschichte der Beowulf Cluster



- **Das Grundproblem:**
 - Es wird immer mehr Rechenkapazität für immer komplexere Probleme benötigt
 - Die bisherigen Lösungen dazu waren Supercomputer z.B. von Cray
- **Die Ursprungsidee:**
 - PCs entwickeln sich wesentlich schneller als Supercomputer
 - PCs haben ein viel besseres Preis-/Leistungsverhältnis
 - Das Ziel ist ein preiswerter Supercomputer aus Standard-Hardware
- **Die Implementation:**
 - Das erste Beowulf System entsteht 1994 in einem NASA-Institut, es besteht aus 16 Intel 486DX4 CPUs mit 10MBit Ethernet; das Betriebssystem ist Linux

4/7/2005

HP

page 5

HP HPC Cluster




4/7/2005

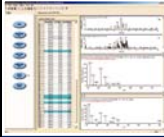
HP

page 6

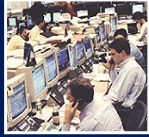
Was ist High Performance Computing




- HPC Probleme sind charakterisiert als:
 - datenintensive, rechnerisch oder numerisch intensive Aufgaben
 - komplexe Berechnungen mit großen Datenmengen bei gleichzeitiger Anforderung an hohe Durchsätze




Forschung




Finanzen



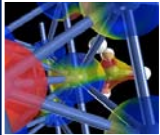
Geo-Wissenschaften




Virtual Prototyping



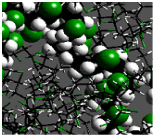
Film, Video und Effekte



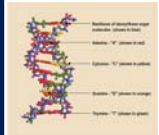
Electronic Design Automation



Regierung und Verteidigung




Product Lifecycle Management



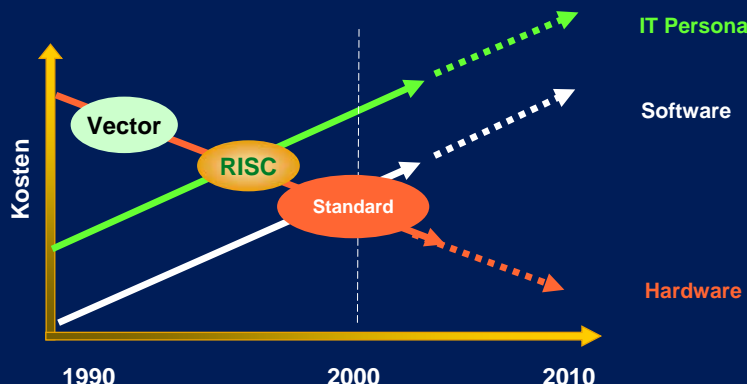
Bioinformatik

4/7/2005
HP
page 7

Veränderte ökonomische Rahmenbedingungen



Vom Supercomputer zum Cluster



4/7/2005
HP
page 8

Beowulf Wachstum



- Zuerst in der Informatik an den Universtitäten und anderen Forschungseinrichtungen, anschließend auch in anderen Forschungsrichtungen;
- heute findet man Beowulf Cluster in allen rechenintensiven Anwendungen innerhalb und außerhalb der Forschung
- stetiges Wachstum seit 1994
- Gutes Preis-/Leistungsverhältnis im Vergleich zu Supercomputern
- Erweiterbarkeit ohne Grenzen
- High Speed Interconnects sind verfügbar
- Linux ist inzwischen "erwachsen" und weit verbreitet

4/7/2005

HP

page 9

Beowulf auf dem Weg in die Produktion



- Seit Beowulf Cluster auch den hohen Anforderungen der Industrie an die Verfügbarkeit genügen, finden mehr und mehr Systeme ihren Platz als Produktivsysteme

- ISVs portieren Anwendungen

Gaussian, Gamess, Star CD, Fluent



- Es werden schlüsselfertige Beowulf Cluster von HP angeboten

4/7/2005

HP

page 10

Beowulf's "Nachteile"



