

Tier 1A Storage Procurement 2001/2002

Andrew Sansum
CLRC eScience Centre



Purpose of Talk

- To make available experience gained during during the 2001 Tier1A procurement
- Thanks to the Tier 1 Procurement team:
Roger Barlow, John Gordon, Roger Jones,
Jones, Dave Kelsey, Ian MacArthur, Dave
Dave Newbold



Procurement Objectives

- Obtain at least 40TB of disk capacity
 - Simple and reliable - no surprises
 - Professionally built/installed
 - Linux O/S installed/configured
 - Commodity prices
- Force solution to:
 - No more than 2TB/server.
 - Require 10MB/s per 100GB storage
- Technology unconstrained by tender



Timetable

-
- 21 September - Planning meeting
 - 12 October - Tenders to EJ
 - 2 January - Tender closed (20 responses)
 - 15 January - Evaluation/Shortlisting meeting.
 - 8 February - Finished benchmarking. Meeting to to select successful bidder
 - 15 February - Order placed
 - 27 March - Physical Installation complete
 - 18 April - Systems installed - Service starts
-



Solutions Offered

- Internal IDE with PCI 3ware controller
- External SCSI/IDE RAID controllers
- PCI based SCSI RAID. External drives
- External SCSI RAID controllers
- Fibre Channel/SAN solutions
- Network attached storage (NetApp)

