

ATLAS Tier-1 Storage Management and Experiences @ TRIUMF

Xinli(Simon) Liu

TRIUMF

HEPIX Fall 2011

@ Vancouver, Oct. 24-28 2011

Storage Services Overview

- **Serving only ATLAS experiment**

In past one year

- 67% are reading, 25-135TB/day, 65-345k file transfers
- 33% are writing, 1-31TB/day, 20-196k file transfers
- 99.79% overall successful transfers rate, including all conditions
- No data lost since we started

- **dCache (1.9.5 in production, 1.9.12 in PPS)**

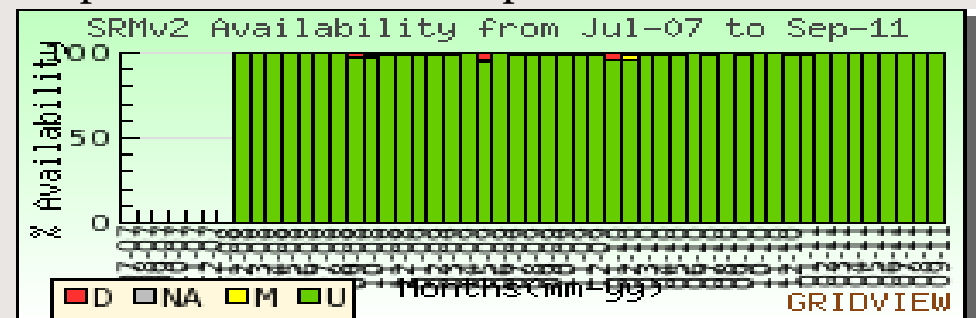
- 1 instance, provides SRM endpoint for both disk and tape (via HSM interface)
- Chimera name space, access protocols: dap, gridftp

- **Disk system (2.1PB usable)**

- SAN storage 700TB (DDN connected with 7 disk pool nodes and 7 hsm pool nodes)
- Direct attached storage 1.4PB(21 SUN 4540)
- Add 4 PB usable into production in Q1 2012

- **Tape Storage total capacity 5.5 PB**

- 14 LTO4 drives, 2000 LTO4 cartridges
- 10 LTO5 drives, 2600 LTO5 cartridges

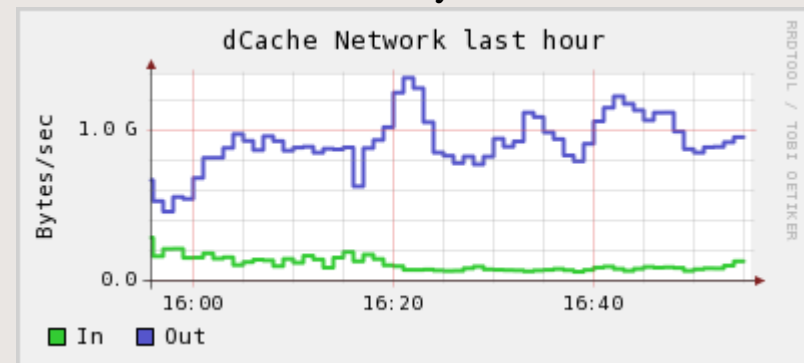


Disk System Hardware (DDN)

- DDN S2A9550(IBM DSC9550) acquisition from IBM in 2007, the first installation in Canada and still performing well
- Mixed drives in two racks (260 x 750 GB and 700 x 1 TB), 52 LUNS (size varies between 6TB-18TB)
- Multipath: priority assigned to each port for performance. XFS
- 14 hosts servers, dual 4Gb FC/server. At least 2 hosts in one soft zone for backup
- 4.5% disk failure rate since installation, but 58% disks failed in 2011 (to Oct. the fourth year)
- Now using log condition to decide on disk replacement instead of waiting until disk fails
- 22% PSU failures, including controller's, possibly due to high temperatures
- Temperature delta between front and rear disks in the same drawer is around 10 degree
- Most strange issues happened in the first year due to fw issue, now it is very stable
- Wonderful support from IBM, plus on site parts

Tested throughput 3.2GB/sec,

normal throughput during reprocessing shown right

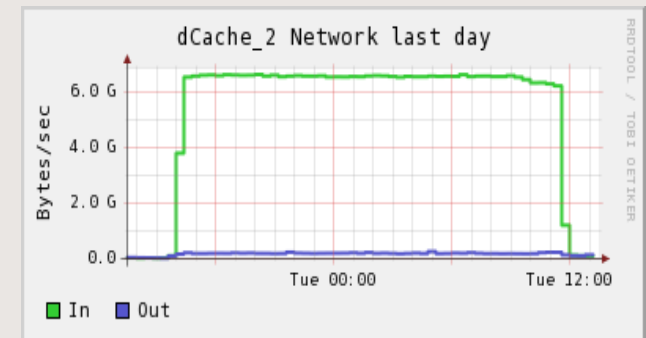


Disk System Hardware (SUN)

