



Minutes

Doc. Identifier:
DataGrid-12-MIN-0236
Date: 12/06/2002

Meeting Object: **Minutes of the PTB Meeting on April 10th 2002**

Editor(s): **Markus Schulz**

Meeting Date: **10/03/2002**

Meeting Place: **CERN**

Attendees: **Olof Barring, Maite Barroso Lopez, Federico Carminati, Mauro Draoli, François Etienne, Steve Fisher, Fabrizio Gagliardi (via phone), Antonia Ghiselli, Frank Harris, Bob Jones (chairman), Dave Kelsey, Peter Kunszt, Julian Linford, Cal Loomis, Laura Perini, Francesco Prelz, Les Robertson, Markus Schulz, Ben Segal, Massimo Sgaravatto, John Gordon, Eric Fedè, Sophie Nicoud, Andrew McNab, Rene Metery, Laurent Bobelin, Gabriel Zaquine, Marias Bubak, Maief Malawski, Jens G. Jensen, Paolo Capluppi**

CONTENT:

1. MINUTES FROM THE PREVIOUS MEETING AND MATTERS ARISING	2
2. REVIEW ACTIONS	2
3. TESTBED SOFTWARE TEST PLAN.....	2
4. TESTBED STATUS	3
5. LICENSE ISSUES FOR EU DATAGRID.....	4
6. ACCESS RULES AND TESTBED SUPPORT	5
7. QUALITY INDICATORS.....	7
8. PUBLICATION POLICY	7
9. INTERACTION WITH RELATED PROJECTS	8
10. DISCUSSION ON THE EU REVIEW REPORT.....	11
11. DELIVERABLES	11
12. TECHNICAL COORDINATOR'S REPORT.....	12
13. OTHER ITEMS	14
14. NEXT MEETING.....	14
15. ACTION ITEMS:.....	14

1. MINUTES FROM THE PREVIOUS MEETING AND MATTERS ARISING

Bob Jones:

There were no comments on the agenda, but it was noted that on the WEB we still have the ASCII version of last PTBs minutes only.

It was agreed that this will be changed.

2. REVIEW ACTIONS

Bob Jones:

- 1 Provide additional storage for the testbed: done
- 2 and 3. Test plan and support. This will is covered by the agenda: 4, 5 and 6 are done.

3. TESTBED SOFTWARE TEST PLAN

by Laurent Bobelin on behalf of the Test Group

see:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t5/transparencies>

Laurent described the motivation for a systematic approach to testing. The middleware is rapidly evolving and has to adhere to a tight schedule. The minor releases that are scheduled for every second month require continuous test activity.

The scope of the tests will be wide and cover the correctness and efficiency of all components from installation, configuration and security to scalability, robustness and fault tolerance.

After this overview of the strategy the organization of the test group was described. It was mentioned that a draft for the test plan is now available in EDMS.

The test bed for the tests will consist of 5 to 6 sites. The resources will be decoupled from production.

The test group will follow the layered structure of the DataGrid architecture and will develop automated test suites for unit and integrated testing. Testing will be done on every release tagged by the IT.

Laurent emphasized that the availability of detailed documentation and pre -release updates is required for their work. The WPs have to provide accurate definition of new functionality.

The sites participating in the test activities are: IN2P3/LPC with 18 PCs, IN2P2/CPPM and CS-SI with 4 PCs and CEA/Saclay with 6 nodes. He stressed that new sites are invited to join.

The responsibilities are split between the initial three sites as follows:

- IN2P3/LPC will work on storage and replica catalog issues,
- IN2P2/CPPM and CS-SI's region of interest are job management on the UI, security and monitoring while CEA/Saclay will cover the testing of job management (RB,JSS,LB,JDL).

The final aim is to cover all functionality of the software on the testbed.

In his conclusion Laurent pointed to the draft of the test plan and invited the collaborators to provide feedback and volunteers.

After the presentation there was a discussion that was centred on a few main subjects.

The relation between tests that have been done by WP8 and the "Loose Cannons" and the proposed tests was not clear to the audience. It was stated that Jean Jacques's test job is not covering all functionality of the edg middleware.

There were some questions about whether a fourth testbed with separate hardware for testing is really required. The argument for a new testbed is that the testbed needs more stability (i.e. should not be updated so frequently) than the development testbeds can provide and that the testing, especially the performance testing is too disruptive for a production testbed.

