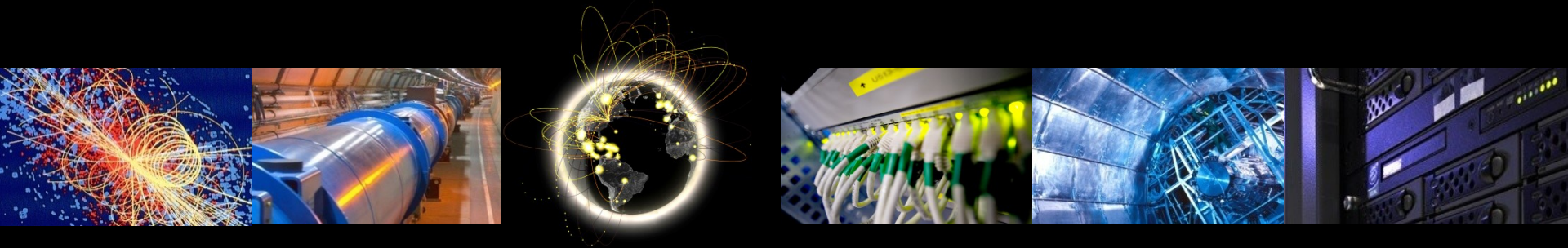


Information System Use Cases and Future Steps

Lorenzo Dini – WLCG Information Officer

June 08, 2011



Use Case Survey

- Contacted

- Experiments



- IT-ES Experiment Links (forwarded request and answers)
- Details from experiments experts
- Previous GDB presentations

- Middleware providers



- Component by component (gLite only)
- ARC and OSG can be added later

- Information Systems experts



- Laurence Field
- Flavia Donno (former WLCG-IO)

Overview

- 6 Use Cases
 - Details, issues and priorities
- Costs for maintenance of the WLCG IS
 - Software, deployment, ops
- Crossroads
 - 4 ways to go
 - The lazy
 - The slow and steady
 - The radical
 - The rocky



1 – Service Discovery

- List available Grid Services: CE, SE, FTS
 - What about VOBoxes, SQUIDs,.... ?
- Real-time automated queries
 - ATLAS and CMSSW Software Installation
 - FTS (cron-job and from API to discover SEs)
 - Dashboard (CEs, SEs, and Latitude/Longitude)
- Manual query to feed configuration systems
 - ATLAS FTS discovery
 - LHCb CE discovery
 - CMS CE discovery for pilot job submission

1 – Issues

- Dynamic architecture used for semi-static information
 - Initially intended for use in WMS
 - What doesn't advertise shouldn't get jobs (Black-Hole-Avoidance)
- Temporary unavailability of information
- Workarounds by the info consumers
 - FTS uses an XML cache
 - Dashboard caches site locations
 - Hardcoded values (the final approach to caching)
 - All are different flavors of **CACHING**

2 – Software Version

- Discover installed software in SubCluster
- HostApplicationSoftwareRunTimeEnvironment
 - Any schema can be mapped to a ‘string’
 - Quite an interesting method to publish information
- Considered fundamental
- Only provided by the WLCG Information System
- Used by ATLAS for Athena version discovery
- Used by CMS for CMSSW version discovery
- Not elegant, but works

3 – SE capacity

- Discover DarkData
 - Data in storage and not in catalog
 - Fundamentally a storage accounting problem
- InstalledCapacity VS TotalSize
 - Information not available in SRM
- Requirement for ATLAS
- Wish for LHCb
- CMS
 - Cannot use SRM as not using space tokens
 - Need more fine-grained sizes (per directory)
 - Not clear weather this should be handled in the IS or SE (du -hs ./)

