



LHC Computing Grid Project

Status Report to Computing Resource Review Board

October 2002

The SC2 and Requirements

The *Software and Computing Committee* (SC2) is making good progress with the work of defining the formal requirements of the project. Ten *Requirements and Technical Assessment Groups* (RTAGs) have been organised of which six, including the important RTAGs on Persistency, and the HEP requirements for Grid technology, have already reported. Projects have been launched to implement the recommendations in the areas of Persistency, Software Process and Mass Storage. In the case of the Persistency project the implementation plan includes committed resources from the experiments. This is very encouraging and it is expected that further involvement of the experiments will be negotiated for future projects in the applications area.

RTAGs in the key areas of detector simulation and software architecture (to make recommendations on ways to ensure an overall coherency of approach in the applications area) are due to complete in October. These will rapidly lead to the definition of further implementation projects.

Project Management & Resources

The project management team was completed in September with the arrival of the Grid Deployment Area Manager. As agreed at the launch workshop in March and later endorsed by the SC2, the Grid Deployment Board has been established to oversee the operation of the LCG grid, providing a forum for the computing management of the Experiments and the Regional Centres to discuss and take, or prepare, the necessary decisions. The Grid Deployment Board met for the first time on October 4th and a specification of the first LCG grid service is expected by the end of the year.

A paper¹ was presented to the LHCC in July, detailing the status of the project planning and proposing milestones to be used by the LHCC in monitoring the progress of the project. This was followed by discussions with the referees at the time of the October LHCC meeting, and the recommendations of the referees are being followed up. Detailed project planning, including the preparation of a consolidated WBS, is expected to be complete by the end of the year.

The special contributions for staff for the project are in line with expectations, and a major effort has been made at CERN and in the member states to recruit rapidly people with the expertise required. About 50 people will have joined the project by the beginning of 2003. The details are reported in the resources paper² presented at this meeting.

Table 1 gives the current breakdown of staff working on the project by funding source and area of work. The unit is FTE, weighted according to the experience of the person. In the applications area, staff are included at CERN and elsewhere. In the other areas, although substantial staffing is provided in regional centres and in grid technology projects, only the staff at CERN is included.

¹ CERN-LCG-2002-12 - *LHC Computing Grid Project (LCG) Phase 1- Status of High-Level Planning*, 21 June 02

² CERN-C-RRB-2002-... - *LCG Phase 1 Project Resources at CERN* –October 2002



LHC Computing Grid Project

Status Report – C-RRB – October 2002

Summary of LCG Personnel Resources					resources at CERN and elsewhere
Status as of: October-02					
experience-weighted FTEs	funding source				
area	LCG	EDG	CERN-IT	EP+Expts	Total
Applications	12.6	1	17.9	7.5	39.1
Fabric	7.9	0	28.8	0	36.7
Grid Technology	4.7	4	4.6	0	13.3
Grid Deployment	10.2	0	1	0	11.2
Management & Administration	3.3	1	7.4	0	11.7
Totals	38.7	6	59.7	7.5	111.9

resources at CERN

Table 1 – Summary of LCG Personnel Resources

EP division at CERN created a central software group on October 1st to bring together software specialists from the different experiments in order to further encourage common, LCG oriented, developments. It is planned to concentrate many of the LCG applications staff, from the IT and EP divisions at CERN and those externally funded, in building 32, close to the offices of the LHC experiments.

As explained in the resources paper presented to this meeting there is a substantial shortfall in the level of materials contributions.

Applications Area

The POOL project, set up to deliver the hybrid persistency solution proposed by the persistency framework RTAG, is now well underway. An initial internal release has been made and the project is on schedule to deliver the first release to developers in mid-November. The Software Process and Infrastructure project is preparing the range of software development services required to support applications development within the project and experiments. The project to provide support for mathematical libraries is getting started.

The first general release of the data persistency system is scheduled for mid-2003. The availability of distributed production environment integrating experiment-specific software with common software components and services, grid portal services and grid middleware is also scheduled for that time.

Over the next twelve months projects are expected to be defined in the following areas: core framework service, simulation, detector description, event generators, analysis tools.

Fabric Area

The focus during recent months has been on work in support of the management automation developments coordinated by one of the work packages of the European DataGrid project (EDG), specifically on installation mechanisms and improved configuration databases and their interfaces. Effort has also been dedicated to the EDG testbed operations, notably after the recent EDG release and during the ATLAS grid tests.

A technology study is under way to review likely developments for processors, disk and tape storage. This is expected to complete within the next month. A workshop has been organised at Fermilab in the context of the HEPiX conference to bring together expertise on fabric management and the deployment of grid middleware from the regional centres.

Over the next twelve months, the first release of the automated fabric management tools will be deployed and the main production service at CERN will be connected to the LCG global grid.



LHC Computing Grid Project

Status Report – C-RRB – October 2002

Computer Centre Upgrade

The former tape vault in the computer centre has been converted into a computer room. During the next twelve months equipment will be moved into this room and the upgrade of the electrical distribution in the existing computer hall will begin. Planning is complete for the new substation and high voltage switching system, and civil engineering will begin next year.

