

More about... Integrating Facilities in US ATLAS

Rob Gardner

8-19-09

US ATLAS Tier 2 / Tier 3 Workshop at U Chicago

What is the Integration Program?

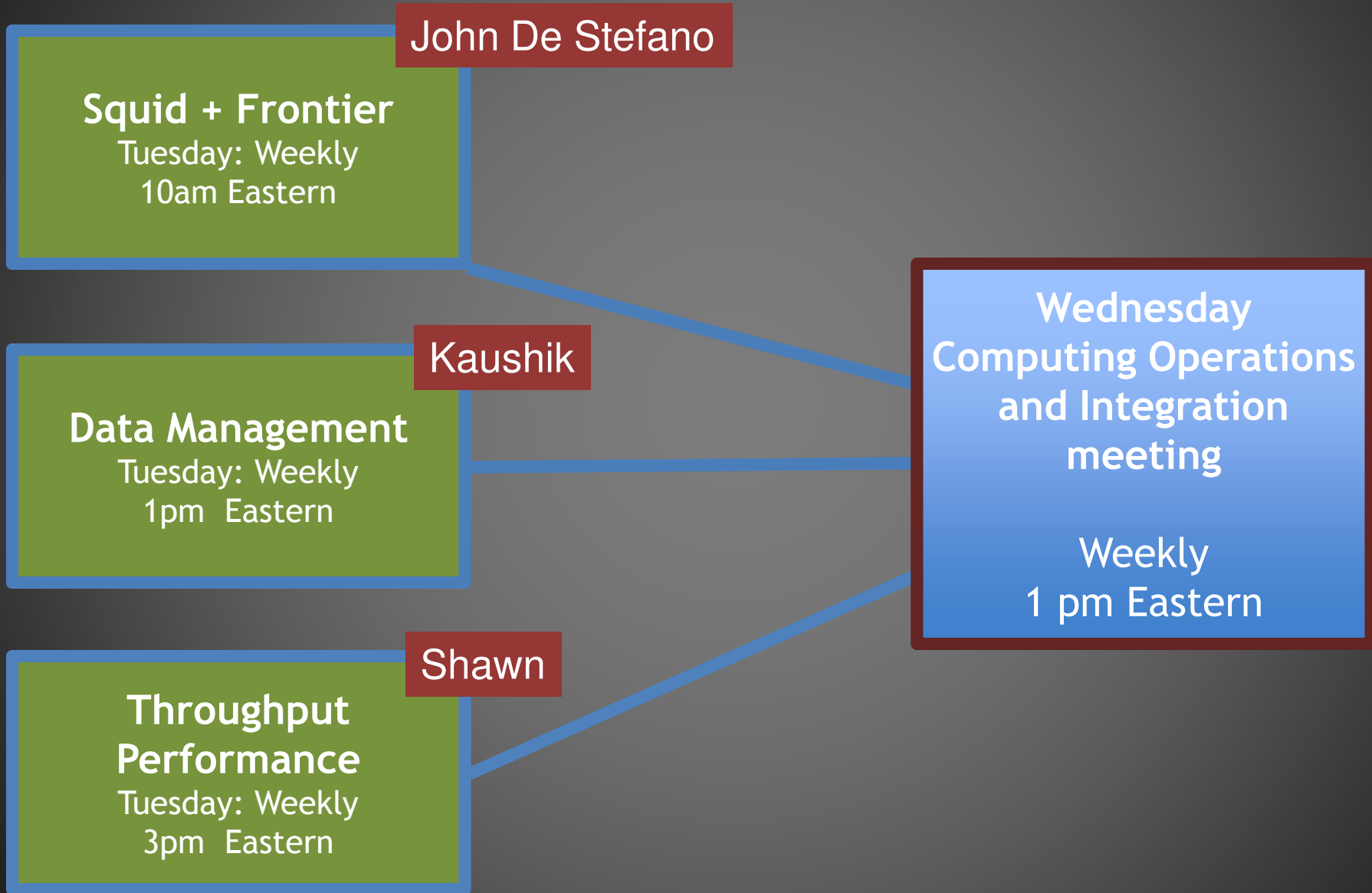
- Phased program of work for the US ATLAS Facility
- Establishing a baseline set of deliverables that integrates fabric resources, ATLAS software and services, grid level services, operational components, and user tools
- Launched June 2007
 - Now in **Phase 10**
- Focus is primarily on US ATLAS Tier 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>

