Experience on site operation in WLCG from Tokyo Tier-2 center

T. Nakamura, T. Mashimo, N. Matsui, H. Matsunaga, H. Sakamoto, I. Ueda

International Center for Elementary Particle Physics
Tiers of ATLAS

Tier-0
- 200Hz RAW: ~1.6MB/evt
- First Pass Processing

Tier-1
- CERN Analysis Facility
- 10 Tier-1 centers
  - RAW data copy on tape
  - Analysis data on disk
  - Reprocessing

Tier-2
- 37 Tier-2 centers
  - Analysis data on disk
  - User Analysis

Event Summary Data (ESD): ~1 MB/evt
Analysis Object Data (AOD): ~100 kB/evt
derived data (dESD, dAOD, NTUP,...)
distributed over the Grid

I. Ueda CHEP2010
Development was started in 2002 as the ATLAS regional analysis center in Japan.

Procurement policy is 3-year lease for the whole computer system, and almost of HW will be replaced in every 3 years.
1st system (2007-2009)
2nd system (2010-2012)

Managed by 6 ATLAS members
+ 3 system engineers (2FTE).
Resources for Tier2

**Worker node**
Blade server (Dell PowerEdge M610)  
Dual Intel Xeon X5560 (Nehalem 2.8GHz, 4 cores/CPU)  
Memory is expanded to 24GiB/node on Jan. 31.  
**160 nodes (more than 1,280 cores)** in LCG-CE and CREAM-CE.  
Operated by Torque/Maui batch queuing system.

**Disk storage**
1.2PB served by 15 file serves.

**Service in WLCG**
Top and site BDII, WMS/LB, MyProxy  
(ActiveMQ based) APEL  
Squid proxy for ATLAS conditions database cache.

**ATLAS software**
Served by NFS.  
To be replaced to CVMFS.
Disk storage configurations for Tier2

File Server: Dell PowerEdge M610
- CPU: Dual Intel Xeon E5530 (Nehalem 2.4GHz, 4 cores/CPU).
- Memory: 24GiB
- Fibre Channel HBA: Qlogic QME2572 (8Gbps)
- NIC: 10Gbps (Broadcom 57710)

Disk Array: Infortrend EonStor S24F-G1840
- 24 SATA 2TB HDDs, RAID6, 2file system
- 8Gbps Fibre Channel Interface

Fibre Channel Switch: Qlogic Sanbox 5800

GridFTP:
- 20 files parallel with 10 stream for each FTS

DPM:
- 60 XFS file systems (1.2PB) in one pool.

1PB pledge
Local system for the Japanese user

**Batch node**
Same architecture with Tier2 part.
Dual Intel Xeon X5560 (Nehalem 2.8GHz, 4 cores/CPU)
Memory is expanded to 16GiB/node on Jan. 31.
496 + 16 nodes (4,096 cores) under the LSF batch queuing system.

**Disk storage**
Same architecture with Tier2 part.
Operating with 10 NFS servers (~PB)

**Tape library**
LTO3/LTO4 drives in Storage Tek SL8500
CASTOR Ver. 2.1.9.
Cache disk + Tapes (~3PB).
Under construction with new hardware.

**ATLAS software**
Served to the batch node by read-only AFS.

x 31 for batch nodes
+1 for interactive node
Network configurations

Local system can read the Tier2 disk, and vice-versa.

Local resources can be easily converted to the Tier2.
Running analysis jobs in TOKYO- LCG2
ATLAS wide
One of the largest sites as a standalone Tier2 site.
2% fraction of the ATLAS analysis jobs have been processed at Tokyo Tier2.

Breakout for Tier3 users
Total 116 active users belonging more than 10 institute are working with the Tier3 local resources.
1 LCG-CE + 1 CREAM-CE

**LCG-CE**
- 112 -> 80 WNs
- 640 job slots (= 640 cores)
- 8GB memory/WN (=2GB/core)

**Cream-CE**
- 32 -> 64 WNs
- 512 job slots (= 512 cores)
- 8GB memory/WN (=2GB/core)

From Jan. 31.
Data distribution and flow

In original ATLAS computing model
Data is distributed to Tier2 a priori. Tier1-Tier2 transfer is restricted in a cloud.

