gLite WMS status & evolution

Fabio Capannini – INFN (presenter)
Marco Cecchi – INFN
Salvatore Monforte - INFN
Alessandro Maraschini – Elsag Datamat

JRA1 All Hands Meeting
6-7 May 2009 Cyprus

www.eu-egee.org

e-infrastructure

WMS release status

- **WMS 3.1**
  - Patch #1841 + Patch #2562 (bugfix patch) - In Production
    - Major **enhancements**:
      - *Enabled submission to CREAM CE (ICE)*
      - *Added recovery procedure for the WM*
      - *Filelist → jobDir communication mechanism*
      - *Added pre-filtered LDAP queries*
      - *Added support for MPI jobs according to the latest specifications from the MPI working group*
    - Significant **fixed bugs**:
      - Previously used CEs are not considered at all in the resubmission (bug #28235)
      - The purger under some particular circumstance segfaults (bug #29182)
      - glite-wms-wm crashes during resubmission (bug #30518)
      - A collection with pending jobs can be processed multiple times (bug #30816)

WMS release status

- Any collection submitted while the WMS is down is not recovered upon WM startup (bug #34508)
- When a collection is aborted the "Abort" event should be logged for the sub-nodes as well (bug #34510)
- Error processing DAG dependencies while generating the ISB for final node (bug #42587)
- The WM terminates unexpectedly handing a cancel request. (bug #42590)

- Patch #1841 hardly converged to certification:
  - *too much bugs fixed ("avalanche effect" triggered)*

– patch #2923: WMS 3.1: fixes for bug #47040 – in PreProduction
- Urgently requested **fix** patch for the following **bug**:
  - *WM terminates upon uncaught exception during recovery (bug #47040)*

WMS release status

- **WMS 3.2**
  - Patch #2597: WMS 3.2 for SLC4 – In Certification
    Significant redesign of the core architecture
    - Major **enhancements**
      - *Parallel match-making*
      - *Match-making with data (reviewed ISM structure)*
      - *Requests to WM are as stateless as possible; prioritized event queue processed by a thread pool; each kind of request processed by the same thread pool as a generic functor*
      - *The jobwrapper template is now cached at each WM start*
      - *Restructured jobwrapper (also removal of perl dependencies)*
      - *ISM can be dumped more often at lesser cost*
      - *added support for IPv6*
      - *WMProxy: gridsite Delegation 2.0 support added; server able to expose both old and new delegation versions (they are backward incompatible)*

WMS release status

- *WMProxy: added support for JSDL; the server translates incoming JSDL into glite JDL (following layers are unable to "talk" in JSDL)*

- Significant **fixed bugs**:
  - *WMProxy should allow jobListMatch to timeout (bug #28696)*
  - *wmproxy stop/restart can leave old processes behind (bug #48172)*
  - *In some circumstances, jobs which are killed by CREAM job wrapper might remain in ICE cache forever (bug #39807)*
  - *A bad handling of delegations slow down dramatically the submission rate of ICE (bug #44604)*
  - *There's a mem leak in ICE that raises in some very rare circumstances (bug #47389)*
  - *Apparent database corruption when ICE exits. (bug #47996)*
  - *...and more which are still in certification*

WMProxy Client Status 1/3

- **WMS-UI 3.2 in certification on SL5**
  - patch #2875
- **Major new features included in patch #2875:**
  - IPv6
    - glite-wms-job-* commands compliant with next-generation protocol
    - Backward compatibility of the code with IPv4
    - Multi-language compliance for Java/C++/Python WMProxy client APIs
  - Delegation:
    - Gridsite Delegation 2.0 support added
    - Client able to perform both old (1.0) and new delegation approach depending on Server version