Phase 10 by Site

- Upgrades to DQ2 site services, LFC server, equipment procurements, Squid, OSG 1.2, SL5, Throughput

Site	DQ2 update	Fabric upgrades	SquidTier2	LFC update	OSG 1.2	NetworkMonitoring	UpgradeSL5
TIER1							
AGLT2							
MWT2_IU							
MWT2_UC							
NET2_BU							
NET2_HU		-					
SWT2_CPB							
SWT2_OU	-		-				
WT2							

Oh.. meetings



US ATLAS & OSG

- OSG serves as integration and delivery point for core middleware components including compute and storage elements for US ATLAS
 - Uniformity, consistency across all sites for job management / gatekeeper (CE) services
- Security infrastructure including site-level authorization service, operational service for updating certificates and revocation lists
- Critical service availability monitoring of OSG services (RSV)
- Service availability monitoring and forwarding to WLCG
- Site level accounting services and forwarding to WLCG

US ATLAS & OSG

- Operational ticketing exchange with GGUS
- Operational online/availability through OIM
- Incorporation of LCG client utilities
 - resolving Globus library inconsistencies
- LCG File Catalog (LFC) server and client packaging
- Bestman and xrootd:
 - SRM and file system support for Tier 2 and Tier 3 facilities
- dCache packaging through VDT and support
- Future - HDFS
- Integration testbed (ITB)
 - for new releases of the OSG software (CE & SE), pre-production deployment testing with Panda

Tier 2 Performance

- Measurements come from RSV probes forwarded to WLCG
- Daily email reports can be used to spot problems
- There have been problems with the OSG information database recently so CPU hours should be checked
- Will be moving to HS06 from SI2K

This report shows USLHC Tier2 reliability and usage during July 2009 as measured by OSG tools.

	Reliability	Availability	CPU Wallclock hours for Owner VO	CPU efficiency for Owner VO	CPU hours for Owner VO	MoU Pledge *	Wallclock hours delivered to all OSG VOs
ATLAS T2 Federations			ATLAS	ATLAS	ATLAS		
US-AGLT2	98%	96%	2,578,248	89%	2,305,967	430,776	2,679,576
US-MWT2	100%	100%	2,643,117	93%	2,447,039	496,396	2,688,654
US-NET2	100%	100%	904,113	94%	851,254	296,856	904,113
US-SWT2	100%	97%	1,393,586	94%	1,311,787	618,710	1,528,571
US-WT2	99%	98%	1,184,153	89%	1,059,310	366,048	1,184,153

Throughput & Networking




























- Led by Shawn McKee - weekly Tuesday meeting
 - Implement a consistent perfSONAR infrastructure for USATLAS (leveraging Internet2 tools)
 - Disk-to-disk tests between BNL and the Tier 2 (Hiro)
 - Work with I2 and Esnet for dedicated T1-T2 circuits
 - BNL-UC, BNL-AGLT2, BNL-BU circuits in place but there have been some problems (eg. UC-BNL link is asymmetric, investigating)
 - Continue establishing performance benchmarks
 - 200/400 MB/s T1→T2, >GB/s T1→multi-T2
- Looking forward to next release of the toolkit to deploy on the facility

Data Management

- Led by Kaushik De
- Storage validation
 - Consistency checking filesystem \leftrightarrow LFC \leftrightarrow DQ2
 - 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

