

TRIUMF SITE REPORT

STEVEN MCDONALD
HEPIX NOV 2007
ST LOUIS

PERSONNEL CHANGES & ADDITIONS

- ✧ **NEW TRIUMF DIRECTOR**

- ✧ Nigel Lockyer May 2007 - May 2012
 - ✧ From Penn State, A former head of CDF



- ✧ **NETWORK & COMPUTING SERVICES**

- ✧ Corrie Kost - retirement June 30th
- ✧ Kelvin Raywood - Corrie's replacement

- ✧ **SCIENTIFIC COMPUTING SUPPORT**

- ✧ Chris Pearson - DAQ system support

- ✧ **ATLAS TIER-1**

- ✧ Andrew Wong - DB Admin
- ✧ Asoka De Silva - user support - root - athena
- ✧ Joe Steele - user support - root - athena
- ✧ offer made - hardware technician

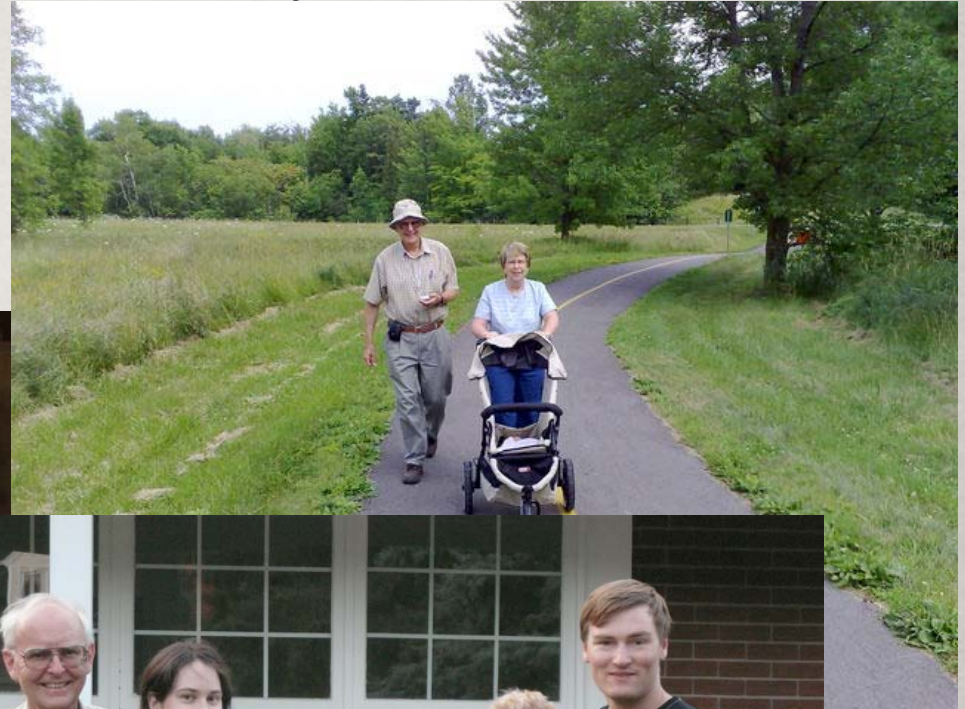
Total 25 people across
4 primary groups

Corrie Kost Retirement

TRIUMF 1971 - 2007 (37 years)