WMProxy Client Status 2/3

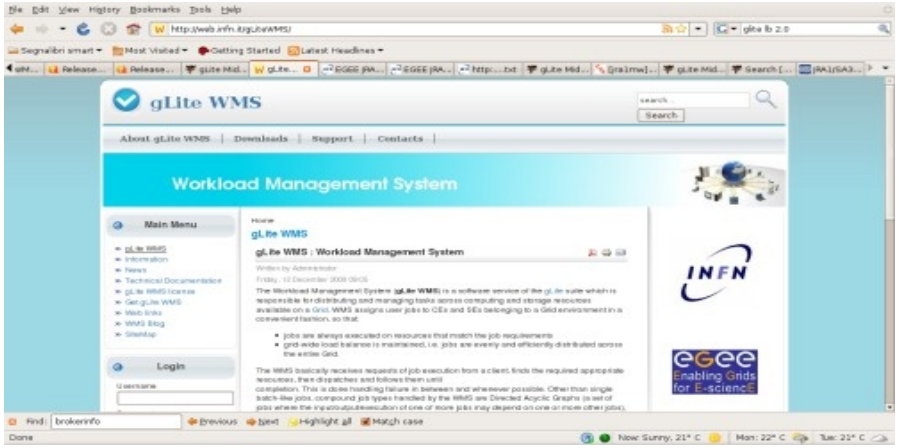
- **Major new features included in patch #2875(follows):**
  - JSDL
    - **J**ob **S**ubmission **D**escription **L**anguage interoperability specification support added
    - `–jsdl` option added to glite-wms-job-submit command
    - Client is able to read/parse/send pure JSDL instances
  - JSON
    - **J**ava**S**cript **O**bject **N**otation output format support added
    - `–json` option added to glite-wms-job-submit/output/cancel commands
  - DNS Load Balancing
    - hostnames resolution from DNS,Alias, IP address
  - Failover mechanism integration
    - next WMProxy endpoint automatically switching upon any client failure

WMProxy Client Status 3/3

- **List of bug fixed for patch #2875:**
  - # bug #17534: unexpected behavior of glite-wms-job-output
  - # bug #29635: bad error message for glite_wms.conf syntax error
  - # bug #34949: Ctl-C'ed submission gives funny state
  - # bug #39890: IPv6 bug: name resolving by WMclients broken under IPv6 (glite-wms-job-list-match broken by IPv6)
  - # bug #41294: IPv6 bug: non compliant name resolving function in source code (gethostbyname_ex)
  - # bug #41295: IPv6 bug: non compliant networking function (gethostbyaddr)
  - # bug #41310: [WMS] IPv6 bug: non compliant name resolving function in source code (gethostbyname)
  - # bug #41322: [WMS] IPv6 bug: non compliant address family type in source code (AF_INET)
  - # bug #41324: [WMS] IPv6 bug: various non compliant functions/data structures in the wmproxy_api.cpp file
  - # bug #43601: glite-wms-job-submit silently adds VO defaults to user configuration
  - # bug #47134: WMS UI: glite-wms-job-{delegate-proxy,submit} fail if HOME is not set
  - # bug #47677: WMS job submission aborts on non-existent endpoint
- **Detailed info on patch #2875**

More on status

- **Documentation**
  - New gLite WMS Website created as a collection point for documentation, guides, news, infos, software, blog. Some guides are still under development (WMS specific installation guide)



