Status and Trends in High Performance Computing & Grids

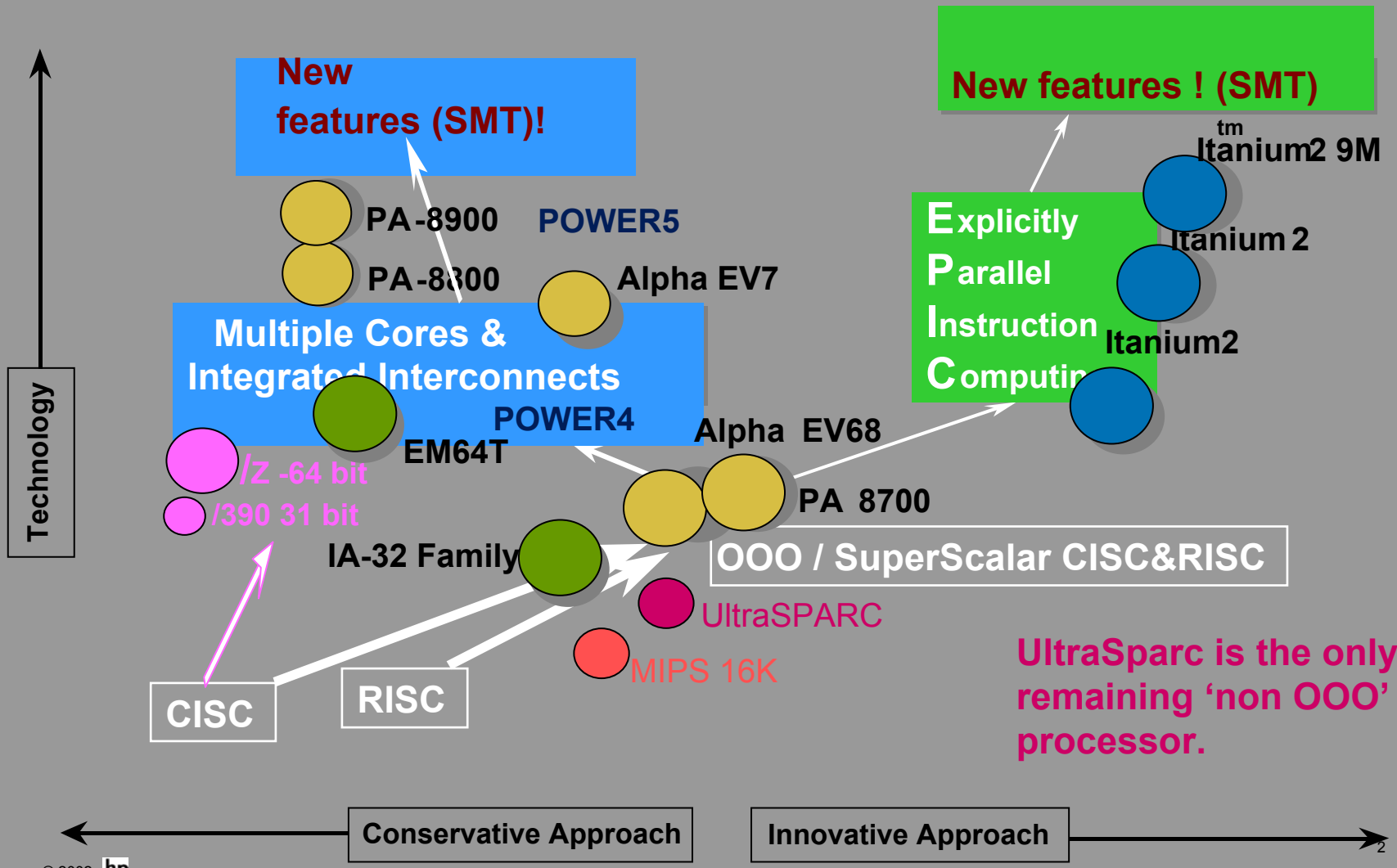
Selected Viewpoints

Dr. Frank Baetke
HPCD Richardson, Tx. & Munich

Processor Trends

Daegu 25.5.2005

the landscape of processor technologies



UltraSparc is the only remaining 'non OOO' processor.

CISC
RISC
EPIC

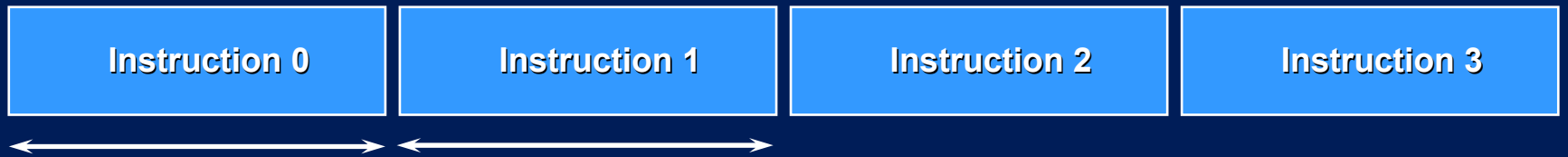
Daegu 25.5.2005

CISC, RISC and EPIC Instructions

CISC (x86,x86-64, vax, z)

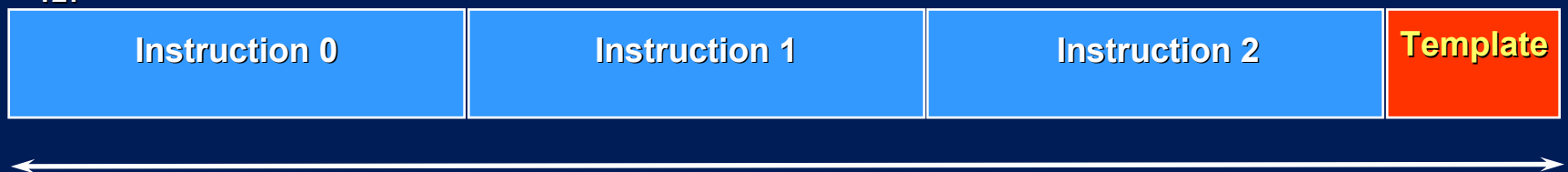
variable length

RISC (alpha, mips, pa-risc, power, sparc)



constant length

EPIC (Itanium)



constant length with template field

Details – not widely known:

IA-64 vs. x86-64

Daegu 25.5.2005

Itanium vs. Opteron (Addressing/Registers)

	Opteron	Madison
Process	0.13u	0.13u
Clock (for this comparison)	2.2 GHz	1.5 GHz
Physical address Space	40 bit	50 bit
Virtual address space	48 bit	64 bit
Int (=GRs) Registers	16	128
I-NaT-bits (for speculation)	-	128
Pre-Regs (for branch elim)	-	64
Float Registers	8	128
SSE2 (SIMD only)	16	-

Itanium vs. Opteron (Address translation)