The applications stated that they would prefer the tests to be run on the testbed where the applications will be deployed.

WP4 wanted to know why the WP4 software was not explicitly mentioned.

The statement from the test group was that the tools will be used anyway by everyone and as a result there will be already good test coverage.

The audience wanted to know about the status of the test harness that the group develops.

Up to now only architectural work has been done. There hasn't been any code development.

It was clarified that Cal Loomis is coordinating all the testbeds.

There was a general consensus that it is good to have a test group, but some concerns were voiced whether a synthetic test can be as efficient as the tests done on the production testbed with many chaotic activities and realistic applications.

In addition it became clear that for effective testing the group needs in depth knowledge of all the middleware.

Action Item

Who are the members of the testing group?

Clarify the relationship between to the release policy

Send a reference to the draft document that has been put into EDMS (has been done).

4. TESTBED STATUS

by Cal Loomis

see:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t6/transparencies>

The production version is currently v1.1.3 this was the first release intended to be used for widespread deployment. However a few minor problems were found.

Core sites are now moving to 1.1.4, some sites have run into problems configuring MDS via LCFG.

The development versions enter now the 1.2 series.

From all middleware work packages RPMs have been received. CERN will move its development site incrementally from 1.1.4 to 1.2.0.

The ITeam will come together at CERN on 10th to 12th April to work on v.1.2

Then Cal gave an overview on the status of the Globus releases. Except for minor fixes we are using the same Globus beta-21 release that was in place for the EU review.

The Globus team has released a beta-22 version which has been announced to incorporate just bug fixes, but no API changes. For the first EDG 1.2 release we will stay with our current version of Globus.

The report on the status of the sites showed that all core sites (CERN, CC-IN2P3, CNAF, NIKHEF, RAL) are in the process of moving to v.1.14. Before all sites upgrade the MDS/II interaction has to be tested.

New sites are joining: Padova, Catania, Torino

At this time it was noted that there is a Spanish site visible via LDAP. They seemed to have joined without asking.

Cal made clear that there is no access control yet, but it could be implemented. This poses no risk since if a site is not configured in the MDS the RB will not use it.

There were a questions by Frank Harris and others whether 1.1.4 is going to fix the problems that have been seen during the demo for the EU review.

Cal summarized the problems seen. The RB crashes have been fixed. Only small problems remain. The II problems should be solved with MDS. The current workaround that is in place is the automatic restart of the Information Index via a Cron job.

The LB crash was a one time event and was not tracked down.

Frank made it clear that for demonstrations it has to be clear which features work and with what stability. An honest statement is needed.

Francesco gave more detailed insight on the problems that happened during the demonstration.

The problem related to fetch output from a job was not a problem, but the job had been removed intentionally before.

Most problems of the RB resulted from a misconfiguration based on a misunderstanding of the effect of a parameter. The LB problem was not understood. Most of the memory leaks are closed now and the RB needs now only to be restarted once a week.

Cal continued his presentation with explaining the WP6 view of stability of the testbeds. He stressed that WP6 desires stability as much as the users and that WP6 is willing to provide this within the limits of the existing software and the given resources. Problems will be addressed as soon as possible, but a 24/7 type of service can't be provided. The deployment schedule will be adjusted to the users needs.

Some questions about the future coordination of the testbeds have been asked centred on the issues of scaling to more sites and the question on transitions between releases.

A group of system administrators will be set up and work coordinated with the ITeam.

In view of LCG and X# we have to reorganize this effort.

No decision has been made on how to do the transition. Several scenarios such as parallel core infrastructure or hard switch are possible.

5. LICENSE ISSUES FOR EU DATAGRID

presented by Andrew McNab on behalf of Anders Wannanen

see:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t7/transparencies>

It is proposed to base the license on Globus's "Open Grid License Agreement" (OGLA) which is close to the Apache License enriched with a clause about implicit granting rights back to the distributors.

The license is BSD-like, but still GPL compatible. This means that commercial apps can use it and do not need to publish the resulting code.

The aim of OGLA is to minimize the overhead for others to contribute. This goal could be achieved by a clear license and a license template form to help projects with an ready made solution.

Then the draft license and the template was shown. See the transparencies for details.

Some questions came up about whom the copyright holder should be. This couldn't be resolved during the discussion and an action item has been assigned.

During the discussion the following issues came up:

Q: How does the transition into a commercial product happen

A: This is covered through the requirement of an acknowledgment in the documentation and in the running system

Q: Does this license "infect" the products that it will be integrated in. (Like the GPL).

