hpc@nottingham.ac.uk
The old and the new

Colin Bannister

Information Services

- New system selection
- New system details
- Jupiter upgrade
- System integration
- Timetable

New system selection

Requirements specification based on user survey:

More memory (RAM) on compute nodes

Better interconnect (internal network)

More scratch disk space

Fairer queuing system

More reliable system

Support commercial software

Run parallel codes better

Successful Tender came from Streamline/HP

Will meet the requirements specification as follows:

"More memory (RAM) on compute nodes"

Old = 1 GB per CPU, New = 4GB or 16GB per CPU

"Better interconnect (internal network)"

Old = 1Gb Ethernet, New = 10Gb InfiniBand

"More scratch disk space"

Old = 44GB , New = 200GB plus a fast parallel filesystem (Panasas)

"Fairer queuing system"

Old = Sun Grid Engine, New = Sun Grid Engine plus MOAB

New System summary
- Streamline/HP/Panasas
- HP compute nodes, Intel processors, quad-core 'Harpertown'
- Panasas storage
- InfiniBand MPI network
- Ethernet Storage network
- MOAB/SGE scheduling, CMA management
- Intel Compilers/Development Toolkit

Streamline/HP/Panasas

- Contract with Streamline

- Hardware from HP and Panasas

- Streamline responsible for managing support for <u>whole system</u>

Compute nodes (1600 compute cores)

HP DL 160, 2x Intel Harpertown (quad-core)

3 GHz 1600MHz FSB

168 with 8 GB RAM

32 with 32 GB RAM

250 GB internal drive

Processors:

Intel Harpertown (quad-core) 3 GHz,

1600 MHz FSB, 12MB L2 cache (2x6MB)

80W

Storage:

Panasas 'parallel storage clusters'

Hardware/Software 'appliance'

