CERN COMPUTER NEWSLETTER

Volume 42, issue 2 April–May 2007

Contents

Editors Nicole Crémel and Hannelore Hämmerle, CERN IT Department, 1211 Geneva 23, Switzerland. E-mail cnl.editor@cern.ch. Fax +41 (22) 766 8500. Web cerncourier.com/articles/cnl.

Produced for CERN by IOP Publishing Dirac House, Temple Back, Bristol BS1 6BE, UK.
Tel +44 (0)117 929 7481.
E-mail joseph.tennant@iop.org.
Fax +44 (0)117 920 0733. Web iop.org.

Published by CERN IT Department

PSS group ensures efficient use of Grid

PSS group's Massimo Lamanna (Arda/EIS project leader), Jürgen Knobloch (PSS leader), Dirk Düllmann (LCG database project leader) and Maria Girone (physics database services).

This interview with Jürgen Knobloch, leader of the Physics Service Support (PSS) group, continues a series this year in which *CNL* is focusing on groups in the IT department that provide key support to the LHC Computing Grid (LCG). PSS is a relatively new group that was formed in a reorganization of department activities in November 2005. The group, of about 50 people, is supporting the rapidly evolving needs of the physics community as a production Grid service comes online and preparations to analyse the first LHC data intensify. PSS has two main sections, one devoted to databases for physics, the other to experiment support and distributed analysis.

What is PSS doing regarding databases to prepare for the start-up of the LHC?

We run the physics databases for the CERN experiments. For LHC experiments we store calibration and alignment data as well as Grid file catalogues. The ORACLE servers, currently more than 200 CPUs, receive the constants from the data acquisition systems and other sources to be ready for the first-pass reconstruction at CERN.

We also support the replication of the databases from CERN (the Tier-0 centre of the LCG) to the Tier-1 centres. This is done for ATLAS and LHCb using Oracle Streams, and for CMS through a web-caching mechanism called FroNTier, which is based on Squid servers. This technology ensures that changed records are transmitted to the other centres so that the Tier-1 databases are synchronized with those in Tier-0. The physics databases at CERN play a central role in the computing strategy of the experiments, and they are designed and operated to be reliable.

One of the challenges we've faced is that this technology hasn't been tried on such a scale before, so we have discovered certain limits. We have made significant improvements through the CERN openlab partnership with Oracle, which is sponsoring two people in PSS to work on this. In particular, in recent months we have overcome a strong latency effect. When shipping data from CERN to places like Taipei this made a difference of an order of magnitude in the rate at which we could refresh distributed databases.

This whole system of distributed databases will be probed soon, when the experiments start their calibration challenges. Until now most Grid work has not relied on local copies of databases, but from April all the experiments will test their ability to use up-to-date calibration constants for the detectors when reconstructing data. We anticipate that

Editorial

this will be a challenging time for us.

We support the persistency framework for LCG, which has been developed in close collaboration with all four LHC experiments. This framework provides a software layer that decouples the user code from the features of any particular database implementation. The persistency framework project focused initially on the development of POOL, a hybrid store based on object streaming into ROOT files and metadata storage into databases. More recently the scope of the project has been extended to provide a generic database access layer (CORAL) and a specialized component for storing and looking up conditions data (COOL).

How does PSS help physicists use the Grid?
Our experiment support and distributed analysis section is involved in certifying new releases of the Grid middleware, gLite. We are also involved in getting the experiment data management systems to work. Furthermore, we support job submission facilities such as GANGA, which was developed by LHCb and has been taken up by ATLAS. We try to foster commonality between experiments, although this is often challenging given their tendency to want to be independent of each other.

Another example of getting experiments to agree on common support is the dashboard used to monitor how jobs behave on the Grid. The dashboard was developed in the ARDA project and was originally used by CMS. It has now been adopted by all the experiments. It's useful because it helps users to understand job failures. Inefficiencies on the Grid can be identified and remedied with this tool.

The role of PSS is to maintain the dashboard and adapt it to the evolving needs of the users. For example, you can now use it to measure the efficiency of individual Grid centres. Each experiment may have problems running jobs on a different set of centres, and the dashboard helps them know where to look for improvements. Since March we have been reporting Grid efficiency statistics regularly to the LCG management board. LCG is now running jobs at nearly 90% reliability at most sites. As we approach LHC start-up the goal is to steadily increase this figure.

As well as high energy physics, what other scientific initiatives do you support?
We receive a lot of support from the Enabling Grids for E-sciencE project for our work with the LHC experiments. In exchange, we are expected to devote some effort to helping other scientific communities adopt Grid technology. This provides our group with useful experience. For example, we have helped to gridify satellite imagery applications for UNOSAT, a United Nations initiative that provides the humanitarian community with access to satellite imagery for use in crises such as earthquakes and tsunamis.

We also supported a series of large-scale data-processing activities carried out by the International Telecommunications Union last year, during a five-week conference to establish a new frequency plan for digital broadcasting in Europe, Africa, the Arab states and the former USSR. By using a few hundred computers on the Grid, the most demanding analysis step could be reduced during the conference from about four hours on a local cluster to less than one hour on the Grid. At the moment we have a doctoral student using Grids and the Geant4 software package to carry out simulation tools for hadron therapy.