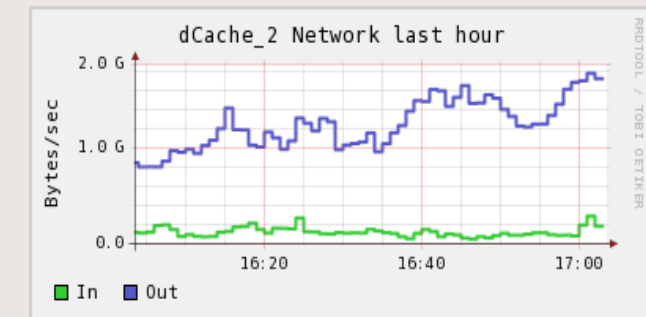
- 21 SUN Fire X4540, acquisition from SUN by end of 2009, in production in early 2010
- 12 cores, 64GB memory and 4x1Gb aggregated connections/server
- 18 servers have 2TB drives, 3 servers have 1TB drives, 20 in production, 1 for PPS (also as a spare node)
- 4 ZFS /server, raidz2, version 22, 3 times scrub/year, arc_max 48GB
- Each array across 6 control cards (the metric blow, each color presents an array, safe for one controller card failure, red is for Solaris UFS)

| | t0 | t1 | t2 | t3 | t4 | t5 | t6 | t7 |
|----|------|------|------|------|------|------|------|------|
| c0 | c0t0 | c0t1 | c0t2 | c0t3 | c0t4 | c0t5 | c0t6 | c0t7 |
| c1 | c1t0 | c1t1 | c1t2 | c1t3 | c1t4 | c1t5 | c1t6 | c1t7 |
| c2 | c2t0 | c2t1 | c2t2 | c2t3 | c2t4 | c2t5 | c2t6 | c2t7 |
| c3 | c3t0 | c3t1 | c3t2 | c3t3 | c3t4 | c3t5 | c3t6 | c3t7 |
| c4 | c4t0 | c4t1 | c4t2 | c4t3 | c4t4 | c4t5 | c4t6 | c4t7 |
| c5 | c5t0 | c5t1 | c5t2 | c5t3 | c5t4 | c5t5 | c5t6 | c5t7 |

- 1.6% disks failed since installation, pretty stable, higher failure rate is expected in coming years
- No on site parts, no parts center in the city, have to wait for 2-3 days, sometimes one week and more. Now using disks from PPS machine as spare



Throughput tested



Throughput during reprocessing

Tape System Hardware (IBM TS3500)

- **Currently (5.5 PB capacity):**

- 1 base frame L53+1*D53+3*S54(high density frame)
- 2000 LTO-4 cartridges(0.8 TB each , native)1.6 PB total
- 2600 LTO-5 cartridges(1.5TB each,native) 3.9 PB total
- 14 LTO-4 drives, 10 LTO-5 drives

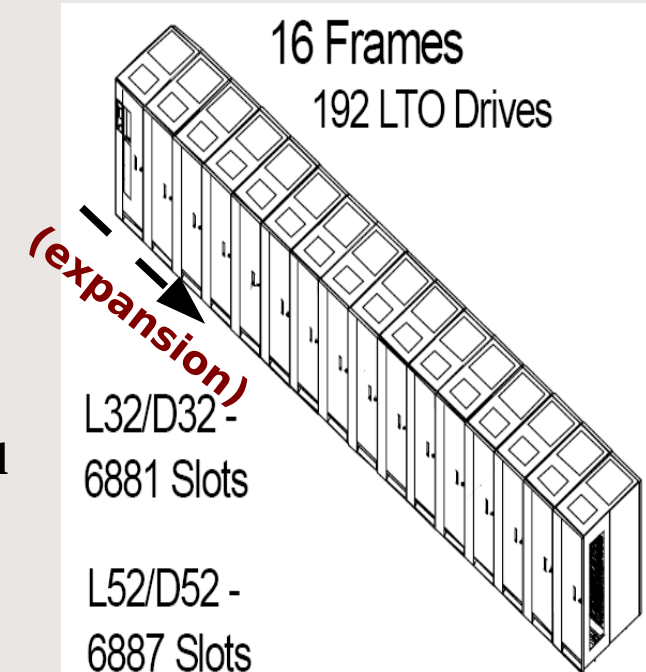
- **Tape operations: non-proprietary software -- tapeguy, fully developed at TRIUMF, gives us total control.**

- **Tape drives accessed by 7 hsm hosts, 100 TB dedicated disk buffer**

- **LTO-5 drives and cartridges have been tested(30TB w/r, including LTO4/5 compatibility test), but not in production yet due to limited SAN connection.**

