Minutes of the CERN-BARC/RRCAT video conference of March/02, 2007

Participants:

Connecting from BARC: A.G.Apte, Kislay Bhat, P.S.Dhekne, R.S.Mundada, Salim Pathan, Phool Chand, Anil Rawat (RRCAT, Indore), K.Rajesh, Dinesh Sarode, R.Sharma.

Connecting from CERN: German Cancio, Tony Cass, Murthy Chandrauiri (BARC), D.D.Sonvane (BARC), Dirk Duellmann, Chris Eck, Giacomo Govi, Gutika Khase (RRCAT, Indore), Jürgen Knobloch, Ioannis Papadopoulos, Zdenek Sekera (minutes).

Minutes of last meeting:
Dirk noticed a small error: in the Fabric Management area, point starting with "Wassh2, a parallel SSH execution tool,…re-engineered by the BARC team…" the RRCAT should be mentioned rather then BARC.

The minutes were **approved** with this change.

Status of current projects:

Fabric Management Area:
The full report together with presentation slides is available on the agenda page.

- **Lemon**

  - **Lemon Security Development**
    Security module developed for both UDP and TCP transports to encrypt data flow between lemon client and servers with RSA asymmetric keys. Extensive set of user options, enabling to turn on/off encryption for metrics at sensor, metric and transport configurations, provide fine grain control on encryption and corresponding reduction in overhead for metric not needing encryption. The module is tested for functionality, stability and delivered to CERN.

  - **Lemon Transport and cache re-engineering**
    The cache and transport mechanism were completely overhauled in order to remove some serious shortcomings of the previous version related to handling of multiple connections and caches. The new client and server code seems to be a big improvement and is being tested for functionality and stability.

- o **Lemon XML API in C++**
  Lemon XML API provides an interface to fetch XML data from remote server and to query the downloaded data. The module is written in C++ to make use of common C/C++ libraries such as 'libxml2' for parsing XML data and 'libcurl' for HTTP fetch. API also makes use of object oriented features of C++ language such as method overloading, Inheritance, etc. The API provides a rich set of methods to query different type of data. First version of the API is ready and is being tested.

- o **Lemon XML API in Perl**
  There is a need to provide Lemon XML API in Perl. Instead of writing a new API in Perl, Software Wrapper and Interface Generator (SWIG) is used to generate Perl interface to the Lemon XML API written in C++. Example Perl code is written to demonstrate how to use the Perl interface. Example code covers use of all the methods available in the API. First version is ready and is being tested.

- **QUATTOR**

  - o **WASSH2 Implementation**
    This task involves the re-implementation of currently existing parallel SSH execution tool called 'Wassh'. After the release of first version in test environment, a few bugs were found and some changes/enhancements were suggested and implemented. The new version of wassh2 is currently running in CERN computer center.

  - o **SWRepSOAP**
    SWRepSOAP is a SOAP based implementation of Software Repository application in Quattor. Several major enhancements were implemented.

  - o **Notification System Re-Engineering**
    Notification System is a framework working in CERN computer center. This is responsible for notifying the individual nodes in the center for some predefined tasks. There is a need to improve the functionality of this system and number of new features was implemented. A detailed document has been prepared to list down the requirements and design issues. First version of this system is ready and is being tested.

- o **CCM (Configuration Cache Manager)**
  CCM (Configuration Cache manager) is responsible for downloading the node profile and caching it on the node. This cache reflects the hierarchical structure of the configuration information. It supports the functionality required by the NVA-API. There is a requirement to extend CCM to cache non-local profiles and allow them to be accessed. First version of changes in API has been committed into CVS and is being tested.

- **CCTracker**
  Version 1.5 of CCTracker is released on 24 Oct 2006. The follow-up version 1.6 with number of new features and enhancements is implemented and is already being tested. The steps required for preparing the CCTracker release are now automated and integrated with the CCTrackerClient netbeans project

In the comments to the presentation Tony Cass stated he is very pleased with the way this project works, producing requested results in time, high quality and in a very productive and pleasant atmosphere.

GridView:

The report available on the agenda page describes more in details number of new developments and achievements in all areas of GridView:

- **File Transfer Monitoring**
  - o New transport mechanism for gridftp data to work around the known loss of tuples and instabilities in R-GMA. It is in production since last 3 months with absolutely no loss of data.
  - o Web Services based transport mechanism as an alternative to R-GMA. Deployed so far for SAM data and it is working reliably.
  - o Weekly and Monthly reports for VO-wise data transfers (Hourly and Daily reports were implemented earlier).
  - o Weekly and Monthly reports for site-wise data transfers.

- **Job Monitoring**
  - o Fixed a few bugs and made some changes based on user's feedback.
  - o Added tool-tips for Job Status graphs.
  - o Developed reports for RB-wise classification of lobs lost from monitoring (due to records missing from R-GMA or other problems).

- **Service Availability Monitoring**
  - o Continue with integrating the original SAM monitoring into GridView, both now share the same database.
  - o All derived metrics will now be computed by GridView but will be accessed by SAM.
  - o GridView will be the primary interface and the entry point to Service Availability Visualization.
  - o Developed Graphs and Reports for presentation of SAM test results with various levels of details.
  - o Implemented traceability from Service Availability Graphs to corresponding test results providing a full transparency in generated availability numbers .

- **Version Monitoring**
  - o Implemented Version Management and Display.
  - o Individual version numbers for the modules.
  - o Tagging modules in CVS.
  - o Overall version numbering for the project.
  - o Stable versions are deployed to production instance.