Storage blades, disk and processors combined

105 TB (raw) total (70 storage blades)

Integrated GE Switch

Battery Module
(2 Power units)
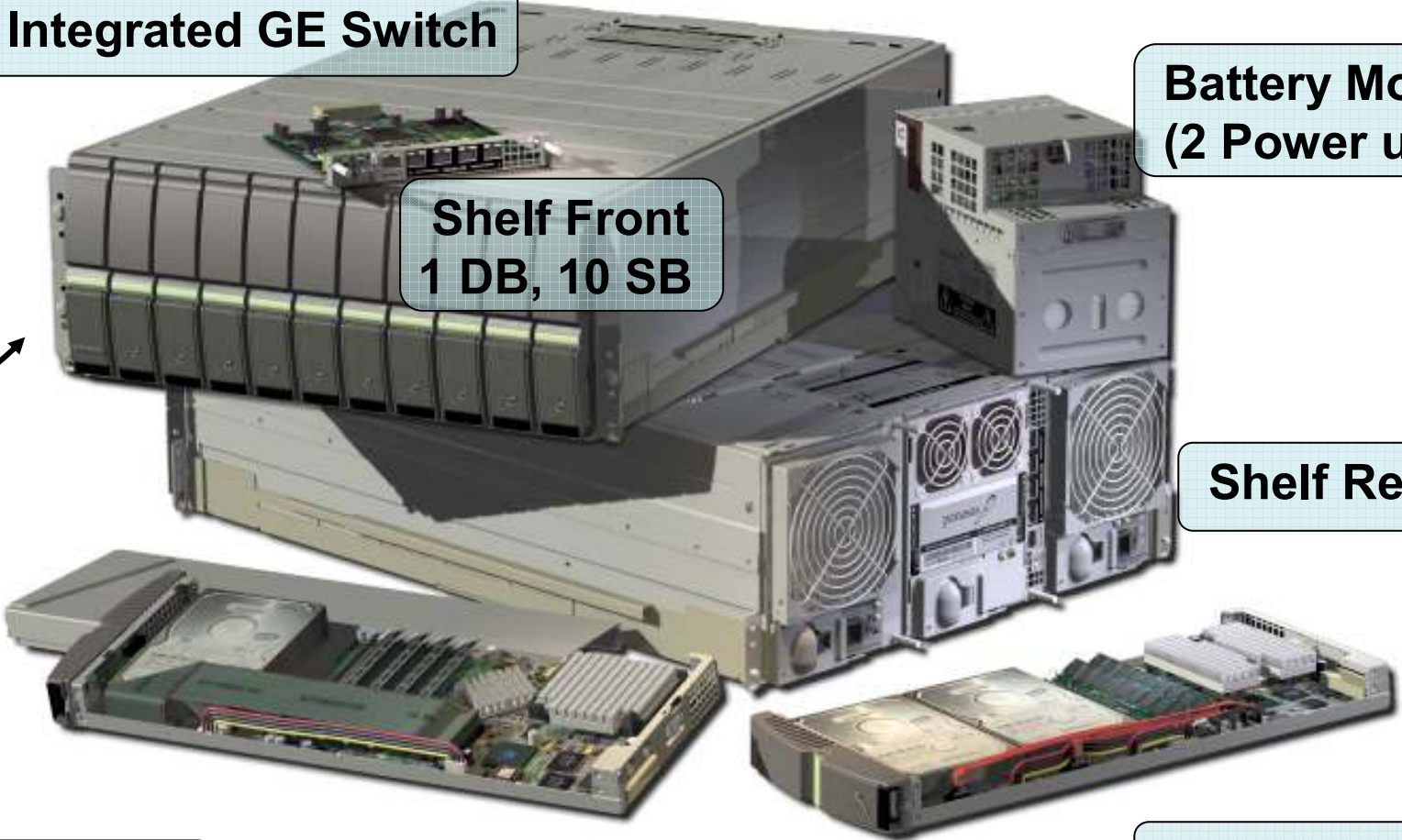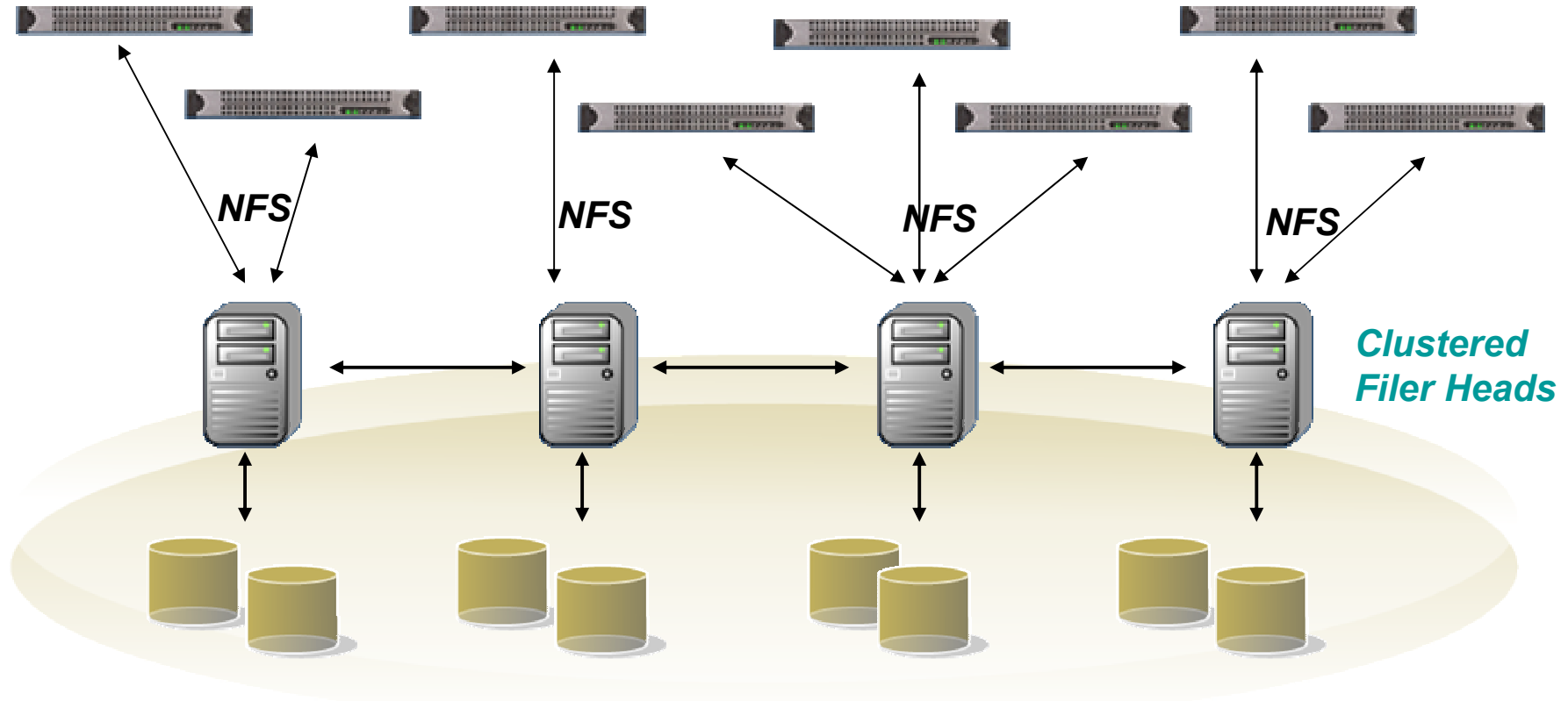
Shelf Front
1 DB, 10 SB