- **Hardware seems quite stable, only 3 drives replaced since installation**

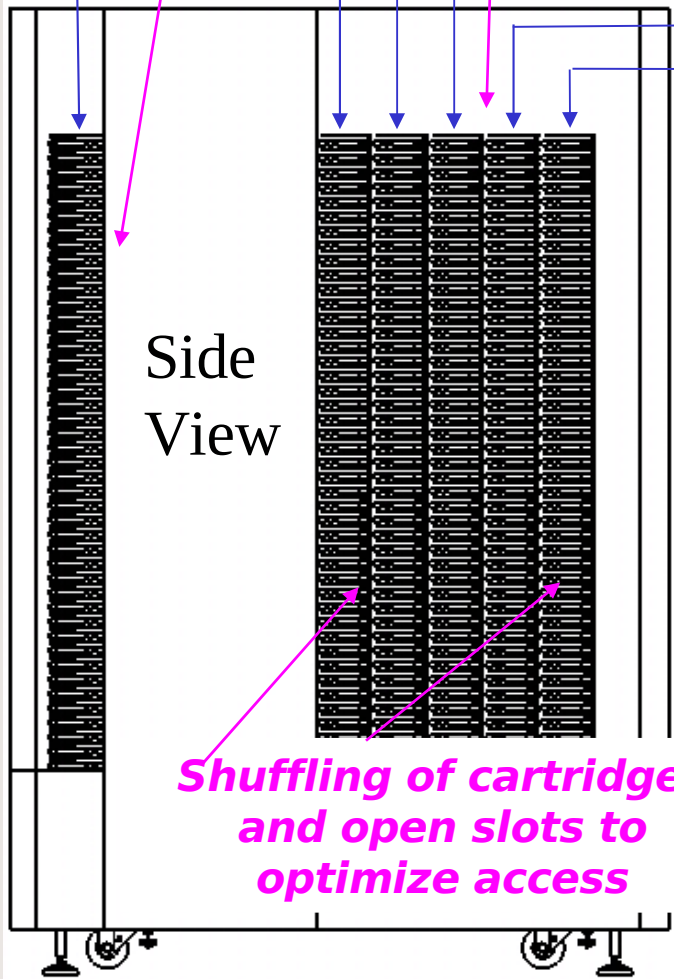
- **Limited by server room space, can't go with traditional frame expansion, we chose high density frame S54**



High Density Frame – Up to 3X Capacity

Nearline cartridge cache (application independent)

Least frequently used



Tier 0 Slots (220 LTO)

Tier 1 Slots (220)

Tier 2 Slots (220)

Tier 3 Slots (220)

Tier 4 Slots (220)

Tier 5 Slots (220)

Library provides automatic cache management

- Cartridges are always demounted to tier 0 or tier 1
- Track most recent access for each cartridge
- As needed, destage Least Recently Used (LRU) cartridge from tier 0 to tier 1

No management required, transparent to end user



1320 LTO cartridges /S54

Compare with

440 LTO cartridges/D53

High density automation daemon:

<http://www.youtube.com/watch?v=L1HqNhhKCCM>

Experience with High Density Frame

- **First one HD frame installed in early 2010**

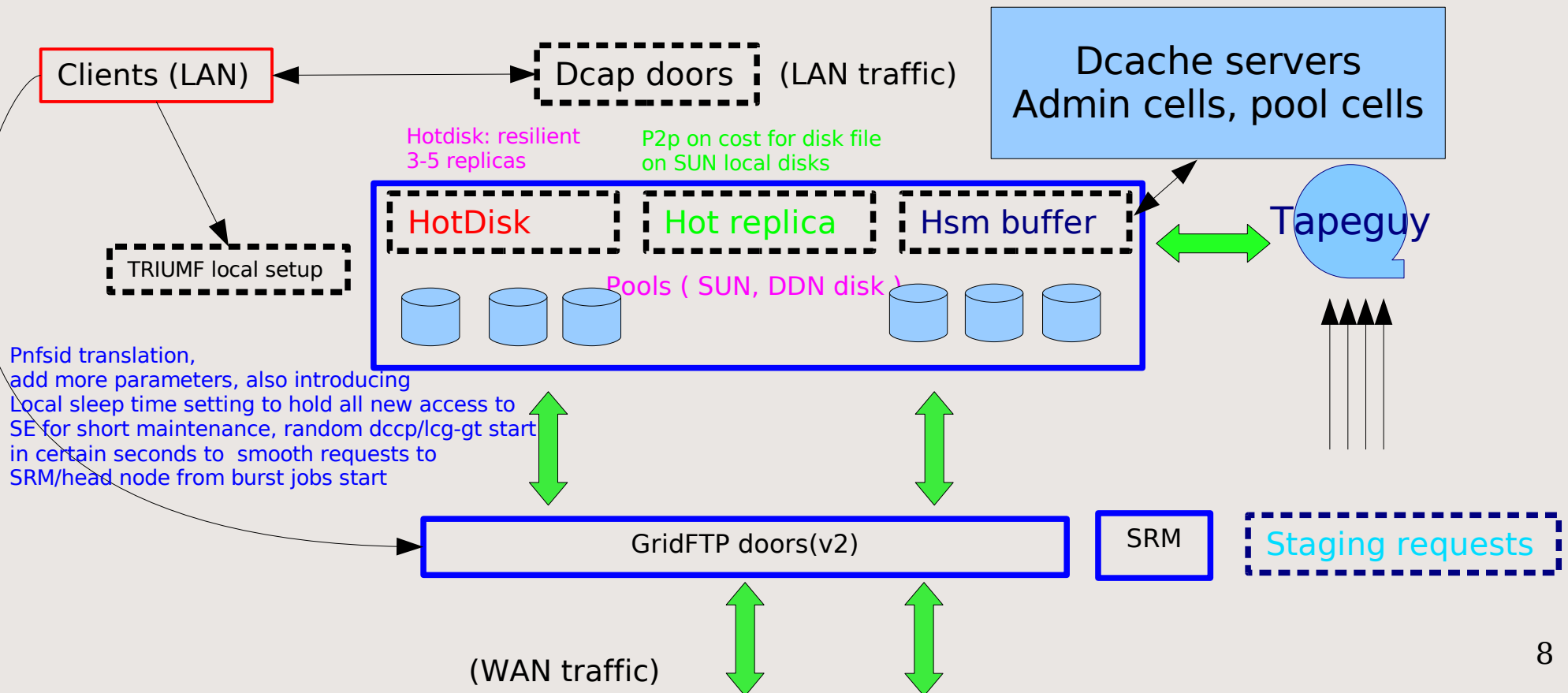
- Serious fw issue, with vio option enabled, some SCSI elements could become 0, therefore library inventory corrupted, some tape volumes were duplicated, library didn't know which is right
- Solution: identified the problem and disabled vio, in the main time, deployed scsi tool to force library to scan the tape label, in addition, in tape system code, we added more strict inventory check
- After several fw releases, IBM resolved the issue, we applied the fw in Oct. 2010