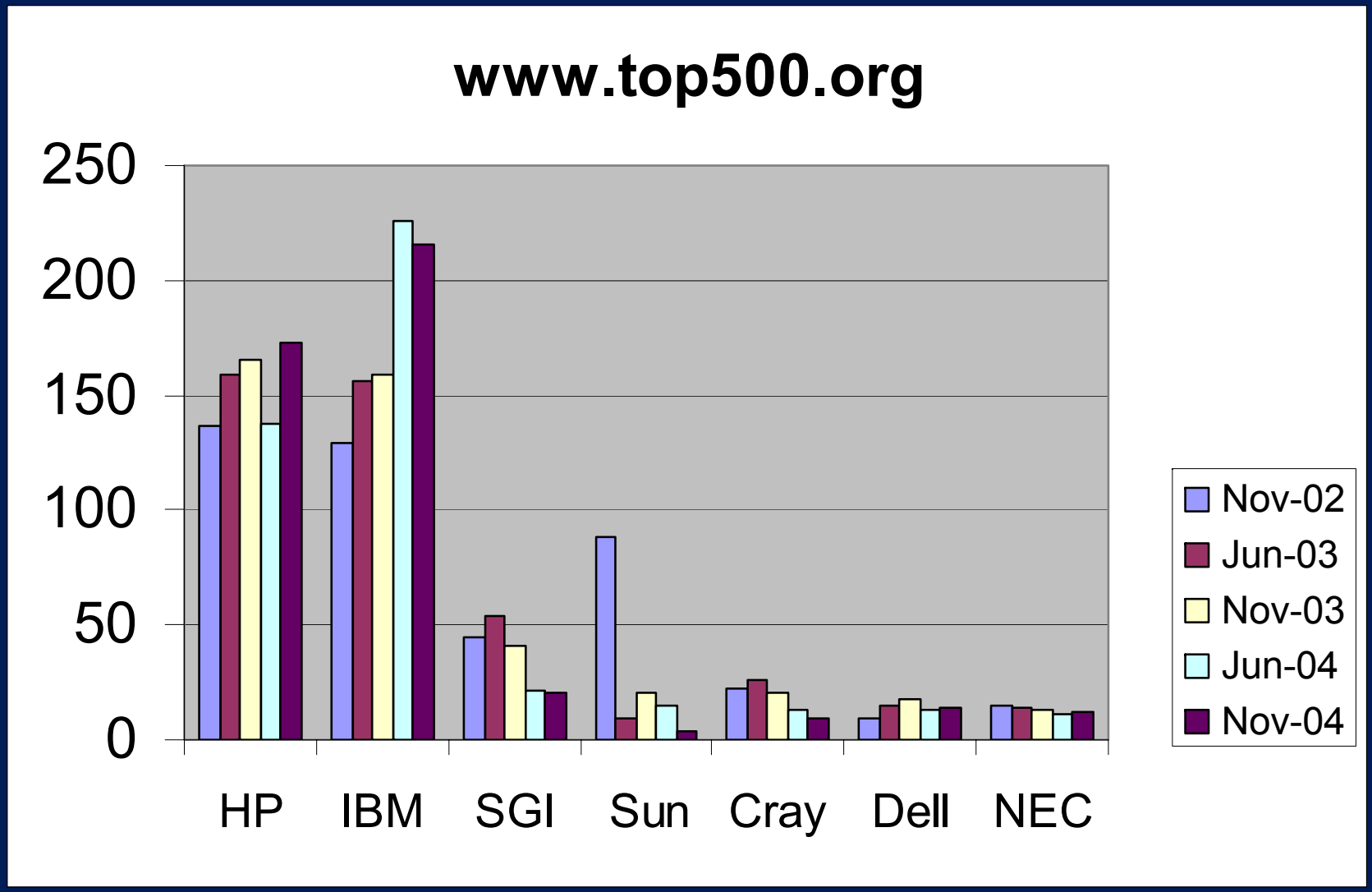
	Opteron	Itanium
TLB entries (instructions)	512	128
TLB-I associativity	4	128 (!)
TLB entries (data)	512	128
TLB-D associativity	4	128 (!)
supported page sizes	4 KB, 2 MB	4 KB … 4 GB
resulting address range with no TLB miss/fault	1 GB	512 GB

Comment: x86-64 heritage is again visible from the supported memory page sizes 4 KB and 2 MB (as in IA-32). Itanium supports variable page size from 4 KB up to 4 GB (!) and a resulting 'fault-less' address space of 512 GB !!