Looking beyond the LHC start-up, what challenges lie ahead in your view?
We must face the fact that on the database side the experiments themselves do not know exactly what their requirements will be. Everything will evolve with experience, once real data starts to flow.

For example, we do not know how fast the "constants" that describe the detectors will vary, because there will always be time dependencies of the calibration. How often you need to update constants remains to be seen. Indeed, the number of constants needed is still not known. For example, when we started to work on the Large Electron Positron collider experiments physicists thought they would only need a few constants. But in the end it turned out that each detector element needed a separate calibration.

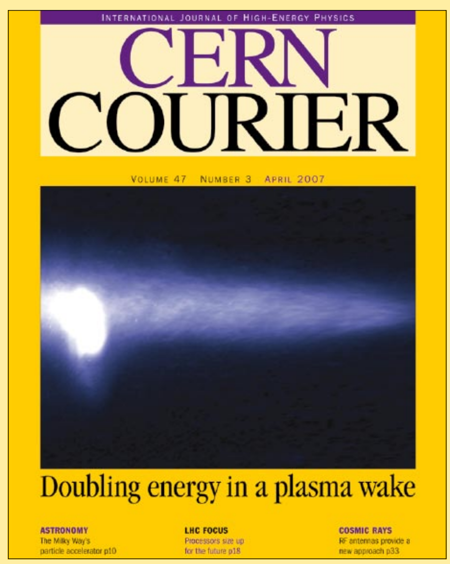
So we have to be flexible when predicting the requirements for detectors. We have now installed about 300 TB raw capacity for the databases, which will be doubled by the time the LHC starts up. We hope this will get us through most of 2008. Beyond that we must be prepared to develop new solutions.

We can also anticipate new applications for databases, which will prove challenging. For example, in ATLAS some physicists are discussing storing all the metadata, the so-called tag data, in Oracle databases. This would be a huge amount of data and we probably wouldn't be able to distribute it to the Tier-1 centres using the streaming techniques we have at the moment, so we would need to study new approaches.

For me the crux of this business is data management. The experiments will need to find the right balance between flexibility and pragmatism. Many researchers would like to see the Grid like a single supercomputer but the Grid cannot provide this sort of functionality. There are limitations due to latency, network throughput and storage capacity. Existing Grid tools will need to be tuned to the needs of the physicists. And the physicists will need to be tuned to the realities of the Grid.

Computing articles featured in this month's CERN Courier

The articles listed below appear in the April issue of *CERN Courier*. Full-text articles and the rest of the issue's contents are available at www.cerncourier.com.



Computing News

- **Worldwide Grid collaboration gets ready for LHC start-up**
Week-long workshop at CERN prepares WLCG members for launch of LHC.
- **Worldwide Grids create worldwide science**
The US science association, AAAS, discusses distributed computing.
- **WISDOM ends second round in the battle against malaria**
Grid analyses 140 million compounds.
- **W3C celebrates 10 years with style**
The anniversary of CSS technology.
- **Commercial quantum chip goes on show**
D-Wave demonstrates the Orion chip.
- **Chip may boost particle-physics Grid power**
Quad-core processor increases speeds.
- **How to avoid a false sense of security**
Advice on protecting PCs from infection.

Calendar of events

Feature article

- **Processors size up for physics at the LHC**
Software challenges in the LHC era and the evolution of PC processors.

CERN upgrades firewall to meet requirements of LHC

In March a new firewall system came into operation to protect the CERN site from threats from the internet. The system combines powerful redundant network devices with a complete management framework, and is making life simpler for end-users, system administrators, network engineers and security experts.

The project to develop a new system began almost one year ago. Limitations with the previous firewall were becoming apparent, and it was clear that it would not cope with the increased traffic load that the Large Hadron Collider experiments will generate. The major limitations of the firewall were limited redundancy, unsupported and nonintegrated management software, and the bandwidth capacity of the "stateful firewall" (a firewall that keeps track of the state of network connections).

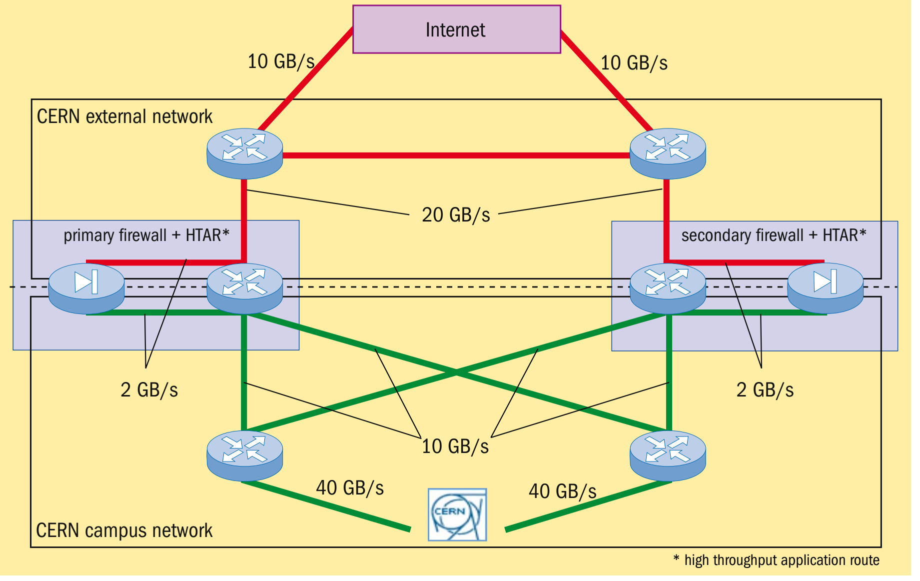


Fig. 1. The new firewall architecture has 10 times the throughput of the previous system.

