Planning in the Integration Program

Rob Gardner
US ATLAS Facilities Meeting, SMU
October 11-12, 2011

Outline

- The Integration Program scope

- Goals for current quarter

- Progress from recent US ATLAS workshops on virtualization & configuration management, and federated xrootd

- Service deployments in this Phase

- Facility meetings & conclusions

Integration Program

- The US ATLAS Computing Facility's framework for employing best practices in distributed computing across federated resources to ensure ATLAS computing resource requirements are met

    - Coordinated procurements of resources

    - Middleware deployments: OSG, WLCG accounting

    - ATLAS services: file catalogs, transfer services

    - Network monitoring infrastructure

- Focused tasks & exercises

    - Production & analysis performance

    - Throughput & latency between sites

    - Storage management & data access services

    - New services & technologies

Current quarter goals
(2b revised during this meeting)

- Finish FY11 CPU & Storage procurements at all sites

- Update OSG worker node & CE to last pacmanized versions; participate in rpm-based installation & config

- Deploy CVMFS uniformly throughout the facility

- Upgrade network monitoring infrastructure; revisit sustained multi-site throughput testing while evaluating new circuits (Esnet, LHCONE)

- Continued testing and development of federated xrootd infrastructure

- Continue to support OSG opportunistic usage

- Phase I integration of newest US ATLAS Tier 2 site at Illinois/NCSA into MWT2

- Engage with and support whole-node-scheduling / AthenaMP and analysis IO performance activities

- Middleware deployment planning & interoperability

- Coordination of Virtualization and Cloud activities as appropriate for the US ATLAS facility

Recent Workshops

USATLAS Meeting on Virtual Machines and Configuration Management

chaired by Michael Ernst (BNL, ATLAS), Robert GARDNER (UNIVERSITY OF CHICAGO)

from Wednesday, 15 June 2011 at **07:00** to Thursday, 16 June 2011 at **18:20** (US/Central)
at **Brookhaven National Laboratory (Berkner Hall, Room B)**

Mana

Description   **Connection information:**

- EVO connection:
  - http://evo.caltech.edu/evoNext/koala.jnlp?meeting=M2MvMB2e2lDIDD9e98DM92
  - Password: 1212
- Phone BridgeID: 356 9753
- Password: 5724
- Phone numbers for the phone bridge : http://evo.caltech.edu/evoGate/telephone.jsp (BNL: 631-344-6100)

Workshop description:
Meeting for site admins to share information about:

- Virtual machine management strategies and experience gained from US ATLAS sites and elsewhere to date
- Consolidation of head node services into virtual machines Looking forward for configuration management approaches

Contributions:
Representatives from each site should be prepared with a few slides addressing the below issues. The workshop format will be very free-formed: we will start with complete description of the full building and configuration management of virtual machines as it is done at BNL. As topics are discussed in detail methods and tools used at the Tier 2's can be compared with the Tier 1 and other Tier 2s.

- How are nodes built? Distinguish: worker nodes, head nodes, storage nodes, use of templates, as needed.
- Which "cluster services" are used (dhcp, dns, ..etc) and network configurations (use of VLANs, NAT, etc)
- How is configuration managed post-install?
- Describe update methods, eg., use of yum repos (which?)
- Describe virtualization frameworks in use, and which services are running in vm's.
- Describe any virtual cluster management frameworks in use (if any) and/or how are virtual machines administered at your site: instance creation, stop/start, migration and/or failover, etc.
- Describe any benchmarks studies of vm's

https://indico.cern.ch/conferenceDisplay.py?confId=141745

6

US ATLAS Facilities Workshop on Virtual Machines and Configuration Management

- Two day meeting with 28 attendees from Tier 1/2/3 sites, Panda, Ops teams to share experiences to date with virtual machine and configuration management tools

- Goal to identify opportunities within the US ATLAS computing facility for collaborating and developing best practices and common infrastructure

- After day of presentations and free-form discussions three areas for future work emerged:

    - Community provisioning and asset management

    - Service configuration management

    - Virtual machine & cluster management

- Second day breakout groups developed recommendations and suggested key areas for future work:

http://indico.cern.ch/conferenceDisplay.py?confId=141745

Community Provisioning and Asset Management

- Sharing of images, configurations, repos

- Is there a need? If so what are the requirements and components?

