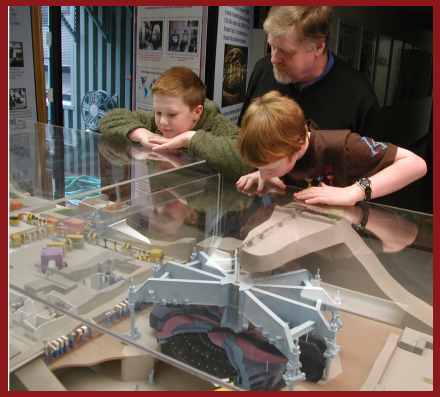
TRIUMF

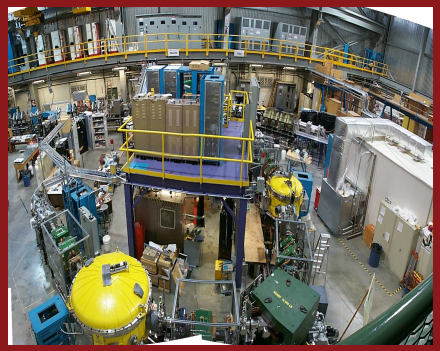
Canada's National Laboratory
Laboratoire national canadien
et en physique des particules

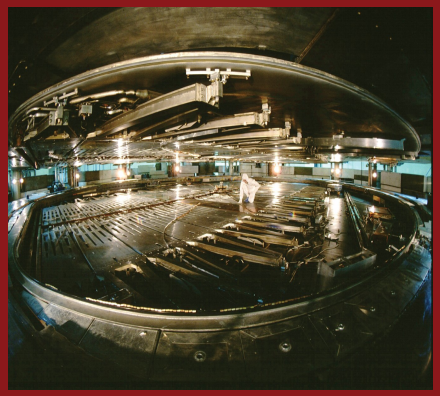
TRIUMF Site Report

HEPix Fall 2009, NERSC/LBNL

Kelvin Raywood
TRIUMF, Vancouver, Canada







Lab update

- York, Guelph and Queens U's joined as full members
  - 11 full-members + 4 associate-members
- 40th anniversary
- SRF e-linac project funded
  - 2nd driver in 40 years
  - Photo-fission for nuclear physics and medical-isotope production ^{99}Mo
  - Contribute to LHC-SPL, ILC & CLS upgrades
- External HPC facilities: WestGrid
  - Bugaboo: SL5, 1280 cores (E5450), DDR/IB, 2GB/core, 350TB Lustre
  - Orcinus: CentOS-5, 3072 cores(E5440) DDR/IB, 2GB/core
  - Checkers: SL5, 1280 cores(E5440), DDR/IB, 2GB/core
  - Storage: 1.5PB GPFS + 800TB (LTO 2,3,4) : Tivoli HSM

External Review of TRIUMF Computing

- Charged to review operations of leader and three of four TRIUMF computing-groups
  - MIS (3), GSC (5), CCN (6) : +1 since 1990
  - WLCG/ATLAS Tier-1 Centre (9) – not reviewed
- Triggered by increased demand for app. devel by MIS
  - modern admin systems, identity management, online work-flow, ...
- and by requirements of experiments supported by GSC
  - DAQ and analysis software, FPGA programming
- Senior management were largely unaware of issues with CCN
  - documentation, response time, high availability, disaster recovery, modernisation, ...
- Plans for resolution
  - app-devel stack, change control, virtualisation/HA, web-site, ...
  - People requirements: MIS +2, GSC +2, CCN +2

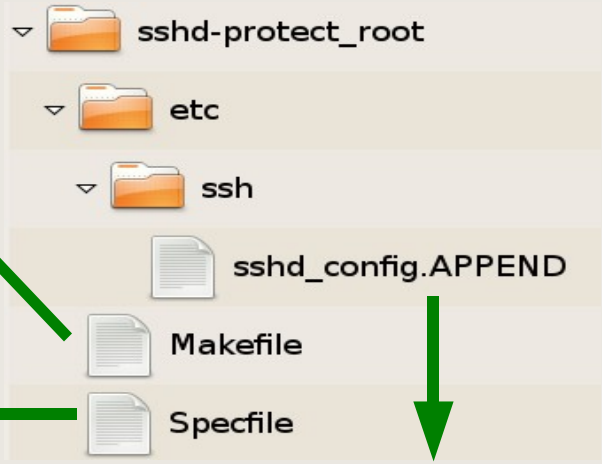
Recommendations of Review

- Strong endorsement of computing leader
  - Increase power for resource allocation, approve user-requests, ...
- Core Computing and Networking
  - Off-load commodity computing to University partners, commercial vendors
  - +1 “short-term staff”, and +1 FTE
- General Scientific Computing
  - Reduce scope of activities
  - Require experiments to fund DAQ development and support
- Management Information Systems
  - Hire external consultant to validate / augment the long-term plan
  - Explore alternatives to in-house developed apps

RPMS for modifying configs

```
# $Id: Makefile 635 2009-06-20 20:26:54Z
YUM_REPO = triumf-server
include ../Makefile.rpmbuild
```