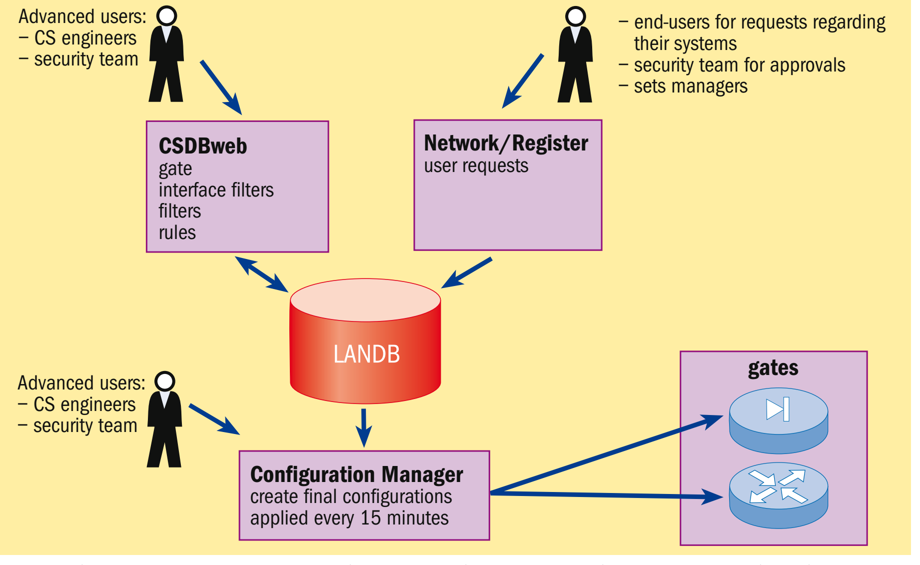


Fig. 2. The management framework features three new applications to work with LANDB.

Technical criteria

The new system, which was launched by the IT department's Communication Systems group (IT/CS) and the Security Team, was designed to meet ambitious requirements and to provide a stable security service platform for years to come.

The firewall had to fulfil the following criteria: to accommodate a growing community of CERN users; to meet the demand for more bandwidth driven by the CERN experiments; to provide a fully redundant system that won't impact any service should there be hardware faults; to offer a flexible architecture that simplifies the integration of third-party security tools; and to provide a fully integrated vendor-independent management framework that can exploit all the data that is already stored in LANDB, the IT/CS network database.

Several challenges had to be tackled. The desired stateful firewall was the high-end product of all the vendors that were contacted, and that meant a high price that exceeded the available budget. This problem was solved with an elegant and affordable redesign of the previous architecture that enables some high-speed data transfer so as not to overload the stateful device and at the same time to avoid duplicating expensive devices. Regarding the management framework, the decision to integrate it with the existing tools and data, and to merge other programs into it, required a significant effort from the team of developers.

Firewall specification

The new firewall system comprises a stateful firewall that can handle 2 GB/s of continuous traffic, an off-load path of 20 GB/s that matches the total bandwidth of CERN's IP network upstream, and an identical set of devices that guarantees an equally capable back-up path (figure 1). The whole system can easily scale to higher throughput when necessary.

The management framework is centred around LANDB, the database that stores all the information concerning the CERN network (figure 2). Three new applications were developed to feed and update LANDB and to use its information to configure the network equipment:

- CSDBweb (Communication Systems DataBase web interface), a tool that is already in use by IT/CS, was expanded to include a module that enables the definition of all the logical parts of any generic firewall and of the filtering policies. It permits information that is already associated with sets of devices to be used to define security policies, in order to simplify the management of large clusters of machines in CERN's Computer Centre.

- Network/Register, a web application that is already known to most users at CERN, now enables any person to quickly submit a request to open the firewall. It also enables the Security Team to seamlessly approve and implement these requests.
- Configuration Manager is the tool that extracts all the information from LANDB and builds the necessary configurations for all the network devices that form the firewall system. The changes are applied every 15 minutes, so a user's request can be carried out in a very short time.

The Configuration Manager plays a critical role; any mistake can impact all the traffic from CERN to the external networks, so it makes many checks before implementing a modification. Special algorithms were developed and validated to estimate and not exceed the actual use of resources.

The new firewall provides much greater throughput, complete redundancy, more secure policies and a simplified management system.

Edoardo Martelli and Nick Garfield, IT/CS

CHEP'07 meeting heads for Victoria, Canada

The next Computing in High Energy and Nuclear Physics (CHEP) conference will be held in Victoria, British Columbia, Canada, from 2 to 7 September 2007.

The conference provides an international forum for the high-energy and nuclear physics community to exchange information, and to review recent, ongoing and future activities. CHEP conferences are held approximately every 18 months. Recent conferences have been held in Mumbai, India (2006); Interlaken, Switzerland (2004); San Diego, US (2003); Beijing, China (2001); and Padova, Italy (2000).