- Architectural issues, factorizing elements

  - by-hand; turn-key appliance; intermediate-templates, tools, best practices

- Privacy and Security

- Libraries of images - drawing on whats already out there; mindful of broader picture

Steps in this direction: DDM puppet forge, Tier 3 puppet SVN

Service Configuration

- Which services are good candidates for configuration management?

- Services already configured within a framework or by a tool: what is works and is natural? What are the best practices?

- Providing feedback to packagers

Virtual Machine & Cluster Management

- Which capabilities are required, and by importance and why

- Consider critical head node services as distinct from worker nodes

- Requirements for (and meanings of) high availability

- Initial virtualization from physical machines

- Live migration & load balancing

- Hardware configurations

- Self-service management system

- VMware, RHEV - any open source alternatives worth pursuing

Summary write-ups available from each group on website

US ATLAS Federated Xrootd Workshop and Bestman Assessment

chaired by Robert GARDNER (UNIVERSITY OF CHICAGO)

from Monday, 12 September 2011 at **08:00** to Wednesday, 14 September 2011 at **18:00** (US/Central)
at **University of Chicago Booth School of Business - Gleacher Center (422)**

Manage ▾

Description **Workshop to focus on:**

- Organized deployments for T3 sites using common installation & configuration patterns
- Coordinated tests across all sites for functionality and performance
- Tests of redirection (federation) options (T3-T3, T3-T2, T3-T1, etc)

In addition, the afternoon of the second day will host that **Bestman Assesment meeting** which has scope and charge of:

- Execute a comphrehensive review of the Bestman software for maintenance, update, and support for the next three years for known usage as required by OSG stakeholders.

https://indico.cern.ch/conferenceDisplay.py?confId=149453

- Technical working meeting in Chicago Sep 12-13

- Participation from T1, T2, T3 sites, USCMS, OSG Storage & Integration groups, US ATLAS Analysis Support teams

- Discussion of deployments and testing to date, technical issues going forward
  (see further http://www.usatlas.bnl.gov/twiki/bin/view/Admins/FederatedXrootdWorkshopSep2011.html)

11

FAX Status monitor

ATLAS EXPERIMENT

USATLAS Federated Xrootd Status - 2011-10-10 13:33:16

Frequently Asked Questions

continuous lightweight probes every 15 minutes

Host: atlas29.hep.anl.gov

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.ping	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-compare	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-direct	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-fax	2011-10-10 13:20:02 CDT	YES	2011-10-10 13:35:00 CDT	OK

Host: atlgridftp01.phy.duke.edu

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.ping	2011-10-10 13:20:05 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-compare	2011-10-10 13:20:07 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-direct	2011-10-10 13:20:05 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-fax	2011-10-10 13:20:07 CDT	YES	2011-10-10 13:35:00 CDT	OK

Host: dcdoor09.usatlas.bnl.gov

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.ping	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-compare	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-direct	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.xrdcp-fax	2011-10-10 13:20:02 CDT	YES	2011-10-10 13:35:00 CDT	OK

Host: dcdoor10.usatlas.bnl.gov

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.grid-xrdcp-compare	2011-10-10 13:20:07 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.grid-xrdcp-direct	2011-10-10 13:20:07 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.grid-xrdcp-fax	2011-10-10 13:20:05 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.ping	2011-10-10 13:20:05 CDT	YES	2011-10-10 13:35:00 CDT	OK

Host: gk06.atlas-swt2.org

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.grid-xrdcp-compare	2011-10-10 13:20:07 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.grid-xrdcp-direct	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.grid-xrdcp-fax	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK
org.usatlas.xrootd.ping	2011-10-10 13:20:06 CDT	YES	2011-10-10 13:35:00 CDT	OK

Host: glrd.usatlas.org

Metric	Last Executed	Enabled?	Next Run Time	Status
org.usatlas.xrootd.ping	2011-10-10 13:20:02 CDT	YES	2011-10-10 13:35:00 CDT	OK

- Global redirector w/
- Tier 1 BNL
- All Tier 2s
- Tier 3GS via N2N, LFC@T1
- 6 Tier 3s
- OSG ITB testing site