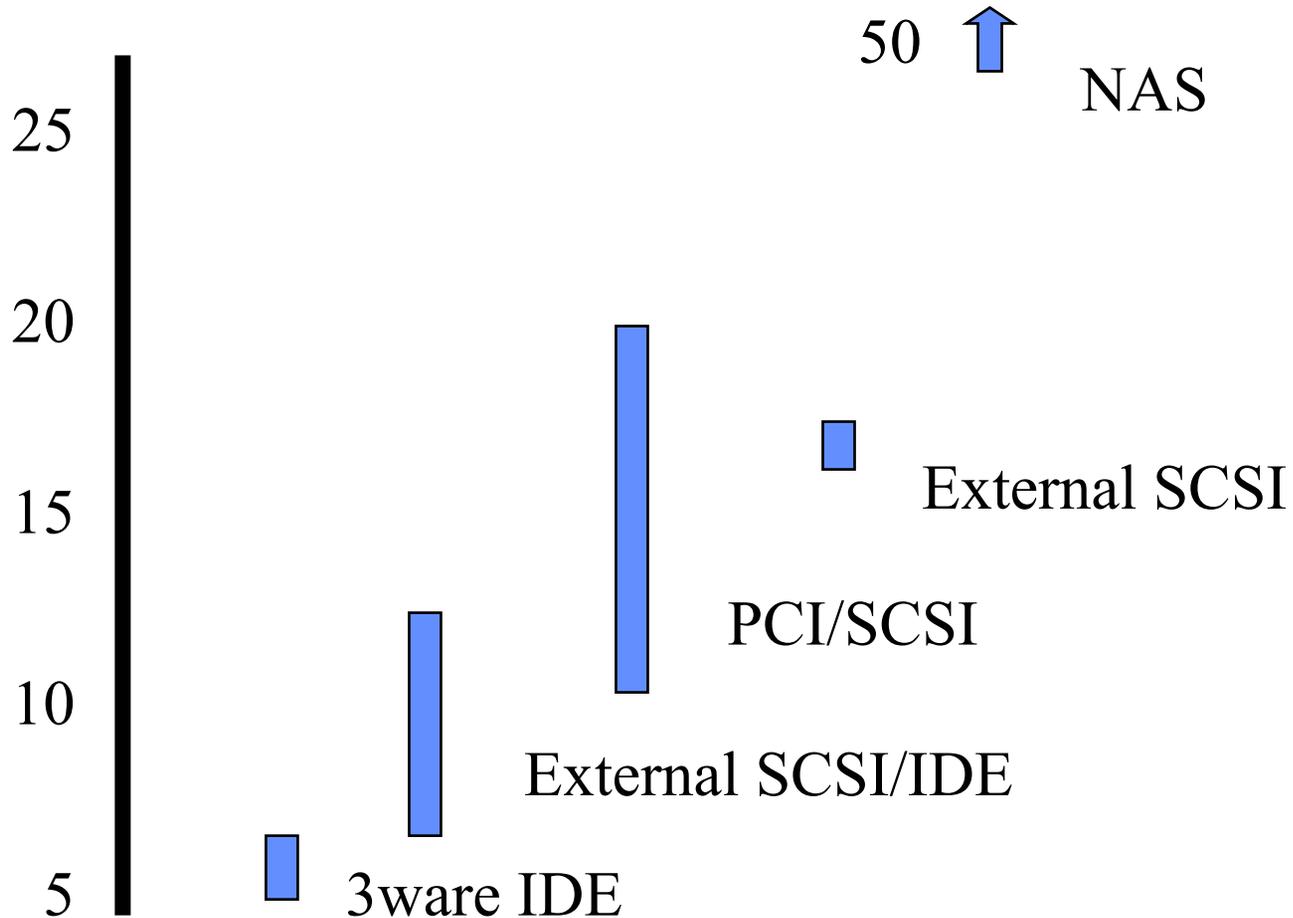


Only Shortlist SCSI/IDE

- Very competitive price
- Avoids complications of Internal PCI based based RAID controllers
- “New technology” - accept some risk associated with this choice. Try and minimise



Price/TB after RAID 5





Choosing a disk drive

- IDE or SCSI?
 - Conflicting responses. No evidence offered that SCSI is more reliable.
- Manufacturer's reputation
 - No objective information available
 - Ensure diagnostic tools/support available
- Drive performance: differences exist, but are they significant? Rely on benchmark.
- Avoid drives with problems/issues (recent IBM drives)
- Very newest drives have additional risk



IDE/SCSI RAID Systems

- Intel 80303 based controllers (12 drive)
 - Zero-D (www.zero-d.com)
 - RAIDstorage (www.raidstorage.uk.co)
 - Transtec (www.transtec.co.uk)
- Intel i960 Based controllers (8 drive):
 - Older and slower - don't buy these (Akhter)
- Infotrend IFT6300 (PowerPC) (8 drive)
 - www.infortrend.com (supplied by a number of companies - eg Streamline)



