egee

Enabling Grids for E-sciencE

IT cluster activity status (Status of WMS & CE)

Francesco Prelz – IT Cluster Manager

INFN Milano

EGEE'06 Conference

Geneva, 28 September 2006

Information Society and Media

Enabling Grids for E-sciencE

- Activity on glite 3.0.*

- Activity on the gLite 3.1 branch

- Progress on TCG actions (or lack thereof)

- CE solutions and CREAM

- VOMS and DGAS status

- Summary and Plans

Enabling Grids for E-sciencE

- Backwards-compatible release

  → **All components of the 'LCG' stack are still there.**

- Since the beginning of EGEE-II, we delivered 6 bugfix tags for `org.glite.ce` and 13 bugfix tags for `org.glite.wms`.

- We were asked to *back-port* several features from the 3.1 branches:

  - Glue 1.2 Voviews used for matchmaking.

  - Forwarding of CE requirements to/via BLAH

  - DAG node resubmits

  - Creation of common-format accounting logs in BLAH and integration with DGAS.

  - Support for job 'prologues', that can check and prepare the environment. Failure of the prologue causes a shallow resubmission. Support for job 'epilogues' was also added.

Enabling Grids for E-sciencE

- 3.1 maintenance branches were forked as requested on March 31st, and are being continuously "internal"-tested since then.

- Remaining "not-yet-backported" features:
  - `g-pbox` (was excluded from glite 1.5 and 3.0)
  - Proxy renewal service with no WMS dependencies
  - ICE-CREAM (see later)
  - Streamlined logging of bulk jobs in WMproxy
  - *WM code cleanup*
  - *Jobwrapper code cleanup*
  - Status and statistics collection on WMS node.

- 3.1 WMS requires at least '3.1' security.

Enabling Grids for E-sciencE

- The scenario for deploying the 3.1 version of gLite has been unclear for a long time, so the (in principle release-candidate) 3.1 branch was used as a testing ground for:
  - Merging changes to support 64-bit builds
  - Merging changes to complete a build against VDT 1.3 (updated OpenSSL caused problems at runtime)
  - Obtaining a build from the new build system

- This caused some interference: moving now from internal test machines to preview testbed.

- Should allow to upgrade the WMS node (WMS + deps) components only.

301. *"Configuration that defines a set of primary RB's to be used by the VO for load balancing and allows defining alternative sets to be used in case the primary set is not available."*

– The appropriate modifications were merged into the 3.1 branch at the end of July, supporting the requested fall-back to Service Discovery. This includes a callout to refuse jobs when an overload condition is detected by a local script.

302. *"Single RB end point with automatic load balancing."*

– This is the "real" high availability solution. The TCG decided it's not a priority.

303. *"No loss of jobs due to temporary unavailability of an RB instance."*

– **We are working on a "hot standby" configuration. Testing of the LCG HAGD and of 'filelist' code replacements just resumed. Completion, commit, integration and test depend on 3.1 plans.**
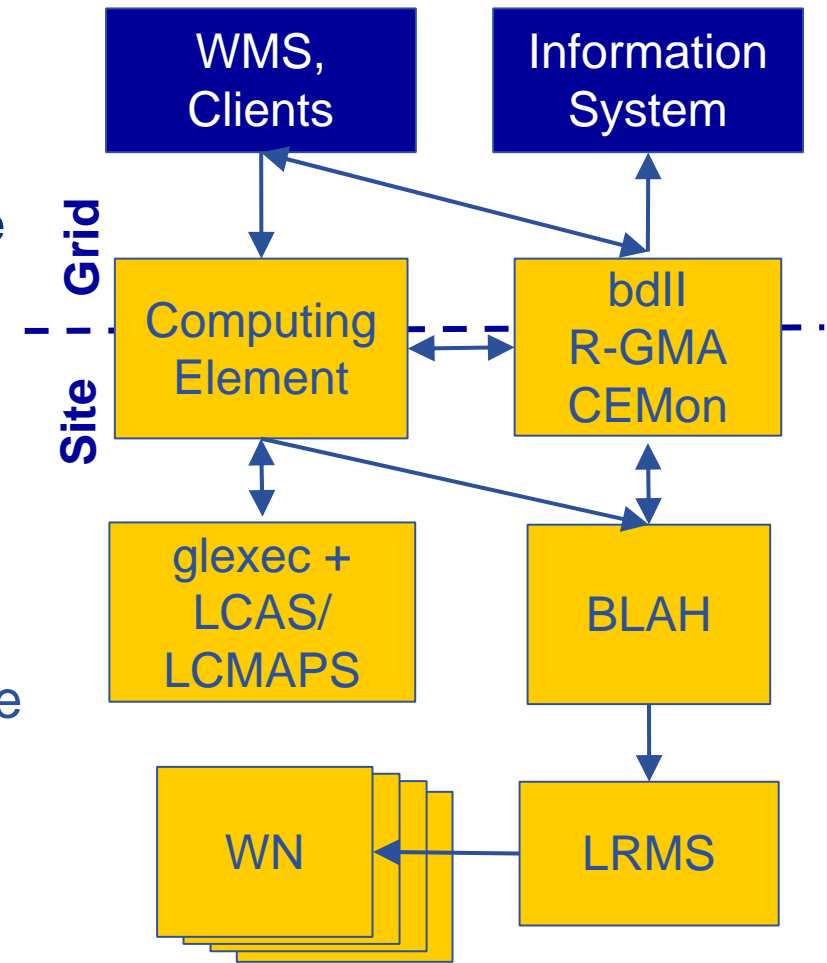
304. "Handling of 10**6 jobs/day."