- **Another two HD frames installed in June this year**

- Inventory issue happened again, but went off after full inventory with audit
- With 3 logical libraries running on one physical library, have only seen once that the first logical library only reports 10 % tapes of its inventory, reboot the library resolved the issue
- IBM suggested to have new fw upgraded, but we are not confident with the latest fw (with new features introduced)
- System working well now, will keep an eye on the issue

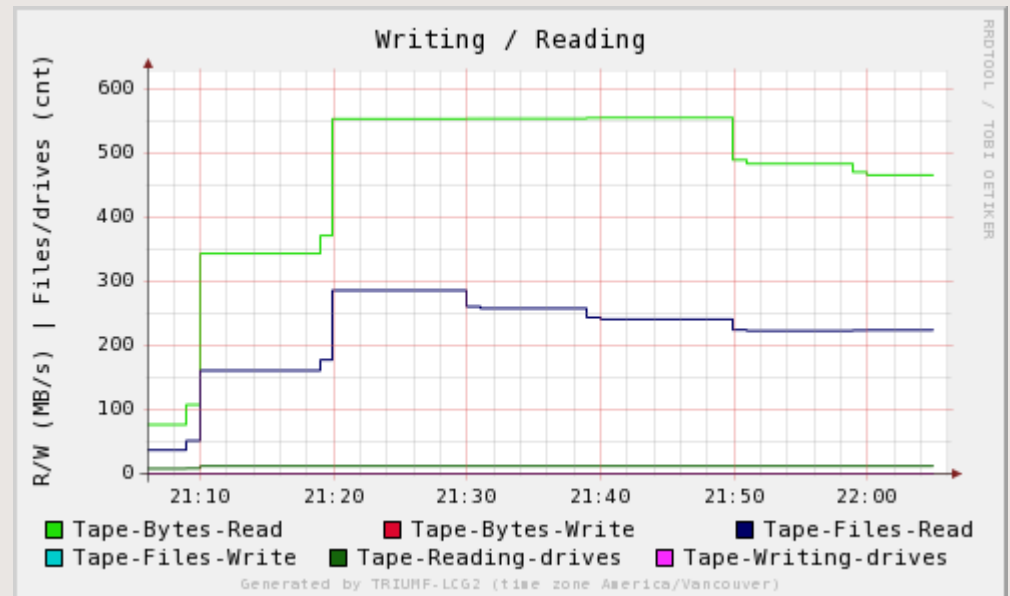
dCache

- One dcache instance: one head node, one SRM node and one dedicated billing node
 - Initially started in spring 2005, with one server, then 3 pool nodes, then 5,7 then DDN storage, then SUN servers
 - SRM/head nodes with pnfs were heavily loaded, migrated to chimera in 02 / 2010
- Chimera is such a wonderful namespace provider, looking forward to test nfs4.1