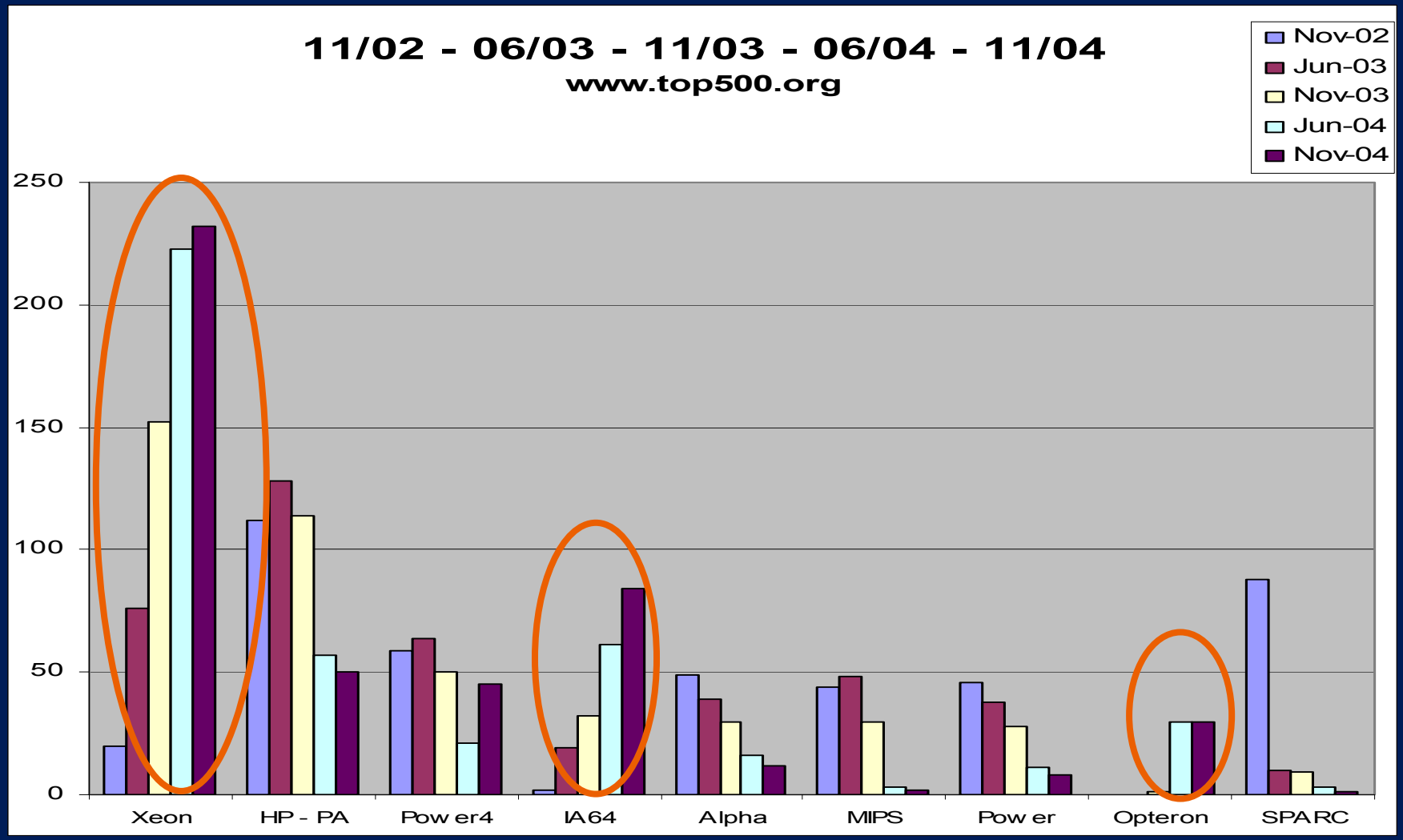
TOP500 Trends

Daegu 25.5.2005

Top500 Supercomputer – Nov 2004
Trends in Vendor Slots



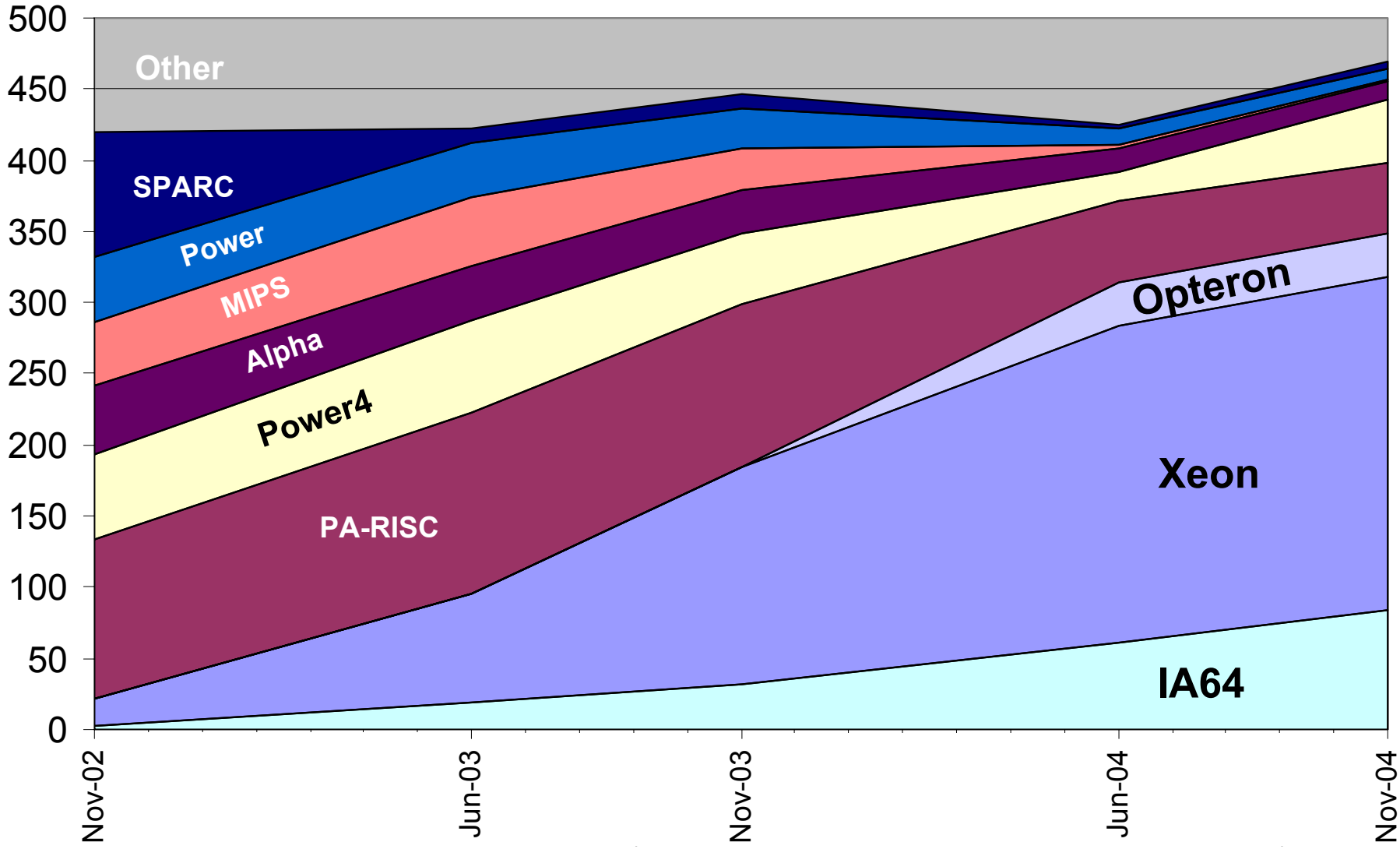
Top500 Supercomputers - Nov 2004
Trends in Processors



Industry Standard Architectures are the only ones growing!

Top500 Supercomputers – Jun 02 to Nov 04

Trends in Processors



OE Trends (The Matrix)

Daegu 25.5.2005

All, system vendors, software vendors and customers are obviously facing a similar question:

Which operating environments will survive ?

Operating Systems vs. Operating Environments II

Nomenclature:

A	=	available and supported by ISVs
a	=	available but (very) limited support
P	=	planned and announced
M	=	migration announced
A, P	=	endangered environments

A Processors / Operating Systems

- Let‘s have a look at the yellow special purpose s

	z/OS	OS400	Linux BlueG	AIX	Solaris	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
/390+	a									
Power400		a								
PowerBG			a							
Power				A			a			
USparc					A					
Itanium						A	A	A		
X86-64							A	A	A	A
X86					a				A	A

OE-Matrix – currently available (A)

		AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A				a			
USparc			A						
Alpha				A					
PA-RISC					A				
Itanium					A	A	A		
X86(64)						A		A	A
X86			a					A	A

OE-Matrix – planned OEs (P)

		AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A				a			
Sparc64			A						
Alpha				A					
PA-RISC					A				
Itanium					A	A	A		
X86(64)			P			A	P	A	A
X86			a					A	A

OE-Matrix – changes / migrations ...

		AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A				a			
Sparc64			A						
Alpha				m					
PA-RISC					m				
Itanium					**A**	A	A		
X86(64)			a			A	a	A	A
X86			a					A	A

OE-Matrix–HPC Node/Server Futures (32bit environments become obsolete)

		AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A				a			
Sparc64			A						
Alpha									
PA-RISC									
Itanium					A	A	A		
X86(64)			a			A	A	A	A
X86								A	A

OE-Matrix (64bit). Next steps …

		AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A				a			
Sparc64			A						
Itanium					A	A	A		
X86(64)			a			A	A		
X86									

… leading to a potential consolidation

	Other VMS	AIX	Solaris	Tru 64	HP-UX	Linux 64	WIN 64	Linux 32	WIN 32
Power		A							
Sparc64									
Itanium					A	A	A		
X86(64)						A	A		
X86									

Server Trends

Fast !
Good !
Cheap !

Daegu 25.5.2005

Server Trends

Fast !
Good !
Cheap !

Pick any two but only two!

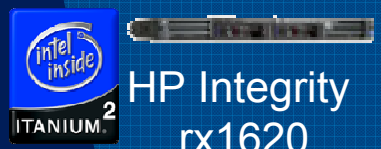
Daegu 25.5.2005

Servers
(HP Examples)

Daegu 25.5.2005

HPC platforms

- Choice
- Performance
- Manageability



HP Integrity rx1620

HP Integrity rx2620

HP Integrity rx4640

HP Integrity rx7620

HP Integrity rx8620

HP Integrity Superdome

HP ProLiant DL585

HP ProLiant BL20p G3
BL25p
BL30p
BL35p

HP ProLiant DL360 G4p
DL385
DL380 G4

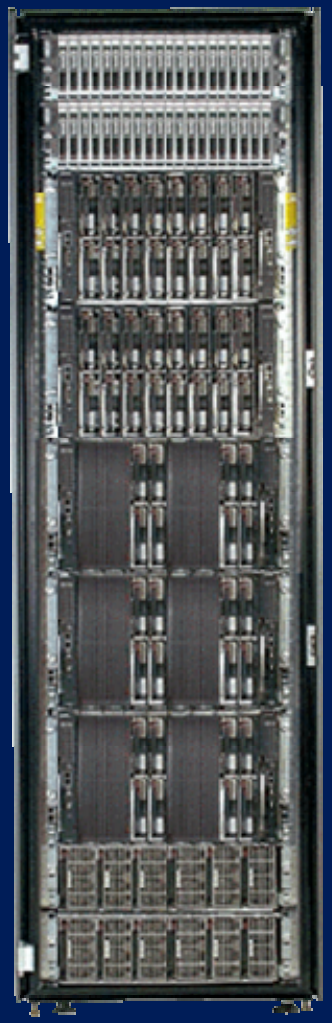
HP ProLiant DL140 G2
DL145 G2

Daegu 25.5.2005

Blades for HPC

The benefits of blades applied to HPC clusters

- Simplified management
- Designed for performance and scalability
- Reduced interconnect and network complexity
- High density
- Centralized power management
- New challenges …