Recent development
Dynamic data replication according to request from user analysis jobs (PD2P). Large Tier2 can contribute to production jobs in the other cloud (Tier2D). Data will be transferred over the cloud.
Data transfer (incoming)

800MB/sec

5 minutes average

1 day average
98% availability and reliability in one year operation of the 2nd system.

Due to the scheduled down time for the annual power outage from Oct. 22 to Oct. 25.
Upgrade plan and Summary

**Upgrade for Tokyo Tier2**
All WN will be imported to CREAM-CE soon.
Effective use and enhancement of the ATLASLOCALGROUPDISK.
Going to apply CVMFS for the ATLAS software release.

**Upgrade for local system**
LTO5 drives and tapes will be installed to the tape robot in this March.
Construction of CASTOR based archiving system is on going.
Test for the large scale GPFS operation was started. It will be used for the user’s disk storage instead of the NFS.
R&D/test for the installation and operation of PROOF for the further data analysis.

**Role of ATLAS Tier2**
Tier2 role is gradually changed with ATLAS computing model.
More steady and reliable bidirectional network (for the incoming data and outgoing data) must be important from now on.
Auxiliary Slides
ATLAS Japan collaboration

15 Institutes,
~110 Collaborators

Kyoto Univ.
Kyoto Univ. Edu.
Osaka Univ.
Kobe Univ.
Okayama Univ.
Hiroshima Inst. Tech.

KEK
Univ. Tsukuba
Univ. Tokyo
Tokyo Metro. Univ.
Tokyo Inst. Tech.
Waseda Univ.
Shinshu Univ.
Nagoya Univ.
SCT(KEK, Tsukuba, TITech, Waseda, Kyoto-Edu, Osaka)

TGC(KEK, Tokyo, TMU, Shinshu, Nagoya, Kyoto, Osaka, Kobe)

Muon TDC(KEK)

Solenoid Magnet(KEK)

Higher Level Trigger(KEK, TITech, Waseda, Kobe)

DAQ(KEK, Shinshu, Hiroshima-IT, Nagasaki-IAS)
Physics analysis activities

• Higgs Boson
  – Vector Boson Fusion (VBF) $H \rightarrow \tau\tau$
  – VBF $H \rightarrow \gamma\gamma$
  – $H \rightarrow W^+W^-$
  – MSSM Higgs

• Supersymmetry and Beyond
  – mSUGRA
  – CMSSM

• Standard Model
  – Top quark
  – Vector bosons (Z,W)
  – QCD jets
  – Onia ($J/\Psi,Y$)

• Extra Dimensions
  – Blackhole
SINET3

- SURFnet
- RENATER
- GÉANT
- Europe
- China
- Hong Kong
- Singapore
- Asia
- SINET3
- CalREN
- L.A.
- N.Y.
- CA*net4
- Internet2
  (Abilene)
- MANLAN
- 10Gbps
- 10Gbps
- Pacific Wave
- AARnet
- ESnet

Europe

Asia

North America

TOKYO
Data transfer (incoming)

1 day average

5 min. average

800MB/sec
Data transfer (outgoing)

1 day average

5 min. average

Request beyond the ATLAS cloud.

2010 - 2011

2011/02/17 09:00 2011/02/18 01:00 2011/02/18 17:01 2011/02/19 09:02 2011/02/20 01:03 2011/02/20 17:04 2011/02/21 09:05

2010/03/01 2010/05/01 2010/07/01 2010/08/30 2010/10/30 2010/12/30 2011/03/01

Tomoaki Nakamura
ICEPP, Tokyo Univ.
Distributed data and usage of LGDISK

The area for ATLASMCDISK was merged with ATLASDATADISK in this Jan.

Demand for LOCALGROUPDISK is increasing recently.