Tape system -- Tapeguy

- **In production since 2007, still performing well**
 - Files grouping
 - reordering
 - Minimize tape mounts
 - Maximize reads per mounts
- **No data lost**
- **No major change since 2009, but did some fixes**
- **Well tested for multi-libraries, no problem for LTO4/5 mixed environment**
- **New release is under testing**
- **Challenge**
 - For each time reprocessing, don't get enough stagings for reordering, which is causing unnecessary remounts(see later slides for more detail)
- **Tape reclaim**
 - Deletion is not real erase, reclaim process need retest

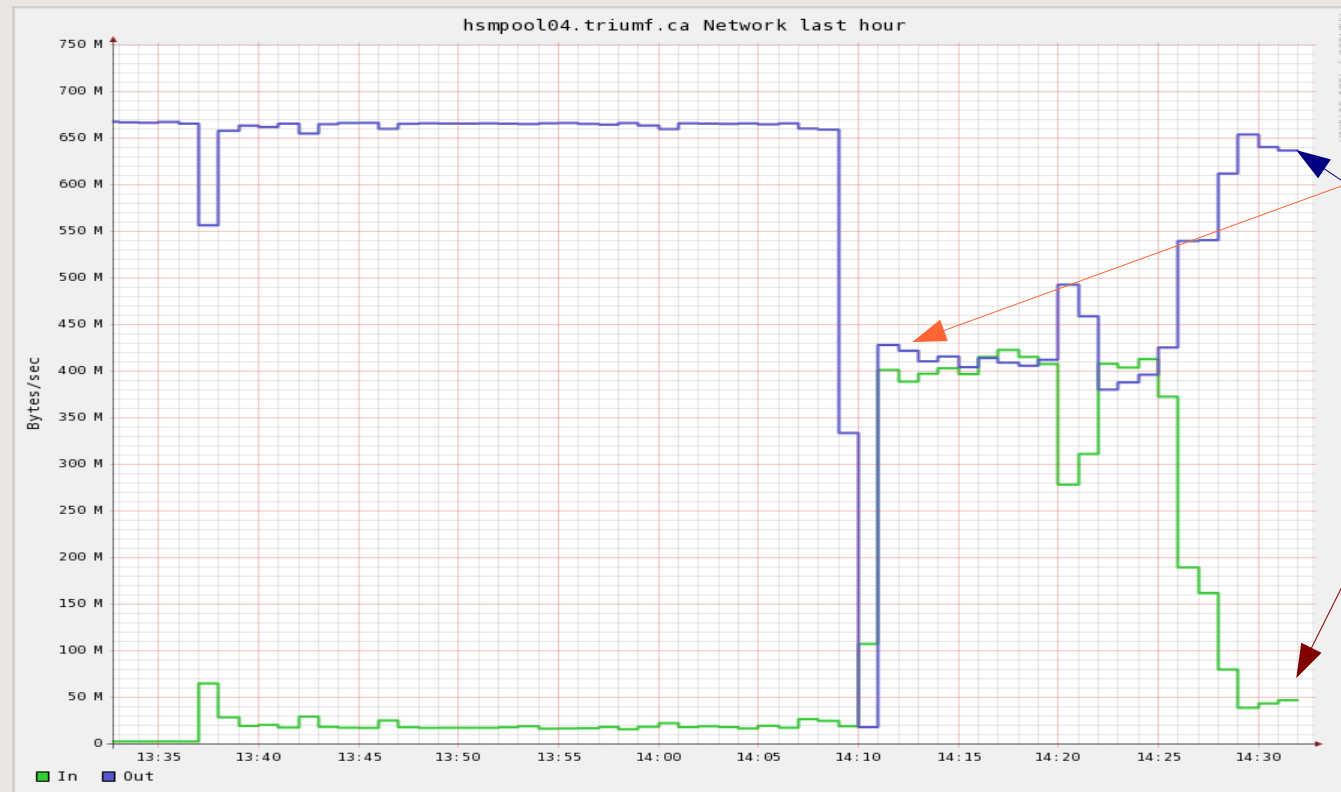


Common throughput during reprocessing
 Peak number >1GB/sec using 12 drives
 for reading

Experience with 10G network cards

- Started with NetXen Incorp. NXB-10gXSR card, on SL4 in 2007
 - Driver issue, we wrote a script to detect the card and reload the module to make it back online
 - Data corruption issue, replaced one card with Myricom
 - The issue got better after Apr. 2010 with new driver, only see the issue occasionally when heavy loaded
- Myricom card
 - Was introduced in when debugging the NetXen card, not performing the best, but stable on both SL4/5
 - We replaced all old NetXen cards with Myricom ones, but Issue is that vendor(ibm) doesn't support the card
- NetXen card support issue appeared after Qlogic bought NetXen
 - Got Qlogic cards when new servers came in, it worked beautifully when write into /dev/null, it was a disaster when write file to file system. Qlogic replaced them with NetXen multifunction card, it works, but again IBM doesn't support it.