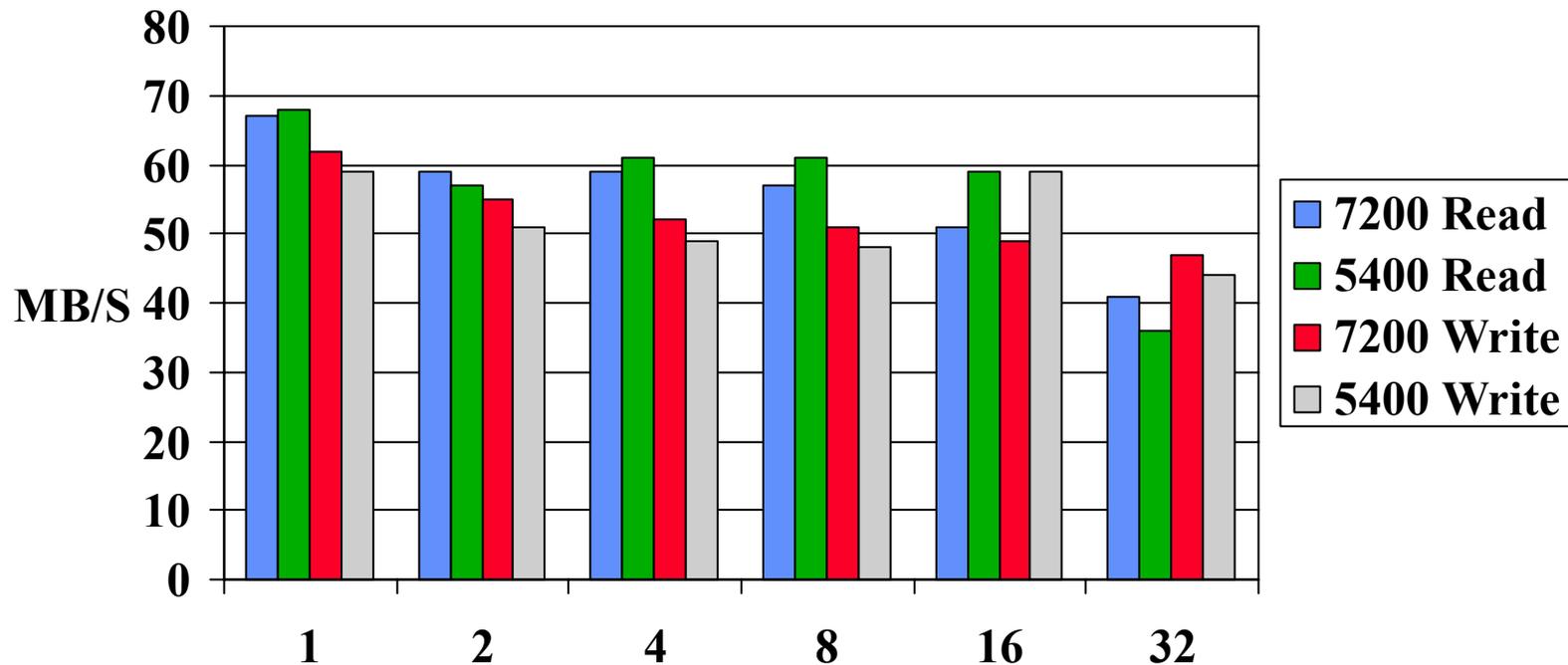
Benchmark

- Mainly used IOZONE but sanity check using Bonnie
- Well defined protocol (stability):
 - 128MB system memory
 - New, empty, 150GB filesystem
 - Standard script
- Careful choice of kernel/filesystem
- Take care of Driver/Firmware updates



