Using Gstat 2 to Check your Published Information

Stephen Burke

RAL

- Why gstat 2?
- What does it display?
  - Demo …
- What to look for
  - and how to fix/report problems

- GLUE 2

  - With help from:
    - Joanna Huang (ASGC)
    - Laurence Field (CERN-IT)
    - David Horat (CERN-IT)

Motivations For Version 2

- The old gstat pages are now too cluttered
  - The EGEE Grid has grown to 320+ sites
  - 1 CE at CERN to 20+ CEs at CERN

- It is a single, centralized instance
  - EGI would like de-centralized, regional-based operations tools

- The information checks are not easily reusable
  - Difficult for use by sys admins and for software certification

- Tight coupling with SAM and GOCDB
  - Requires high-availability operations and notifications

- High-maintenance backend
  - Due to the gradual evolution of the code base

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- **Consolidate** the existing code base
  - To give a **low-maintenance** solution
- Isolate the testing component
  - To ensure that the tests are **reusable**
- Remove the dependency on the GOC database
  - To enable **de-centralized** deployment
    - Bootstrapping should be achieved by querying a BDII
- Redesign the displays to address specific use cases
  - Generally improve the presentation
- Ensure that components are modular
  - And that GStat is **extensible**

Science & Technology Facilities Council
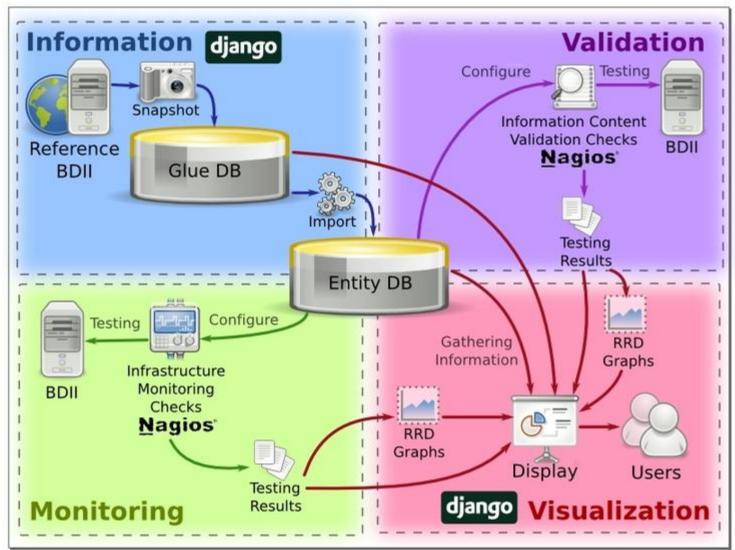Rutherford Appleton Laboratory

- Nagios
  - Manages the execution of tests and test results
  - Probes can be re-used by other OAT applications
  - Used for:
    - Information Content Validation Tests
    - BDII Service Monitor

- Django
  - Web application framework to simplify page generation
  - Object-relational mapper simplifies database integration
  - Significant experience already exists within the OAT
  - Used for:
    - Snapshot and topology import scripts
    - Web page rendering and management

Science & Technology Facilities Council
Rutherford Appleton Laboratory

gstat 2

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- How to obtain the information content?
  - By querying a BDII

- How to ensure the information integrity and quality?
  - Do different checks based on the different entries
    - Does the information agree with the schema?
    - Are the entities we expect to see published?
    - Are there things published that we don't expect?
    - Is the information logically self consistent?
      - Is the number of free CPUs less than or equal to the TotalCPUs?
    - Is there agreement with external information sources?
      - Is the host registered in DNS?
      - Does it match the GOC DB?
    - Is there conformance with extra project constraints?
      - Valid gLite version?
      - LCG installed capacity document

gstat 2

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- Drop-down lists to filter by Grid, Country, VO, EGEE_ROC, WLCG_TIER
    - Currently one name space so watch for clashes, e.g. VO == Grid
    - No registry for most of these names, just convention
    - No EGI-related info defined yet
- Information source is mostly your published GlueSite info (VOs excepted) – so make sure it's right …
- ldapsearch -x -h lcg-bdii.cern.ch -p 2170 -b o=grid gluesitename=UKI-SOUTHGRID-RALPP
- GlueSiteLongitude: -1.3163
- GlueSiteLatitude: 51.5721
- GlueSiteWeb: http://www.ppd.clrc.ac.uk/public/ppd.html
- GlueSiteLocation: Oxfordshire, UK
- GlueSiteOtherInfo: EGEE_ROC=UK/I
- GlueSiteOtherInfo: EGEE_SERVICE=prod
- GlueSiteOtherInfo: GRID=EGEE
- GlueSiteOtherInfo: GRID=WLCG
- GlueSiteOtherInfo: GRID=SOUTHGRID
- GlueSiteOtherInfo: GRID=GRIDPP
- GlueSiteOtherInfo: WLCG_NAME=UK-SouthGrid
- GlueSiteOtherInfo: WLCG_PARENT=UK-T1-RAL
- GlueSiteOtherInfo: WLCG_TIER=2