A: No

Q. How do we have to distribute (include) the copyright statement? By including or reference (URL).

A. The license file has to be distributed. Inside the source a reference to the distributed license file can be made. A single line has to be put in each file

Q: What is the relationship between the Globus license and the EDG scheme?

A: Globus will use the same scheme. And discussions are ongoing with Ian Foster about synchronization with Globus.

Q: Have CERN lawyers seen it and what are their reactions?

A: They see problems with clause (6) which deals with implicit licensing of contributions. These could contain IP violations against which the license doesn't protect EU DataGrid well. In addition they are concerned about what should happen to the copyright statement if sources from different projects are mixed. This seems to be problematic, but is common practice in open source projects.

Q: If changes are published under GPL then all the source would be moved to GPL!!! Isn't this problematic?

A: We can use GPL code via binary libraries (link only).

Q: Condor is using a different licensing scheme. Does this affect us?

A: Code snippets of Condor can be released under a different license.

Q: Which release do we start to put under the license?

A: This will be 1.3. 1.2 is too early to include the finalized license.

In addition some comments were made:

Cal: Please make sure that (it is understood) that GPL compatible means that people can take our code and release it under GPL, but that we can not use GPL sources.

Peter K: It is very urgent that we clarify the license issues soon since already now this is inhibiting collaboration on database integration between Spitfire and IBM.

In addition there was a short discussion about trademarks that lead to no clear conclusion.

Action Item:

Check who has to be put in to the form as the Copyright Holder.

Clarify how this has to be handled if multiple groups contribute to the same package.

Bob: Audit of the licenses used by the external packages has to be one by release of 1.3 and the text has to be finalized.

The template has to be made available for discussions.

6. ACCESS RULES AND TESTBED SUPPORT

by Sophie Nicoud

see:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t8>

To manage access of users to several sites of the edg testbed(s) a user should know the rules for the use of the DataGrid Testbed. They should not register with all the sites.

The sites should accept all users of the DataGrid testbed according to the usage rules. The sites are responsible to enforce their local rules and national laws.

Sophie presented an outline of the rules.

The first draft was issued July 11th 2001 and now the actual version is V10 from September 15th 2001.

On April 9th there was a discussion with CERN's legal service.

As a result a few modifications have been made:

- The title was changed from "Guidelines" to "Rules"
- The part addressing the application of partner, site rules and national laws has been clarified
- It has been specified much more clearly who has access to the information stored in the Testbed computing facilities

In addition the wording was changed in a minor way. These changes lead to the acceptance of the document by the CERN legal service. Version 12 will be published on the 9th of April 2002.

Q: Does every user need to sign the access rules of all the sites?

A: No, sign one form only and only once.

It was discussed on how to preserve the national laws and what is the role of the VOs in this domain. It was not clear how the access rights (accounting) for users should be restricted on future production systems.

Actions:

We have to describe how a site joins the Testbed

Then Sophie gave a presentation on User Support.

The overall goal is that all users get answers to their questions and have access to the current DataGrid documentation. She defined a user as a scientist using the DataGrid Testbed or a system administrator of a DataGrid Site.

The scope of questions range from questions about certificates, installation to forgetting passwords.

As a solution to this a system was proposed that is based on a knowledge base to gather experience, an self-help system and access to the Testbed documentation. The support should be organized with hierarchical levels of support. It was stressed that there should be only one help system. The contact to the user support can be by www, phone or email. The support team needs an interface to the system that allows them to handle the flow of questions answered, link persons to areas of expertise, and update the knowledge base and documentation.

Then the two levels of support have been presented:

The first level answers most of the questions that are already part of the shared knowledge base. Second level will deal with critical questions and seek advice from experts.

There will be a one on one mapping between persons and problems. It was stated that it is very important that the support team knows the software and are users of the software.

A dedicated person is needed full time.

There are many commercial and free tools around, to pick the right one is essential.

New people are needed but they have to be experienced. In summer a user support team should be in place.

A comment was made that X# is working on the organization of support too.

7. QUALITY INDICATORS

by Gabriel Zaquine

see:

<http://edms.cern.ch/document/341708>

After giving an overview of the role of Quality Indicators (QuI) in the project as control and monitoring tools. Different sets of QuIs will be used during the development and production phase.