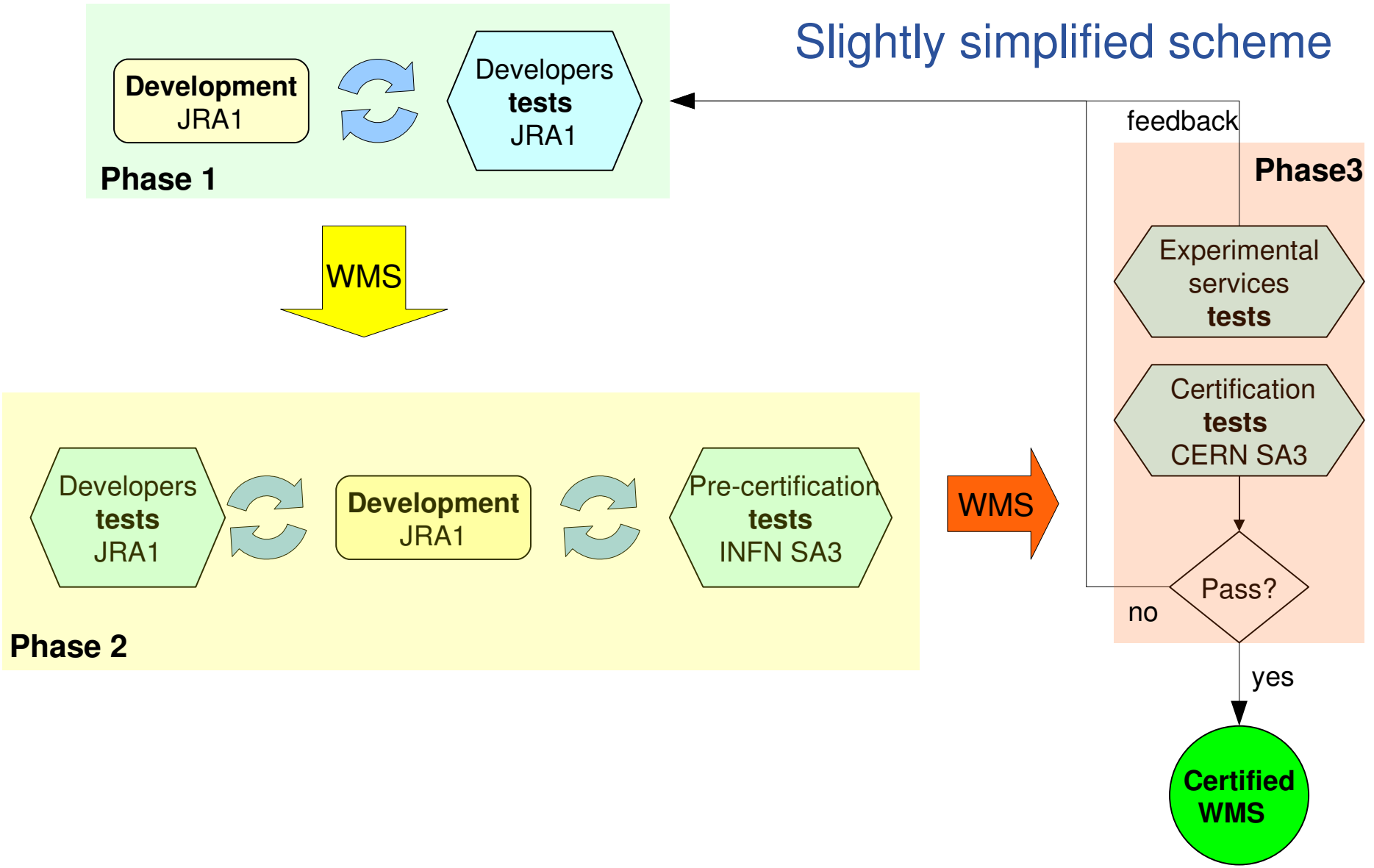
http://web.infn.it/gLiteWMS

- **WMS-UI Interfaces**
  - Multi-language (java-cpp-python) WMProxy APIs provided for requesting WMProxy Server operations

More on status

  - CLI C++ Command Line provided for performing Job life-cycle operations (submission/output retrieval, cancel, etc.)
  - CLI Python Command Line provided for requesting Job information (status/logging info)
- **WMS interface**
  - Web Service Interface (WMProxy) exposing Job life-cycle needed operations (submission/output retrieval, cancel, etc.)
- **Error/logging**
  - A common and organic specification for error/logging information must still be agreed upon and implemented
- **Platforms**
  - WMS is currently available on SLC4 (gLite 3.1)
  - Porting to SL5 is already in progress (gLite 3.2) (see later)