Technical issues identified (cf Wei's presentation tomorrow)

Benchmarks & testing underway

↓ Following up on Open Isssues from Workshop
   ↓ N2N2
   ↓ Intergrated checksumming
   ↓ Proxy server
   ↓ 3.1 release schedule
   ↓ sss module development
   ↓ xprep warnings
   ↓ Logging, alerts
   ↓ FRM script standardization
      ↓ cmdsd+dcache/xrootd door
      ↓ Authorization plugin
   ↓ ANALY queue
   ↓ D3PD example
↓ Sharing Configurations
↓ New benchmarking results
↓ Monitoring
↓ Ganglia monitoring information

Service Deployments in this Phase

Site level service deployments in this Phase

- CVMFS

    - complete this - already in production at Illinois, AGLT2, MWT2 (new queue), ...

    - Requires close coordination with Xin and Alessandro

- OSG CE

    - Last pacmanized version

- OSG Worker Node

    - Has updated LFC python bindings

- PerfSonar

    - Release imminent

    - New boxes

Middleware Deployment Planning

- As discussed at WLCG-Hamburg, need to provide input to gLite 3.2/SL5 → EMI/EGI deployment planning
- OSG undergoing packaging/deployment reorganization this year (pacman to rpm+yum based)
- As we have seen ATLAS services particularly impacted by WN enviro, CE info & monitoring systems, UI
- In ATLAS' interest to keep capabilities in synch ahead before deploying onto production resources
- Discussing in WLCG context creating EGI-OSG interoperability testbed
  - Info sys, monitoring and accounting
  - Panda queues for pilot and transformation testing

Need to build into our program

15

Site Certification

- Defined for each quarter

- Identify key performance drivers

- Contributions from T1, T2 admins & others

- Below was Phase 17, April-June 2011

Focused tasks

Notation:	completed	work is in progress	not applicable	defer to next phase	table to be updated

Site	CPU/Storage upgrades	Network Monitoring	AnalyMilestones	FAX	CVMFS	OSG CE update	OSG non-ATLAS access
TIER1	(work in progress)	(completed)	(4950)	(completed)	(completed)	(completed)	(completed)
AGLT2	(completed)	(completed)	(3160)	(work in progress)	(completed)	(completed)	(completed)
MWT2_IU	(defer)	(completed)	(not applicable)	(not applicable)	(completed)	(defer)	(completed)
MWT2_UC	(defer)	(completed)	(3900)	(completed)	(completed)	(defer)	(completed)
MWT2_Illinois	(not applicable)	(completed)	(256)	(not applicable)	(completed)	(completed)	(not applicable)
NET2_BU	(completed)	(completed)	(600)	(completed)	(defer)	(defer)	(completed)
NET2_HU	(not applicable)	(not applicable)	(750)	(not applicable)	(defer)	(defer)	(not applicable)
SWT2_CPB	(defer)	(work in progress)	(completed)	(completed)	(completed)	(completed)	(completed)
SWT2_OU	(completed)	(completed)	(completed)	(defer)	(defer)	(defer)	(completed)
WT2	(completed)	(work in progress)	(1400)	(completed)	(table to be updated)	(completed)	(work in progress)

16

Site Certification FY11Q4

Phase 18: July - Sep

Notation:	● completed	● work is in progress	● not applicable	● defer to next phase	● table to be updated	

Site	CPU/Storage upgrades	Network Monitoring	FAX	CVMFS	OSG CE & WN update	OSG non-ATLAS access
TIER1	🔴	🔴	🔴	🔴	🔴	🔴
AGLT2	🔴	🔴	🔴	🔴	🔴	🔴
MWT2	🟢	⚪	🟢	🟢	🟢	🟢
MWT2_Illinois	⚪	🔴	🟢	🟢	🟢	⚪
NET2	🔵	🔵	🟠	🟠	🟠	🟢
SWT2_CPB	🔴	🔴	🔴	🔴	🔴	🔴
SWT2_OU	🟠	🟠	🟢	🟠	🟠	🟢
WT2	⚪	🔵	🟢	🟠	🟢	🟢

Reminder updates due Oct 19 :)

17

Meetings, joy of...