- Applikationen müssen für die verteilte Architektur angepasst werden
- Algorithmen für verteiltes Rechnen benötigt entsprechende Kenntnisse in paralleler Programmierung
- Bis vor kurzem gab es kein adäquates Dateisystem, dies wird durch das Lustre File System beseitigt
- Es gibt keinen "Single System Image View" des Clusters
- Es wird Management Software gebraucht, um die Administrationskosten niedrig zu halten

Prozessor/Speicher-Architektur



- Shared-Memory Parallel (SMP)
 - Gemeinsamer Speicher wird von mehreren CPUs genutzt
- Distributed-Memory Parallel (DMP)
 - Cluster von Einzel-CPU-Rechnerknoten vernetzt durch ein low latency/high bandwidth Netzwerk
- Combined SMP/DMP
 - Cluster von SMP-Rechnerknoten
 - SMP-NUMA (Non Uniform Memory Access)

Beowulf Aufbau Rechnerknoten (Nodes)



- Control Node(s)
 - Typen: Front-End, Management Node, Service Node, Administration Node, Utility Node
 - Zugriffspunkt zum Cluster
 - Datei-Server
 - kann gesichert werden
- Compute nodes (Application Nodes)
 - für die (Benutzer-)Applikation

4/7/2005

HP

page 13

Beowulf Aufbau Netzwerke



- Out of Band Network
 - auch: Management network
- In Band Network
 - auch: Administration network
- Application Network
 - auch: data network, high speed network
- External Network
 - auch: outside network

4/7/2005

HP

page 14

Management network out-of-band network

Ethernet 10 / 100

Administration node

Administration Node:

- Zugriff auf BIOS und Konsole
- Nodes ein- und ausschalten, neu starten
- Management der einzelnen Compute Nodes
- Benötigt eine spezielle Management Kard (ILO, MP, Zircon)
- Benötigt eine TCP/IP Adresse (fest oder DHCP)