The QuIs will be followed up during the whole lifetime of a project and updated monthly. For each QuI one WP will be responsible to produce it.

For the development phase there are indicators that deal with time spent on various tasks compared to the planned time and indicators dealing with the number of requirements covered by the released software.

These indicators will concern all WPs.

A second main set handles the production of reports on deliverables. There the number of PTB cycles until approval and the time needed to deliver them to the EU will be measured. WP12 will produce these indicators. During the production phase the number of new anomalies and the time for resolution is tracked together with the number of pending anomalies. WP6 will produce these indicators using the bug-tracking tools.

For the testbed the testbed operability will be monitored. Quantities to watch are the number of sites in production, the number of CPUs and the number of declared and active users.

In addition there is an "Availability testbed rate" defined that describes the ratio between successfully processed and failed jobs. WP3 and WP1 should produce these numbers.

At this point some discussion started why WP3/WP4 should collect these informations and how to measure these quantities in a meaningful way. It was suggested that WP8s standardized jobs may be helpful. In the discussion it was mentioned that for the development phase the daily auto-build result could be used too. Another quantity that might be monitored is the ratio of computing cycles available/those used which will give a measure of the utilization of the nodes.

8. PUBLICATION POLICY

By: Gabriel Zaquine

See:

<http://edms.cern.ch/document/329638>

In his proposal different procedures for the publication types General DataGrid overview publications, Overview publications on WP or subsystems and publications on specific subjects which are names DataGrid Official Technical Notes. The technical notes represent work of a restricted number of

individuals. The author list will depend on the type of publication. In case of the overview publications the author list will contain the editor and a link to EDG's author list. For the WP or subsystem related documents the editor and the WPs members will be authors. For technical notes the main contributors should appear as co-authors.

Gabriel proposed an approval process in which the Project Manager will centralize publication proposals. The review process starts within the WP after which the PTB has one week for reviewing. The approval will be given by mail.

Some questions came up concerning the global overview articles. It was not clear which work package should do the first review (all??).

Another problem was how all the smaller dissemination publications should be handled. WP11 will in this case replace the WP and then the documents are forwarded to the PTB.

The Project Manager's role in the process of centralization of publications proposals was discussed and needs more clarification.

There was a broad consensus that the main results should be published in journals.

9. INTERACTION WITH RELATED PROJECTS

9.1. CROSSGRID (X#)

by Marian Bubak and Bob Jones

see:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t12/transparencies>

It was described what part of the middleware and tools CrossGrid covers compared to other grid projects. CrossGrid is located closer to the application level and more centred around scientific computing. There are links with the European efforts and some US projects. The project has 21 partners from 11 countries. Most of which are not in edg now. The three year project was started on 1st of March and is in the process of gathering requirements.

The main difference from edg is that there is a strong focus on interactive and data intensive applications and the development of an grid application programming environment which will include debugging and benchmarking tools.

The architecture relies on MPI (MPICH-G) as the link layer between the application layer and the Grid services. Apart from this the architecture is quite similar and the project expects to inter operate well with different Grids. It is planned to reuse Grid components. A good link between EGD and X# is insured by the participation of EDG's and X#'s chairs of ATF in meetings.

For the testbeds it is planned that they can be inter operable. EDG will profit in the area of benchmarking and application monitoring from X#. In the domain of testbeds a close collaboration between edg and X# is planned. To cooperate on dissemination issues a joint industry and research forum is set up. Deliverables will be exchanged.

On the level of the management regular contacts of the project management and the exchange of operation procedures and reviewers will ensure close collaboration.

After the presentation it was stated by edg that the first discussions on joint dissemination projects have already started and that visibility is crucial for both projects.

In June there will be a meeting to organize the operation of the CAs.

EDG's technical coordinator mentioned that cooperation in testing is very interesting for the edg.

9.2. LCG

by Les Robertson

See:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t11>

Les gave an overview of the scope and status of the LCG project. He stressed that LCG is not a grid technology project, but a grid deployment project. The goal is to prepare and deploy the LHC computing environment. This includes tools, frameworks and persistency for the applications on one side and the development of the computing systems. There especially services, automated fabric management and collaborating computer centers. In this the data challenges of the experiments have a central role.

The project has two phases. The development and prototyping phase which has been approved by the CERN Council in September 2001. For this funding is provided by special contributions from member and observer states. In the second phase of the project which will run from 2006 -8 the initial production GRID will be installed and operated world wide. For this the costs are included in the LHC cost to completion estimates.

