

Integration Program Update

Rob Gardner

3-2-09




US ATLAS Tier 3 Workshop
OSG All Hands @ LIGO












Overview

- Phased program of work for the US ATLAS Facility
- Establishing a baseline set of deliverables that integrates fabric resources, ATLAS software and services, grid level services, operational components, and user tools
- Launched June 2007
 - Now in Phase 8
- Focus is primarily on US ATLAS Tier 1 and Tier 2 facilities with some participation from (Tier 2-like or Tier 2-satellite) Tier 3 sites
- <http://www.usatlas.bnl.gov/twiki/bin/view/Admins/IntegrationProgram.html>

Integration Program - Tasks




Phase 8 (Jan 1 - Mar 31, 2008 - FY09Q2)


Notation:  completed  work is in progress  defer to next phase

WBS	Deliverable	Description	Major Milestones	Status	Resource
1	P8	Integration Phase 8: deployment, validation			Staff, Tier1 + Tier2 centers
1.1	ATLAS releases	ATLAS releases via Pacball & DDM	Mar 1		Xin, Fred, Torre
1.2	Fabric upgrades	Compute and storage upgrades	Mar 31		Tier2
1.3	Network monitoring services	US LHC network monitoring infrastructure: validation	Mar 1		Networking group
1.4	Bestman validation	Validate OSG/VDT based Bestman	Feb 1		OU, SLAC, BU, UTA
1.5	Squid servers at Tier 2	Conditions data access at Tier 2	Mar 1		SLAC, ...
1.6	Load tests 200/400 MB/s and multi-site 1 GB/s	Read/write sustained BNL-Tier2	Mar 31		Networking group
1.7	Storage Validation	Disk, LFC, DQ2 catalog validation	Mar 1		all sites
1.8	Site certification Table	Site certification and benchmarks registered	Mar 31		all sites
1.9	Analysis Benchmarks	Boxed analyses across sites	various		FacilityWGAP
1.10	US ATLAS Facility clients	US-specific support for various ATLAS and OSG clients for admins and physicists	on-going		Marco
1.11	Summary Report	Summary report	Mar 31		all