4/7/2005
HP
page 15

Administration network in-band network

Ethernet 10 / 100 / 1000

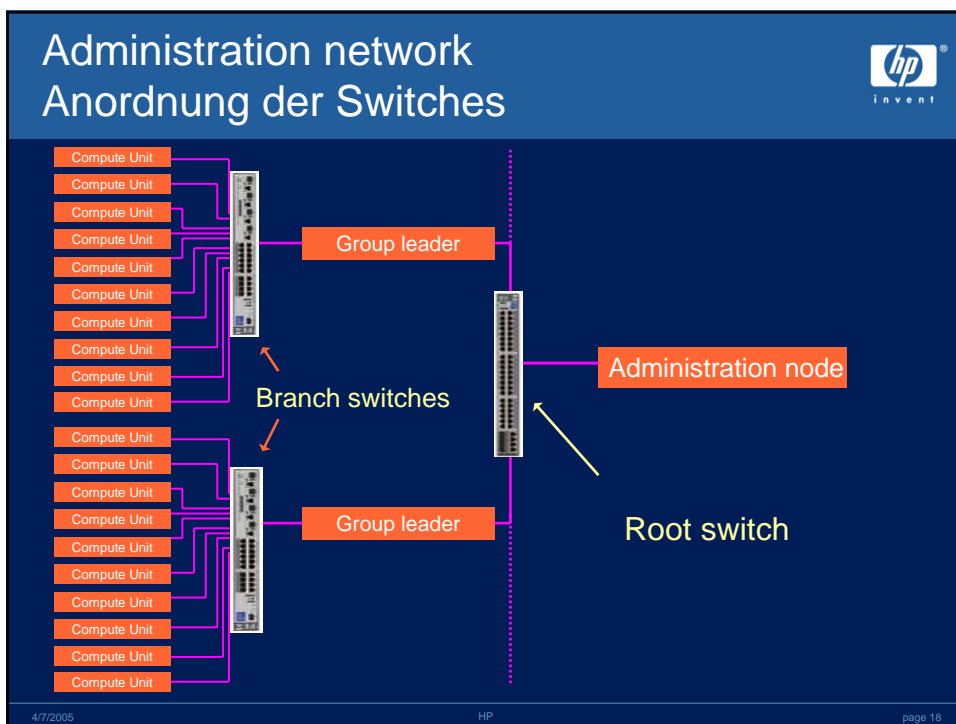
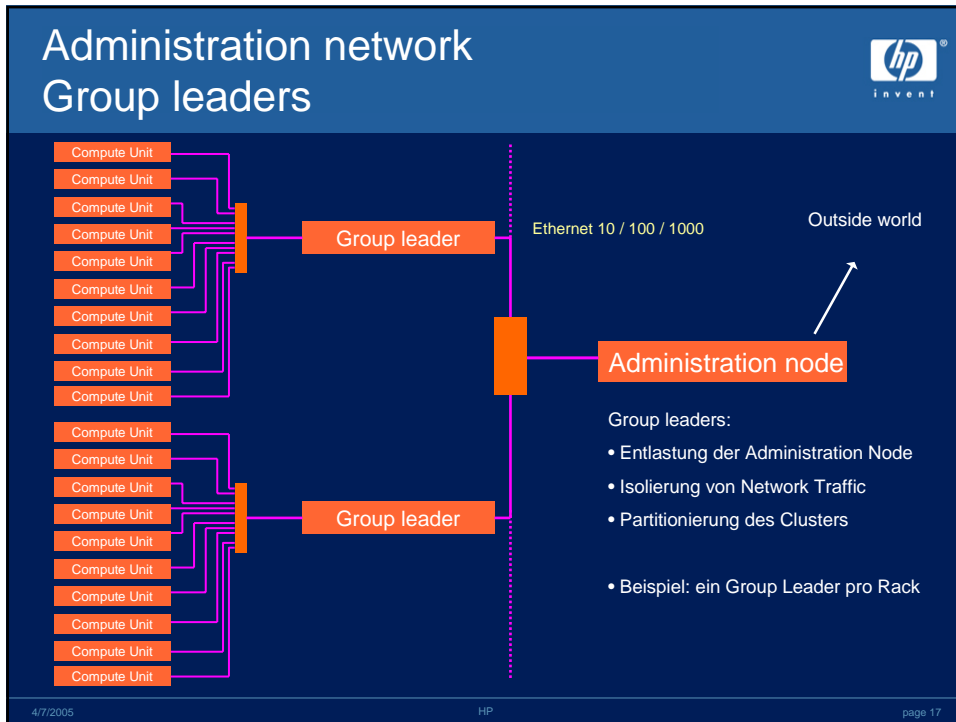
Administration node