One of the original HEPiX Members

Corrie & Lndia first time Grandparents one week before he retired

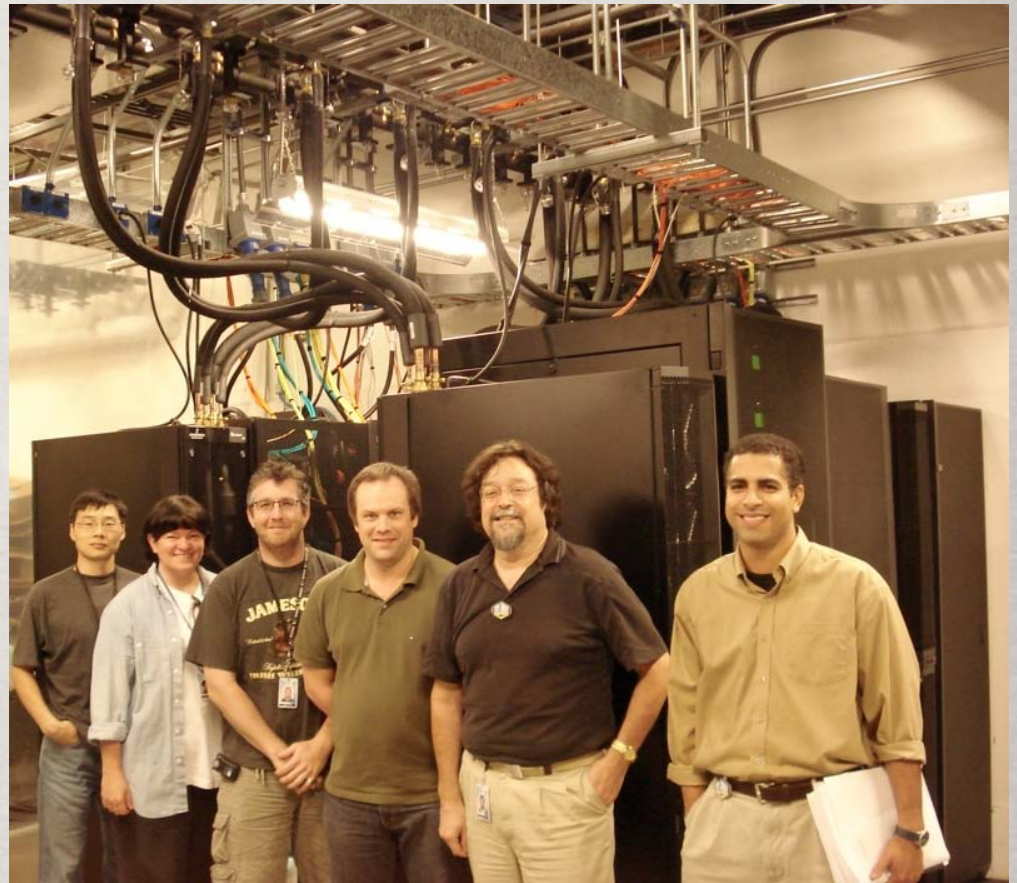


Corrie, Kiera, Lydia



TRIUMF ATLAS TIER-1

- ☼ Dedicated facility - funding approved in early 2007 for 23.5M over 5 years.
- ☼ RFP out in May 07, Installed in a newly furbished data center, fully operational by end August
- ☼ ~5% of Atlas ~7% of computing resources
- ☼ 9 new positions since 2005 **all fully dedicated** to ATLAS TIER-1 operations
- ☼ Room capacity can meet our commitments up to 2011



Some of the Tier-1 Team
Simon, Denice, Chris, Rod, Mike, Reda ,
DB Admin, User Support (2), Technician recently hired

TRIUMF TIER-1 COMMITMENTS

Cumulative numbers (Canadian contribution only)

<i>Year</i>	<i>2007</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>	<i>2011</i>
<i>CPU (kSI2k)</i>	<i>228</i>	<i>1304</i>	<i>2046</i>	<i>3569</i>	<i>5092</i>
<i>Disk (TB usable)</i>	<i>156</i>	<i>716</i>	<i>1418</i>	<i>2844</i>	<i>4096</i>
<i>Tape (TB)</i>	<i>111</i>	<i>554</i>	<i>1077</i>	<i>2067</i>	<i>3235</i>