- http://gstat-prod.cern.ch/ or http://gstat2.grid.sinica.edu.tw/
  - Also links from the gstat 1 page, and information on the gstat web site

- Geo View – plots sites on a map
  - Not that interesting, but check your co-ordinates
  - You can also jump to the site view – click on a site for a popup

- LDAP View – LDAP browser
  - See directly what your site is publishing
  - Need to pick a top-level BDII – but in theory they should all be the same!
  - URL includes the BDII (and base DN) – so can be bookmarked, and can query a site BDII directly

- Service View – shows top and site BDII monitoring status
- VO View – tree view of jobs and storage by VO

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- Site View – starts with a whole-Grid summary, drill down to your site
  - Can filter the table on various criteria
  - Popup help over the field labels
- Select your site to get a detailed view
  - URL can be bookmarked
- Overall summary of your site information
  - Does it look right?
  - BDII name is checked for DNS aliases
  - VO tabs show resources per VO
    - Storage not yet separated by space token, but I have suggested that as an enhancement

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- Most links on the site view take you to a tree view of your site services
  - Can get graphs of measured quantities
    - Look for graph icon
  - "GLUE" link displays raw published info, "LDAP" link goes to LDAP browser

- BDII content validation – click BDII name for details, then click to expand each test section
  - "WARNING" means "probably wrong", "CRITICAL" means "definitely wrong" – may or may not be serious
  - Most DPM sites get a lot of warnings for non-compliant "legacy SAs" - will be off by default in the next DPM release
    - And can be turned off by hand now
  - Some other things may be due to middleware bugs, but should mostly be fixed
    - AssignedJobSlots = 0
  - Please look at what is being flagged and either:
    1. Fix it
    2. Report a bug in the middleware
    3. Report a bug in the gstat test

- GGUS ticket or mailto:project-grid-info-support@cern.ch
  - Developers are responsive
  - Relatively few problems reported so far – either gstat is perfect (unlikely!) or people aren't looking much yet
    - Daniela and Duncan reported problems

- Can also ask for enhancements – further development will be driven by demand

- You can look at the source code for the information system checks to see what it's doing
  - And indeed run it yourself

gstat 2

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- GStat 2.0 is in production
  - http://gstat-prod.cern.ch/
  - It can replace the functionality of the original version
  - gstat 1 will be switched off on June 17th

- Please
  - Provide feedback,
    - both problems and suggestions
      - project-grid-info-support@cern.ch
  - Look at the views important to you!
    - Take ownership of the information you see!
      - Submit a GGUS ticket if there is something not right!

YOUR COUNTRY NEEDS "YOU"

And now for something slightly different …

gstat 2

Science & Technology Facilities Council
Rutherford Appleton Laboratory

- Just a quick overview
  - For more info see my talks at CHEP09 and EGEE09

- Abstract schema was approved in March 2009
  - Complete redesign, not backward compatible, hence must deploy in parallel
  - All services published in the same framework, much more flexible

- LDAP rendering defined and implemented in the BDII in glite 3.2 update 5 (September 09)
  - Query on the same port (2170) but base DN is o=glue

- Generic service info provider has been written
  - First to production is CREAM, update 12 (May 2010)
  - More to follow

- Site BDII needs to aggregate from resource BDIIs and add the site (AdminDomain) info
  - Already certified, but unrelated deployment changes were controversial and have been rolled back – release is imminent
- Top BDII needs to aggregate info from site BDIIs
  - To follow soon
  - Can then see the whole Grid in GLUE 2
  - gstat will be extended to monitor it
- Storage info providers need to come from SRM developers
  - No timescale yet
- CREAM info provider needs to be extended
  - Probably it will appear incrementally – end of the year?
- Client tools
  - Service discovery on the way, lcg-utils and WMS not started

Science & Technology Facilities Council
Rutherford Appleton Laboratory