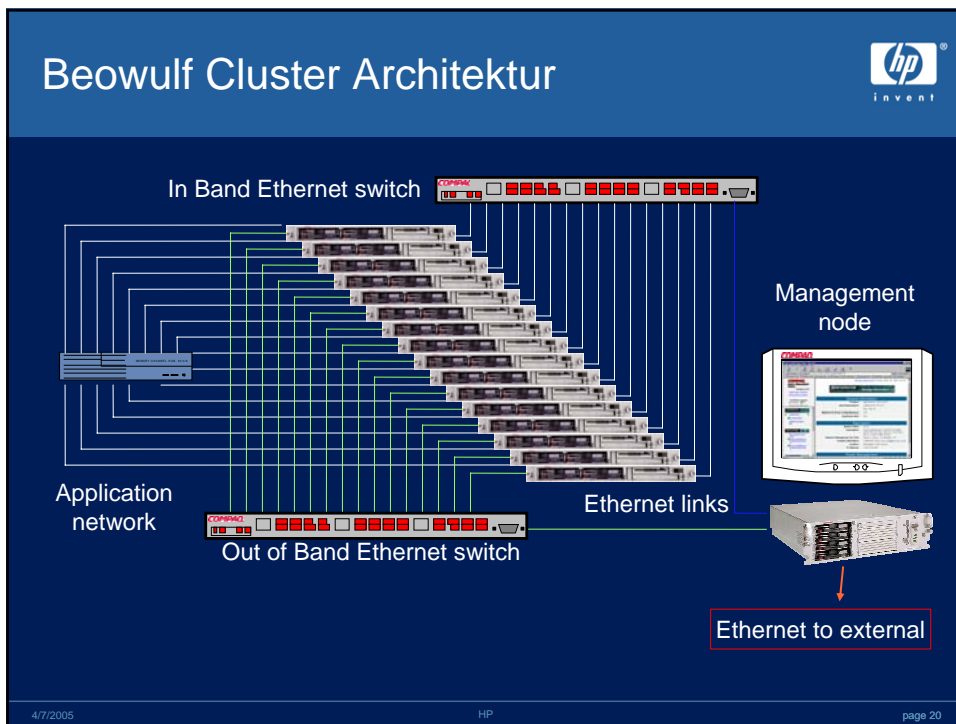
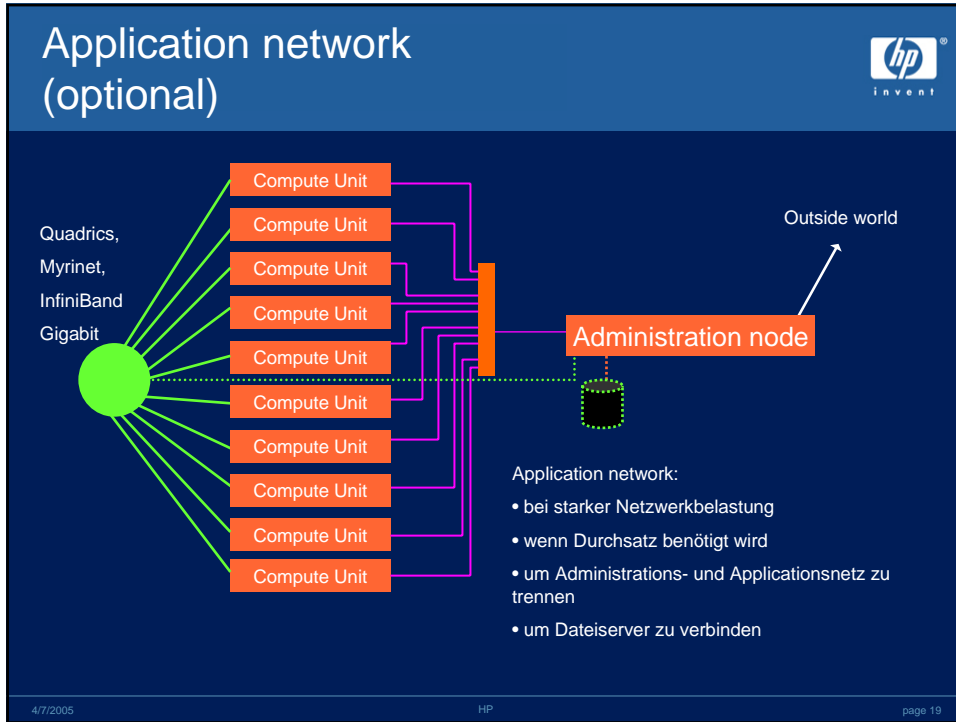
Outside world