Devel/Test/Release 1/4



Devel/Test/Release 2/4

- **Phase1 (JRA1)**
  - Design and implementation of the new release. Code committed to a dedicated CVS branch.
  - Up to the developer whether to create Etics configurations. Builds are performed locally. Continuous process of test → code adjustment → build → test
- **Phase 2 (JRA1, INFN SA3)**
  - The product is ready for the Pre-Certification tests; developers tests go on in parallel; CNAF SA1 takes part in the the process
  - Typically intense bug fixing activity with tight interaction between testers and developers.
  - Each new version is rebuilt with a new tag. Etics configurations are modified accordingly (not locked). Testing installations are updated by means of the newly generated yum repository
  - Builds are performed remotely

Devel/Test/Release 3/4

- **Phase 3 (CERN SA3, Experimental service users)**
  - The product is mature for certification
  - The corresponding Etics configurations are locked
  - artefacts are published in the registered repository
  - Concurrently with the certification process, an experimental service is set up for the VOs in order to collect precious feedback from production-like use cases
  - If the certification is passed, still the product might not go to production, depending on the experimental service results (and pre-production)

Devel/Test/Release 4/4

- **CVS**
  - The centralised model proved to be usable and sufficient for the needs of the WMS team (small group, no big problems with policies, usually connected to the network). Open to benefits of a distributed approach though
- **Etics**
  - Configuration management, build and deployment of artefacts are now usually straightforward. Some solvable problems encountered (e.g. m4 files management and multiple build-common-cpp configurations within the same build)
- **Savannah**
  - Main downside is in the management of work plans (no synoptic view of all tasks/subtasks i.e. gantt-like, nor complex accounting of resources). So far we adopted a spreadsheet based work plan

Evolution 1/4

- **Porting WMS to 64bit SL5 glite3.2**
  - Savannah task #9429
  - Ultimate code baseline will hopefully be 3.2
    - Depending on patch #2597 outcome
    - Most recent bugs fixed in 3.2 only
    - 3.1 still suffers from some stability issues
    - The whole picture suggests to pick 3.2 for porting
  - Porting of WMS 3.2 is already in progress. Some issues already addressed. Major activities to be completed:
    - Migration to org.glite.jobid (in progress)
      - *Jobid is widely adopted throughout all WMS components*
    - Compatibility check with boost-1.33.1 (1.32.0 on GL3.1)
  - Porting of ICE should also be started

Evolution 2/4

- **WMS 3.2 code baseline**
  - Code hardening and consolidation
  - Implementation of some missing functionality
    - Implement “Cancel” request for collections (task #9426)
    - ISM/Purchaser restructuring
      - *Split of ISM component decoupling purchasers code i.e. new stand-alone component for each purchaser (org.glite.wms.ii-purchaser , org.glite.wms.ism)*
      - *Move to the restructured purchasing code: removal of the ldif2classad/ii_attrutils wrapper libraries versus direct ldap_api usage within the ii-purchaser code; removal of globus ldap library dependency vs use of system ldap library*
      - *Move to a flyweight implementation for the representation of the resources within the ISM, changing internal ISM data structure.*

Evolution 3/4

- **Submission to CREAM CE via ICE component**
  - An intense testing activity is expected in production-scale environments. See also Sgaravatto's presentation
- **IPv6**
  - Thorough testing of IPv6 compliance
  - Ipv4 backward compatibility already under test (patch #2597)
- **Submission to ARC CE**
  - Work with the Nordugrid is expected to shift to a more operational phase, also requiring user support
- **DNS-based load balancing and client failover**
  - Savannah task #9435
  - A better understanding and implementation is required
    - bug #43633 (even if all the proposed suggestions need to be carefully evaluated in terms of man power)

Evolution 4/4

- **Integration with the new gLite authorisation framework**
  - Savannah task #9430
  - Analysis and design for the integration with the **job submission** process (WMProxy) are already in place. Implementation phase will follow soon.
  - Investigations for the integration with the **matchmaking** process are still to be performed
- **Mechanism that forwards generic parameters from the JDL to the LRM of the selected CE**
  - via RSL for lcg-ce; directly to ICE/BLAH for CREAM based CE
  - tune up the execution environment at the LRM side by adding #BSUB / #PBS directives that handle these parameters properly (LFS/PBS)
  - recommended solution by the MPI WG to give hints on cores allocation for an effective execution of some MPI applications