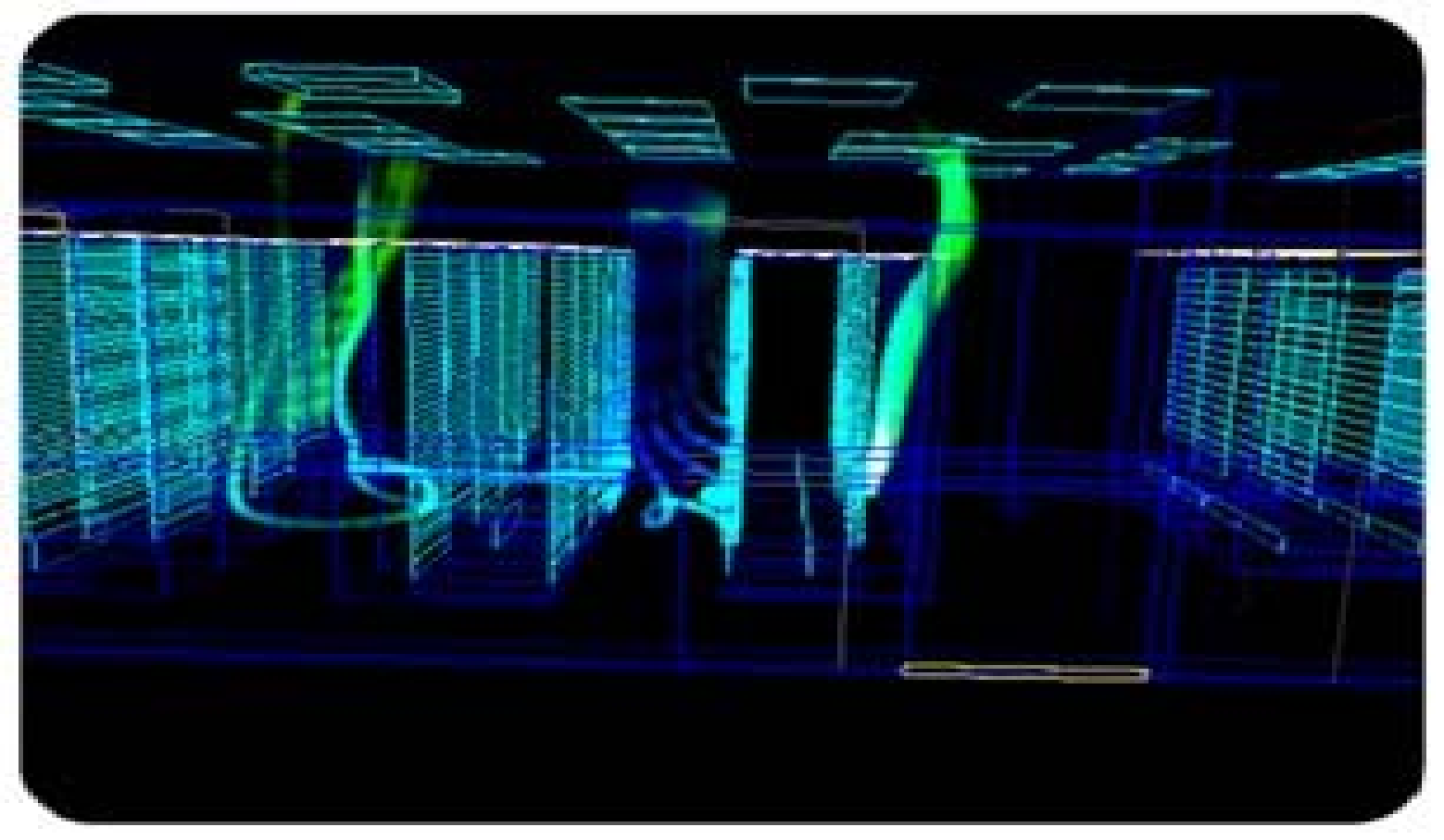


Front

Rear

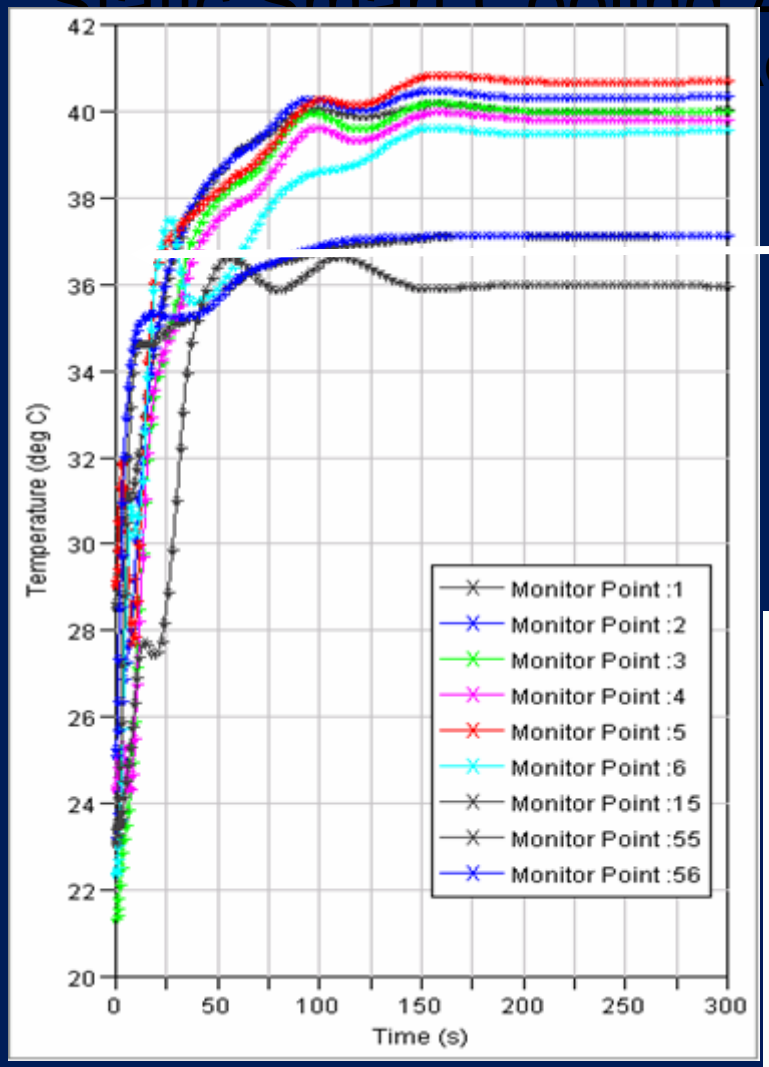
CFD Modeling of Machine Room

HP used CFD and 3D modeling to understand thermal characteristics of data centers