CHEP will be held at the Victoria Conference Centre, close to the inner harbour, above.

Programme highlights

There are eight programme tracks, and the subjects to be covered are as follows:

- Online computing. This will include CPU farms for high-level triggering; farm configuration and run control; describing and managing configuration data and conditions databases; online software frameworks and tools; and online calibration procedures.
- Event processing. This will cover event simulation and reconstruction; physics analysis; event visualization and data presentation; toolkits for simulation and analysis; event data models; and specialized algorithms for event processing.
- Software components, tools and databases. Topics include programming techniques and tools; software testing; configuration management; software build, release and distribution tools; quality assurance; documentation; foundation and utility libraries; mathematical libraries; detector geometry models; component models; object dictionaries; scripting; and graphics.
- Computer facilities, production Grids and networking. Subjects for discussion are Grid operations, monitoring and user support; management and operation of Grid-wide services; experience with operational Grids; technology evolution; global network status and outlook; networks and their relation to Grid systems; and the digital divide and issues of access, readiness and cost.
- Grid middleware and tools. This will include Grid software and monitoring tools; global use and management of resources; Grid technology and its exploitation in distributed computing models; Grid middleware interoperability; Grid middleware reliability; data challenges; evolution and perspective of Grid middleware; and experiment-specific middleware applications.
- Distributed data analysis and information management. Topics for discussion include development of the distributed computing models of experiments; real experience in prototypes and production systems; interactive analysis over wide-area network; mobile computing; tools for supporting distributed analysis; and remote access to and control of data acquisition systems and experiment facilities.
- Collaborative initiatives with other sciences. This will cover areas of mutual collaboration in computing between HEP and other disciplines; and shared computing facilities and Grids.
- Collaborative tools. Subjects to be covered are collaborative systems, progress in technologies and applications; advanced videoconferencing systems; and experience in using videoconferencing tools.

A meeting of the Worldwide LHC Computing Grid will be held before CHEP'07, on 1–2 September. More information can be found at http://indico.cern.ch/conferenceDisplay.py?confId=3578.

For further details about the CHEP'07 conference and to register online, visit www.chep2007.com.

School of Computing will be held in Croatia

This year's CERN School of Computing, organized by CERN in collaboration with the University of Split, will be held from 20 to 31 August in Dubrovnik, Croatia. The school is aimed at postgraduate students and research workers with a few years' experience in scientific physics, computing or related fields. The deadline for registration is 15 May.

Special themes this year are:

- GRID technologies. This track will deliver unique theoretical and hands-on education on some of the most advanced GRID topics.
- Software technologies. This will address some of the most relevant modern techniques and tools for large-scale distributed software development and handling and for computer security.
- Physics computing. This track will focus on informatics topics that are specific to the high energy physics community. After introductory lectures it will address data acquisition and ROOT.

Grants from the European Union Sixth Framework Programme are available to cover part or all of the cost of the school.

Further information can be found on the CERN School of Computing website at www.cern.ch/CSC. This includes more details about the programme, practical information about this year's school, and details of how to apply, including the web-based application form.

Fabienne Baud-Lavigne, IT/DI

A new engine for CERN Search

Much information and documentation is available at CERN and more is produced every day. This includes scientific, engineering and administrative documents, webpages, and minutes of meetings. Depending on their nature, the documents are stored in various collection systems such as CERN Document Server (CDS), Electronic Document Handling (EDH), Engineering Data Management Service (EDMS), websites and databases. Being able to search and find information easily can benefit every user.

The document collections and information systems provide varying advanced functions to search within each collection. This is fine if you know where the document you want is located. But what if you don't? Then you may have to repeat your search in several collections.

CERN Search streamlines document retrieval

In early 2006 the IT department began a project to address the need for a CERN-wide search service. The new search engine, CERN Search, is meant to be a single entry point for users to access the various documentation sources.

IT worked in collaboration with the Norwegian firm FAST, one of the leading companies that provides enterprise search technology. The result is a flexible, scalable and redundant solution that has been designed and implemented to fit CERN's information retrieval needs for now and the future. To access the search engine, visit http://cern.ch/search.

CERN Search replaced the old Intranet Web Search engine in early February 2007. About 1 million indexed webpages are now available to users, and more data sources are being indexed (figure 1). The CDS data became available in March, and EDMS documents will be added in April.

There are plans in the future to make other document collections available, such as all administrative documents in EDH. Since some of these documents are confidential an anonymous search will only display public documents. Authenticating yourself to the search engine will enable users to search within protected content, and the results displayed will depend on the user's access rights to these documents.

CERN Search is providing a service to users and also to service providers who want to offer advanced search functions within their applications and services.

Users can perform a general search in the global index of all available documents, as indicated by the arrows (1) in figure 2. At the same time, some document and information services will offer a more specific and specialized search within their local documents, as indicated by the arrows (2). This localized search can be highly customized to the specifics of the collection and may offer users additional advanced search features. CERN Search is designed to make it possible for service providers to create custom search interfaces, enabling feature-rich subsearches within parts of the global search index.

Further information

For more details consult http://cern.ch/search, and the help pages of CERN Search at http://cern.ch/search/help.

If you have any questions please contact helpdesk@cern.ch. You may send suggestions and feedback to project-search@cern.ch.