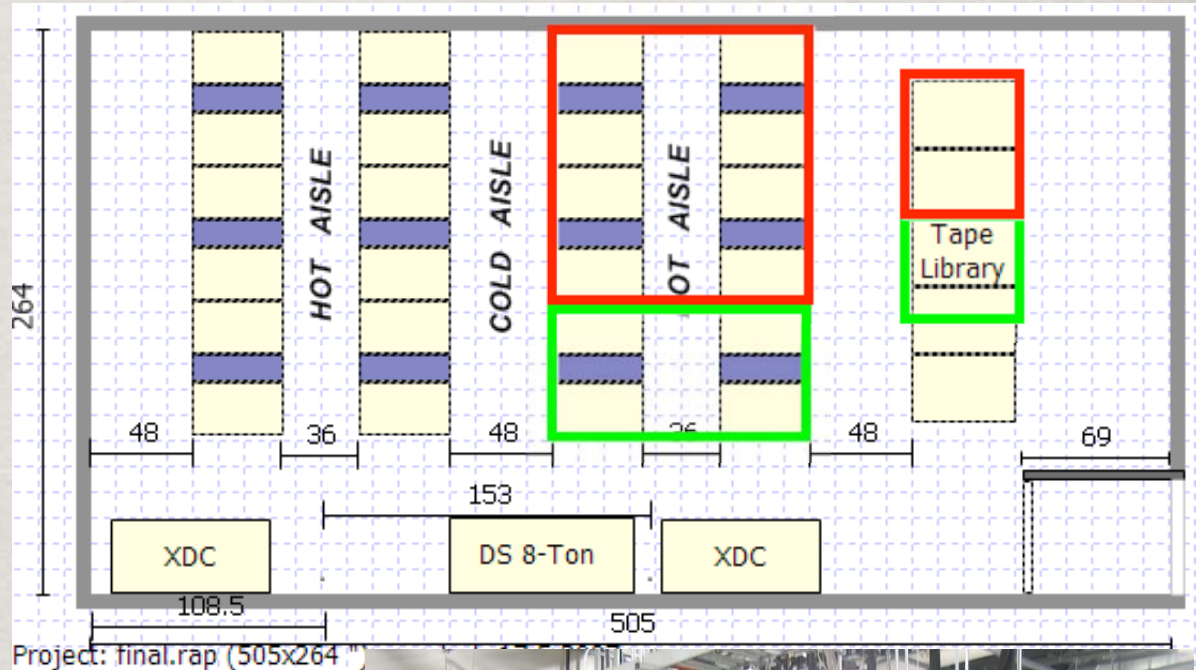
Based on Nov 06 Computing model &
7.2% of ATLAS computing resources

Commissioned last week of August

Scheduled to arrive first week November

TIER-1 DATA ROOM

- ✿ Very limited floor space only 950 sqft
- ✿ No false floor
- ✿ Rack Optimized for high density using Hot & Cold aisles
- ✿ Power estimate, 0.4MW up to 2011 (includes cooling)
- ✿ Cooling Liebert XD system, liquid cooled in row coolers/ heat exchangers 340kW
- ✿ 225 kVA UPS
 - 2007 + 2008
 - 2009 assuming quad core, 1TB disk