3 – Issues

- Different values in WLCG IS and SRM

SITE	SPACETOKEN	SRM free	BDII free	SRM used	BDII used	SRM total	BDII total	BDII guaranteed	BDII installed
INFN-GENOVA	HOTDISK	0.9	1.0	1.0	1.0	1.9	1.9	1.9	1.9
	LOCALGROUPDISK	10.2	10.1	3.0	3.0	13.2	13.2	13.2	13.2
	PRODDISK	1.9	2.1	0.3	0.1	2.2	2.2	2.2	2.2
	SCRATCHDISK	4.0	4.0	0.1	0.1	4.1	4.1	4.1	4.1
INFN-MILANO-ATLASC	DATADISK	173.4	37.8	210.3	122.2	383.7	160.0	160.0	160.0
	GROUPDISK	75.2	29.8	32.6	12.6	107.8	42.4	42.4	42.4
	HOTDISK	1.0	0.2	1.0	0.8	2.0	1.0	1.0	1.0
	LOCALGROUPDISK	1.9	7.9	52.0	37.2	53.9	45.1	45.1	45.1
	PRODDISK	3.2	7.2	2.6	1.9	5.8	9.1	9.1	9.1
	SCRATCHDISK	2.6	5.7	5.4	4.5	8.0	10.2	10.2	10.2
INFN-NAPOLI-ATLAS	CALIBDISK	24.4	24.7	47.1	46.8	71.5	71.5	71.5	71.5
	DATADISK	198.3	198.4	164.5	164.5	362.8	362.8	362.8	362.8
	GROUPDISK	69.7	69.7	40.3	40.2	110.0	110.0	110.0	110.0
	HOTDISK	1.3	1.3	0.9	0.9	2.2	2.2	2.2	2.2
	LOCALGROUPDISK	17.7	17.7	48.3	48.3	66.0	66.0	66.0	66.0
	PRODDISK	11.4	11.6	4.0	3.8	15.4	15.4	15.4	15.4
	SCRATCHDISK	3.0	3.3	19.0	18.7	22.0	22.0	22.0	22.0
INFN-ROMA1	CALIBDISK	13.6	13.7	17.2	17.0	30.8	30.8	30.8	30.8
	DATADISK	164.6	164.0	154.3	154.8	318.9	318.9	318.9	318.9
	GROUPDISK	95.6	85.1	14.4	24.9	110.0	110.0	110.0	110.0
	HOTDISK	2.4	2.4	0.9	0.9	3.3	3.3	3.3	3.3
	LOCALGROUPDISK	22.0	22.0	33.0	33.0	55.0	55.0	55.0	55.0
	PRODDISK	14.5	14.7	2.0	1.8	16.5	16.5	16.5	16.5

4 – CE Status and Queues

- Ranking CEs for job submission
- LHCb Dirac uses WMS for pilot submission
 - Plan to migrate to direct submission to Cream-CE as soon as LCG-CEs are decommissioned
- CMS CRAB, ProdAgent and WMAgent
 - Plan to increase pilot jobs (100% pilot in 2012-13)
 - Freedom of Choice fundamental for black-listing
 - Form of VO based information annotation
- ALICE queries the Cream-CE Resource-BDII
 - From a VOBOX within the same site
- ATLAS is planning to use CEStateStatus

4 – Issues

- Wrong information published (occasionally happens in small sites)
 - 0 waiting jobs in queue
 - 4444 published too often
- EstimatedResponseTime
 - Often appears to VOs to be random
 - One number to lead them all
 - Not per-user (high priority users skip queue)
 - Not critical as main decision is based on CE-SE link

5 – Configuration

- Augment information with experiment-controlled annotations
- Add experiment-specific configuration
 - Not a new concept: Installed Software and FCR
- Single, central, independent services:
 - ATLAS is migrating from a Python dictionary to AGIS (4 months timeline)
 - CMS site-contacts keep the SiteDB updated
 - ALICE has AliEn LDAP
 - LHCb has Dirac Configuration System

5 – Issues

- Duplication of effort
 - Several information in common
 - Topology, Site info, CE list, SE list, FTS list, Admins
 - Authentication
- Different information representation
 - Makes it difficult to share tools
- Not leveraging the experience that led to the GLUE-2 Information Model
 - Developed under the participation of IS experts, experiments, sites
 - Implementation independent, extensible schema
 - <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4736858>

6 – Installed Capacity

- Accounting report generation based on the actual status of the WLCG infrastructure.
- REBUS (based on GStat2)
 - Requires detailed information about SubCluster, CE, SE, SA, [Access|Control]Protocol and Service
- Issues
 - Occasionally sites do not publish information according to the agreed specifications
 - https://twiki.cern.ch/twiki/pub/LCG/WLCGCommonComputingReadinessChallenges/WLCG_GlueSchemaUsage-1.8.pdf