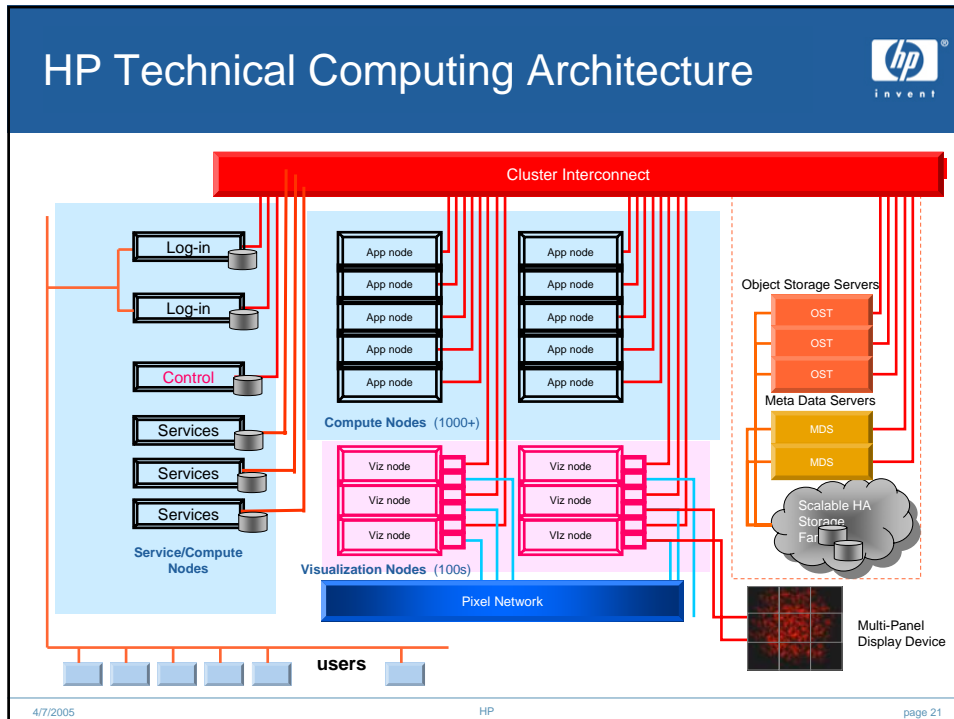
- Installation und Management des Clusters
- Beobachten der Compute Nodes
- log on für Benutzer
- Gateway
- Kontrolliert die Jobs auf den Compute Nodes
- Kann auch für Applikationen genutzt werden

Generally private TCP/IP network

4/7/2005
HP
page 16








Voraussetzungen für HPC Cluster

Die drei Hauptkriterien von HPC sind:

1. Benötigt sehr viel Prozessorkapazität
2. Benötigt sehr schnellen Zugang zu riesigen Datenmengen
3. Benötigt sehr schnelle Verbindungen zwischen den Rechnereinheiten

At the bottom of the slide, the text "4/7/2005" is on the left, "HP" is in the center, and "page 22" is on the right.


HPC Clusterkomponenten



- Plattform / Prozessor
 - Server
 - Management card (ILO, MP)
 - PDU
- Netzwerk
 - Ethernet 10/100/1000
 - Myrinet
 - QSW
 - InfiniBand
- HPC software
 - Protokolle
 - Betriebssysteme
 - Dateisysteme
 - Middleware
 - Management Software
 - Applikationen
- Storage
 - Local disks
 - NFS, Lustre, Sestina
 - MSA500, MSA1000 ..
 - EVA3000, EVA5000

4/7/2005 HP page 23

HP Server für Beowulf Cluster



HP Compute nodes:

Intel/AMD, x86


- ProLiant (Intel Pentium 4) (32bit)
- Blade Server BL20p (32bit)
- Opteron Server (64bit)

Intel Itanium II (IA64)



- Integrity Server rx1620, rx2620
- Integrity Server rx5670
- Integrity Server rx4640



4/7/2005 HP page 24



HP ProLiant servers for HPC



- DL300/500 Series
- DL360
 - Maximum performance power per 1U of rack space for dense compute environments
- DL380
 - The world's #1 selling server combining flexibility, leading 2P performance & manageability
 - Superior uptime with hot plug PCI, redundant drives, power supplies & Online Spare Memory
- DL585
 - Maximum compute power per 4U of rack space for dense environments with technical robustness
- DL100 Series
- DL140
 - High-performance, Xeon DP 2P/1U compute node optimized for HPTC environments.
- DL145
 - High-performance, AMD Opteron 2P/1U compute node optimized for 32-64 bit HPTC environments.
- BL series
- BL20p/BL35p(1Q05)
- BL30p(May 04)
 - Xeon DP and AMD Opteron 2P blades designed for HPTC
 - GigE connectivity
 - SAN Attachability
- Betriebssystem: Linux (und Windows auch für HPC möglich)