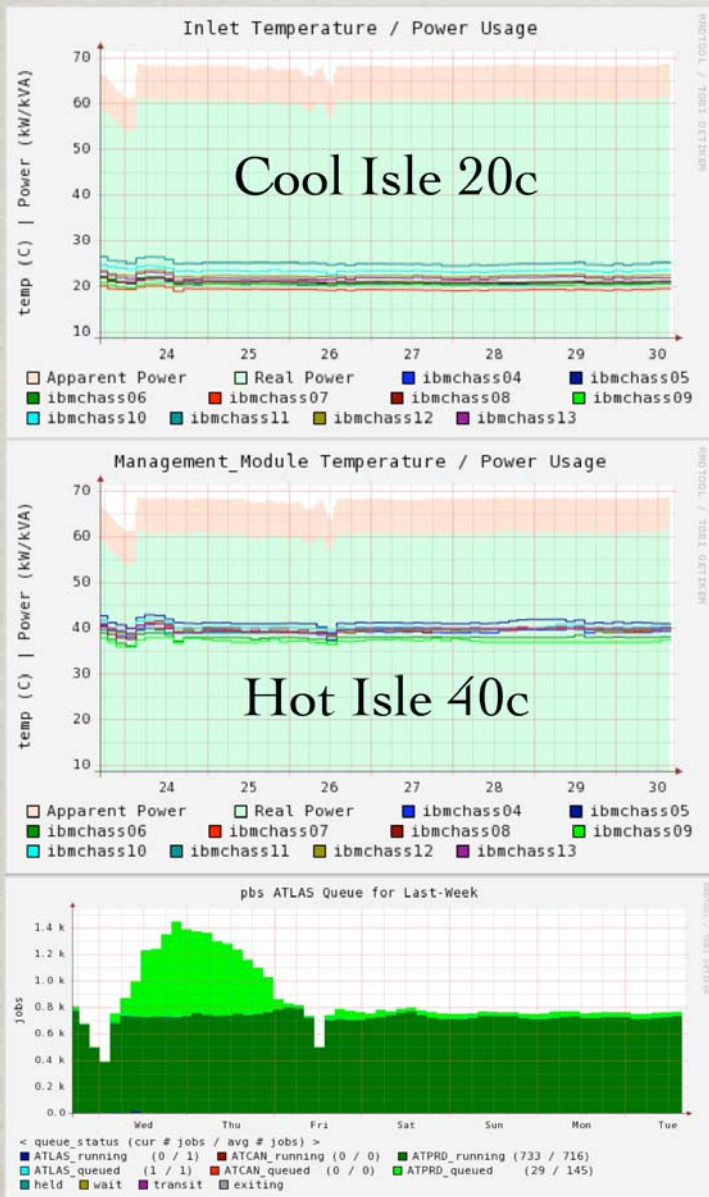








ATLAS POWER & COOLING



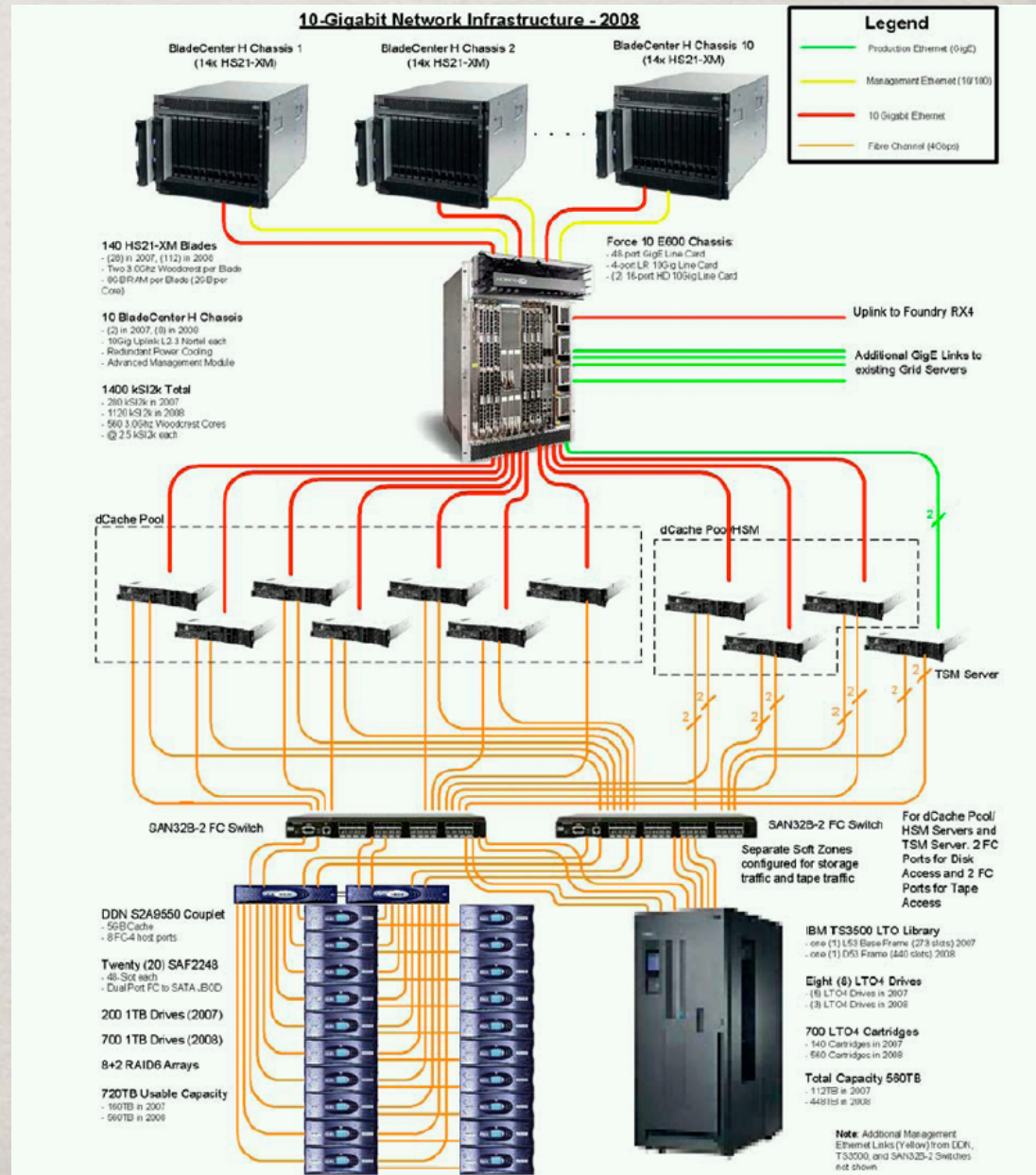
XDH coolers 32kw each

XDV 10kw spot coolers available for hot spots can be mounted on top of racks

Air condensers mounted on roof, 2 per XDC

TIER-1 CLUSTER DESIGN

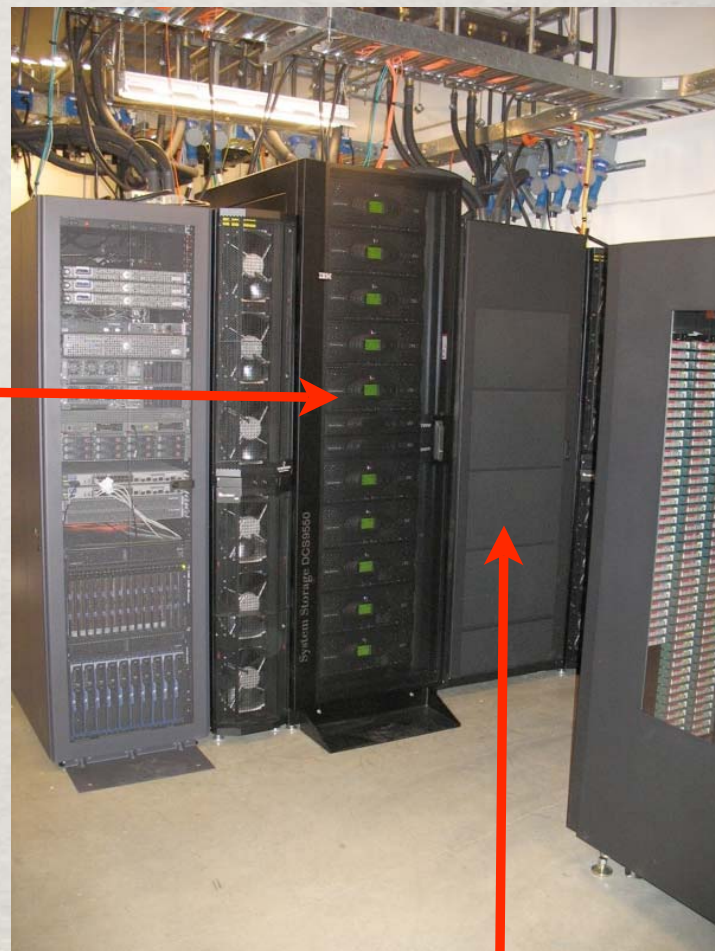
- ☼ Contract awarded to IBM for 2007-2008 resources
- ☼ CPU ~1400 kSI2K
 - ☼ 280 3.0GHz woodcrest processors 560 cores
 - ☼ 12 Blade chassis
- ☼ Disk 720TB usable
 - ☼ 7 dcache 3650's SAN disk
 - ☼ 3 dcache 3650's SAN tape
- ☼ Tape 560TB native
 - ☼ LTO-4 800GB/tape native
- ☼ Network Force10 E600
 - 36x 10GbE data
 - 48x 1 GbE control
- ☼ SAN for storage with 4 GB/s
 - FC 2x 32 port brocade switches
- ☼ GRID nodes not shown