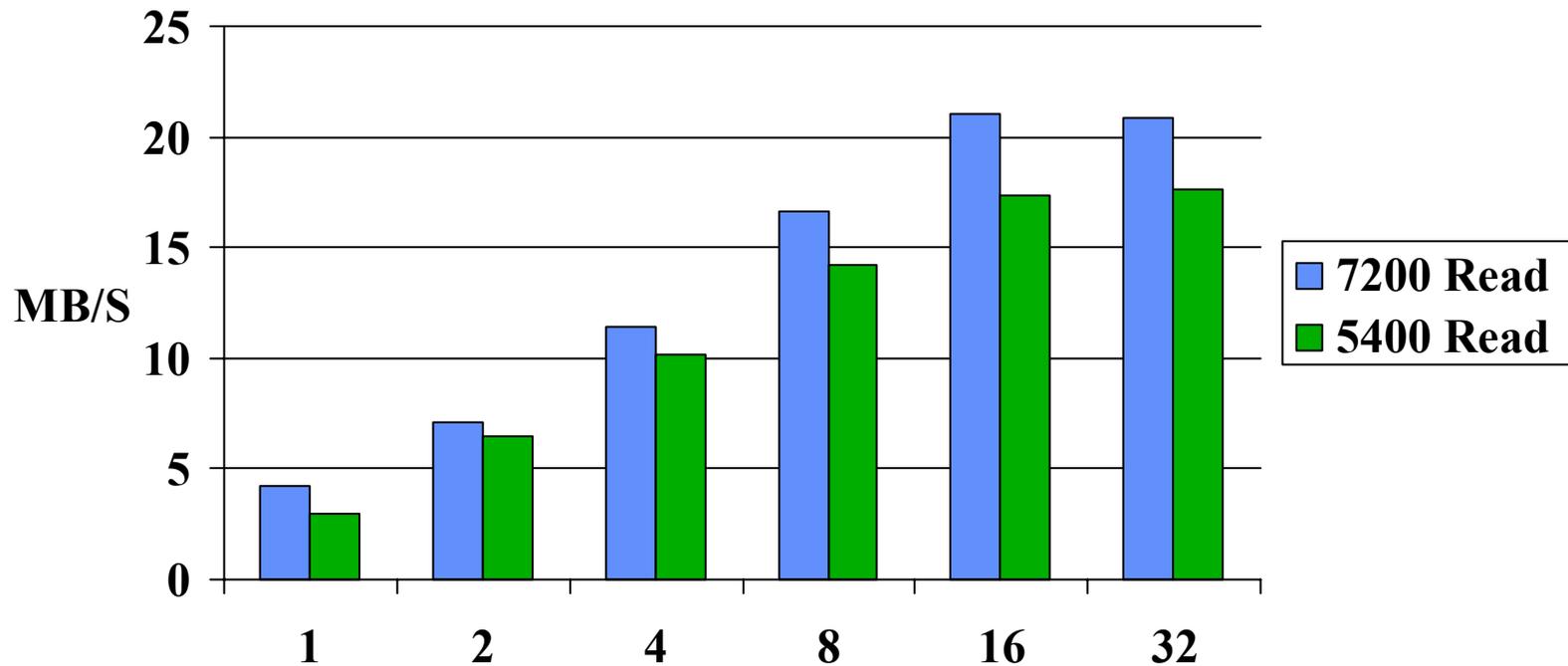
5400v7200 Sequential

2.2.19-ext3





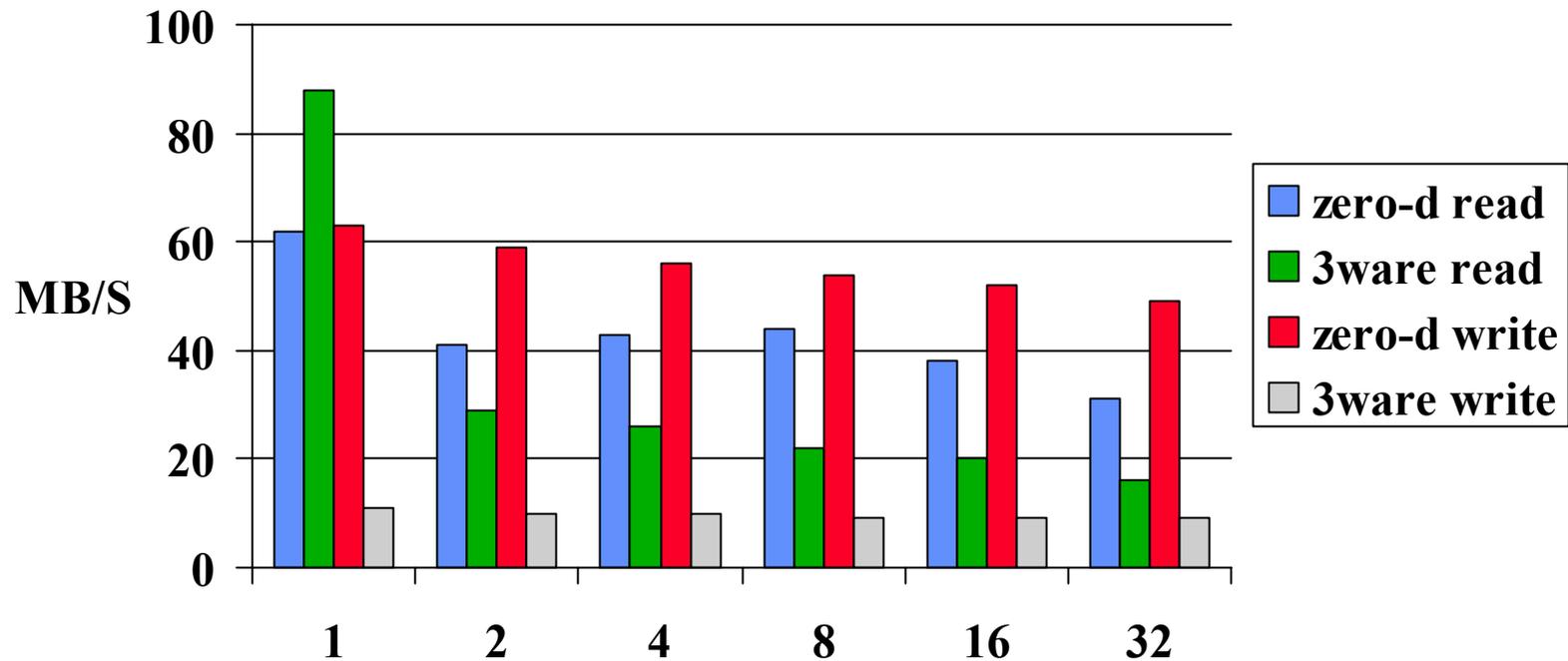
5400v7200 Stride Read





ZeroD v 3ware (6800)