4/7/2005
HP
page 25


Application Network



Low Latency / High Bandwidth network

- Gigabit Ethernet
- Myrinet 2000
- Quadrics
- Infiniband



4/7/2005
HP
page 26

Software Architektur

Software auf HP Beowulf Cluster

- Betriebssysteme
 - Linux (Red Hat oder SuSE)
 - HP-UX (nur auf HP Integrity Server)
 - Windows Server 2003 (nur auf bestimmten ProLiant Systemen)
- Cluster Management Software
 - Cluster Management Utility - CMU
 - Scali Manage
- Communication Library für das Application Network
 - Message Passing Interface – MPI

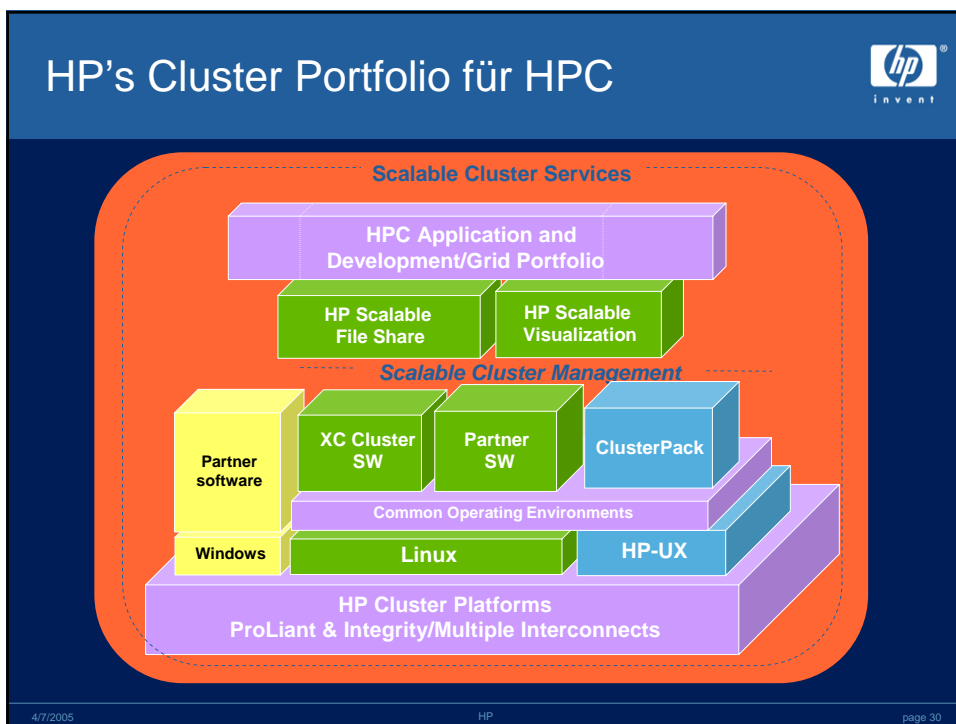
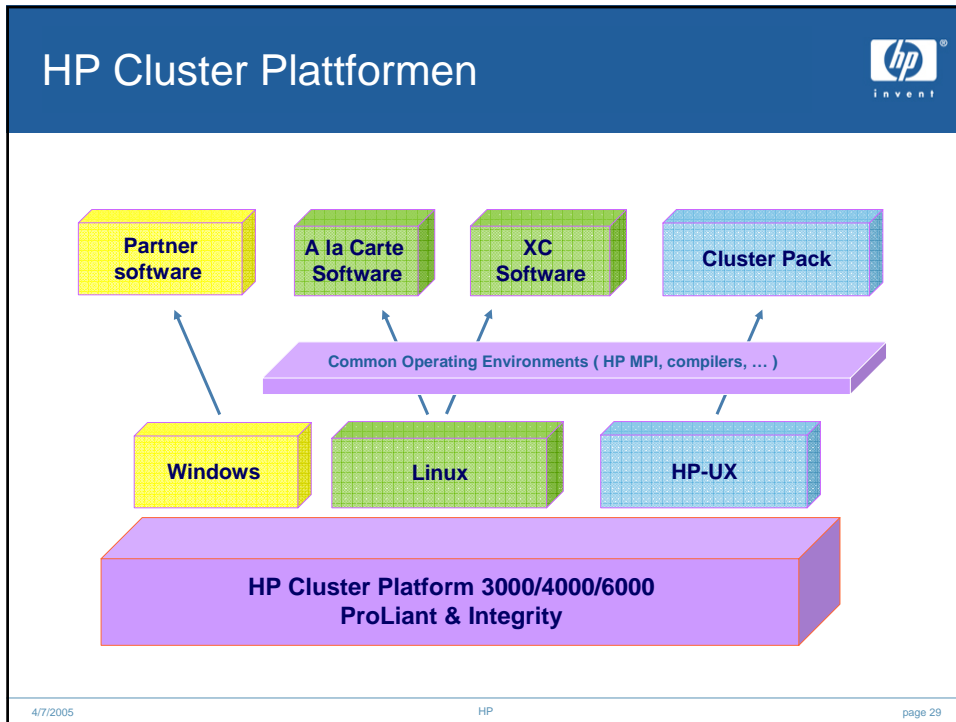
4/7/2005 HP page 27