US ATLAS Facility Meetings

- Data Management

    - Tuesday, noon Central, Kaushik & Armen

    - Moving to bi-weekly schedule

- Throughput & networking

    - Tuesdays, 2pm Central, Shawn (always biweekly)

- Computing Integration and Operations

    - Wednesday, noon Central, Michael & Rob

    - Meeting has transitioned to bi-weekly

- US ATLAS federated xrootd

    - Fridays, 1pm Central, bi-weekly, Rob & Wei (biweekly)

Today's workshop themes:

Session	Speaker	Time
0 - Welcome *Heroy 153, Southern Methodist University*	ROSS, Justin 09:00 - 09:10
24 - Cloud Activities at CSU Fresno *Heroy 153, Southern Methodist University*	LIN, Cui 09:10 - 09:20
1 - Review of Facilities Requirements and Overview *Heroy 153, Southern Methodist University*	ERNST, Michael 09:20 - 10:00
2 - Planning in the Facilities Integration Program *Heroy 153, Southern Methodist University*	GARDNER, Robert 10:00 - 10:20
8 - US ATLAS Analysis Support and the Facilties *Heroy 153, Southern Methodist University*	COCHRAN, Jim 10:20 - 10:50
9 - CPU, Disk, Pledges and Requirements Discussion *Heroy 153, Southern Methodist University*	11:10 - 11:40
7 - Zeus Racks & PDU Infrastructure & other highlights for compute node, s *Heroy 153, Southern Methodist University*	GOFF, Roger 11:40 - 12:10
10 - Virtualized Worker Node *Heroy 153, Southern Methodist University*	STRADLING, Alden 13:40 - 14:00
3 - Program of work for Cloud developments *Heroy 153, Southern Methodist University*	HOVER, John 14:00 - 14:45
23 - Cloud Computing Discussion *Heroy 153, Southern Methodist University*	14:45 - 15:30
4 - DYNES and LHCONE *Heroy 153, Southern Methodist University*	MCKEE, Shawn 15:50 - 16:20
18 - Facility Networking, PerfSONAR monitoring, etc *Heroy 153, Southern Methodist University*	ZURAWSKI, Jason 16:20 - 16:40
6 - Analysis Performance and IO Optimization *Heroy 153, Southern Methodist University*	CRANSHAW, Jack 16:40 - 17:00
16 - Whole Node Scheduling Discussion *Heroy 153, Southern Methodist University*	17:00 - 17:20

Requirements, resources
Facility organization

Analysis & the Facility

Fabric Deployments, Planning

Virtualization & Cloud

Networking

Analysis optimization

20

... tomorrow

21 - Staus of AutoPyFactory: features, implementation, deployment in US and beyond		*HOVER, John*
Heroy 153, Southern Methodist University		09:00 - 09:20
12 - PD2P Caching and Job Brokering Update		*WENAUS, Torre*
Heroy 153, Southern Methodist University		09:20 - 09:50
15 - Analysis Queues: Performance & Capacities		*GARDNER, Robert*
Heroy 153, Southern Methodist University		09:50 - 10:20
11 - Tier 2D Operational Issues		*SOSEBEE, Mark*
Heroy 153, Southern Methodist University		10:20 - 10:40
22 - Federated Identity Management		*HOVER, John*
Heroy 153, Southern Methodist University		11:00 - 11:20
13 - Data Federation with Xrootd		*YANG, Wei*
Heroy 153, Southern Methodist University		11:20 - 11:50
14 - dCache Locality Aware and Federation Prospects		*MCKEE, Shawn*
Heroy 153, Southern Methodist University		11:50 - 12:20
17 - Tier 3 Storage Survey and Status of Tools for Tier 3 access, futures		*BENJAMIN, Doug*
Heroy 153, Southern Methodist University		14:00 - 14:30
5 - Deploying OSG with yum and rpms		*ROY, Alain*
Heroy 153, Southern Methodist University		14:30 - 15:00
0 - Open Discussions (various, TBD)		
Heroy 153, Southern Methodist University		15:00 - 17:00