- *After massaging bulk submission in, we are now working on a bulk match-making feature, based on user-specified "significant" attributes. At the same time, we are decoupling bulk-submitted jobs from DAGs. Work just stabrted, in parallel with #303, but these features will be added in a tag after the tag for #303..*

304. "Better input sandbox management (caching of sandboxes)."

- *The few corrections needed to support Input Sandbox files via HTTP (and HTTP proxies) were committed into the 3.1 branch at the end of July. `htcp` (from gridsite), which provides proxy support, was used.*

- Three flavours available now:

  LCG-CE (GT2 GRAM)

  gLite-CE (GSI-enabled Condor-C)

  CREAM (WS-I based interface)

- How to deal with them:

  – LCG-CE is in production now but will be phased-out by the end of the year

  – The gLite-CE still needs thorough testing and tuning. Being done now

  – CREAM is being deployed on the JRA1 preview test-bed now. After a first testing phase will be certified and deployed together with the gLite-CE

- BLAH is the interface to the local resource manager (via plug-ins)

  – CREAM and gLite-CE

  – Information pass-through: pass parameters to the LRMS to help job scheduling

Grid / Site

WMS, Clients	Information System

Computing Element	bdII R-GMA CEMon

glexec + LCAS/ LCMAPS	BLAH

WN	LRMS

Enabling Grids for E-sciencE

- Implemented functionality
  - Job submission
    - Supported job types: simple, batch jobs, MPI jobs, interactive jobs (for submissions via WMS-ICE)
  - Proxy delegation
    - Delegation 2.0 integrated
  - Job cancellation
  - Job status
  - Job list
  - Job suspension and job resume
  - Proxy renewal
  - Disabling of new job submissions
    - Can be explicitly invoked by CE admin and also possible to define policies on waiting/pending/running jobs to disable new job submissions
      - E.g. disable new submissions if the number of active jobs is > 3000
- Interfaces: CREAM CLI and Java client available
  - Of course everyone can implement his own interface to interact with CREAM (Web service)
- Comprehensive (at least we hope so) documentation available in the CREAM web site (http://grid.pd.infn.it/cream/)

Enabling Grids for E-sciencE

- Performing stress and performance tests
  - See "Preview Testbed" presentation and also the CREAM web site (http://grid.pd.infn.it/cream/) under "Test Results"

- Some issues
  - Problems with VOMS based AuthZ (bug #18244)
  - Thread safe issues in the BLAH-glexec interaction

- Current activities
  - Fixing the problems found in the tests
  - Support for bulk jobs
  - Starting JSDL support (with the OMII hat, see http://grid.pd.infn.it/omii/)

- Future activities
  - BES support (again, with the OMII hat)

Enabling Grids for E-sciencE

- ICE (Interface to Cream Environment) is responsible for WM—CREAM interaction
- It implements the functionalities currently done by JC+LM+Condor
  - Job submission, Job cancellation, Proxy renewal
  - Job status monitoring (and then taking the appropriate actions)
    - Receiving notifications about jobs from CEMon and doing an active polling if/when needed
- A single ICE instance can interact with multiple CREAM servers
  - ICE uses a thread pool to interact with CREAM servers
- Everything is user-configurable
  - ...with sensible defaults which should be OK for most installations

Enabling Grids for E-sciencE

- Tests done so far in a small and controlled environment
- Starting now with stress and performance tests in the Preview testbed
- Some next steps
  – Debugging and fixing of problems found during testing
  – Better approach for proxy renewal? (to be discussed and agreed with Daniel K.)
  – Better integration with the WM? (to be discussed with Francesco G.)

    currently ICE runs as a separate process, interacting with the WM via filelist;

Enabling Grids for E-sciencE

These updates were applied to the '3.1' branch of VOMS:

- Insertion of AA Certificates into the AC
  - Requested by SA1 (easier system management)
- VO-specific directories in vomsdir
  - Better separation of duty among VOs.
- Generic Attributes
  - (name, value) pairs could be added to VOMS credentials
    Requested by many of people. Very useful indeed.
- Updated Java APIs
  - Fixed several standing problems.

Enabling Grids for E-sciencE

- **Support for PKCS12-formatted credentials**
  - As you get them from CAs.

- **Support for GT4**
  - Verified in the context of the GIN activity.

- **Globus-free C/C++ APIs**

- **The Future**
  - New version 2.0 of voms-admin

    More stable, faster, and more maintainable.

Enabling Grids for E-sciencE

- **Site-level HLR available**
  - irepository for detailed (with user, VO and VOMS attributes) and aggregatejob information for jobs executed on a given site, or reported from multiple sites.

- **DGAS sensors are available for BLAH and globus-jobmanager based CEs**
  - PBS and LSF, with Condor and SGE being implemented.

- **VO-level HLR also available**
  - mostly for evaluation purpose, not foreseen to be deployed

- **CLI interface available**
  - query for accounting records or aggregate reports

- **Site HLR, receiving records from the "unified" accounting log via DGAS sensors, being deployed on preview testbed.**

Enabling Grids for E-sciencE

- Software from the 3.1 maintenance branches is being deployed on the preview testbed.
- TCG actions commits and other progress on the 3.0 branch (DGAS addition and VDT transition) that was supposed to happen at the end of july were sidetracked by the "escalated" developer attention requested on rb102.cern.ch :
  - **A few scale issues were identified and addressed.**
  - **Hardware and software configuration was performed.**
- Program of work to address TCG items is now resuming.
- New planning for 3.1 certification, deployment (which components, when?) and for addition of these pending changes is needed.