Since the project relies on important external resources the project has to manage CERN resources and coordinate the overall activities.

The coordinators for certain areas have been appointed:

- Torre Wenaus for Applications, Wolfgang von R"uden for the Computing Fabrics.
- Fabrizio Gagliardi will coordinate the Grid Technology. In the area of
- Grid Deployment Mirco Mazzucato will be responsible for the policies forregional centers and grid deployment.

For the management of the operation of the data challenges and the grid operation there is still an open position.

Les made it clear that LCG is a Grid deployment project that is a consumer of Grid technology developed by the various Grid technology projects. Since LCG must deploy a global grid it has to have compatible middleware between all the different Grid projects.

Then Les introduced the concept of John Gordon's testbed trident. The grid technology projects will run the developer's and development testbeds which will be used for development, beta testing and demonstration. The third prong of the trident the production grid. LCG will be responsible for operating this on 24x7 base.

Les then went into the description of development of LCG over time:

- April 03 The first 24X7 service will be deployed based on the toolkit recommendations from GLUE. There will be about 10 sites from Europe, Asia and North America. The size will be about 1500 CPUs and 1500 disks. The service will be available for all experiments at all times.
- In October 03 reliability and performance targets have to be met and the scale of the deployment will be extended,
- The grid services will then evolve slowly through 2005 and will provide services for the LHC experiments.

The proposed split of responsibility between LCG and DataGRID was explained. At the end Les summarized the dependencies on DataGRID, DataTag and iVDGL.

In summer 2002 it has to be demonstrated that the testbed can do real work for the data challenges of the LHC experiments. In October 2002 the agreement on the LCG-1 middleware has to be achieved. By end of 2002 the middleware for LCG-1 has to be implemented and available.

After Les' presentation several issues have been raised and discussed.

Q: What is LCG's role in the Data Challenges?

A: LCG will not manage the Data Challenges, but provide resources.

Frank pointed out that LHCb and CMS want to use stable middleware and that he thinks that this implies that for the current year this has to be edge.

Marco noted that for CMS this will be more problematic since they have to support two test beds.

Federico voiced concern about progress made with GLUE. Concerning production support he made clear that 12X5 for the time being would be sufficient for the experiments.

Les responded that 24X7 is a LCG requirement.

Olof was worried about the way GLUE went. He suggested that CERN should send a cluster expert who will cover the site issues.

9.3. DATA TAG AND GLUE

By: Antonia Ghiselli

See:

<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t10/transparencies>

The main subject of DataTAG is to reach interoperability between the EU and US grids and this especially for the HEP applications.

The current objectives are to provide input to a common LHC grid architecture, plan EU-US integrated grid deployment and to provide a test environment to the LHC applications to help expand the existing use-cases to test the interoperability of the grid components.

Currently there are the following funded resources available: 8 fte at INFN, 1 each at PPARC and UvA.

Then Antonia summarized WP4 tasks and named the persons being responsible for them. These tasks cover resource discovery, security, interoperability of collective services between EU_US grid domains and test applications.

The DataTAG framework and the relationships between different projects was described in some detail. There are now established contacts between EDT/WP4 and iVDGL for Security, Information Discovery and Grid middleware.

Then Antonia described the status of GLUE. Glue stands for Grid Laboratory Uniform Environment (www.hicb.org/glue/GLUE-v0.04.doc).

The collaboration of grid projects is focused on the interoperability between Grid software projects.

Glue is managed by iVDGL and DataTAG. Several HEP experiments and grid projects are participating.

GLUE plans to evolve through three phases:

- The first provides an authentication infrastructure and a unified service discovery and information infrastructure.
- The second will build a data movement infrastructure and establish an authorization service.
- During the last phase computational services are set up.

Each of the three phases has at its end a test phase. The first steps in the project are already done. In Feb. 2002 the interoperability of US and European certificates has been tested. Now the review of common resource discovery schema is in progress.

Antonia gave a very detailed time-line for the project which is included in the slides. This time-line includes the required changes in edg components to have by release of the EDG testbed 2 in September a working inter-operable testbed.

It was reported that it is still open who will maintain and own the glue schemas and the MDS and R-GMA information providers.

It was proposed by Antonia to include C.Vistoli in the EDG WP managers meetings representing EDT/WP4.

There are hardware resources for testing in CNAF, Padova, Milano, Torino, Trieste and Bologna.