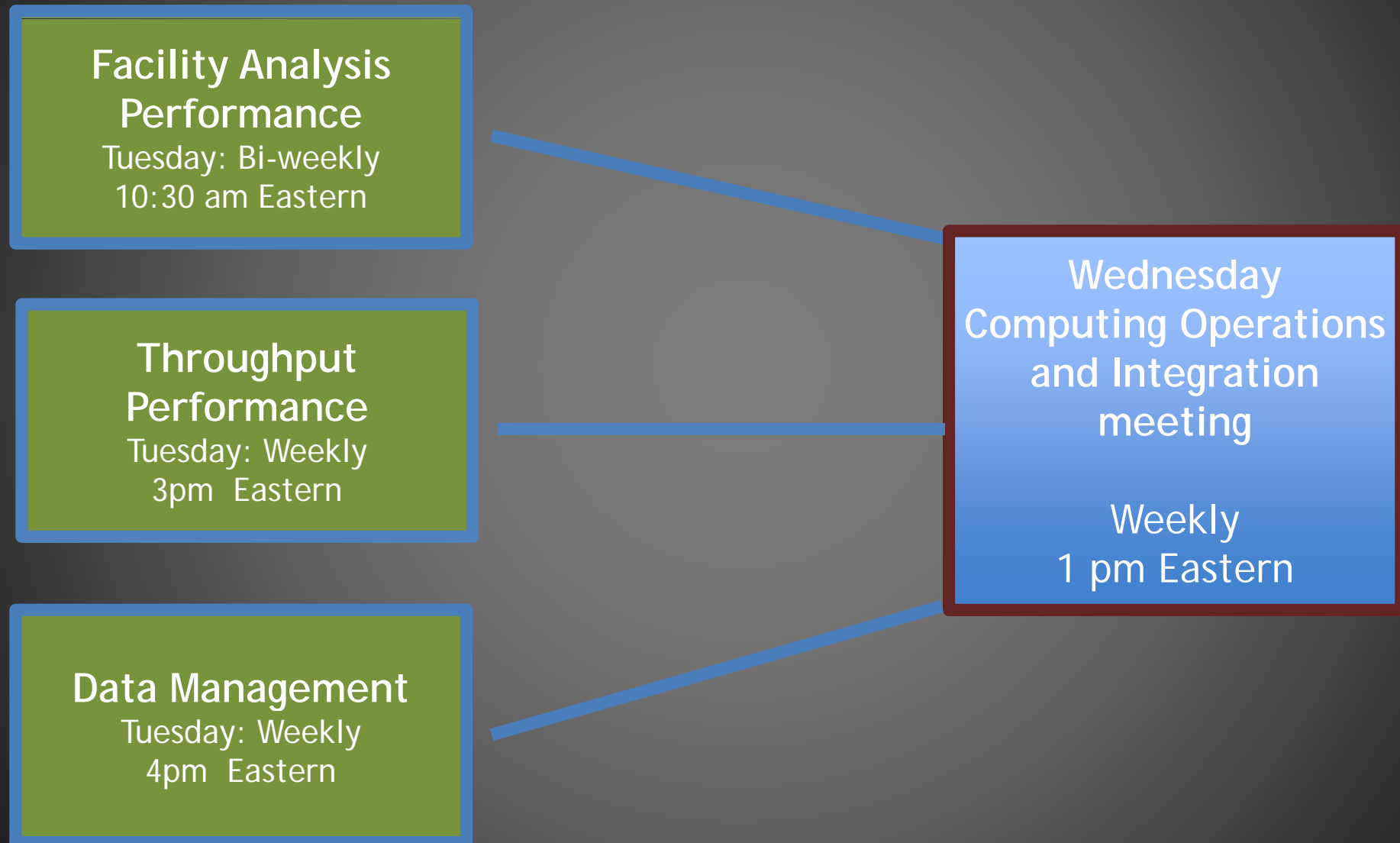
Integration Program - Sites

- Phase 7 site certification table

Notation:  completed  work is in progress  defer to next phase

Site	Space Tokens (Oct 15 MS)	Fabric upgrades	NetworkMonitoring	LT: 200 MB/s	LT: 400 MB/s	AB 1	AB 2	AB 3	AB 4	FileCatalog	TuneRSV
TIER1											
AGLT2											
MWT2											
NET2											
SWT2											
WT2											

Coordination



Analysis Queue Performance

- Background
 - We are beginning to see Tier 2 analysis jobs that stress storage systems in ways unlike production jobs
 - Many analysis workflows
- General objectives
 - Support testing and validation of analysis queues
 - Address issues uncovered during functional tests in ATLAS
 - Measure analysis queue performance in a number of metrics
- Overarching goals
 - is to assess the facility readiness for analysis workloads in terms of
 - scale: response to large numbers of I/O intensive jobs
 - stability and reliability of supporting services (gatekeepers, doors, etc)
 - Physics-throughput (user's latency and efficiency)

Working group

- Organized bi-weekly meeting with representatives from sites, software, operations, physics
- Not a general analysis user-support meeting
 - However will hear reports of user problems and track those indicating systematic problems in the infrastructure (facilities or distributed services)
- Feedback to site administrators, ADC development & operations

Queue Metrics

- For a specified set of validated job archetypes...
 - DnPD making for Tier 3 export
 - TAG-based analysis for Tier 3 export
 - Others as appropriate
 - Data access methods
 - direct read from storage
 - copy to local work scratch
- ...define queue metrics such as:
 - Completion times for reference benchmarks
 - Local I/O access performance
 - Elapsed user response times for a reference job set
 - Submission to queued
 - Time in queue
 - Total elapsed to completion

Measurement

- Measurement modality:
 - Regular testing (robot-like)
 - Stress testing
 - Response versus scale (for example, queue response to 10/100/200/500/1000 job sets of various types, I/O intensive) to determine breaking points
- Deliverables (preliminary)
 - March 15: Specify and test a well-defined set of job archetypes representing likely user analysis workflows hitting the analysis queues.
 - March 31: Specify set of queue metrics and test measurements made
 - April 15: Facility-wide set of metrics taken

Throughput & Networking

- Led by Shawn
- Implement a consistent perfSONAR infrastructure for USATLAS
 - Include “mesh-like” testing between Tier-n sitesb.
 - Improve access to perfSONAR data (new frontpage for perfSONAR, Jay’s additions for graphing/display/monitoring)
- Document existing monitoring (Netflow, FTS, ATLAS dashboard, dCache monitoring)
- Implement standard dataset for disk-to-disk tests between BNL and the Tier-n sites
- Work with I2 and Esnet for dedicated T1-T2 circuits
- Continue establishing performance benchmarks
 - 200/400 MB/s T1→T2, >GB/s T1→multi-T2
- Iterate on all the above as needed

Data Management

- Led by Kaushik
- Storage validation
 - Consistency checking filesystem \leftrightarrow LFC \leftrightarrow DQ2 catalog
 - Standardize cleanup tools & methods
- Storage reporting (ATLAS, WLCG)
- Data placement policies, token management and capacity planning
- Data deletion policies
- Data transfer problems
- Dataset distribution latency

Challenge Summary (incomplete!)

- Analysis
 - Establishing performance at scale
 - Very dependent on storage services: developments in dCache, xrootd will be very important
- Throughput
 - Monitoring infrastructure and performance benchmarks
- Data management
 - Standardizing management tools and placement/deletion policies, capabilities, performance
- Tier 3
 - Exporting datasets to Tier 3s efficiently