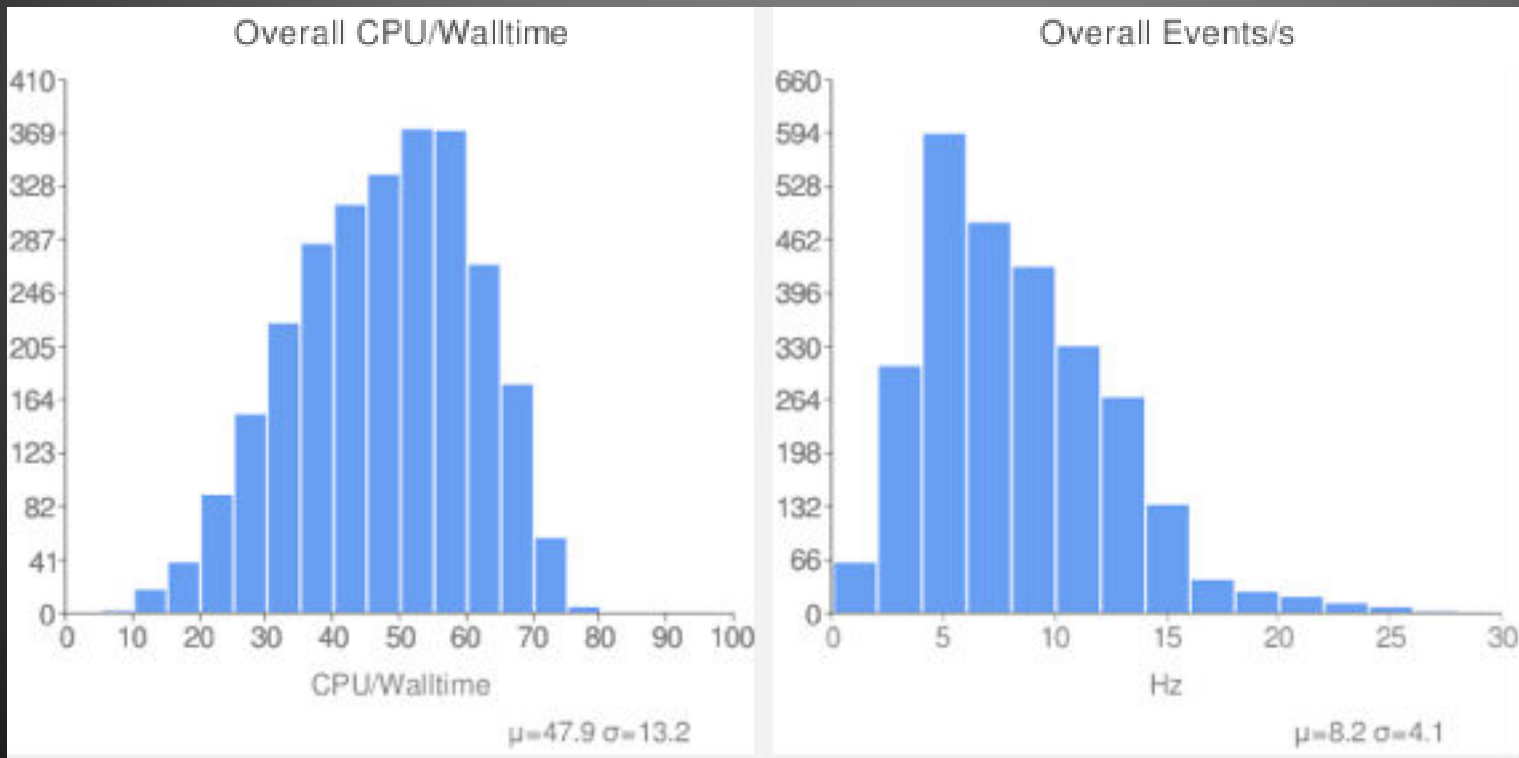
Analysis Validation for Sites

- Nurcan Ozturk, Rik Yoshida created a set of job types to validate in advance of STEP09
- Nurcan painstakingly check all the sites and followed up with admins
- We should “institutionalize” this procedure for the fall as new job types come into the queues - this would be big service contribution for ATLAS!

Notation:  completed  work is in progress  defer to next phase  table to be updated								
Job type	Athena release	Input	T1	AGLT2	MWT2	NET2	SWT2	WT2
SusyValidation	14.5.0	AOD						
D3PD making with TopPhysTools	14.5.0	AOD						
TAG selection	14.5.0	TAG						
AthenaRootAccess	14.5.0	AOD						
Data reprocessing	14.5.0.5	DPD						
ANLASC1	14.5.1	AOD						
ANLASC2	14.2.23	ESD						
ANLASC2 Jet sampling	14.5.1	ESD						
More jobs types to come ...								
HammerCloud								

Automatic Performance Issues

- Failure rates were low in Step09
- Sites varied by CPU/walltime efficiency & event rate
- Two causes - storage access & local disk contention



ANALY_MWT2
(post Step09)

Summary, Concerns, Questions

- In Phase 10 we have a lot of upgrades to the facility, and these need to be made well in advance of the LHC restart in November
 - SL5, DQ2, LFC, OSG, new equipment, 12 tools, ..
- Supporting analysis at scale is biggest concern
 - Performance (storage access)
 - User jobs and data
- What new will come with 30 Tier 3 sites?
 - T2gs integration should be straightforward
 - T2 → T3 data export - need to explore best options here
 - Storage systems @ T3