

Minutes of the IT-CZ cluster of EGEE JRA1 Milan, March 24-25, 2004

Attending: F. Prelz, A. Krenek, F. Pacini, S. Beco, F. Giacomini, D. Rebatto (only first day), M. Marchi, M. Pappalardo, L. Zangrando, A. Guarise, R. Piro, G. Patania, M. Sgaravatto, M. Mordacchini, A. Gianelle, M. Mezzadri, V. Ciaschini (only second day), S. Andreozzi (only second day), E. Ronchieri (only second day), A. Di Meglio (only second day)

F. Prelz summarizes the goal of the meeting: the activity planning for the first phase of the project.

New persons (D. Rebatto, L. Zangrando, M. Mordacchini, M. Marchi) are introduced.

Arda

F. Prelz explains what ARDA is, what is its goal, what is the ARDA prototype. There will be a very first ARDA prototype, followed by a 'real Prototype'. In the ARDA prototype document, and in presentation shown at the LCG workshop, it appears that in the first prototype the EDG WMS won't be integrated (to implement the TaskQueue component Alien will be used instead). According to F. Prelz this is not an issue, since it is important that the EDG WMS can be integrated in the 'real Prototype'.

S. Beco and M. Sgaravatto stress that it is important to be involved since the beginning on the ARDA activities.

It is decided that F. Hemmer and E. Laure will be contacted to make clear that the IT-CZ cluster plans are to reengineer the EDG WMS to have it integrated in the ARDA architecture.

Milestones and deliverables

F. Prelz shows the list of deliverables and milestones foreseen for the first project year.

Status of the DAGMan release

The status of the merge (between rel. 2 and DAGMan branch) for the various components is reported:

- JC e LM (A. Gianelle): merge done. Everything compiles. To be tested
- DGAS (A. Guarise): the same
- Job checkpointing and partitioning (A. Gianelle): the same
- WM (F. Giacomini): the same
- Purger (M. Pappalardo): the same

- Interactive jobs (M. Mezzadri): the same
- “WP2 interface” stuff (E. Ronchieri): the same
- UI (F. Pacini): the same
- JA (E. Ronchieri): the same
- Testing (A. Guarise): the same
- LB & proxy renewal (A. Krenek): merge done. A problem with the LB testing stuff is reported. Should be easy to address.
- NS & matchmaker (M. Pappalardo): merge done, but there are still some problems, which should be easy to fix

When the merge will be done, tests will be performed first on the WP1 internal testbed (grid012g, bbq, lxde01, etc.). When happy with these tests, RPMs will be produced. They will be deployed in the INFN development testbed for further tests.

The status of the INFN development testbed is not too clear. L. Gaido will be contacted to see which machines/sites can be used.

Guarise and the other TO people will have the responsibility to run tests. They will basically consider the LCG test suite, which will have to be augmented at least with tests for the DAGMan stuff. Also the new people will be involved in testing activities. The idea is to consider a level of testing similar to the one performed by LCG.

When the software has proved to be good enough, the management of the GRID.IT grid infrastructure (Antonia ?) will be contacted to have this software deployed in this Grid environment.

Support issues

Massimo S. reports that in particular 2 issues have been reported by LCG:

- They don't like that all the GRISes of the matching CEs are queried, since this slows down the process (LCG bug #2682). They are implementing a new BDII, updated quite often, so they just want to query this server without querying the GRISes
- There is a wrong use of the Glue schema information: the matchmaker assumes that the gatekeeper name is the same of the cluster name (LCG bug #2701). S. Andreozzi already proposed a solution which according to Salvo should be ok. F. Prelz will have a look at the problem, since Salvo is not yet back.

Ideas for improvements of the current code

Ideas to improve the current WMS software are proposed and discussed:

- Instead of querying the Information Service for each matchmaking operation (actually right now for each of these operations there are multiple queries to the Information Service), there should be a sort of repository of resource information

(which could be called 'Resource Information Supermarket'), populated asynchronously, and where this information is already stored in the desired format (e.g. class-ads).

This Resource Information Supermarket should be able to refer to multiple Information Services 'sources' (so, in case an Information Service is not available, the other one(s) can be considered).

- Some profiling of the code should be done, i.e. to remove unused code. This could be assigned to Martelli, seasoned programmer who will work in Milano.
- The matchmaking mechanisms when data have to be taken into account should be improved, considering in particular the inputs received by users. For example users should be allowed to specify if all files are needed, if all files must be closed to the target CE, if all files must be accessible with Posix system calls, etc.
- Fault tolerance in the LB server
- Address race conditions in the LB interlogger (bug #1938 in the EDG bugzilla). There is already an aggressive proposal to address this issue, which would foresee to have the synchronous logging operations bypassing the interlogger and locallogger, contacting directly the LB server instead.
- Bidirectional communication between NS and WM, needed for the edg-job-list-match
- A LB proxy server, running on the RB node is proposed (see slides presented by A. Krenek). The idea is that the services running to the RB node logs to this proxy server, in a synchronous way, without SSL. This logging would therefore be reliable and fast. The logged event will then passed to the 'real' LB server. When the WMS services running in the RB node need to query the LB server, this LB proxy server would be queried instead. F. Prelz is concerned about having a DB in the RB node.
- See if the outbound connectivity restriction can be removed (e.g. CondorG now
- allows transferring other files besides executable and standard streams)
- The implementation of the RB strategy library should be much easier. A sort of toolkit to implement it is needed.
- Submitting a job takes a lot of CPU in the client side (problem in socket++ ?)
- Thread safety of client library
- Zamboni: this has already be assigned to Perugia. They have already produced something (implemented in Perl)
- ClassAd function to require a certain network bandwidth between the CE and the SE. This has already being addressed by people in Bari
- Improve the MPI support. For example users are not happy that an executable must be submitted (they would like to be allowed to submit a script with issues mpirun)

It is reminded that NO new functionality or code reengineering should be committed in the CVS main trunk before releasing the DAGMan rpms.