Static Smart Cooling Analysis

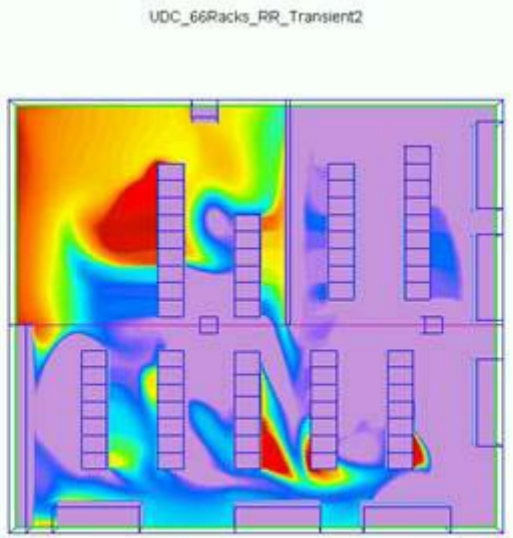
C #1 failure

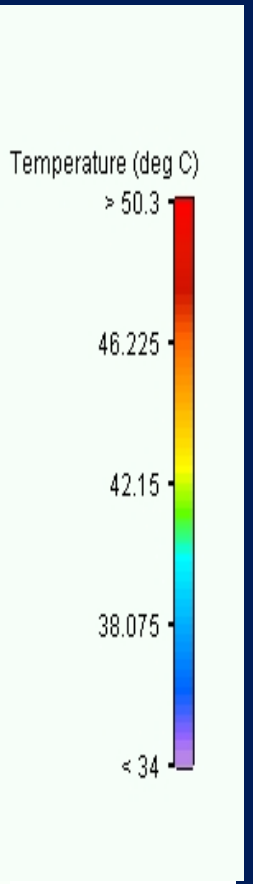


Unacceptable temperature level

Develop Strategies for Failure Mitigation

- Workload Migration
- "Smart" Redundancy of AC equipment



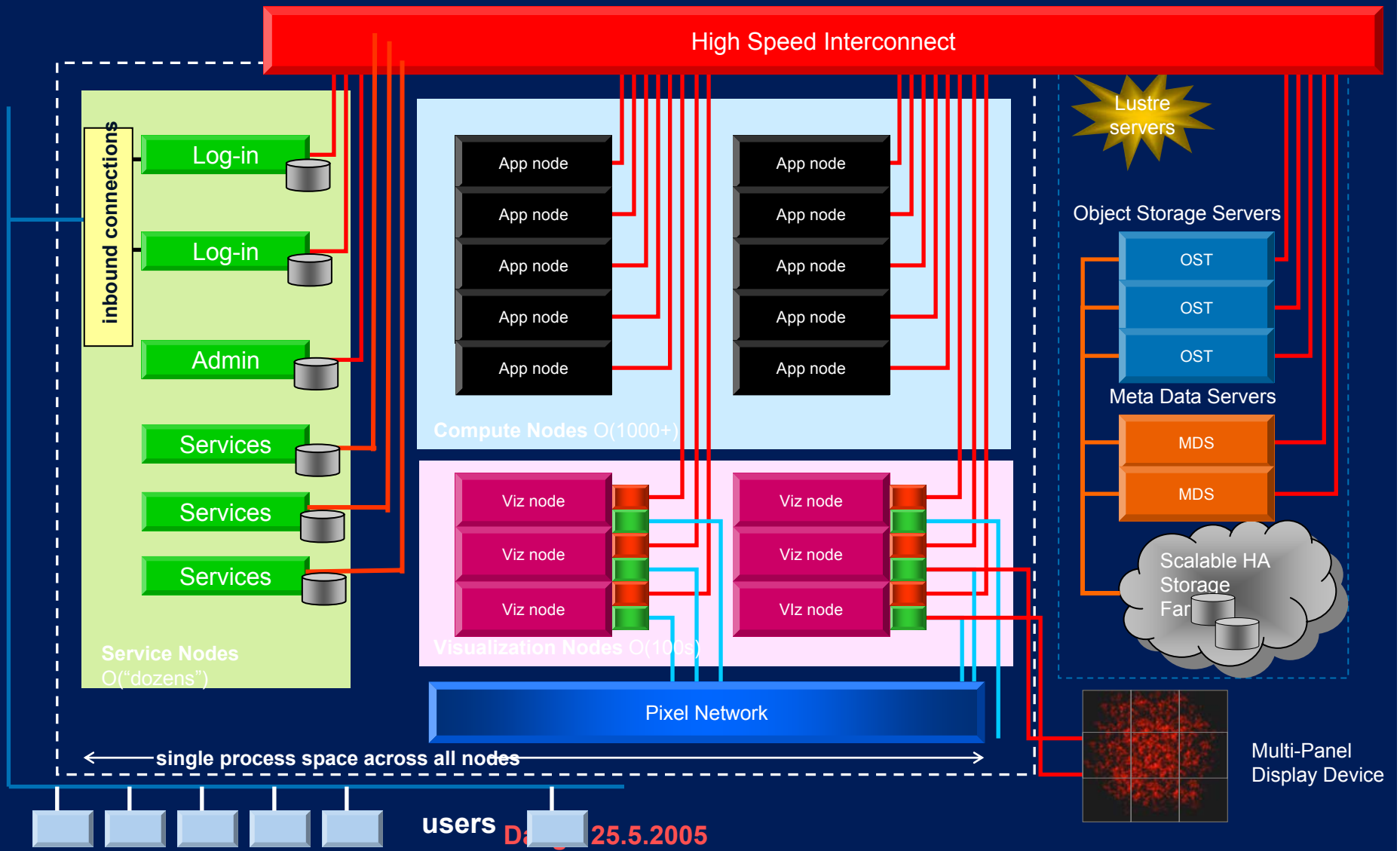


** It took 80 seconds to reach unacceptable temperature level.

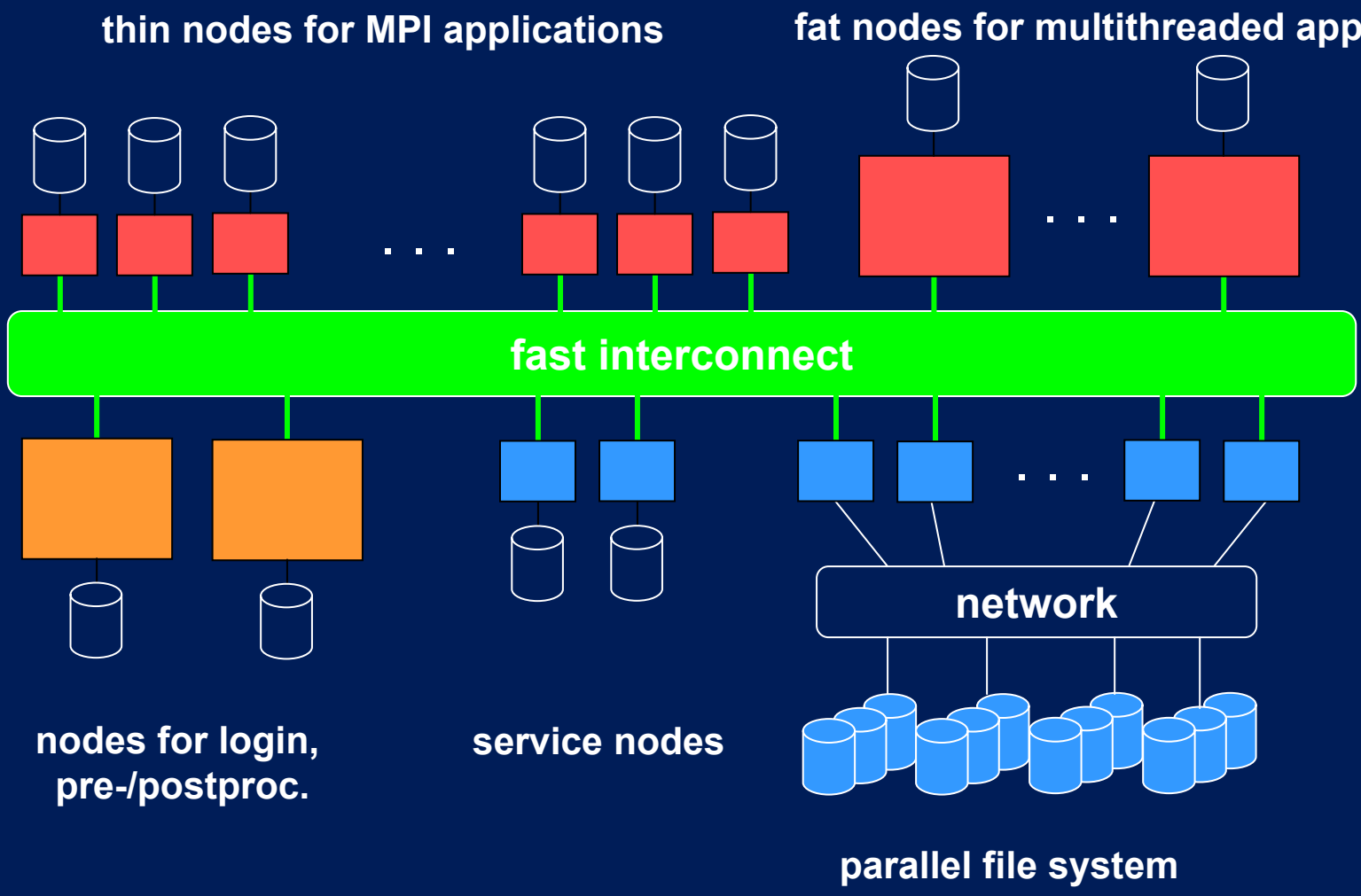
Cluster Trends

Daegu 25.5.2005

A Generic Cluster Vision … strong tendency towards standards!



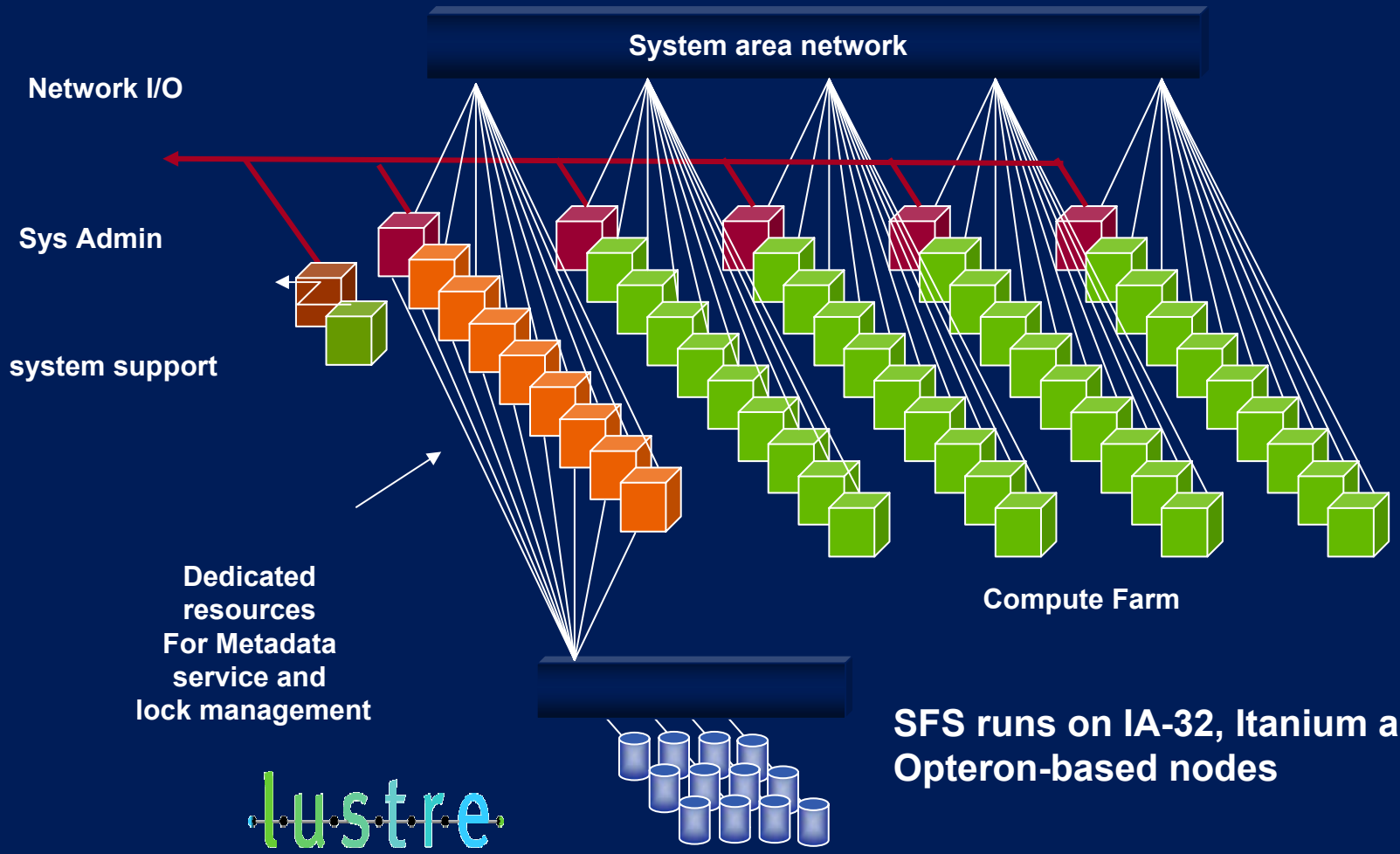
SSCK – new 10+ TFLOP/s System (University of Karlsruhe)