ATLAS - STORAGE SAN

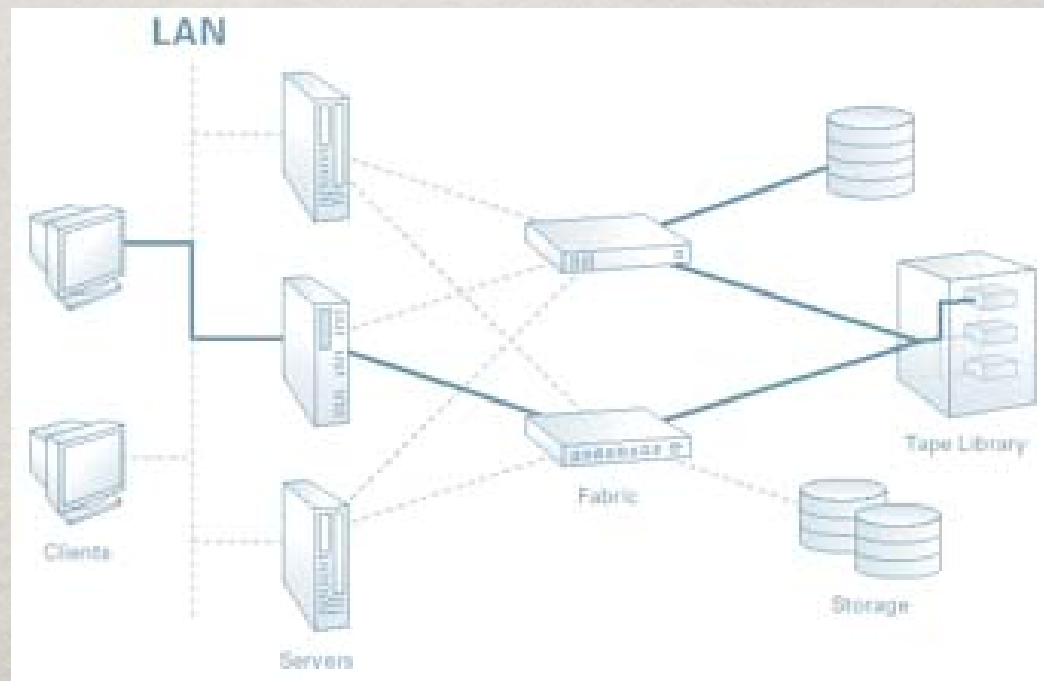
- ✿ DDN 9550 SAN Disk System
 - ✿ Dual Controller
 - ✿ Hot swap power/cooling/disk
 - ✿ SATA Disks
 - ✿ 48 disks per tray
 - ✿ RAID-6, vertical across 10 shelves
 - ✿ Dual San switch zoning
 - ✿ Performance
 - ✿ 2.4 GB/sec achievable throughput
 - ✿ 400MB/sec single transfer
 - ✿ Dual 32 port Brocade FC switches



Space for 2nd DDN rack
Arriving this week 480 1TB drives

DCACHE POOL NODES

- ✿ Connected to FC SAN via two 32-port Brocade switches
- ✿ 2 HBA'S in each pool node
- ✿ HSM pool nodes have 4 HBA's, 2 to the disk SAN and 2 to the tape library
- ✿ 7 dcache pool nodes and 3 HSM pool nodes separated into 4 groups and 4 zones
- ✿ Any nodes goes down, the other nodes in the same group can take over the running job



