



Accounting Update

Dave Kant
Grid Deployment Board
Nov 2007



Overview

- User Level Accounting
- VOMS Groups/Roles
- APEL Status
- Tier2 Accounting and Reporting
- Issues
- Suggestions



User Level Accounting

- User Level Accounting Delivered
 - UserDN captured from CE log files (grid-jobmap logs)
 - APEL uses the data to build accounting records
 - Data published to GOC with on-the-fly encryption using APEL public key (1024 bit RSA)
 - At the GOC data are extracted from RGMA and stored in a Central Accounting Repository.
 - Data decrypted using APEL private key
 - User Level summary table created
 - On-the-fly encryption using EGEE Portal certificate
 - Encrypted table pushed to CESGA portal
 - Portal decrypts data and provides SSL based access to the summaries.



User Level Accounting

- Testing
 - We encouraged sites to publish UserDN.
 - Sites had to manually configure APEL to perform on-the-fly encryption.
 - Our 2007 sample contains 1067 distinct UserDNs from 33 sites.
 - No problems seen decrypting UserDNs at the GOC.
- Observations
 - NIKHEF publishing their own encrypted UserDN strings
 - Example LCGUserID: HPfh56sbc3AYKDn1Yusxgg
 - Can only attribute usage to the VO



VOMS Groups and Roles

- UserFQAN
 - Capture UserFQAN from grid-jobmap log on CE
 - FQAN chain processed at the GOC to derive Group and Role from the primary part of the chain.
 - If UserFQAN present, we can use the Group to derive the VO of the user submitted job (otherwise we use the local unix group).



UserFQAN

- Testing

- Our 2007 sample shows the following Groups and Roles for ATLAS

PrimaryGroup	PrimaryRole
/atlas	Role=lcgadmin
/atlas	Role=NULL
/atlas	Role=production
/atlas/ca	Role=NULL
/atlas/lcg1	Role=NULL
/atlas/nl	Role=NULL
/atlas/soft-valid	Role=NULL
/atlas/soft-valid	Role=production
/atlas/usatlas	Role=production

- Matrix looks reasonable.



APEL Status

- Production
 - New release of APEL UPDATE 35 for gLite 3.0.
 - Main features are:-
 - UserDN and UserFQAN support
 - Joins match SpecInt2000 of the ClusterID to the ResourceIdentity of the CE.
 - Additional accounting table to provide a high-level checksum view of sites accounting database.
 - This is used by GOC to verify if the site has published all of its accounting data for SAM.



APEL Status

- Development

- Bug fixes and improvements

- <http://goc.grid.sinica.edu.tw/gocwiki/ApelDevelopment>

- Critical bug fixes

- Support for multiple SpecInt200 per CE (Savannah Bug # 28593)

- Tested at CERN and IN2P3

- Impact:

- » Can lead to wrongly assigned SpecInt2000 in accounting

- » Sites with multiple CE's that re-evaluate their SI2K numbers

- Log rotation of grid-jobmap logs based on UTC (Bug # 28592)

- Tested at CESGA and CERN

- Impact:

- » Can lead to missing accounting data



APEL Status

- Development

- Enhancements

- Bug # 29400 - Improved LSF Log Parser (Tested at CERN)
 - Bug # 29401 - Identification of Gatekeeper logs
 - Bug # 29577 - Joining data between Condor and grid-jobmap logs
 - Bug # 24083 - Gap Publisher

- These bug fixes have been written to the CVS but

- We have some issues to address concerning ETICS and SLC4
 - Bug # 29510
 - Location of log4j, bouncy-castle and mysql-connector-java libraries in APELs build.xml files and the startup scripts.
 - Consequently, we have not yet produced a new release tag or patch.



APEL Future Work

- New things in the pipeline (not implemented)
 - Accounting Local Usage
 - What information do we want?
 - Attribute usage to the VO
 - Do we care about local users identity? Or Normalised usage?
 - For a High-level Anonymous summary at VO level
 - Bug # 26995
 - Evaluate a summary of all batch log data that did not get included in a grid join.
 - Attributes usage to the VO, but not at the User-level
 - Summaries are evaluated on-the-fly every time the publisher executes after the grid join process.
 - Issues
 - Normalised Usage
 - » Matching the local job to the SpecInt2000 may be problematic as there is no information about the CE in the batch log (?)
 - » This is not a problem for sites that run CE / batch servers on the same node.
 - Mapping the local unix group to a known VO
 - » Mapping table at the GOC ... is there a better method available?



APEL Future Work

- Support for MPI Jobs
 - Bug # 31027
 - IN2P3 have sent some Torque/PBS logs for MPI jobs and verified that APEL does not support MPI.
 - » The total CPU time is correct
 - » Wall time derived is underestimated because PBS publishes this per CPU.
 - » Efficiencies > 100%
 - APEL needs to take into account the Number of CPUs to get total WCT
 - What information do we want to capture from an MPI job?
 - CPU and Wall time
 - » Only need to determine the number of CPUs to fix the WCT
 - Categorise Grid Jobs
 - » Requires modifications to accounting record in order to distinguish them from POGJ.



Resource Trees / MOU Pledges

- Tier2

- Sue Soffano Spreadsheet

- Tree representing the Tier2 structure has been delivered.

http://www3.egee.cesga.es/gridsite/accounting/CESGA/tier2_view.html

- Report showing the “Tier2 MOU SI2K Pledge” against the actual usage delivered.

<http://www3.egee.cesga.es/gridsite/accounting/CESGA/reptier2.html>



Issues

- User Level Accounting
 - On-the-fly encryption would be better controlled by YAIM (Bug # 31015)
 - Specify options to publish UserDN in site-info.def file.
 - Not Implemented.
 - Use of Service certificate for encryption.
 - APEL client uses an RSA public key, but not a certificate.
 - UserDN decryption chain used in production should be implemented by CESGA for the PPS service.
 - Work has started
 - What happens if the User changes their UserDN? How does the User access their data if they no longer have the old certificate? Do we need a mechanism to track the UserDN history?
 - Case Study; changed institutes and the CA issued a new certificate when the old one expired.
 - `/c=uk/o=escience/ou=clrc/l=ral/cn=dave kant`
 - `/c=uk/o=escience/ou=queenmarylondon/l=physics/cn=dave kant`



Issues

- Tier2 MOU Pledges
 - We need to make sure that we can distinguish between Tier2s that pledge against all LHC VOs, and those that pledge against a specified VO.
 - Particularly important as some sites appear in multiple Tier2's because they pledge on a VO-by-VO basis.
 - Assume that if the VO is not in the Tier2 name, then the pledge represents a total for the entire LHC
 - Are there any MOU pledges for storage?



Suggestions

- More Spreadsheets?
 - Can we have a Tier1 spreadsheet?
 - What about VO specific spreadsheets?
 - Clouds-of-Atlas?



Farewell

- Leaving the project at the end of the week.
- Thankyou for all your help and support.
- Goodbye and Good Luck!