Clusters will dominate: what else is of interest

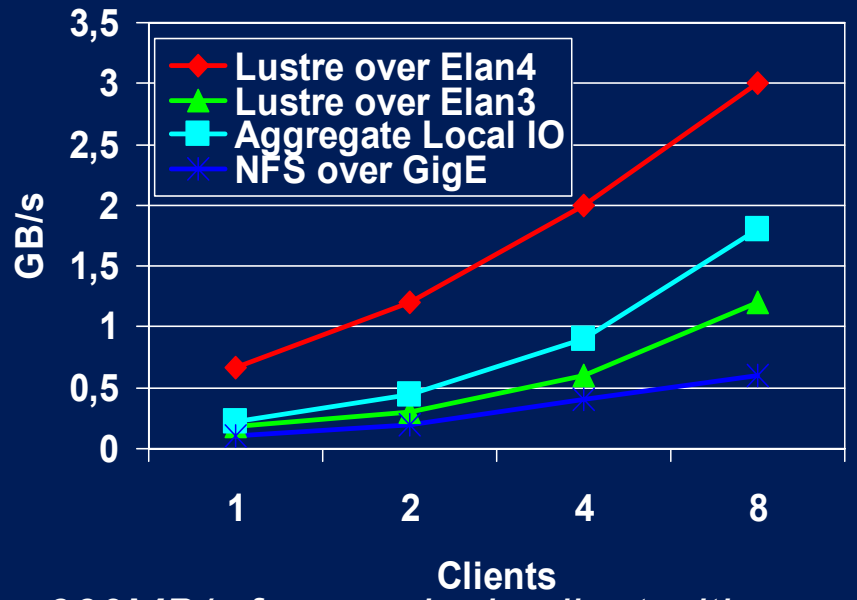
Daegu 25.5.2005

… A Scalable Cluster Filesystem



SFS runs on IA-32, Itanium and Opteron-based nodes

Increasing filesystem technology to meet scalability and performance metrics



660MB/s from a single client with a simple "dd" is faster than any local or global filesystem we have tested.

- I have over a year's experience running a data center using Lustre for aggressive IO from a supercomputer.
  - Highly stable
  - Still hard to manage

  (Scott Studham, PNNL)

We are finally in the era where global filesystems provide faster access

Making Lustre a Products

Daegu 25.5.2005

HP SFS: delivers value by providing an unmatched combination of features

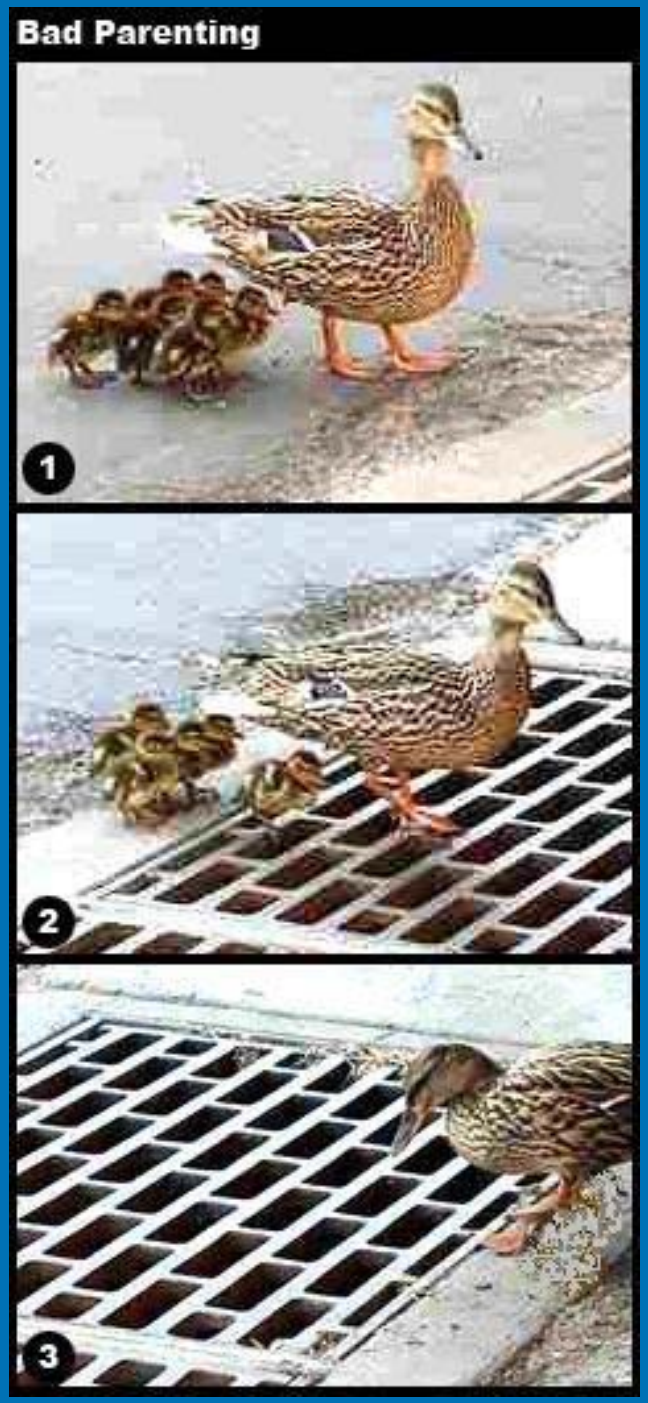
The Lustre-based HP SFS product is the result of a joint resarch and development project (Hendrix) between the HP, the US DoE and Cluster File System Inc.

	SFS	AFS	NFS
Network-Neutral	Yes	No	No
Cluster-Wide POSIX Semantics	Yes	No	Yes
No Single Points of Failure	Yes	Yes	No
High Performnace Direct I/O	Yes	No	No
Seperate Metadata and Data Servers	Yes	No	No
Scalable to Thousands of Nodes	Yes	Yes	No

Grid Trends

Daegu 25.5.2005

Many definitions of Grid ...



Sep 2004

Mark Linesch Named Chairman of the Global Grid Forum

In September, HP's Mark Linesch, a member of the Enterprise Servers and Storage (ESS) Chief Technology Office (CTO Office) was named Chairman of the Global Grid Forum (GGF). The GGF is the preeminent Grid standards body, with participation from over 30 countries and 500 organizations worldwide. Linesch, who will serve a full-time, three-year term, replaces founding Chairman Charlie Catlett.

Mark Linesch

Grid technologies are distributed computing technologies that provide the foundation to large-scale efforts utilizing the Internet to build distributed computing and communications infrastructures. They are important to HP because they are one manifestation of the utility computing vision on which the Adaptive Enterprise (AE) strategy is based. Many infrastructure technologies and standards core to the grid are also core to HP's AE portfolio.