Software Configuration Management plans

A. Di Meglio, responsible for the integration and configuration activities within EGEE JRA1, presents the proposed plans (see his slides), reported in the document already circulated to the IT-CZ cluster.

These ideas have already been discussed with the CERN cluster and with S. Fisher. Alberto will go to RAL to discuss them with the UK cluster members at large.

Elisabetta then presents issues and comments on this proposal collected among the members of the IT-CZ cluster (see her slides).

The following issues are discussed:

- Ant and autotools
Basically all the IT-CZ people are very concerned about the use of ant and the “suppression” of autotools.
They will like to be allowed to keep using the autotools stuff within the WL (workload) directory. Many developers report that Ant is not suitable to manage the C++ based WMS code, while Alberto doesn’t agree.
V. Ciaschini also reports that with the proposed plan, there would be big problems in making available the VOMS sw also outside of the project, as it is already required.
No conclusions and agreements are reached. Alberto is trying to rely on his proposed system to manage the existing WMS software, to better understand if and how the proposed plan is suited for the existing structure of the WMS software. The item will be therefore rediscussed when this exercise is completed.
- A CVS server at Cern will be used. This requires having a AFS account at Cern, which can be created without problems even for non-CERN users.
It is asked to allow also GSI access to CVS: Alberto will investigate if this is possible.
- Authorizations to commit: it is up to the cluster manager to decide who is allowed to commit for the various modules (e.g. a DGAS developer is allowed to commit in the LB module ?).
- Perl. Alberto reports that, while in the beginning the SA1 wouldn’t have liked to support Perl, now it seems that this is not true anymore. Not clear yet.
JRA1 IT-CZ people stress that limiting the number of languages can make sense for development made from scratch, while this would have big impacts when the goal is reengineering existing code (as it is the case for EGEE).
- Supported platforms. Alberto reports that this is still being discussed with SA1. At the moment it looks like that Linux RH7.3 won’t be the main platform (even if it will have to be supported). It is likely that the main platform is going to be Linux

RH Enterprise. Probably some part of the software (e.g. clients) should also run on Windows.

- Migration of namespaces from `edg.*` to `org.egee.*`. Alberto clarifies that this is not mandatory.
- Directory structure. Long discussion. The concluding remark is that we first need to understand what is a module for the WMS software. Is a module a directory under 'workload' in the current CBS structure ? To be clarified by the next meeting ?
- External packages. Long discussions if CVS is suited to be used also for binary files. Alberto doesn't see particular problems on this, but he agrees that these external packages could be stored in a "normal" repository (e.g. a FTP server) as it happened in EDG.
For what concerns external packages modified (i.e. patched) by EGEE developers, if the patched version is accepted by all middleware groups it is managed as a normal external package.

Unit testing & test coverage

See slides presented by A. Guarise

The only raised issue (by F. Prelz) is that TeamB should be free to change the unit tests implemented by team A.

The envisioned testing model will be feasible when all the new people will be hired.

C++ design and coding guidelines

See slides presented by F. Giacomini

These inputs will be given to Alberto as contribution for the definition of the EGEE coding guidelines.

Experience with Web Services

See slides presented by L. Zangrando

After the discussion it turns out that the integration of web services within the WMS can take place in 3 places:

- At the NS level
Catania and Datamat will work on this item
- In the UI-LBserver interface
CESNET will work on this item, with Catania as "guardian angel"
- In the resource (CE) access

Padova will work on this item

Policy Framework Proposal

See slides presented by V. Ciaschini

Datamat presentation

See slides presented by S. Beco

About Resource Broker work to be supposed to be performed by Datamat, it is proposed that they could address the problem of providing a “builder” for the RB strategy module.

Cesnet presentation

See slides presented by A. Krenek.

- Direct communication for synchronous logging.
- The Holub service: CESNET is planning for a prototype in April, possibly followed by a better, second version at the end of 2004.

Date for next meeting

As no specific IT-CZ meeting will happen at the Cork conference, we plan a next meeting in Milano for Monday and Tuesday, May 3/4.

People are welcome to travel to Cork, especially if they want to hear about project-wide planning from the live voice of higher levels of management, but are warned that probably not many technical discussions will be allowed to occur as part of the conference programme.

AOB

A 'JRA1-Workload management' website is needed. P. Prelz started to clone the 'Data Management' website already, trying to rip away dependencies from MS Frontpage. As a form punishment for ending the meeting 10 minutes late, the website will be hosted in Milano, under the care of M. Mezzadri.