Audun Ostrem Nordal and Andreas Wagner, IT/IS

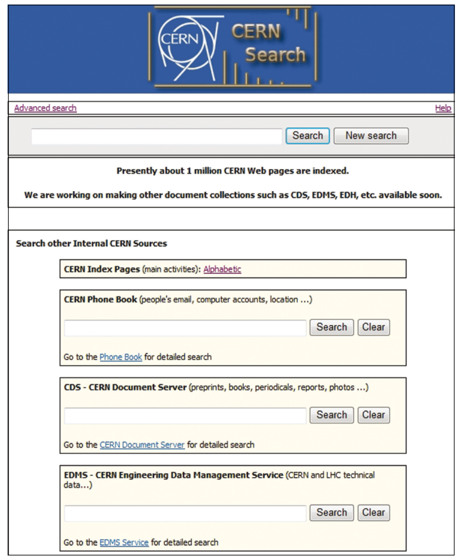


Fig. 1. Various collections are available.

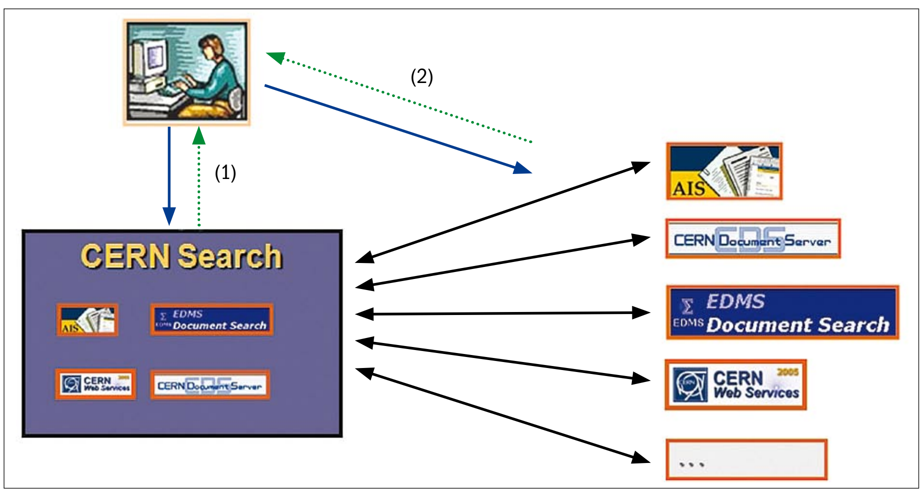


Fig. 2. Users can perform a global search or customized searches of specific collections.

Office 2007 has arrived

The new version of Microsoft Office is now available on all NICE computers through Computer Management Framework (CMF). Office 2007 is deployed in "pilot mode", however. This means that users can choose whether or not to upgrade, and anyone who has problems with this new version can roll back to Office 2003.

You will find detailed documentation about "Using Microsoft Office Enterprise 2007" at https://cern.ch/winservices/Help/?kbid=070195.

A website dedicated to the Office 2007 deployment can be found at http://cern.ch/Office2007. There you can send us feedback and also submit questions and comments to a web forum. To benefit from all the features of Office 2007 (such as the new user interface and Desktop Search), 1 GB of RAM and a fast processor (1 GHz) are strongly recommended.

The following actions have already been taken on all NICE computers at CERN:

- The Microsoft Office 2007 Compatibility Pack has been installed on each computer that runs Office 2003 SP2. This will enable all Office 2003 users to read Office 2007 documents.
- Office XP has been upgraded to Office 2003 SP2. Office XP is no longer supported.
- Project 2003 has been upgraded to Project 2003 SP2, and Visio has been upgraded to Visio 2003 SP2.
- Any missing Office 2003 patches have been applied.

The NICE team

Distributed database project ensures replication to LCG sites

Physics metadata stored in relational databases play a crucial role in the Large Hadron Collider (LHC) experiments and also in the operation of the LHC Computing Grid (LCG) services. A large proportion of non-event data such as detector conditions, calibration, geometry and production bookkeeping relies heavily on databases. Also, the core Grid services that catalogue and distribute LHC data cannot operate without a reliable database infrastructure at CERN and the LCG sites.

The Distributed Deployment of Databases (3D) project is a joint activity between the IT department's Physics Services Support (IT/PSS) group, the LHC experiments and LCG sites to coordinate database services that are coherent, scalable and highly available.

The 3D project and service architecture

Most of the LHC data can be stored and distributed as read-only files. Nevertheless, a significant proportion of data from the central experiment and the Grid requires database services such as:

- consistent and highly available storage for data that is simultaneously accessed or updated;
- recovery to a consistent state after hardware, software or human failures;
- support for efficient *ad hoc* queries.

The 3D project started in mid-2004. Database representatives from 11 Tier-1 sites and three LHC experiments (ATLAS, CMS and LHCb) are profiting from a close cooperation with Oracle as part of CERN openlab. The project also involves two teams from IT/PSS group: physics database services and LCG persistency framework, which also provide the main LCG database services and applications. From the start the project has focused on using database clusters as the main building-blocks in the database service architecture (figure 1).