Jan 2005

Jim Hughes Elected

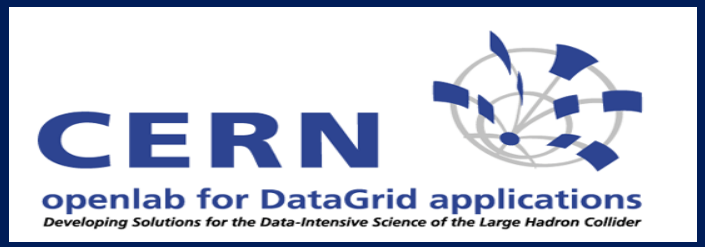
HP's Jim Hughes Elected Chairman of OASIS Board

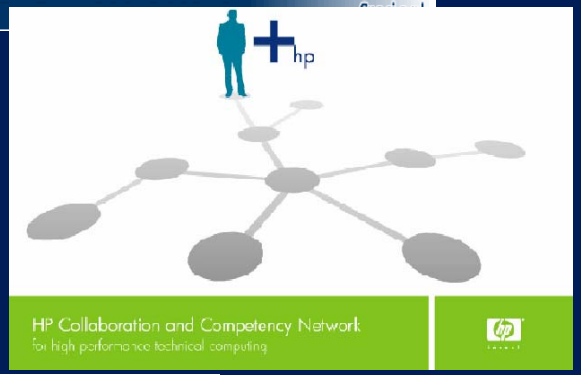
At its January 2005 meeting, the OASIS Board of Directors elected Jim Hughes of HP's Software GBU as Chairman. Hughes had previously served as Treasurer and earlier as Secretary since his election to the Board in 2001.

OASIS (Organization for the Advancement of Structured Information Standards) is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. The consortium produces more Web services standards than any other organization along with standards for security, e-business, and standardization efforts in the public sector and for application-specific markets. Founded in 1993, OASIS has more than 4,000 participants representing over 600 organizations and individual members in 100 countries.

24 Jan 2005: HP, IBM, Intel and Sun Launch Globus Consortium

(some of) HP's Grid Collaborations



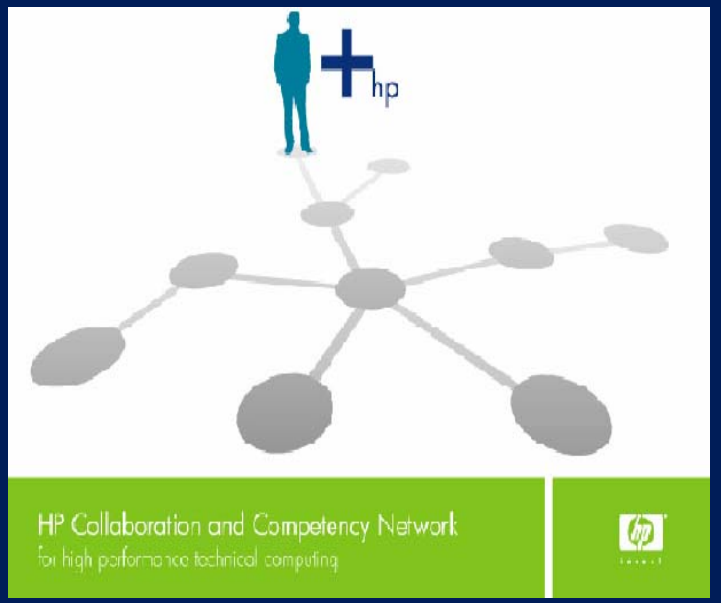


HP Collaboration and Competency Network

HP CCN is a forum to facilitate collaboration, innovation, discovery, and competency sharing between HP and our HPC customers and partners

Collaboration areas:

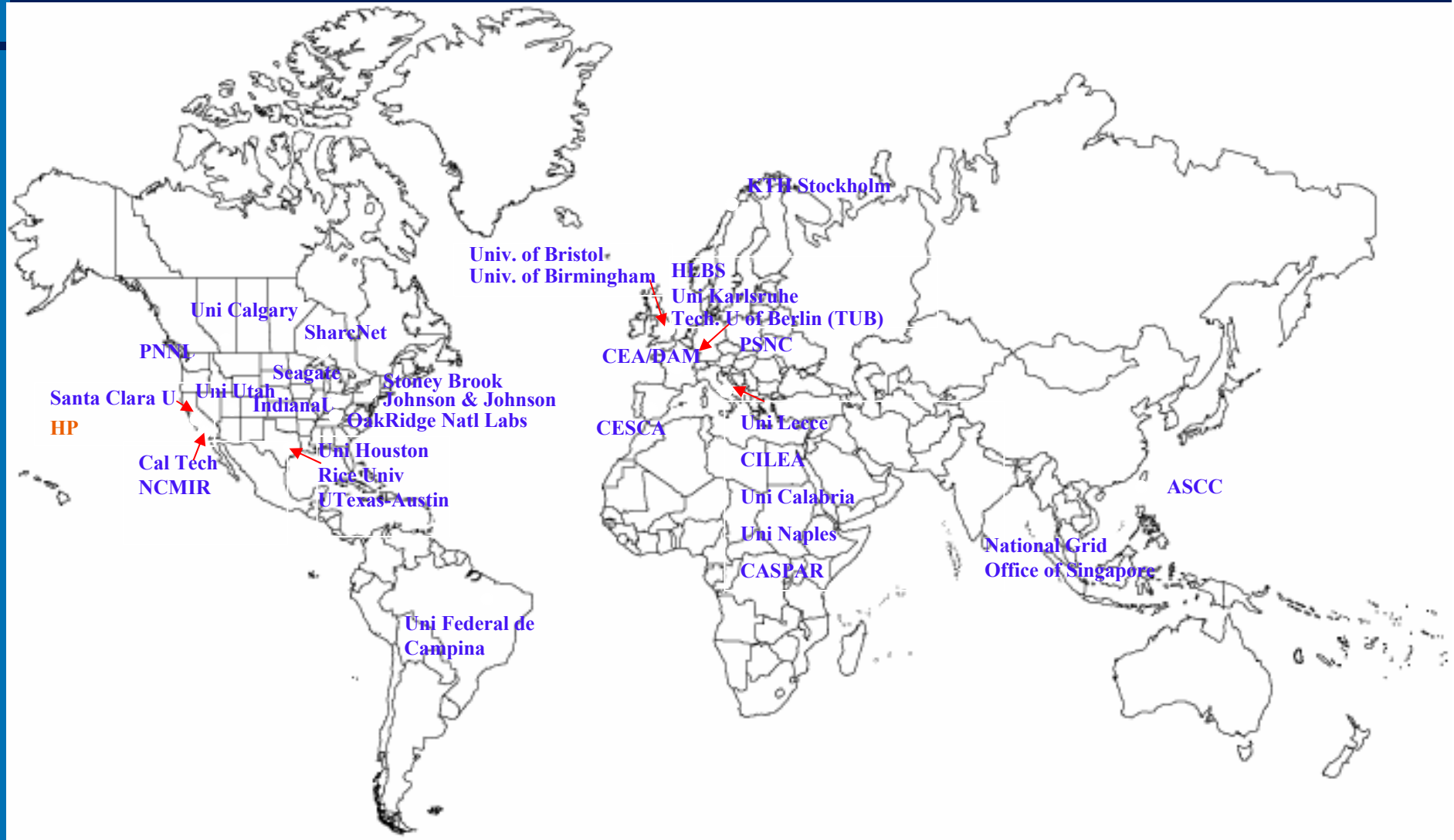
- Computational and data grids
- Global file system for Linux (Lustre – HP SFS)
- Scientific visualization (SEPIA)
- Linux SMP scaling
- Algorithms & Methods



Why HP CCN?

- Persistent framework for partnership with high-performance computing customers
- Enable customers to interact with HP's R&D community and with each other in a structured manner
- Facilitate exchange of expertise and experience among like-minded people world-wide
- Avoid the disadvantages of creating ad-hoc relationships on a case-by-case basis

HP CCN Members



HP CCN Grid Collaboration

- Objective

  Advance the state of the art in creating, running, and using grids

- Member Engagement
  - Regular technical teleconferences and discussions
  - The HP CCN Grid, an operational grid that includes systems at member sites and at HP
  - Exchange of applications, middleware, and tools
  - Experiment and test new grid technology
- Expected Outcome

  A worldwide community of leading-edge research institutions working together to produce a body of understanding and best practices for grids