The agenda for the joint meeting on 26.11 in Chicago was shown.

After the presentation several questions were asked.

Antonia answered that there are only small differences for EDG. Mainly the changes required will affect the Information Provider.

It was noted that there are still some differences in how Globus and EDG sees a CE.

Jean Jacques raised the point that he has problems to identify any US resources.

Antonia listed resources and named persons involved on the US side, which she estimates to be more than those committed from Europe.

JJ sees a large overlap between edg people and DataTag and wonders what this will mean for the EDG software.

Bob clarified that if we agree on the schema this will be the only one that needs to be maintained. If edg uses the schema it will maintain it.

Olof remarked that David from NIKHEF is taking part on the security and authorization discussion.

Bob raised the point that he sees a mismatch in the timing of LCG and GLUE.

Les answered for LCG that he agrees that this has to be addressed.

Olof wanted to know what Flavia will report on packaging and deployment tools?

It was answered that the tools will be based on LCFG.

Olof wanted to know what LCG plans to use as installation tool for the releases?

Answer: There will be a document at the end of autumn.

A discussion started if GLUE is a solution to Les' hope for a similar orientational software? It became clear that this is not necessarily the case since the schema can be used by different software.

Frank suggested that the component people should speak to each other.

It was clarified that both technical coordinators are already in touch and that the major differences are in the region of data access.

10. DISCUSSION ON THE EU REVIEW REPORT

for details see Bob Jones reply to the EU.

11. DELIVERABLES

D7.5: Security Requirements and Testbed-1 Security Implementation

presented by Dave Kelsey

The deliverable is due in M-15, it is now a few weeks late. The document has reached about 90 pages by now. There has been a lot of input. V2.1 will be released by the editor Linda Cornwall on 12th April. Most of the many comments received are in now, but two areas are not covered in this release. The

removal of the duplicates and the process to reformulate the non-formal requirements. In addition more work is needed on chapter 5 about the TB1 and the chapter6 which covers the plans of the WPs. The V3.0 will be ready for PTB review in a few weeks from now.

There was a discussion on how to proceed and how to address the problems with chapter 4 (Requirements). It was discussed if it was worth to do a major revision or just maintain a list of requirements. Bob remarked that test cases for security should be added. The current state of chapter 4 with the comments added might be good enough, but 5 needs still some work. Dave expects this to take a week. It was agreed that no major revision to chapter 4 should be done and the moderator should decide at the end of the next week on submission.

D11.5: First annual conf. and grid forum

Maite Barroso Lopez (moderator)

The last draft is not ready to be reviewed. There are many empty chapters and the document is at an early stage. Especially the comparison with D11.4 is missing. The document needs to go through a major restructuring.

Mauro didn't agree with the moderators view of the document. The moderator stated that most of the information is there, but the comparison with the plan (11.4) is missing. The reviewers didn't give many comments because the document isn't yet in a state to be reviewed.

The document will not go to the PTB, but first again to the reviewers. The PTB will give approval by email.

D12.9 Quarterly Report

by Gabriel Zaquine

The report is expected for the 15th. Up to now only WP4 has been submitted

12. TECHNICAL COORDINATOR'S REPORT

by Bob Jones

see:<http://documents.cern.ch/cgi-bin/setlink?base=agenda&categ=a02113&id=a02113s1t17/transparencies>

Bob reported on the following subjects: Architecture Group, Release Plans, Port to Solaris, plans of a Tutorial and OGSA impact.

12.1. REPORT FROM ARCHITECTURE GROUP

He summarized the meeting of the Architecture Group 9th of April.

The given short presentations lead to in-depth discussions.

From the applications WP10 discussed data management and security requirements.

WP8 presented the common use cases gave the prioritized requirements.

WP9 presented their requirements and use cases.

The Cross Grid Architecture was discussed.

Bob presented the medical data infrastructure that has been developed by WP10.

By splitting normative and anonymous data it allows data replication on sites which are not trusted.

Then Bob summarized the WP8 technical requirements these are now ranked by priority. The most important is the availability of a reliable dg-job* suite. This is a prerequisite for realistic large scale tests.

Reliable gdmf-* commands and file-transfer commands are needed for data management.

There is a strong need for a lightweight UI that could be installed on a standard desktop node.

The other technical requirements are mid term which is defined as being for testbed 2.

They are portability, a better way to handle scratch space and ways to handle output files with the help of JDL.

