New Types of Accounting

*Beyond CPU*

John Gordon, STFC

EGI Technical Forum,

Prague, September 2012
Outline

• Requirements
• Readiness to Receive
• Records under development
• Next Steps
Requirements for New Types of Accounting

- Storage
- MPI
- Applications
- Cloud/VM
- New versions of cpu UR (CAR)
Readiness to Receive New Records

- The APEL repository is in the middle of a migration but the capability to receive usage records over SSM is already in production for both *Job Records* and *Summary Job Records*. It is being used by Open Science Grid and CERN, with a number of other publishers set to migrate.

- Consumers are in place on the test service for CAR and StAR.

- Adding others is a lightweight process. Much simpler than defining the record schema.
CAR

- CAR is an EMI revision of the OGF UR v1.0 rationalising some issues and including some common extensions deployed in a number of implementations.
- Not UR 2.0 which is a bigger revision started by OGF UR-WG
- XML document
- CAR v1.0 was agreed but now some issues have arisen.
- APEL new schema incorporates CAR
  - But there might be some tweaking.
The EGI Federated Cloud Task Force has adopted a VM UR based on OGF UR:
- Many values like memory and network were defined in the UR but seldom if ever implemented.

Sensors to collect the information and construct usage records were written:
- OpenNebula and OpenStack so far.

SSM used to send to RAL.
• StAR (Storage Accounting Record) is a version of the OGF UR 1.0 adapted for accounting of storage utilisation

• Developed by EMI, submitted as a public document to OGF, revised in the light of public comments.
  – Added Sitename, allocated space, and revised time definitions.

• Status – being implemented by EMI storage providers (dCache, StoRM, DPM) in EMI-3
EMI StAR timeline

• May 2011: StAR definition agreement
• May 2012: Each SE provided an implementation plan
• June 2012: Each SE to implement accounting sensors using SSM to publish
• June 2012: Accounting repository (APEL) ready to receive storage records
• September 2012: Each SE to test and deploy their accounting sensors (for EMI 3 release)
• January-April 2013: Bug-fixing in EMI 3 RC’s
• April 2013: Accounting sensors released in EMI 3
Storage Solutions

- dCache (EMI)
- StoRM (EMI)
- DPM (EMI)
- Castor
- EOS
- **BestMan**
- Gratia (already collecting)
- xrootd
- hadoop
- Gstat
- Cloud Storage
- (Unicore)
- (ARC)
- iRods
- ?????

- The middleware in scope for EGI and WLCG
- Anything that collects storage information could also publish usage records
- eg Julia Andreeva talk at CHEP
• EGI Virtual Team on MPI has just reported.
• Agreed that UR nodecount and ncpus(ncores) are sufficient.
  – Not in production APEL. Will be in CAR and emi3 client.
• Asked for a flag that a job is deliberately parallel, rather than just accidentally.
  – This will allow visualisation of the parallel subset.
  – RT ticket slipped through the net. Will need CAR addition.
Applications

- Little progress recently.
- CESGA produced a plan for VOMS attributes for specific application.
- Hannover have a local prototype based on process accounting.
- No UR proposal drafted yet.
Data Centre Accounting

- StAR is about storage allocated and used.
- What about who used the data? And how often?
- What would the big sites like to see about the use made of their data centre?
  - Bandwidth?
  - Source/destination?
  - Identity of users?
Next Steps

• CAR
  – Ready for direct testing today
  – Apel client in EMI-3

• Storage
  – Records for testing
  – Visualisation strawman
  – Engage other storage suppliers

• Cloud
  – Extend range of infrastructures
  – Visualisation strawman
Next Steps (2)

- MPI
  - Test data today from DGAS
  - Visualisation strawman
  - Parallel flag in CAR?
  - Apel client in EMI-3
- Applications
  - Kickstart
- Data Centres
  - Requirement Gathering
  - EU-DAT?
  - Resource Centre Forum