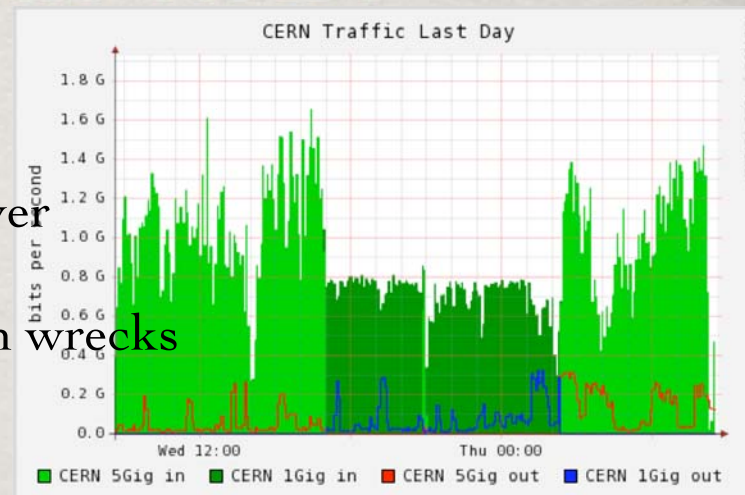
TIER-1 TAPE LIBRARY

- ✿ IBM TS3500
 - ✿ Presently using two Frames - 8 drives
 - ✿ Can be extended to 5 frames in our available space
 - ✿ Uses LTO-4 800GB native/cart
 - ✿ achieve 100MB/sec write/drive
 - ✿ Achieve 120MB/sec read/drive
 - ✿ Present capacity 560TB
 - ✿ can be expanded to 1616TB in available space
 - ✿ can meet our 2009 commitment of 1077TB but NOT our 2010 2067TB. Need LTO-5 by then or a bigger room - in the planning stages



ATLAS NETWORK WAN

- ✿ T0 <-> T1
 - ✿ 5GbE primary CERN BGP
 - ✿ 1GbE Secondary CERN BGP - auto fail over
 - ✿ Not as diverse a paths as we would like
 - ✿ Really notice the number of floods and train wrecks
 - ✿ ~12,000 km
 - ✿ Several instances of both paths unavailable



- ✿ T1 <-> T1
 - ✿ 1GbE BNL this month circuit already provisioned across TRIUMF - CA*net4 - ESnet - (BNL?)
 - ✿ SARA Tier-1 peer with TRIUMF - still in pipeline, hardware available, just need circuit to be provisioned.
- ✿ T1 <-> T2
 - ✿ 1 GbE dedicated lightpaths - backup path is routed network UVictoria, UAlberta, UToronto, UMGill, SFU

IF TIME PERMITS

EMAIL @#\$*) 1

- ✱ **TRIUMF does not impose quotas on e-mail services**

- ✱ 1000 users, ~500 regularly active

- ✱ **Several issues have arisen**

- ✱ Many users with large mail folders 100's of MBytes some even in GB's
- ✱ Storage issues 95% utilized ~300GB
- ✱ MBX format makes Backups difficult - a single 1k new message results in the entire folder having to be backed up 100's MBytes
- ✱ High system loads due to file IO to large files

- ✱ **Mailbox formats changed to Mix format**

- ✱ Hybrid mailbox format - cross b/w single file per mailbox folder and single file per message, breaks a file up into 5MB chunks
- ✱ **Significant improvement in access speed and backups**

```
time mailutil check /home/andrew/mail/spam
78725 new message(s) (78720 unseen), 78725 total
mix - real    0m0.407s
mbx - real    0m18.731s
```

```
time mailutil check /home/andrew/mail/cron
24876 new message(s) (24819 unseen), 24876 total
mix - real    0m0.159s
mbx - real    0m1.572s
```

EMAIL - 2

- ✿ Email volume - 60k per day 50k identified as spam or containing viruses
- ✿ Move to implementing Milters (Mail Filters) to allow earlier spam rejection

Present system

Incoming email -> Sendmail -> Antivirus -> SpamAssassin -> Procmail -> Dmail

Problems:

Some mail may be silently discarded due to spam filters

Spam and virus forwarded offsite

~50% of our users collect spam and do not remove it from junk folders

Miltered system

Incoming email -> Sendmail -> Antivirus Milter -> Sendmail -> SpamAssassin
Milter -> Sendmail -> Procmail -> Dmail

Advantages:

No legitimate email is lost since sender receives notification of rejection

Forwarded email is filtered through antivirus and spamassassin

Rejection becomes the default - lazy users do not collect spam - save space

EMAIL - 3

✻ **Barracuda incident - bad bad bad**

TRIUMF's main mail server got on their blacklist - reason unknown despite multiple requests - not very pleasant to deal with unless you are a paying customer. Number of collaborating institutes (including SFU Tier-2) using Barracuda's service, ~300 emails rejected over 36 hrs.

One week later the same happened again, a tool was used to check 233 ONLINE Blacklists, none of them list .triumf.ca

Any HEP Sites using Barracuda network spam firewalls ?

THANK YOU

?

