THREE TIERS OF ATLAS
An Integration Program for The US ATLAS Computing Facility

Rob Gardner

Meeting of the American Physical Society’s Division of Particles and Fields
August 2011
Acknowledgements

• For this presentation I’ve borrowed material from several colleagues:
  • Artur Barczyk, USLHCNet
  • Ian Bird, CERN
  • Eric Boyd, Internet2
  • Michael Ernst, Brookhaven
  • Eric Lancon, Saclay
  • Shawn McKee, UMich
• Mistakes are all mine however (!)
Introduction

• The Worldwide LHC computing Grid (WLCG) has been well described

• In the US context we implemented WLCG using an “Integration Program”

• Phased program of activities across the sites in US ATLAS, viewed collectively as a “facility”

• Can be seen as a framework for maintaining best practices in a distributed computing collaboration
US Tier 1 Profile

~ 5000 job slots
62 kHS06
7.4 PB disk

M. Ernst
US ATLAS Tier2 Profile

- 4/5 sites are federated
- Commodity hardware
- 350-450 compute nodes (of various vintages, largest component is now Dell R410, Intel X5660, 48G, HT sched @ 24 jobs/node)
- 10-15 storage servers (10G, eg. 150 TB useable each)
- 10-15 head nodes
- Most using virtual machines (KVM or VMWare hypervisors)
- 2GB/job
- 1000-2000 TB storage

- Various storage systems:
  - dCache (2)
  - Xrootd (2)
  - GPFS (1)
  - Lustre (1)
- Grid middleware services from OSG
- Cluster (SL5.5) management: Cobbler+Puppet or Rocks
- Services: Ganglia, Nagios, Cacti, Sysview, lsm-view, dCache monitor, CVMFS, Squid, memcached
- 10G WAN & 2 L2 virtual circuits between sites, and to BNL Tier 1

~ 15000 job slots
137 kHS06
7.4 PB disk
Facilities Organization

Facility comprising two principal lines
- Production and Analysis Operation Coordination
- Facility Deployment, Integration and Operation
Integration Program

• Goal is to employ best practices in distributed computing across federated resources
• Computing resources: coordinated procurements
• Middleware deployments: OSG, WLCG accounting
• ATLAS services: file catalogs, transfer services
• Network monitoring infrastructure
• Focused tasks & exercises
  • Throughput validation
  • Analysis queue validation
  • Data management
• New services (eg. CERN virtual machine file system)
• Operations
Site Certification

- Defined for each quarter
- Identify key performance drivers
- Contributions from T1, T2 admins & others
- Example for April-June 2011
Focused Task: perfSonar

- Work begun 2.5 years ago with perfSONAR-PS TEAM, coordinated by S. McKee (Michigan)
- Goal to instrument network connections between US Tier 1 and Tier 2 centers uniformly
- Aid in problem diagnosis and location identification
- Differentiate end-site issues from network problems
- Archive measurements
- Implemented two endpoints at each site (one for latency, the other for throughput testing)
- Defined a mesh of connection tests between all sites
Centralized status at a glance

perfSONAR-PS: Throughput graphs

Standalone Dashboard

The Experimental Independent perfSONAR Dashboard

Services for USATLAS Cloud

Throughput

Latency

but one can also retrieve older time slices from archive, or can also zoom in on a particular time within such graph.
Focused Task: Analysis Performance

- In preparing site for user analysis we ran several coordinated analysis stress tests across sites.
- Examined wall time efficiency and factors affecting performance across all sites (storage, access mode, ..).
Stable Operations

- Continuous, automated monitoring for availability and reliability (must be >95%)
- EU-CERN based and US operations teams watch for failed services, issue tickets quickly
- Analysis jobs robots periodically submit test jobs and blacklist site if repeatedly fails
- Data transfers continuously monitored; failures generate emails to broadly viewed lists
- User jobs can be destructive (unintentionally) as framework is permissive (eg. open 2000 input files; run parallel jobs on host)
- Capacities, pledges, performance reviewed monthly by WLCG project office at CERN
Tier 2 Delivery to WLCG

- US ATLAS Tier2 sites compared to top 50 T2 federations worldwide
- Production Jan 2009 through July 2011

US-NET2 does not include Harvard U contributions due to an accounting error
Processing at Tier2

- Goal was to enlarge model to allow processing of all types at Tier 2 centers

Overall CPU consumptions All Jobs (Cumulative Graph)
3477 Hours from Week 00 of 2011 to Week 21 of 2011 UTC

First 6 months 2011
Typical Tier 2

Running Jobs by Activities
144 Days from Week 00 of 2011 to Week 21 of 2011

- simulation
- analysis
- pileup reprocessing
- validation
US Global Analysis Share

- Where analysis jobs ran during three months in 2011
- Driven by preplacement, dynamic placement, site reliability
- 26% of total on US Tier 2 sites
- 32% total on US Tier 1

T2 centers

~50% of analysis jobs in ~10 sites

T0+T1 centers

E. Lacon
Over 500 users last year, most are international users.
Efficiencies by Job Type

Mitigate by improving access: worker node caching direct access IO
New Challenges & Initiatives
Better than expected Networks: more options

Evolution of computing models

Hierarchy vs. Mesh
Evolving computing model requires network infrastructure changes
Enable Tier 2, 3 to easily connect to any Tier 1 or 2
Use open exchange points
Do not overload general research networks with LHC data
First prototype (in production) this Summer (with AGLT2, BNL connecting)
Data Federation

- Provide new access modes & redundancy
- Share storage resources
- Examining in T1-T2 context and also Tier 3
- Tier 3 approach is to remotely replace missing files from jobs
Performance Testing

- Processing direct access to a Tier2 site over WAN
- Walltime efficiencies 10-15% lower
- Depends on job type
- To study:
  - Direct access (streaming over WAN)
  - File caching
  - Sub-file caching

Remote processors 5 ms RTT
Resource Challenges

- Multicore trend poses significant challenge to business as usual in terms of low cost commodity clusters (IO bottlenecks)
- Increase in machine intensity implies more memory per job slot to properly simulate pileup
- Today: 24 job slots / 1U (dual 6 core Intel hyperthreaded)
- Soon: 128 job slots / 2U
  - 2U Dell C6145
    2 x 4 socket motherboards
    8x16 core AMD Interlagos
  - Implications for a worker node?
    - disk (SAS/SSD), RAID
    - memory (at least 256G)
    - NIC (10G)
- Pursuing multi-process ATLAS code and whole-node scheduling to mitigate both memory & IO
Summary

- US ATLAS implementation of the WLCG has been very successful in delivering required resources to ATLAS
- "Owning" resources a particular advantage for US LHC
  - Full site, control w/ vested interest
  - Innovation and scalability testing
- New challenges lie ahead for data access, IO throughput, networking
THANK YOU!