Recent Teleconference Topics

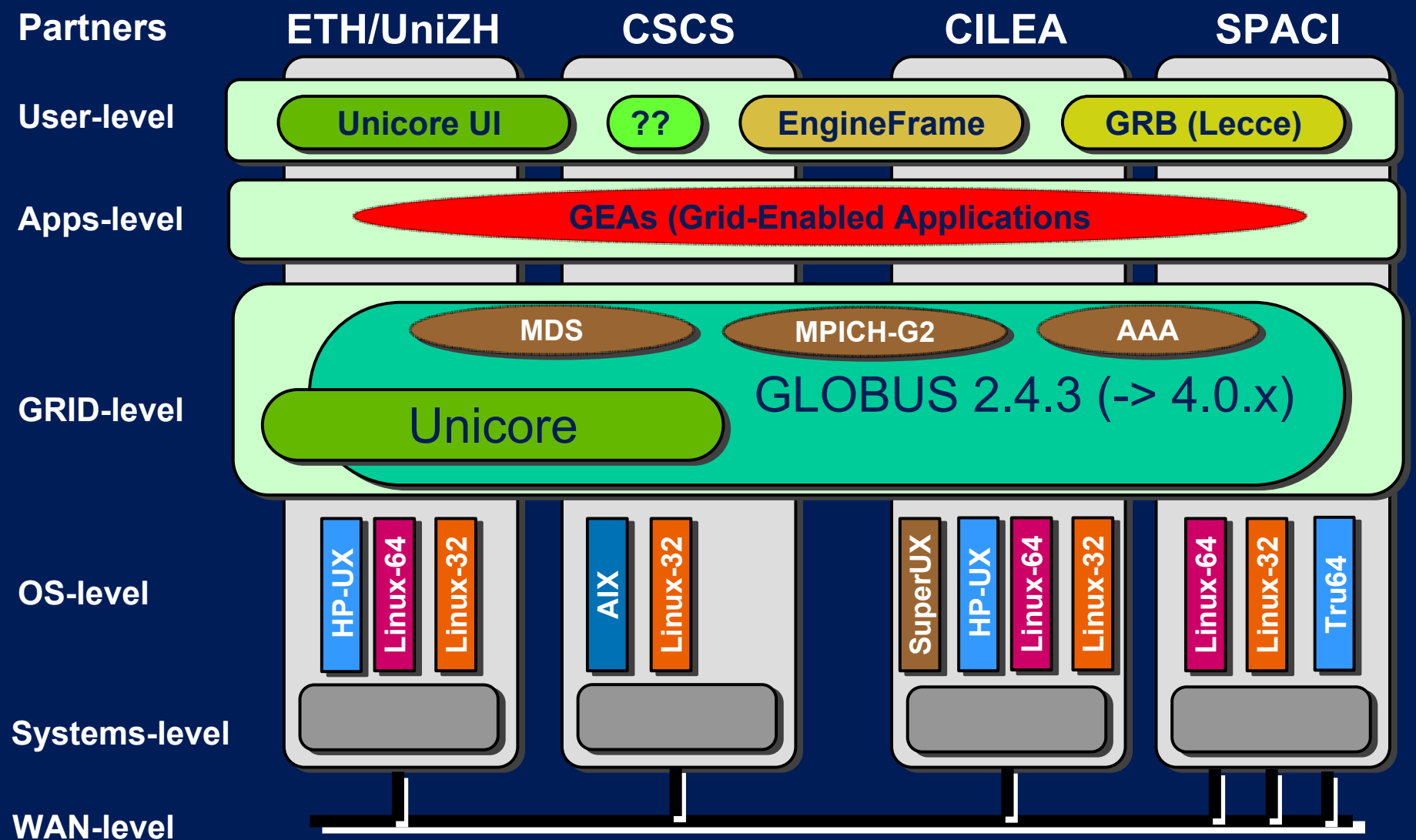
- Use of Realm Specific IP (RSIP) in conjunction with MPICH-G2
  - National Grid Office of Singapore and HP
- “OurGrid”, a grid for smaller and mid-size research labs
  - Universidade Federal de Campina Grande
- Monitoring of Grid environments
  - University of Calgary Grid Research Center
- Data replication management system using WSRF
  - University of Calgary Grid Research Center
- Requirements for Grid management
  - HP

SEPAC

South
European
Partnership
for
Advanced
Computing

Daegu 25.8.2005

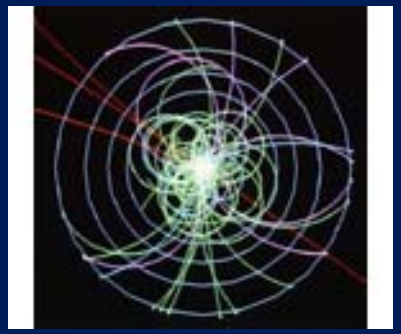
SEPAC Partner Sites & Logical Grid Structure



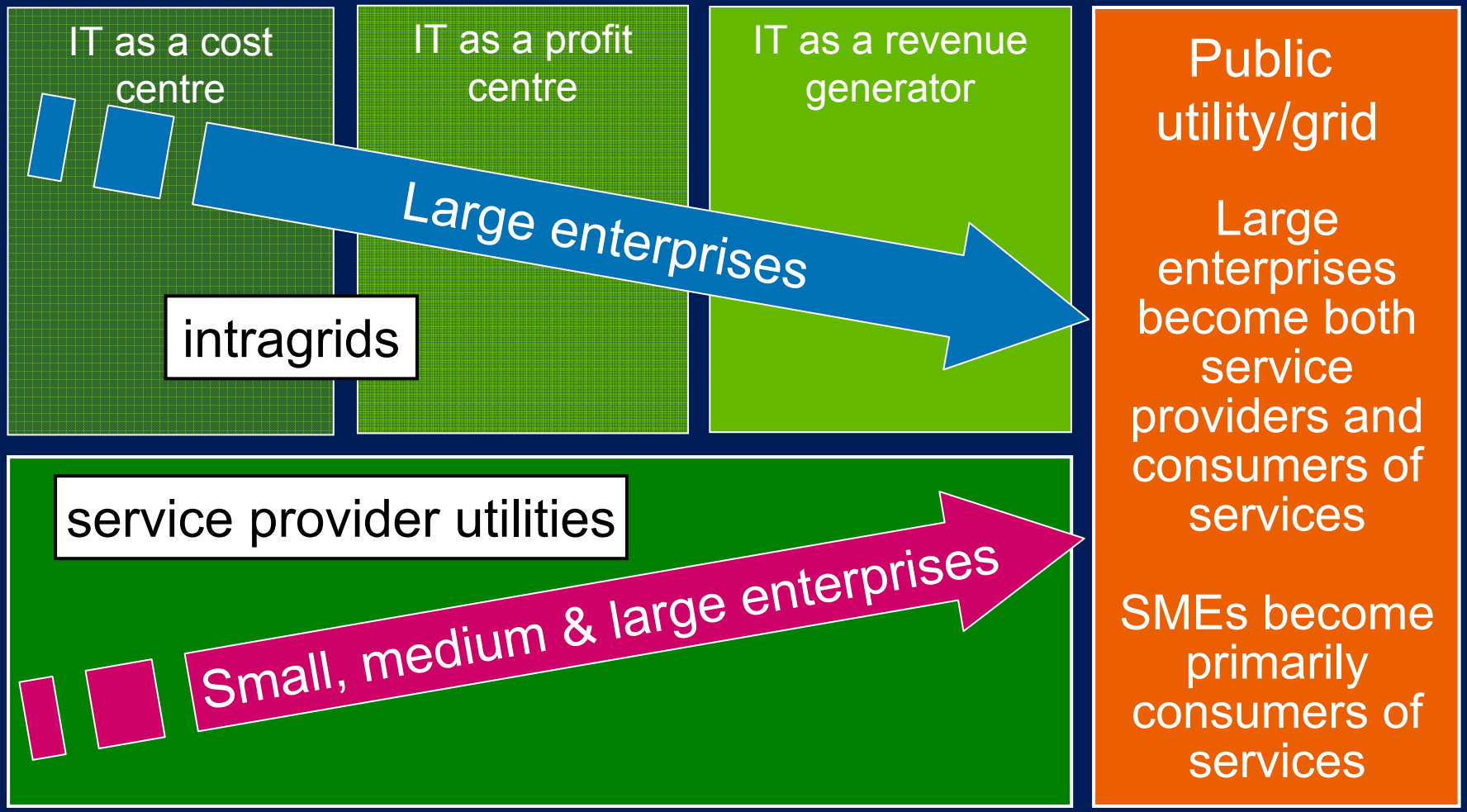
Deploying the LHC Grid Service

Today
- > 30 countries, 130 sites
- > 14,000 CPUs
- > 5 Pb storage

Today
- 2 HP sites – Bristol and Puerto Rico
- HP 1st commercial member of the LHC grid
- IA64 resources on LCG ... 2nd to Poland



Roadmap to HP Labs Grid Vision



Closing Remarks

- HPC applications are the major driver of Grid computing – and the best fit for the early adoption of a utility computing business model
- As usual, HPC learnings and advances will influence mainstream enterprise IT
- Broad industry agreement on vision of utility/ grid/service-oriented/adaptive/on-demand computing
- The roadmap to the vision is unfolding: lots of different starting points, similar destinations

Thank You

Daegu 25.5.2005