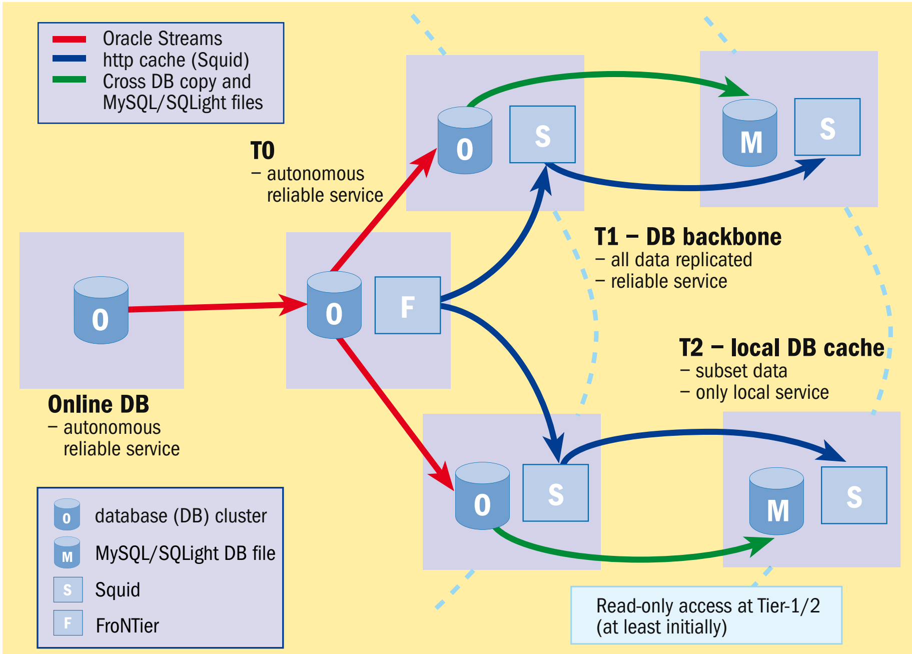


Fig. 1. Database clusters form the main building-blocks of the service architecture in the LCG Distributed Deployment of Databases project. Oracle clusters are used up to Tier-1.

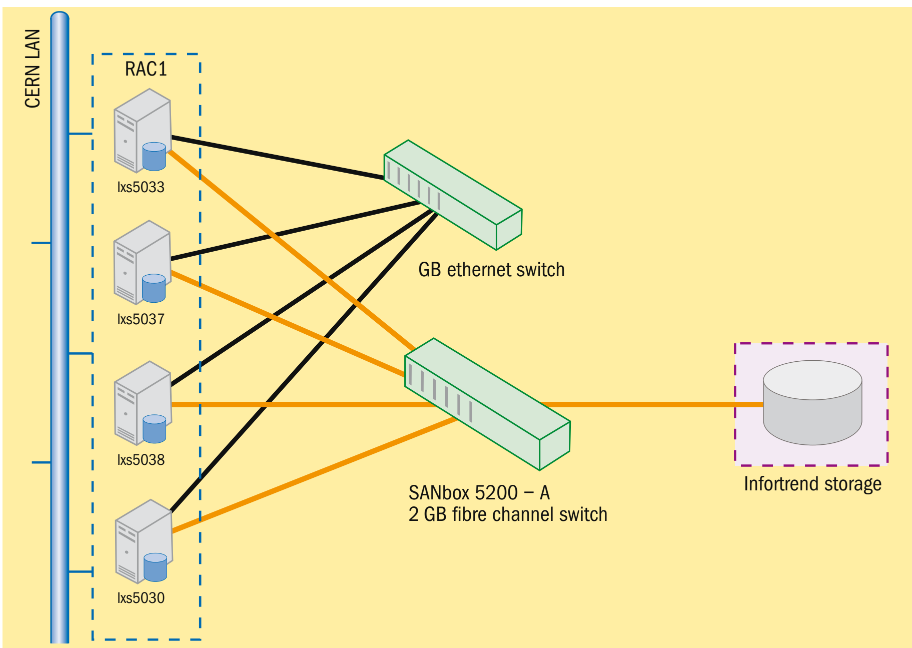


Fig. 2. An example of a database service provided by an Oracle Real Application Cluster (RAC). Several processing nodes access data shared in a storage area network (SAN).

Building-blocks for a distributed service

Oracle Real Application Clusters (RACs) are used to implement reliable database services at the different stages of the main data flow, from online to offline up to Tier-1 sites. Each database cluster (figure 2) consists of several processing nodes that access data shared in a storage area network (SAN). Today for each experiment there are typically eight nodes at Tier-0 and two to three nodes at Tier-1. The cluster nodes use a dedicated network to share cached data blocks to minimize the number of disk operations. A public network connects the database cluster to client applications, which may execute queries in parallel on several nodes. The set-up provides important flexibility to expand the database server resources (CPU and storage independently) according