Usage Summary

	SERVICE DISCOVERY	SOFTWARE VERSIONS	SE CAPACITY	CE STATUS QUEUES	CONFIG	INSTALLED CAPACITY
ALICE				With Issues	Future Plan	
ATLAS	With Issues	Satisfactory	Not satisfactory	Future Plan		
CMS	With Issues	Satisfactory	Not satisfactory	Satisfactory		
LHCb	With Issues		Future Plan	With Issues		
WLCG						

 Satisfactory  With Issues  Not satisfactory  Future Plan

Open Questions

- Is there “hidden” or “unconscious” usage?
 - GStat2
 - SAM tests, Gridview
 - LCG-utils
 - Gfal
 - Apel
 -



Costs

- Activities only, effort is limited by available resources
- Software maintenance (last year figures)
 - GGUS tickets: 33
 - Bugfixes: 17
 - LDAP problems investigation
 - Build, Release, Testing/Certification process
 - 5 Top-level and Site-level releases (gLite-3.2, EMI-1)
- Site deployment and configuration
 - Hardware requirements
 - Configuration of information providers
- Network use
 - Due to required “freshness” quite high for some sites
 - Not an issue for any T1/T2

Crossroads

- Current IS has been developed with WMS as the main driver
 - Very detailed information on computing resources
 - Queues, number of active jobs, open slots etc.
 - Very high availability requirements for all data
 - Avoidance of stale data by all means
 - Updates on the scale of (2-5) minutes
- Use has shifted:
 - Service discovery, annotation and reporting (monitoring)
- Very different requirements and characteristics
- Crossroads
 - 4 ways to go
 - the lazy, the slow and steady, the radical, the rocky

The Lazy – No Action

- Pros
 - + No additional effort required



- Cons
 - High maintenance effort
 - Several issues on most of the use cases
 - Impossible to achieve other use cases

- Not optimal

The Radical - Decommission



- Pros
 - + Zero maintenance effort
- Cons
 - Software Versions for CMS and ATLAS
 - CE status and queues for CMS, LHCb and ALICE
 - Freedom of Choice of Resources
 - Service Discovery
 - Installed Capacity
- Currently not a solution without a serious disruption of services

The Slow and Steady - Improve

- Try to fix current usage with the current technology
 - + Caching system to improve service discovery reliability
 - + Improve information providers for the SEs
 - + Nagios probes to monitor information quality
- Started already (Proposal by Flavia and Laurence)
- Cons
 - Fixes the symptoms, not all problems
 - Still maintenance effort
 - More a bridge than a solution



The Rocky – Evolve – I

- Separate the use cases
 - Mix, match and extend our building blocks
 - When unavoidable add new components
- + Centralized and reliable service discovery system
 - + Site registration service
- + Make resource BDII publicly and directly accessible
- + Activity already started as EMI Registry
- Cons
 - Requires development
 - Requires infrastructure changes



The Rocky – Evolve – II

- Provide a single system for experiment annotation and configuration
 - + Elimination of duplication of effort
 - + Leverage the experience that went into GLUE-2
- Requires collaboration between the experiments, the IS developers and projects
 - Many open questions
 - Maybe a mini WLCG-WorkShop on IS ???

My 2 cents

- Improve where possible
 - + Caching system to be deployed
 - + Information providers for the SEs
 - + Nagios probes to monitor information quality
- Evolve only with a strong commitment from the experiments
 - Work together on the specifications
 - Use the system once completed



Thank you!

- **ALICE:** Maarten Litmaath, Maria Dolores Saiz Santos
- **ATLAS:** Alessandro Di Girolamo, Simone Campana
- **BDII, EMI Registry:** Laurence Field
- **CMS:** Andrea Sciaba, Stefano Belforte
- **Dashboard:** Pablo Saiz
- **FTS:** Michail Salichos
- **LHCb:** Stefan Roiser, Roberto Santinelli
- **SAM:** Wojciech Lapka, Pedro Andrade
- **WLCG:** Flavia Donno