Experience with 10G network cards (II)



Simultaneous W/R start

- Read – 10 dccp
- Write – 10 lcg-cp with 10 streams/each

Problem appears 20 minutes after

- Beside the support issue, With current NetXen multifunction card, it works well for most situations, but it still suffers a bit when simultaneous w/r (multi-thresholds transfers)
- Transfer still on going without data corruption, just writing performance gets suffered, <5MB/sec/w in LAN
- 7 hsm pool nodes are using this type of card, with limited grdftp movers

Experience with 10G network cards (III)

Tools and Experience for a pool node test

- **Iperf (basic functionality test)**
- **Iozone to make sure file system is not the bottle neck**
- **Gridftp server**
 - Write to /dev/null and read from /dev/zero, then to/from file system
- **Scaling tests with dcache setup in PPS**
 - Get work nodes(1-2 node/chassis) involved for the test, enough to saturate the node resources.
 - Write, read and w/r. Sometime use different subnet combination for further investigation
 - Data corruption checks for every step
 - Practical enough to understand how a disk server behaves under different situations, then tuning can be applied to prevent the worst-case from happening

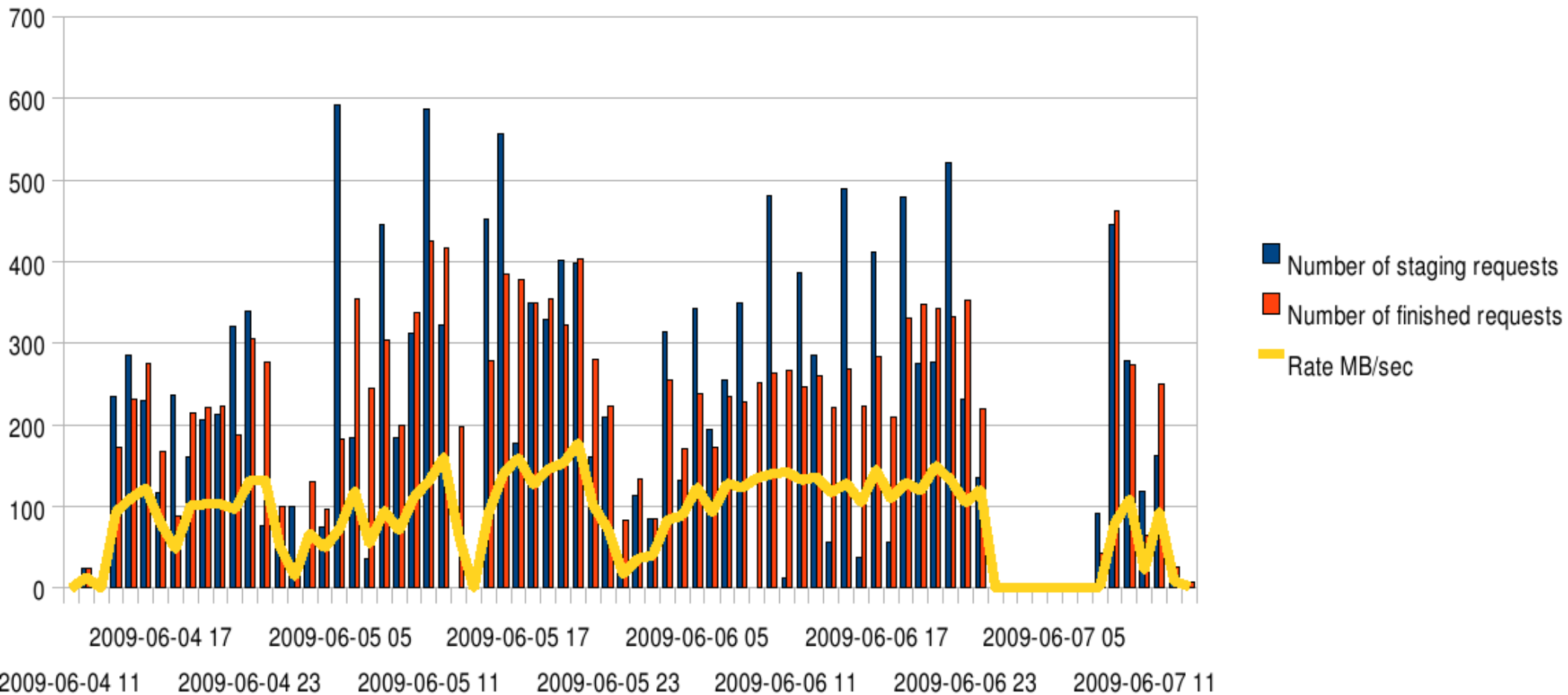
Recent tests for next storage server purchase