- **GridView future plans**
  Gridview is now widely used in WLCG/EGEE. Many requests for new features have come from different user groups like
  - o Site Admins
  - o VO Admins
  - o WLCG Management
  - o Service Challenges
  - o Monitoring Working Group
  - o EGEE

  The discussions are being held with users, understanding/analysing/prioritizing new requirements.

  The immediate plans are as follows:

  - o Deployment of Web Services Based Transport for collection of Job Status data in GridView.
  - o Exploration of ways by which we could collect data for Jobs submitted directly to CE (possibly from CE logs).
  - o To design and implement common controller interface in order to integrate SAM portal with GridView.

- To improve navigation in GridView Service Availability pages and across GridView and SAM portal components.
- To provide Wiki pages for GridView Documentation/FAQs.
- Participating in the WLCG Monitoring core Workgroup (member of core working group). Monitoring Workgroup is working on standardization of Grid Sensors, Transport, Repository/Schema, Visualization, Interfaces between monitoring tools/components etc. GridView will. implement the recommendations from the monitoring workgroup
- Implementation of **WLCG management requirements**, in particular:
  - To compute and visualize a few new metrics like Site Reliability, Scheduled Downtime etc.
  - To improve the Service Availability computation by taking into consideration some additional factors like
    - Scheduled Downtimes for sites
    - Occasional unavailability of SAM test results
  - To provide PDF generation option in Gridview pages.
  - To automate the generation of "Site Reliability/ Availability Report" circulated to LCG Management Board.
  - To automate the Generation of LCG MB report about "Data Transfer Performance Targets achieved by Tier 1 sites" in order to verify 2007 targets.
  - To generate reports about data transfers from Tier1s to their associated Tier2s.
  - To Export Data in CSV (Comma Separated Values) or Excel format for all data transfer Graphs in Gridview.
  - Modifying Data Transfer GUI in order to enable selection of T2 sites as per their associated T1s and VOs.
- Implementation of **requirements from VO admins**:
  - To compute and plot success rates for every individual SAM test (whether critical or not) aggregated by sites and also by duration as hourly, daily, weekly, monthly.
  - To display test results for VO specific tests.
- General requirements**:**
  - To make all Gridview pages bookmarkable so that sites should directly view their relevant pages (Site Admins).
  - To develop gridview sensor for DCache SE (Service Challenge).
  - To visualize SE statistics like the Used Space/ Free Space etc.
  - To visualize some EGEE metrics for service Availability.
  - Visualization of FTS Statistics.

Jurgen comments that GridView has become nowadays a very important and very visible part of the Grid project. Zdenek added that there is no better proof of the quality of the results of this project than the fact people want to use it, they like it and this is why they demand more and more features. Zdenek also noted the excellent collaboration spirit, professional results delivered in time and highly positive attitude of BARC developers. Mr.Dhekne asked how we shall manage the quantity of requirements required by users, Zdenek replied that a special meeting will be setup to sort all requirements and define appropriate priorities and timeline for the development. This list will be distributed to all interested parties for comments when appropriate. Zdenek also pointed out it may be desirable to add one or two developers to the project, all this should be discussed during the GridView special internal meeting.

Distributed Databases

The full report is available on the agenda page.

Several experiment applications are based on LCG database abstraction layer CORAL for accessing relational data from several supported back-ends (*Oracle, MySql, SQLite*). To simplify deployment of CORAL based applications a set of associated tools, called **Coral-tools**, are required to allow copying of individual tables or complete schemas between existing databases and technologies. The tools were developed in Python, they are targeted to copy big volumes of data.

The CoralTools have been developed and tested using the test cases with Oracle, MySQL and SQLite.

The work continues with:

- o Validation with acceptance tests for Coral Tools
  - case with COOL application
  - case with POOL-ORA application
- o Testing
  - Scalability test
  - Performance analysis
  - Optimization
- o Documentation compliant with CORAL doc framework.

Dirk commented that he is highly impressed with the professional approach to problems and highly positive attitude of developers. The results are delivered in time and with great commitment of developers to the testing and in general to the quality of delivered software.

AOB:

o Mr. Dhekne mentioned that MOU extension was approved by BARC, which brings the additional 23 man/month to the whole CERN-BARC collaboration. Mr.Dhekne would like to create a new team with new team(s) working on new projects such as perhaps the code review (C & C++) and others. The meeting took note of this request and Jurgen stated that it would be preferable to discuss it off-line to make an appropriate decision on how to best use the additional resources.

o Mr.Dhekne mentioned that for the record the certificates of the job done should be issued.

o Chris Eck announced that the money transfer to BARC has been completed, there should now be sufficient amount in Indian account till end of 2007.

o Chris Eck mentioned he is having discussions with prof.Gurtu about the situation with the T2 site (TIFR) in India. The site apparently has manpower problems, replacements will be active as of Apr/1, 2007, however new people will have only very little knowledge of the Grid. The question raised by Chris was if it wouldn't be possible to find some people to lend to TIFR for a short period of time to help to bridge this period. Mr. Dhekne stated there would be no problem with help. Chris will discuss it off-line.

o Chris Eck also mentioned he is going Mar/13 to Mumbai and asked a few questions about the organization of his stay, Mr.Dhekne said he will take care of it.

o Mr.Dhekne noted that he would like to change little bit the organization of CERN-BARC collaboration in the future, he wishes the same framework but more flexibility. It was decided to take up this subject during Chris Eck's upcoming visit to Mumbai.

o There was a request to send two more people from BARC to CERN, request has been sent to Chris Eck as well as to Jamie Shiers.