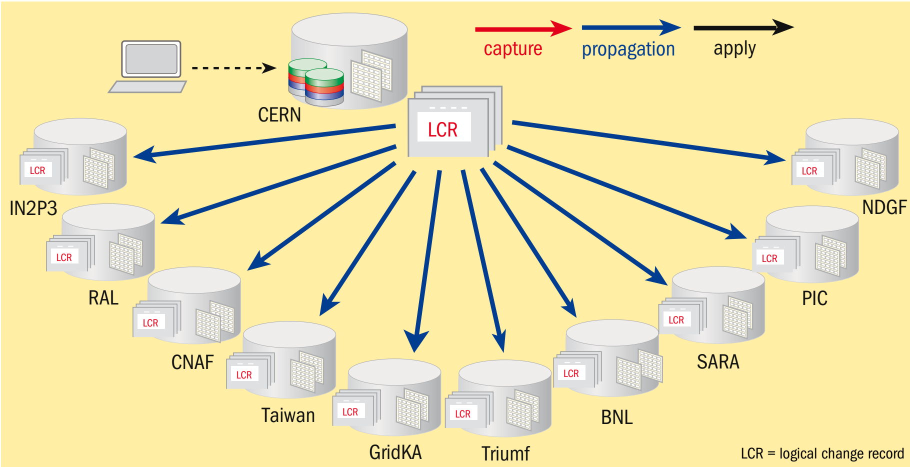


Fig. 3. Oracle Streams forms a replication backbone between Tier-0 and Tier-1 sites.

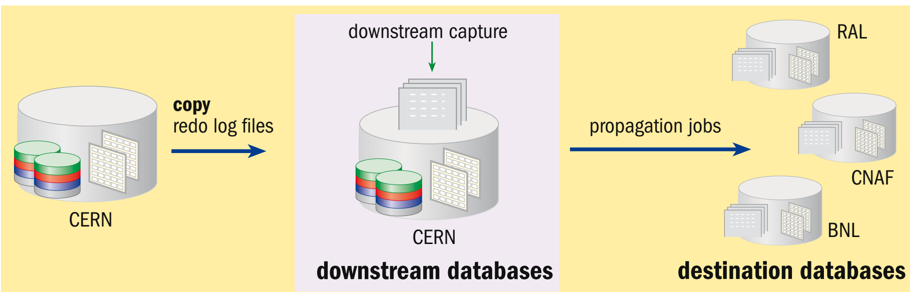


Fig. 4. Data is captured downstream at Tier-0 before being replicated to Tier-1 databases.

to users' needs. This is particularly important during the early phases of the LHC operation, since several applications are still under development and data volume and access patterns may change.

In addition to its intrinsic scalability, the cluster also increases significantly the availability of the database. Should individual nodes fail, applications can failover to one of the remaining cluster nodes and many regular service interventions can be performed without database downtime node by node.

During the last year the physics database service run by IT/PSS has undertaken major preparations for the start-up of the LHC and is now fully based on Oracle clusters on Intel/Linux. More than 100 database server nodes are deployed in some 15 clusters serving almost 2 million database sessions per week. The positive experience with this new architecture at CERN and other sites has led to the setting up of similar database installations at the Tier-1 partner sites worldwide, forming one of the largest Oracle RAC installations.

Connecting database servers via replication

To enable LHC data to flow through this distributed infrastructure, Oracle Streams, an asynchronous replication tool, is used to form a database backbone between online and offline and between Tier-0 and Tier-1 sites. New or updated data from the online or offline database systems

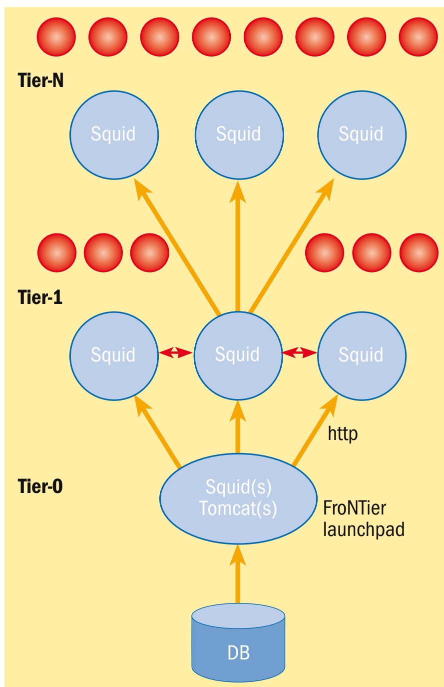


Fig. 5. CMS is using a FroNTier server and connected Squid servers to cache data.

are detected from database logs and then queued for transmission to all configured destination databases. Only once data has been successfully applied at all destination databases is it removed from message queues at the source (figure 3). A dedicated downstream capture set-up is used for replication via the wide-area network to further insulate the source database in case of problems in connecting to Tier-1 replicas (figure 4).

The replication process can be applied either to individual tables or to whole application schemata and takes care of transactional dependencies between several related data changes. In the case of network outages or overloading, the destination databases may experience a time lag to the source database at Tier-0 until the connection is re-established. Data consistency, however, is preserved at each site and at any point in time. This property greatly simplifies the deployment of applications against replica databases, as the semantic of state changes remains intact.

Tests performed by the application developers have confirmed that the experiment and Grid database applications (e.g. the LFC catalog and LCG COOL) work correctly against replicated databases without significant changes in code.

Even though the Oracle Streams replication mechanism supports more complex configurations (e.g. updates at replica databases), its initial deployment will be based on the model that all data changes happen at the replication source at CERN, and replica databases offer read-only access.

Distribution and caching of database results

Complementary to database replication, the CMS experiment is deploying a second technique for its Tier-1 and Tier-2 sites, based on distributing and caching database data via a web protocol. This approach, which is based on the FroNTier package developed at Fermilab, encodes database queries as url requests and transfers the corresponding query results as xml documents. Using http and html enables standard web cache servers such as Squid to enhance the scalability of the distribution system by caching frequently requested data close to the client location at each Tier-1 or Tier-2 site. Figure 5 shows the FroNTier server and a hierarchy of connected Squid cache servers. At CERN a load-balanced set-up with three FroNTier servers was used successfully as part of the CMS challenges in 2006.