EXTRAS

REMOTE TEACHING

- ✿ TRIUMF is offering an accredited Nuclear Structure course to graduate students across Canada
- ✿ Now in its second year - Byron Jennings
- ✿ 9 students this year 3 local, 6 remote as far as Ontario - Guelph & McMasters Universities
- ✿ Students Participate via Polycom, VRVS, Evo

VIRTUAL SERVICES

- ✻ Starting to explore virtual services in production environment
 - ✻ Presently
 - ✻ Primary DNS and DHCP are VMware instances
 - ✻ NIS
 - ✻ LTSP Server
 - ✻ LDAP
 - ✻ Future
 - ✻ elog - web logbooks for experimenters
 - ✻ email - services, webmail , imap, smtp
- ✻ VMware is used extensively by the ATLAS Tier-1 Group for testing dcache, and upgrades etc. Also used in Production Services such as top bdii, site bdii, monbox, oracle enterprise manager

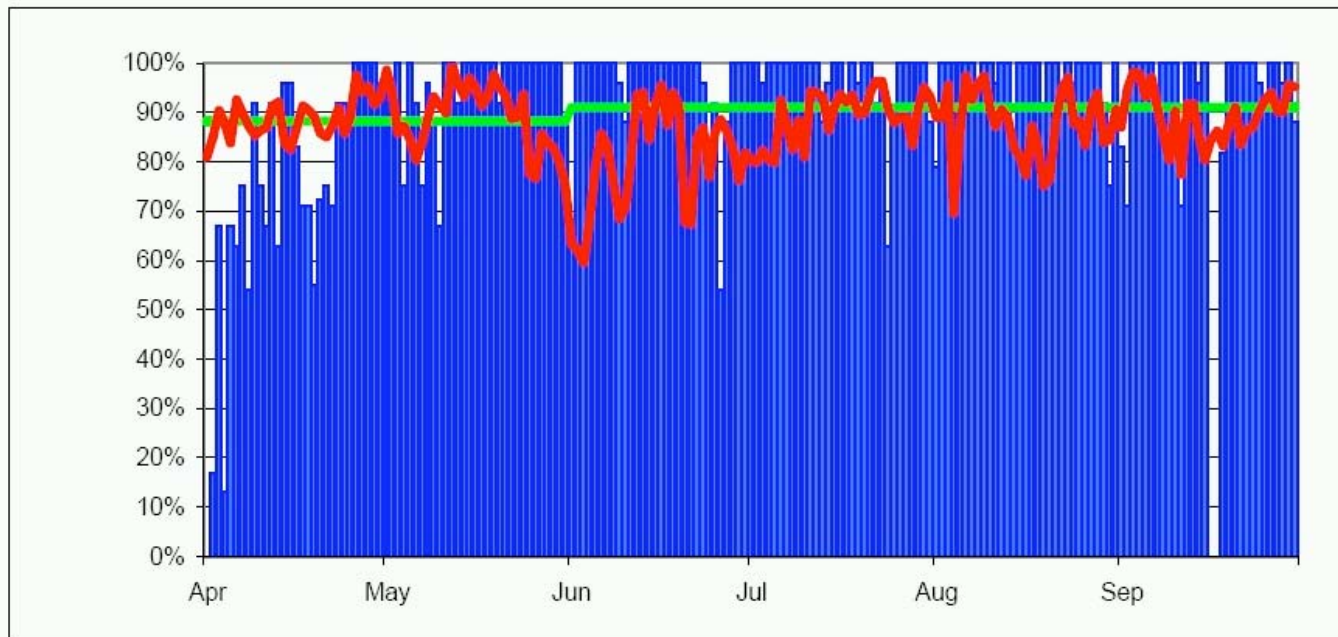
MANAGED POWER

- ✿ All core servers, routers and network gear are now on managed power
- ✿ No breaker trips since using metered/monitored power
- ✿ Rare sub-panel trip - but it has happened in past
- ✿ Managed power allow to distribute power across two sub-panels and still reboot equipment
- ✿ CAN\$35/port ~300 at present



TIER-1 AVAILABILITY

- TRIUMF's availability
- 96% last 3 months
- Average availability of the 10 ATLAS Tier-1's



TRIUMF-LCG2

av.reliability last 3 mths **96%**