- **Test node sponsored by vendor**
- **Dual intel x5650(12 cores), 48GB memory, 2 SSD disks, 4 x 10Gb network cards**
 - Ethernet controller: Brocade Communications Systems, Inc. 1010/1020/1007 10Gbps
 - QLogic Corp. 10GbE Converged Network Adapter
 - Emulex Corporation OneConnect 10Gb NIC
 - Mellanox Technologies MT26448 [ConnectX EN 10GigE, PCIe 2.0 5GT/s]
- **Two DDN luns mounted, XFS file system**
- **On both SL5 and SL6, using dcache 1.9.12-10/1.9.5-28**
- **Tests are nearly done , we are happy with the test result, sorry, no result can be released here**

MSS performance (I)

- The following three slides data comes from step09 in 2009, still applicable to current situation

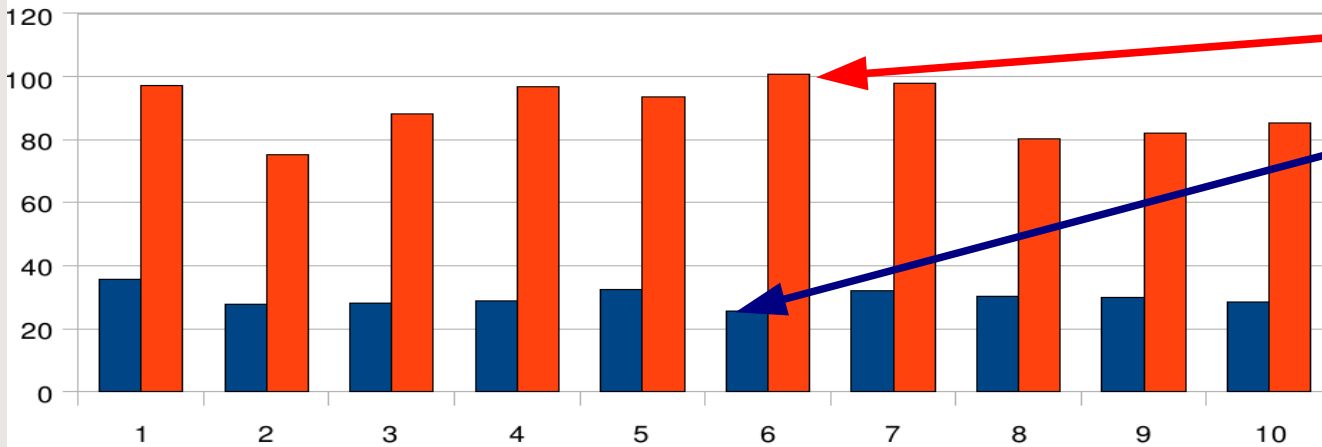
Number of staging requests, finished requests, Rate(MB/sec) (Jun04-Jun08)



Total 8 LTO4: 2 for writing and 6 for reading

MSS performance (II)

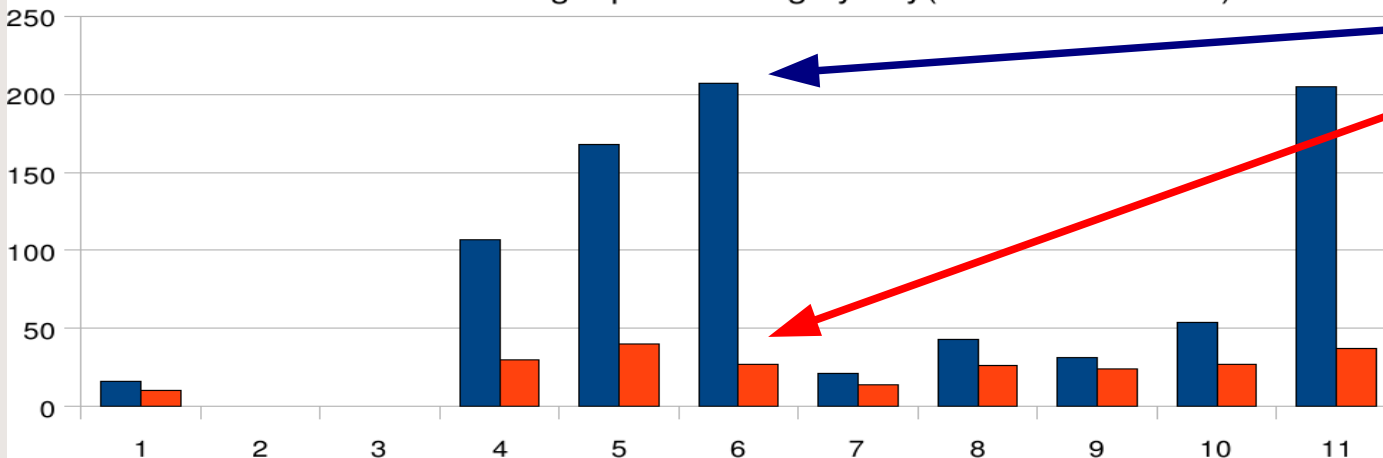
Reading rate: overhead vs wo overhead(start from Jun2th tp Jun12)



Read rates per drive:
80-100MB/s without overheads

Down to ~30 MB/s overall rate,
with mount, seek, etc. overheads

Reading tape mounting by day(start from Jun 2th)