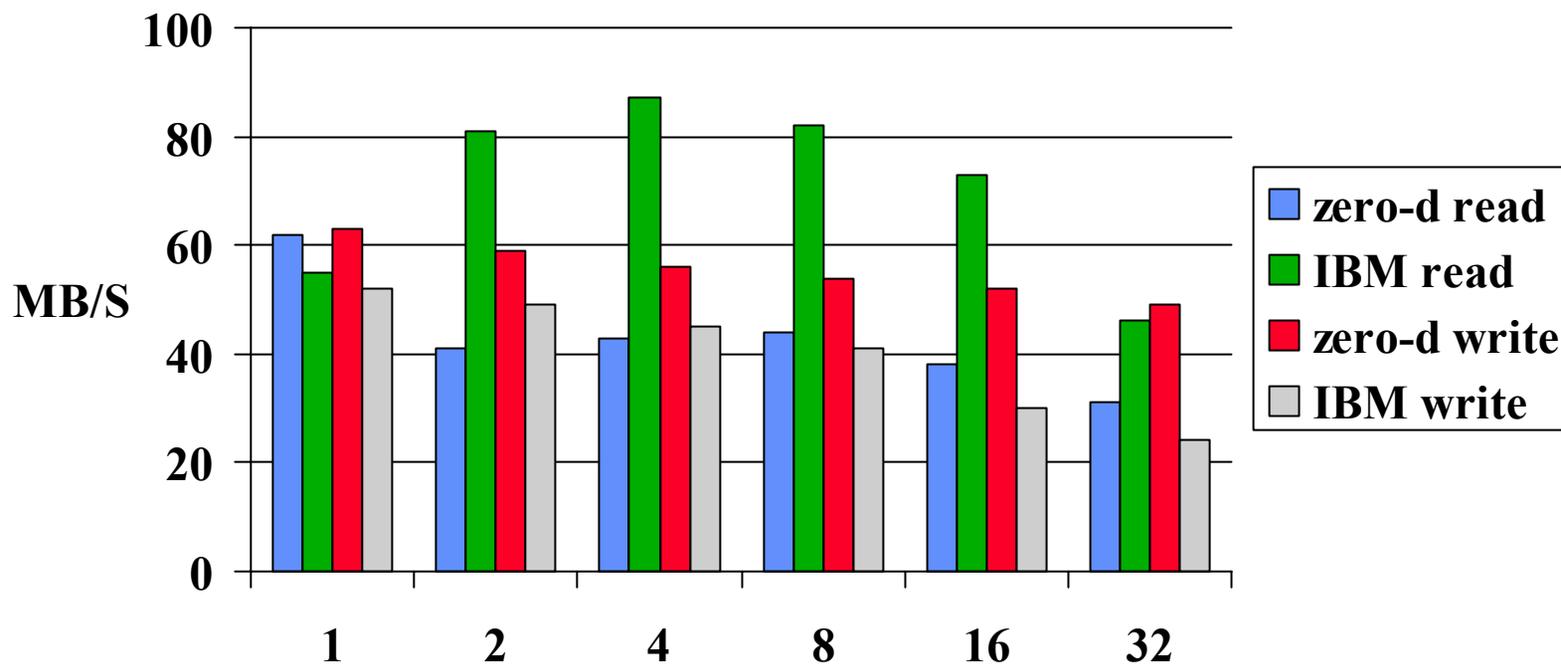
2.4.17-ext2





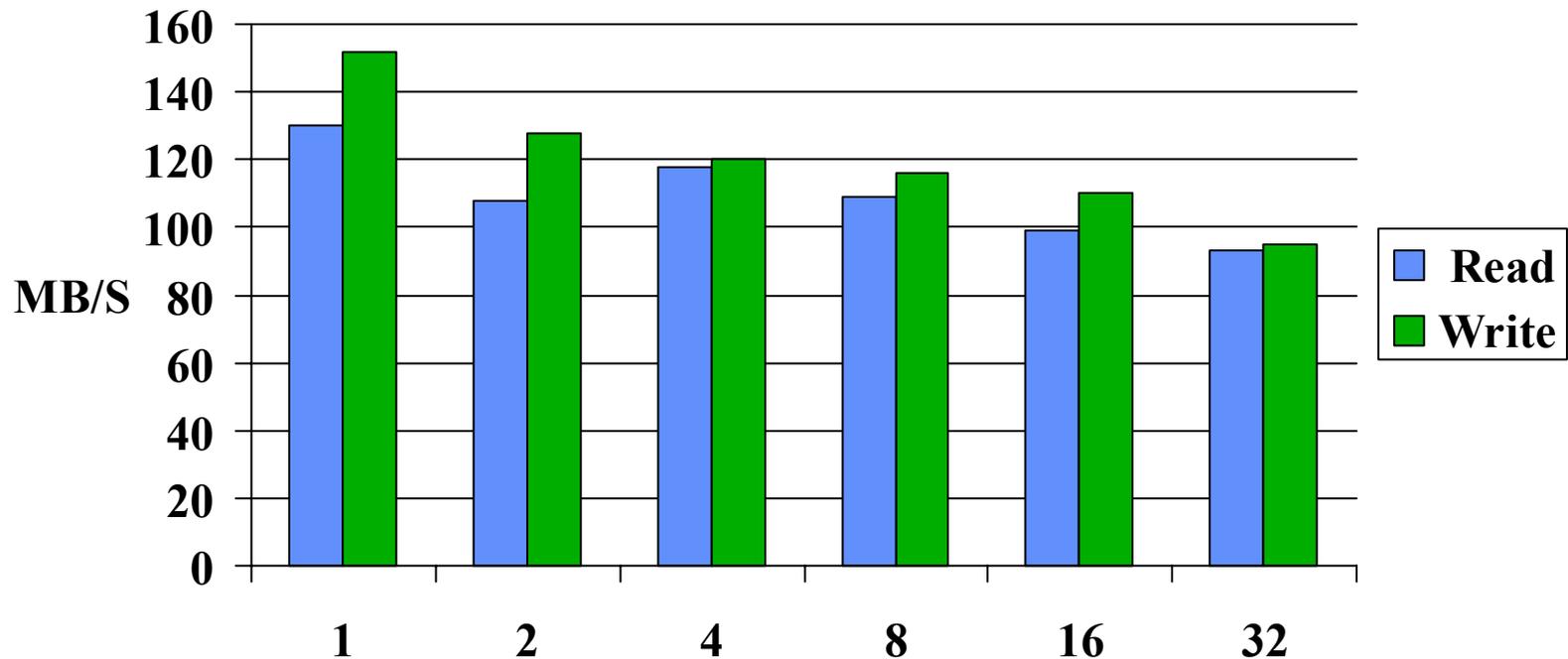
ZeroD v IBM FAST

2.4.17-ext2/2.4.5-ext2





2 RAID5 Arrays+S/W Stripe





NFS Performance

- Not fully tested:
40MB/s write
30MB/s read



Post Tender Experience

- 6 Months of running experience
 - Initially many drive read errors - reported as remapped remapped blocks on console. Bad batch of disks - now now being replaced by Maxtor
 - Concern over method and algorithm used by controller controller to handle block read errors (when drives not drives not ejected promptly remap area can overflow). Working with vendor.
- Since August reliability is much better. Solution Solution can probably be made to work well but but some way to go yet.



Latest Concerns(1)

In preparing for this year's purchase the following issues arose:

- 1) Most IDE manufacturers are downgrading warranty from 3--> 1 3--> 1 years. Some (eg Maxtor) are introducing enterprise enterprise class IDE disks at what is presumably a premium premium price
- 2) Short term disk capacities:
 - We bought 80GB*7200rpm
 - Maxtor's new Enterprise class 7200 rpm is already available up to 250GB.
 - Not sure about availability though. 320GB 5400 rpm drives also available.
- 3) Short term RAID controller enhancements
 - 16 drive units available for Acusys controller
 - 12 drive units now available for Infotrend.
 - Unlikely to get significantly faster controllers in the next few months. months.



Latest Concerns(2)

- 4) RAID controller speed Big concern because controllers are controllers are no faster than last purchase, bigger disks disks plus bigger RAID arrays means less performance per performance per GB storage. Do we care - maybe.
- 5) Only really two manufacturers of external IDE RAID controllers (cf lots of PCI based solutions).
- 6) Serial ATA RAID controllers will be available very soon now. IDE is becoming more mainstream for the enterprise - possibly we will see more external RAID controllers in the future.



Conclusions

- Tender process was very successful
 - Wide range of solutions offered
 - Wide range of skills identified
 - Some competitive prices
- SCSI/IDE has good characteristics for larger HEP HEP installations
- Too early to say if technology choice correct
- Establish a relationship with RAID supplier