Grid Technology and Deployment Areas

Planning is under way for the first LCG GlobalGrid Service, LCG-1, scheduled to begin operation in mid-2003. In preparation for this the grid deployment teams are being set up, in close collaboration with the integration and test teams of the EDG project, with shared staffing and closely coordinated management. A large fraction of the staff is located outside of CERN in member states. It will be important to maintain funding for these external resources when the EDG project stops at the end of next year.

LCG-1 is targeted to provide part of the capacity for computing data challenges during 2003, and to provide the major part of the capacity during 2004. As part of the planning for LCG-1, a schedule will be prepared containing the requirements of the experiments for computing data challenges and the resources available for these in Regional Centres.

The specification of the requirements for grid technology takes the form of a report, “Common Use Cases for a HEP Common Application Layer (HEPCAL)”, documenting 43 use cases to be addressed by the grid middleware. An analysis of these by specialists from European and American grid projects has shown that well over half of these will be met by available tools or developments due to be completed in 2003. There are still issues of quality and usability to be tackled, but this is an encouraging sign for the LCG-1 service.

Nevertheless, it is clear that many projects, both R&D and industrial, are competing to deliver Grid middleware and it is difficult to predict now which software will be successful, easy to use or maintainable long term. At this point, the most appropriate strategy is for the LCG project to use the influence of the HEP community to negotiate the support of our requirements by these projects, and to collaborate with new projects being defined now in Europe, America and elsewhere to ensure that they also take account of the LCG requirements.

Taking a longer-term view, operating a real, production grid for HEP will be a major undertaking, requiring much cross-site co-ordination and user support. During 2003 first experience will be gained in operating a production grid and a long-term operations plan will be developed.

Level 1 Milestones

The following table summarises the high-level milestones of the project. These will be further elaborated for discussion with the referees at the time of the November meeting of the LHCC.

While there are specific high level milestones for the Applications Area, the progress of the work of the other three areas is marked by a single set of milestones concerned with the deployment of the LCG Grid service which integrate deliverables from all four areas of the project.

M1.1 – Availability of the first LCG Global Grid Service (LCG-1)

June 2003

LCG-1 is a reliable Global Grid Service offering 24 hour, 7 day a week availability, including about ten Regional Centres in Europe, Asia and North America that are ready to take part. The grid provides a batch service for all four experiments for event production and analysis of the simulated data set. This milestone is a functionality and existence test, with a related milestone



LHC Computing Grid Project

Status Report – C-RRB – October 2002

(M1.3) 5 months later for which throughput and reliability will be the key measures. A level 2 milestone will define the target performance levels for this and milestone M1.3 six months before the due date of this milestone. All of these targets must be sustained during a 7-day period for the milestone to be considered met.

M1.2 – General release of hybrid event store June 2003

The first release of the hybrid data store capable of providing event data persistency in a production setting, and supporting also non-event data persistency. This will be a release offering robust basic hybrid persistency services, properly documented and packaged for general use by the experiments, but it will not yet have full functionality.

M1.3 – Fully operational LCG-1 Service & distributed production environment November 2003

This milestone corresponds to the LCG-1 service achieving defined target levels of capacity, throughput, reliability and availability. The targets, which will have been defined in the specification of the LCG-1 service, must be sustained during a 30-day period for the milestone to be considered met.

The milestone includes the availability of the technical service specifications and user documentation, and an appropriate user support infrastructure. The service will by this point include all of the Regional Centres that are participating in LCG-1.

It also includes the availability of a distributed production environment integrating experiment-specific software with common software components and services, grid portal services and grid middleware.

M1.4 – Distributed end-user interactive analysis from a Tier 3 Regional Centre May 2004

Extension of the distributed production environment integrating experiment specific software with common software components and services, grid portal services and grid middleware. This extension will support the analysis environment, enabling distributed end-user interactive analysis down to Tier 3 in the LHC grid. A related level-2 milestone, 6 months before the due date, will provide a full specification of what this milestone includes.

M1.5 – Fully operational LCG-3 Service December 2004

LCG-3 will include all essential functionality required for the initial LHC production service. LCG-3 will be used as a practical demonstration of the LHC computing model, including Tier 0, 1, 2 and 3 Regional Centres, providing input for the computing service TDR. LCG-3 will use the LHC Grid Toolkit, will have 50% of the components required for the 2008 production service for CMS or ATLAS, and will be used for the "20% milestones" of the experiments.

This milestone will be met when specified levels of performance and reliability have been met for a period of 30-days. These target levels will be defined 6 months before the due date of this milestone by a related level-2 milestone.

M1.6 – Full function release of persistency framework March 2005

Completion of the fully functional persistency framework



LHC Computing Grid Project

Status Report – C-RRB – October 2002

M1.7 – Completion of the Computing Service TDR

June 2005

The Computing Service TDR will specify the requirements for the Grid that will be used for the first production services for the four LHC experiments. It will include details of the architecture, functionality, capacity, performance, throughput and availability. It will include the Regional Centre plans that will have been developed to meet these requirements, and will provide cost estimates and an overall installation and verification schedule. It is assumed that the TDR will be approved by the LHCC within three months following its availability, and may be used to provide data for the Memorandum of Understanding for Phase 2 of the project. The full process from acquisition to service verification is expected to take 12-18 months (according to the administrative procedures of the Regional Centres). The initial service must be in full production by September 2006 (6 months before data taking). The TDR will therefore be approved after the acquisition procedures have started, but before orders are placed.