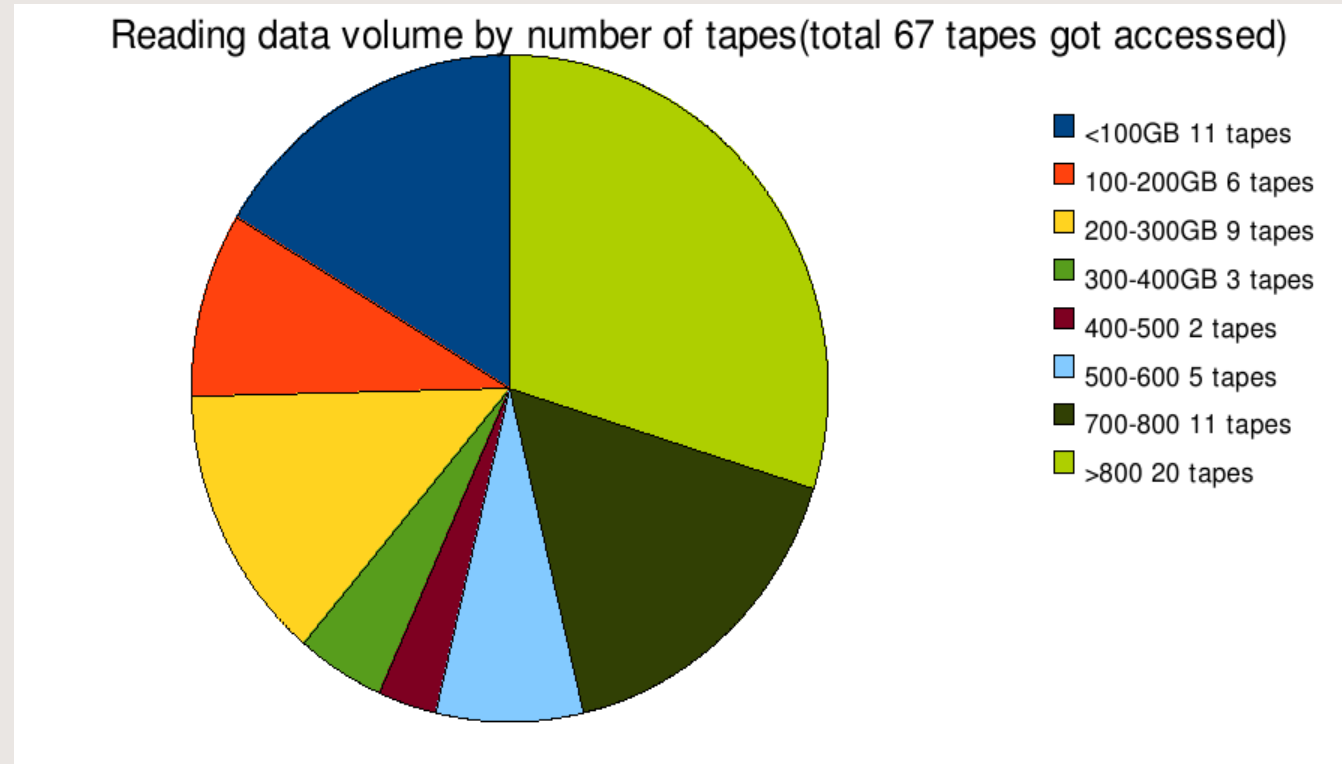
Number of mounts ~200
Number of tapes read ~25

Issue correlated with:
nprestage in Panda conf
Number of Assigned jobs

Tape system never got a long queue of requests to be able to optimally sort it to minimize tape mounts

MSS performance (III)

The right plot shows that for most of the tapes, almost the full tapes are read in the end



Indication that there is room for improvement, if reading full tapes in one mount, could push rates up to x3

Conclusion and Future Plans

- Current storage system for both disk and tape is performing well according to current use cases of the ATLAS computing model
- dCache is a very nice solution for us and scalable
- We could go with either direct attached storage or SAN storage for our future expansion
- Will do large scale tests on dCache 1.9.12 and possibly nfs4.1 as well during the commissioning of new storage in early 2012.
- Closely monitoring ATLAS jobs file access patterns and tuning accordingly.
 - transitioning to dcap io access for analysis jobs
 - One CE is using CVMFS, another two are waiting in queue
- Keeping an eye open for other storage solutions out there

Thank You , Merci !

Acknowledgements:

Canada Foundation for Innovation (CFI)
British Columbia Knowledge Development Funds (BCKDF)
National Research Council (NRC)
National Science & Engineering Council (NSERC)
CANARIE
BCNET
HEPNET



4004 Wesbrook Mall

Vancouver, B.C. Canada V6T 2A3

Tel: 604 222-1047 Fax: 604 222-1074

www.triumf.ca

Extra Slides

Overall Storage Usage Evolution