Pilots, caching, brokerage, operations

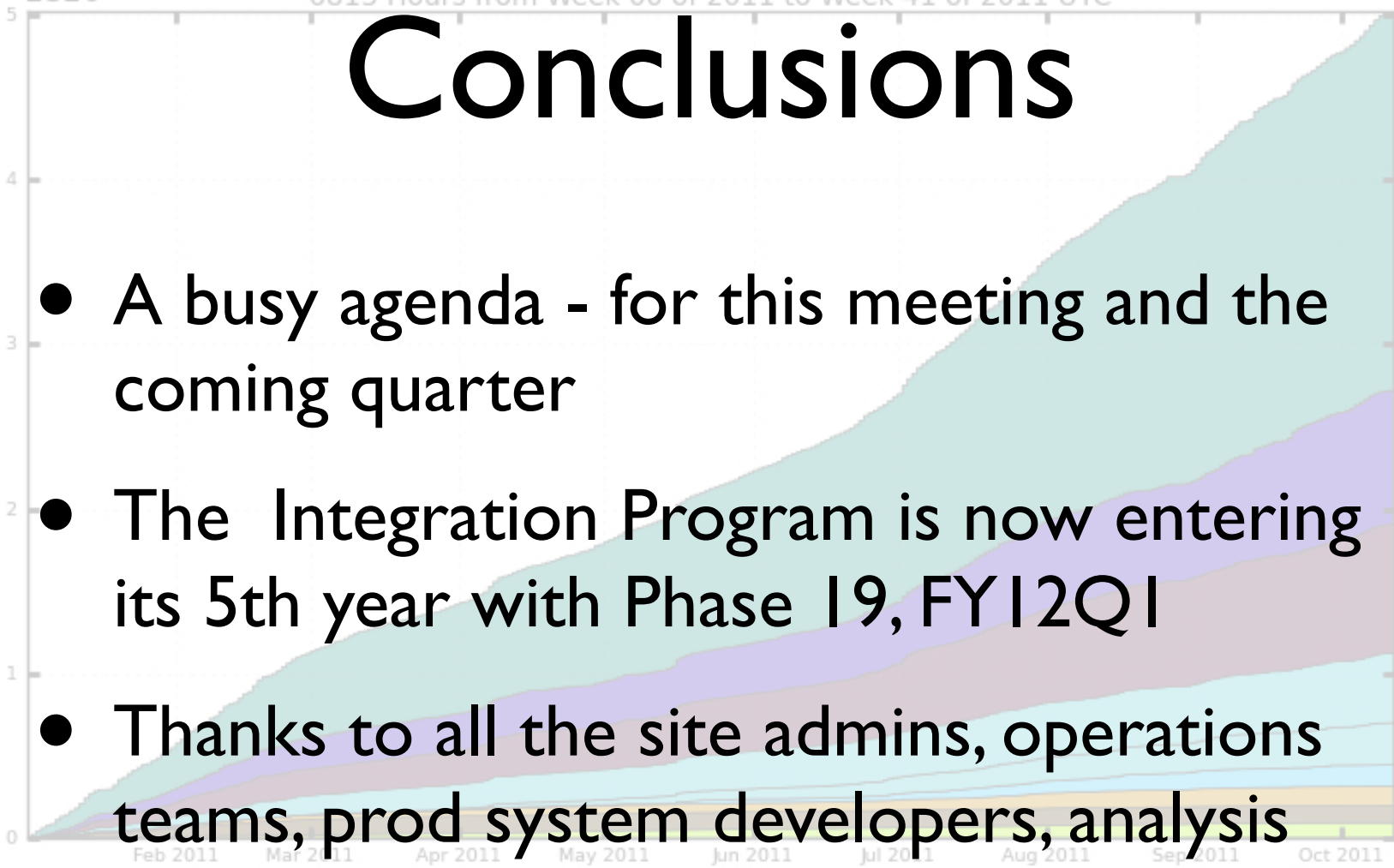*Federating... identity & storage*

Tier3 integration

Service Deployments

Conclusions

- A busy agenda - for this meeting and the coming quarter

- The Integration Program is now entering its 5th year with Phase 19, FY12Q1

- Thanks to all the site admins, operations teams, prod system developers, analysis support, collaborators from OSG and WLCG, & users, for making this possible