Shelf Rear

DirectorBlade
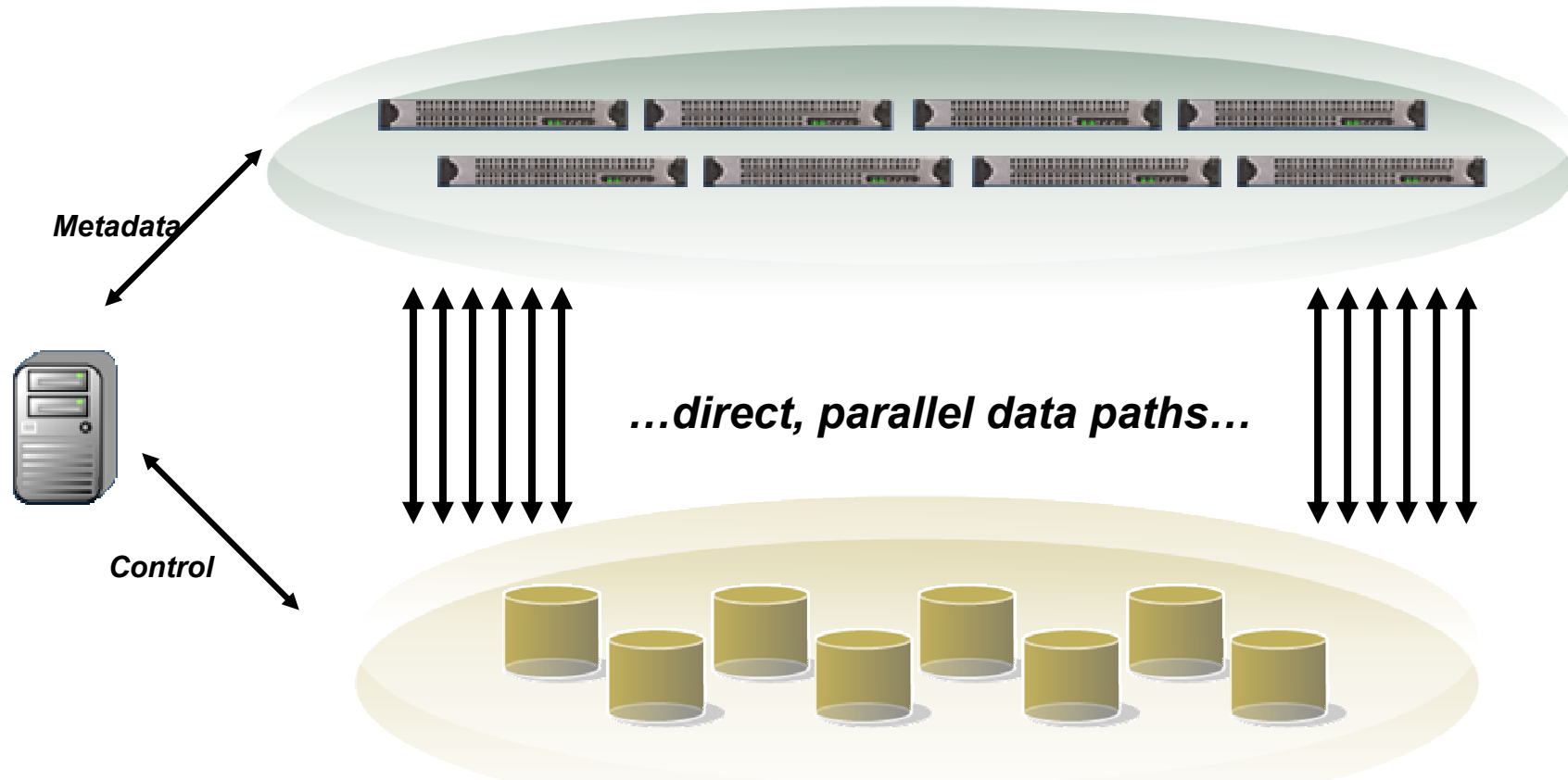
StorageBlade

Midplane routes GE, power

Clustered NAS

NFS **NFS** **NFS** **NFS**

Clustered Filer Heads

"Bridged Islands of Storage"

Filer Heads still create I/O bottlenecks

Load balancing becomes an issue

Parallel Clustered NAS

Metadata

Control

...direct, parallel data paths...

"Pool of Parallel Clustered Storage"

*I/O Bottlenecks and Management
Challenges Solved*

Panasas parallel FS

Is a network FS that is faster than local disk

Can support 100's of concurrent I/O's

Offers multiple levels of data redundancy

Can expand storage by simply plugging in more blades

Fluent 12 beta

Supports Panasas parallel FS

Ansys are claiming a 10-15 times performance increase on large simulations due to parallel I/O (compared to v6.3)

Interconnect

InfiniBand for MPI

Ethernet for I/O

N.B. separate networks!

Jupiter upgrade:

- New cluster head-nodes

- CMA management

- Compute nodes upgrade (Suse 10, Linux 2.6 kernel, compiler upgrade etc.)

- New storage – phase out ufs,qfs!

System integration:

- Same files on new & old systems
- Same scheduler (MOAB/SGE)
- Separate AMD/Intel development environments, and queues
- Jupiter for serial/task farm jobs?

Timetable:

- Panasas/network installation w/c 18th February
- Upgrade Jupiter w/c 18th February
- Handover updated jupiter/storage w/c 25th February
- Login/compute nodes installation w/c 10th March
- Acceptance tests w/c 17th March
- Handover 31st March

Information:

- www.nottingham.ac.uk/hpc
- www.nottingham.ac.uk/is/support/knowledgebase/guides/IS1309.pdf
- www.nottingham.ac.uk/is/services/software/is-machines/cats/sys_hpc.phtml
- uonlists.nottingham.ac.uk/mailman/listinfo/

Help:

- Clone admins
- hpc@nottingham.ac.uk
- IS-ResearchApplications-Team@nottingham.ac.uk
- Helpline