WP8 prepared a list of commands that are hard to use and often lead to mistakes. These will be collected on a "commands black-list".

The design choices should be driven more by Use Cases than before.

There was some skepticism that a single Web Portal could be produced by edg that would be useful for all applications.

A set of well-documented middleware interfaces accompanied by expandable examples would be preferred by the experiments.

Bob reported on the presentation of Spitfire by WP2.

Spitfire is a secure database for meta-data, scenarios have been presented on how it could be used as a meta data catalog replica manager and how it could be used within WP10.

The security layer was designed in collaboration with the EDG security group.

Since all application groups have the need for application specific meta data multiple instances of Spitfire could be used in these cases. It has to be analyzed how closely the applications use of meta data should be linked with Spitfire. Consistency and scalability issues have to be addressed.

WP5 : Currently it is difficult to define interfaces that cover all SE aspects and can be implemented for all MSS. The required support for "local files" complicates the issue. The architecture of the SE is still not entirely clear.

Many questions have been raised concerning traceability that effect most middleware components.

Then Bob gave a short overview of the CrossGrid architecture (see talk by Marian).

Then the next steps of the architecture group were presented.

The D12.4 document needs to be updated. This will clarify the testbed 1 architecture. This should be done during May 2002.

Taking into account the merged prioritized requirements and Use Cases the architecture for testbed 2 has to be developed until July 2002.

The architecture group needs to spend more time together to meet the goals. Meetings are scheduled for May 6&7, June 12&13 and in addition they will convene during the GGF in July and the EDG conference in September.

12.2. SOFTWARE RELEASE PLAN

Then Bob moved on to present the Software Release Planning. The scheduled release dates are for 1.3 May, for 1.4 July and for 2.0 October. The WP's details for the releases are updated now. After the release of 1.4 the plans for releases beyond 2.0 have to be reviewed. Items like the release interval of 2 month have to be addressed and release dates for 2003 have to be picked.

12.3. STRATEGY FOR PORTING EDG SOFTWARE

The strategy for porting edge is the following: The versions up to 1.2 will be available on RedHad 6.2 only. 1.3 will be released on RH6.2, but to prepare for migration the WP6 auto-build tools will be used to produce the release. This will simplify future migration. Next step will be to set up build servers for different platforms like RH 7.x and Solaris. The goal for 1.4 is to have it running on RH 6.2 and 7.x with different versions of the compiler.

The WPs responded on the request to assess the work associated with porting to Solaris. WP1 will release details on April 16/17. WP2 assumes that the new Java based code should port easily. WP3 sees no major obstacle since part of their components is already running under Solaris or is Java based.

WP4 sees some difficulties to port the installation tools. The sensor framework is being investigated. For WP5 there is no serious blockages foreseen. rfio is not problematic and the gridFTP server will be tested on Solaris soon. WP6 has produced mostly script based software which should be trivial to port. The main concern is that external tools and packages might not be available on Solaris.

12.4. TUTORIAL

There have been many requests for tutorials, overviews and hands on introductions covering the EDG software. The material has been produced by many project members. The aim is to collect this and produce a tutorial that can be adapted to cover different presentations and tutorials. There will be three main versions. For users, developers and system administrators. The development of the tutorial is coordinated by Mario and Elisabeth a. A draft is available, this will be extended taking input from all WPs including the application groups. The material is expected to be completed by the release of EDG 1.4. The tutorial needs to be kept up to date with each release.

12.5. OGSA

The presentation was followed by a short discussion on OGSA and its influence on our architecture. It was agreed that an evaluation should be done by the WPM. The WPMs have to provide resources for this and make sure that we do not fall too far behind the development.

13. OTHER ITEMS

14. NEXT MEETING

July 3.

15. ACTION ITEMS:

15.1. FOR THE TESTING GROUP

Who are the members of the testing group?

Clarify the relationship between to the release policy

Send a reference to the draft document that has been put into EDMS (has been done).

15.2. LICENSES

Check who has to be put in to the form as the Copyright Holder.

Clarify how this has to be handled if multiple groups contribute to the same package.

Bob: Audit of the licenses used by the external packages has to be done by release of 1.3 and the text has to be finalized.

The template has to be made available for discussions.



Minutes

Doc. Identifier:
DataGrid-12-MIN-0236
Date: 12/06/2002

15.3. TESTBED ACCESS RULES:

We have to describe how a site joins the Testbed.