Special vars: YUM_REPO, PKG_ARCH, DIST_RELEASE, ...




```
# triumf-sshd-protect_root
# $Id: Specfile 640 2009-06-20 20:47:17Z
#
Version: 1.1
Release: 1
PreReq: openssh-server

%description
Disables ssh to root via password

%post
service sshd condrestart > /dev/null || :

%postun
[ "$1" = 0 ] && \
   service sshd condrestart >/dev/null|| :
```

```
# //-----------------------------
# rpm: triumf-sshd-protect_root

# Allow only ssh-key auth for root
PermitRootLogin without-password
# ------------------------------//
```

Special extensions: .ADD .APPEND .ED .REPLACE .SED .SYMLINK

```
make install
```

- Builds rpm
- Installs in repo

Linux configuration management

- Deploy single-purpose servers, preferably virtual
- Start with minimal installation and record extra packages
  - Small number of kickstarts and VM gold-masters
- Use well behaved rpms for common configurations
  - Nagios client, RAID monitoring, ssh keys, yum, syslog, ntp, ...
  - Config changes undone on removal
- Keep everything in a version-control system (svn)
- Separate software, configs, data
- Push custom configs to server
  - Makefile checks svn status
  - Reloads service

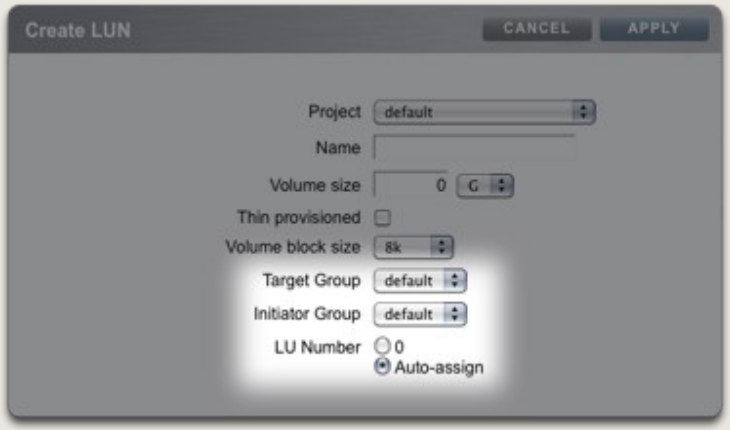
```
# $Id: Makefile 2665 2009-02-28 18:33:11Z
# Install nagios configuration files

SERVER       = trnetmon.triumf.ca
CONFIGS      = $(wildcard *.cfg)
INSTALL_ROOT = /etc/nagios/conf.d
RELOAD_CMD   = /sbin/service nagios reload
CONFIG_OWNER = 82
CONFIG_GROUP = 82

include ../../Makefile.server
```

SUN Unified Storage Server

- Acquired in spring
  - Based on ZFS, RAID-Z DP
  - Provides iSCSI, NFS, in kernel CIFS
  - Hybrid storage, DRAM, SSD, SATA
  - File & block level snapshots
  - Scales to 500TB
- Locked up on system-disk failure
  - Known bug
  - fixed in system update that we had not yet applied
- Q3 update contains new iSCSI stack
  - Required rebuild of iSCSI config
  - client config needed modification



ATLAS Tier-1 Upgrades

	In production (IBM)	New capacity (SUN)	Total
CPU / HEPSPEC-06	6300 656 cores, Woodcrest 3GHz	7000 560 cores, Nehalem 2.53GHz	13300
Disk / TB	720 RAID-6, DDN SAN	1400 RAID-Z2 / ZFS, SUN x4540 DAS*	2120
Tape / TB	560 8 LTO-4 drives, IBM TS3500	240 frame expansion, +6 LTO-4 drives	800

* 8% 1TB drives, 92% 2TB drives (Nov.)

- New capacity is installed and being commissioned
- Will provide ~7% of world-wide ATLAS resources
- Upgrade Oracle RAC, ~30TB, +2 nodes