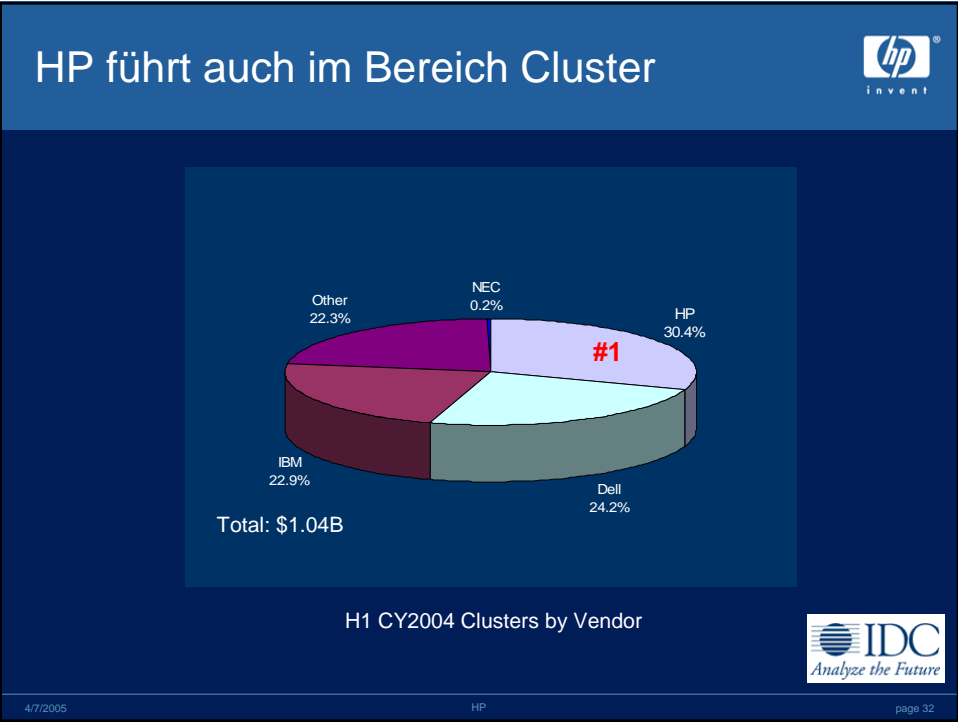
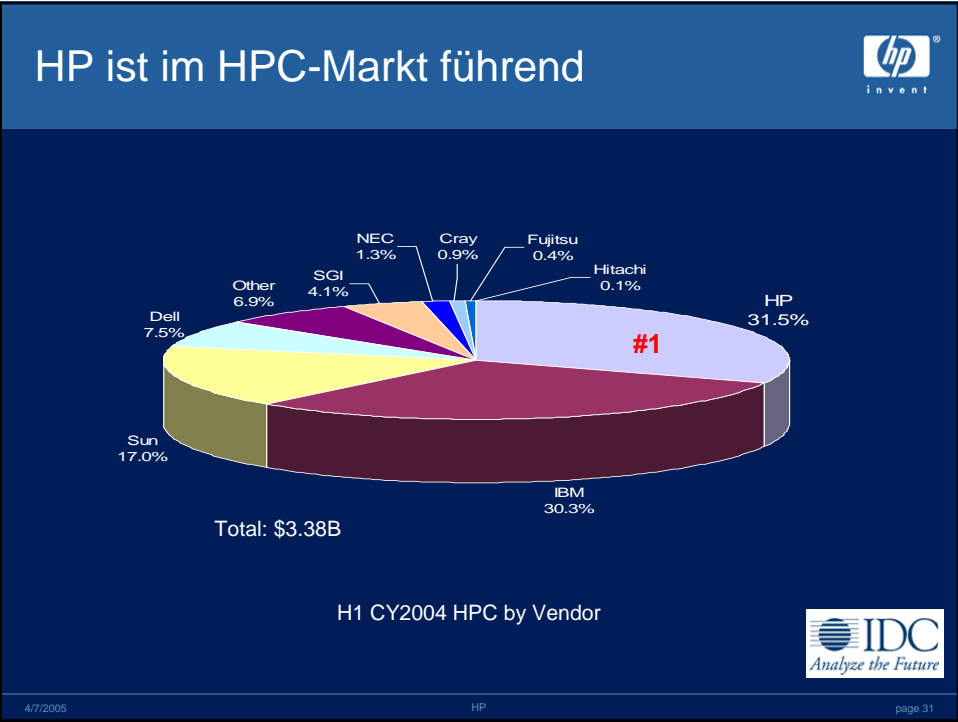
Design-Auswahl bei Cluster-Lösungen