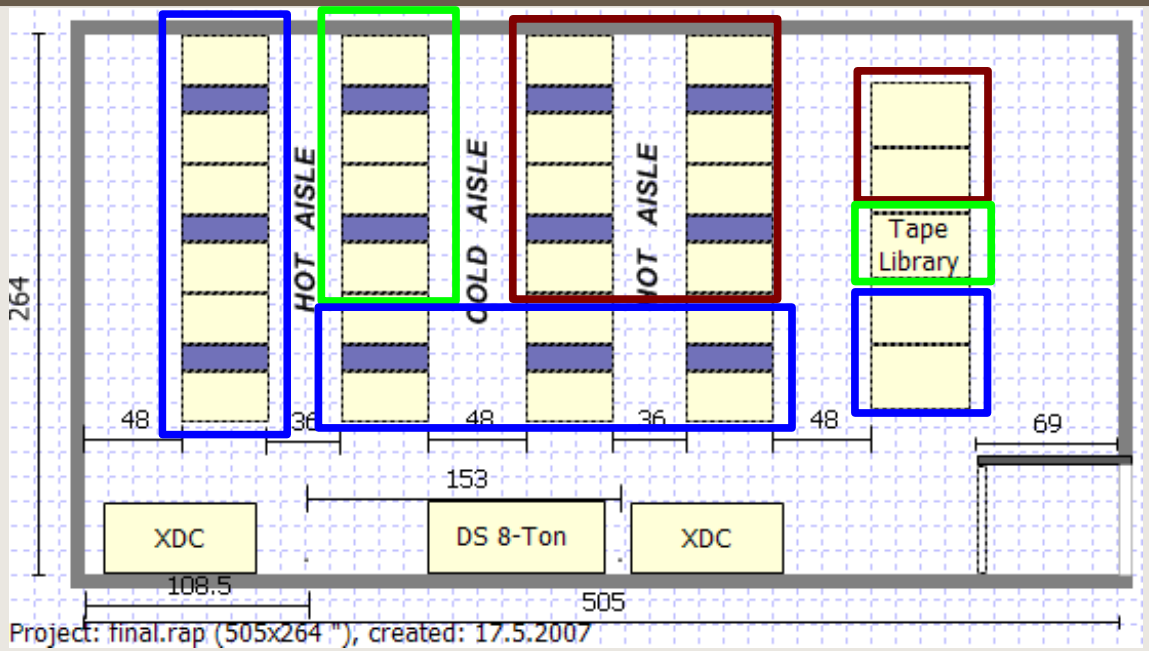
ATLAS Tier-1 Infrastructure

Limited floor space:
43' x 22'

No false floor

Rack optimization:
- high density solution
- hot & cold configuration

Power estimate:
~0.4 MW (up to 2012)
(including cooling)



Cooling solution: Liebert XD system (very flexible)
340 kW total capacity (~1/4 used)

UPS: 225 kVA (in the future CPU racks on regular power or expand UPS capacity) (~1/3 used).
(no diesel backup except for core network)

Expansion beyond 2012: new infrastructure will be required (TRIUMF next 5YP)

2007+2008

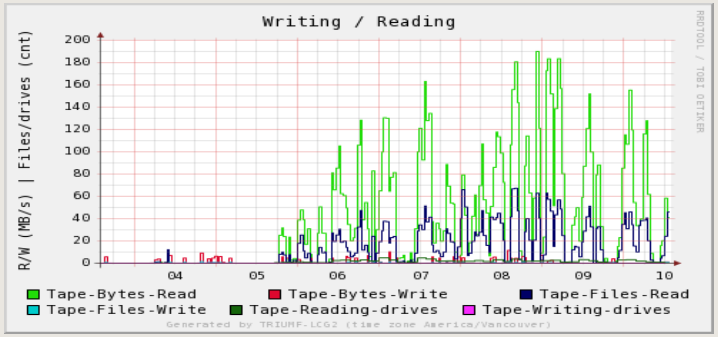
2009+2010 (higher density / LHC delay)

2011+2012

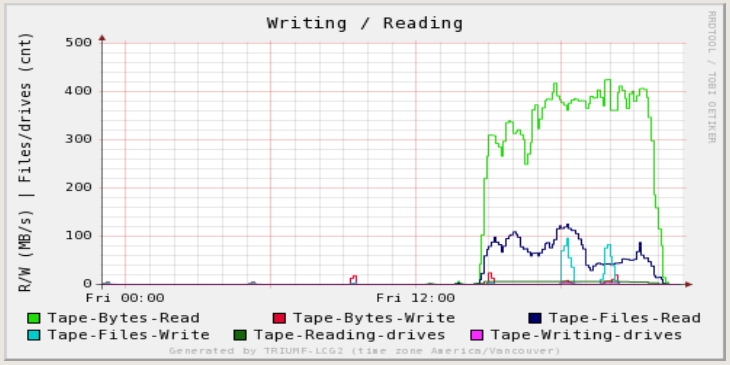
ATLAS Tier-1 HSM

- High performance HSM
  CHEP09 (Denice Deatrich, Simon Liu, Reda Tafirout)
  - In production
  - interfaced with dCache storage
- developed at TRIUMF
  - No proprietry code for tape-drivers or changer
  - TCP socket based server daemon
  - backend MySQL
- Design goal: Efficient reprocessing
  - Reorder and prioritise requests
  - Control over file and tape grouping
    - Minimses tape mounts
    - Maximises reads per mount

No prestaging ~0.2TB / h



Prestaging ~1TB / h



LHCOPN Meeting at TRIUMF

- 1st LHCOPN meeting at T1 outside Europe
- Dante perSONAR multi-domain monitoring installed at all sites
- Monitoring T0-T1 traffic
- Considered extending LHCOPN to T1-T1 and T1-T2

- Traffic patterns unknown
- Need traffic-pattern specs from WLCG before designing a network topology

Network Status – the last km

All TRIUMF external network connections and ATLAS lightpaths pass through this shack.

Lac TRIUMF Lake

Thank You!

Merci!

4004 Wesbrook Mall | Vancouver BC | Canada V6T 2A3 | Tel 604.222.1047 | Fax 604.222.1074 | www.triumf.ca