Überblick über die SW- und HW-Komponenten eines HPC Cluster

Applikation	Bioscience, Engineering, GeoScience, ...		
Cluster Mgmt	XC/CMU/Scali/OSCAR/ROCKS		
Middleware	Job control, cluster admin, monitor, MPI, MLIB, App restart files migration, Dev tools (compilers, debuggers, profilers)		
Dateisystem	Local disks, NFS, SAN, Lustre, ...		
Betriebssystem	Linux, HP-UX, Windows		
Protokoll	Elan	MPI	TCP/IP
Netzwerk	Myrinet	Ethernet	Quadrics Infiniband
Platform/Processoren	IA32/IA64/Opteron/Alpha/		

4/7/2005 HP page 28





Die TOP-500-Liste




- TOP500
 - Listet die derzeit 500 schnellsten Computersysteme auf
 - Basierend auf LINPACK Benchmark Ranking
 - verlässliche Basis, um Trends in HPC-Bereich zu verfolgen
 - HP ist die Nr. 2 in den TOP 500 bei der Anzahl der Systeme, aktuell: 173
 - schnellstes HP-System mit Standard-HW auf Platz 16 (schnellstes HP-System, Platz 6)



<http://www.top500.org/>

4/7/2005
HP
page 33

HPC @ HP



HP C&I and Support

HP Engineering

HP Project Management

- Kundenpräsentation
- Projektqualifizierung
- HPTC Benchmarking
- Proof-of-Concept
- Integration der ISV-Produkte
- Event support
- Training

- ISV Support : PLM, MCAE, Oil & Gas etc.
- « Itanium Packaged Solutions »
- Geprüfte Konfigurationen
- Weltweite Forschungsprogramme

- Projektmanagement der Lösung
- Integration der Kundenkonfiguration
- Hardware- und Softwareanpassung
- Integration der Drittanbieter
- Test und Validierung
- Start-up der Konfiguration
- Kundentraining

4/7/2005
HP
page 34