A key advantage of FroNTier/Squid distribution is that it requires less effort to install and operate the cache servers, compared with deploying a replica database server. Since all cached data are still available at Tier-0 there is no need to recover the cache state after a hardware failure. Today some 30 Tier-1 and Tier-2 sites have been successfully integrated into the CMS distribution set-up.

To fully benefit from caching, care must be taken when designing the application to avoid subtle inconsistencies that may be caused by stale cache content after changes have been made at the source database. CMS is dealing with this potential problem by invalidating the cache once a day. The experiment is also investigating more effective mechanisms that exploit the knowledge of the application access pattern.

Service monitoring and optimization

The monitoring and diagnostic system is crucially important for the successful deployment of a complex distributed database infrastructure.

The database administrators at CERN and the Tier-1 sites need a detailed view of the server status to spot resource shortages or application misbehaviour before they affect the service. Users and the Grid coordination team also need an overview of the global operational state of the LCG databases, to correlate Grid user problems with database failures.

For this purpose the 3D project has set up a monitoring infrastructure based on Oracle Enterprise Manager (OEM). This is a web-based application that collects many operational metrics from agents running on each database node worldwide. The system can produce various graphs and enables a qualified administrator to quickly identify problems and intervene.

A central repository for the metric data is hosted at Tier-0 and contains information from all of the participating sites. OEM agents can coexist with other monitoring frameworks so that 3D database servers can be integrated into the fabric monitoring in use at a site, such as the Lemon system at CERN.

The 3D Streams Monitor is an additional web application that complements the OEM system by providing a higher-level overview of all databases with a few key metrics. It collects and visualizes information about the streams replication state, which is not yet covered by standard Oracle monitoring. The monitor provides those in charge of experiments and the Grid with summaries of replication data flow, topology and server information, either by site or by experiment on a single screen (figure 6). This information has recently been integrated into the experiment dashboards.

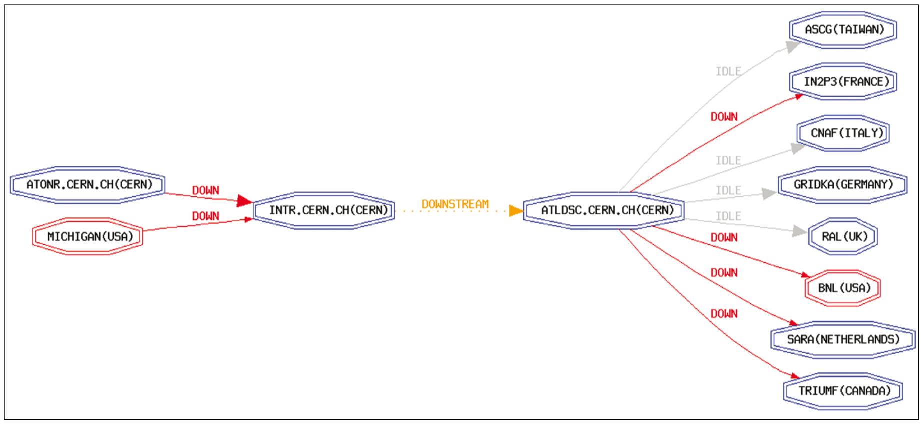


Fig. 6. The 3D Streams Monitor provides data flow, topology and server information.

The monitor also uses an automated system to inform those in charge of sites and experiments of any database problems. This system is being integrated into the Global Grid User Support problem-tracking system deployed by the LCG.

Summary

The 3D project has collaborated with the LHC experiments and LCG Tier-1 sites to implement one of the largest distributed database services for the LCG. Extensible database clusters and streams replication are the main components that ensure a consistent flow of data between online, offline and Tier-1 databases. At CMS this set-up is complemented by database access via web protocols and the caching of local results at Tier-1 and Tier-2 sites.

The initial database is now available at 10 Tier-1 sites around the globe, and it has been used for several months for testing by the ATLAS and LHCb teams. FronNTier servers at Tier-0 and the associated cache servers are in place at some 30 Tier-1 and Tier-2 sites, and have been tested successfully by the CMS experiment. With the ATLAS and LHCb conditions challenges starting now the Tier-1 database replicas will also move into full production, and experience will be gained of accessing database data from Grid jobs on a larger scale. The new service will enable the experiments to complete their preparations for the distributed reconstruction and analysis of detector data in the LCG.

Further information

- Distributed Deployment of Databases project: http://lcg3d.cern.ch
- Physics Database Services at CERN: http://phydb.web.cern.ch
- LCG Persistency Framework: http://pool.cern.ch, http://pool.cern.ch/coral, http://cool.cern.ch
- the recent *CNL* article "Persistency framework manages LCG databases" at http://cerncourier.com/articles/cnl/3/9/7/1.

Dirk Düllmann and Maria Girone, IT/PSS

ALICE tests its digital chain

ALICE successfully tested the infrastructure of its data acquisition, transfer and storage system during its seventh data challenge. The following article has already been published in the *CERN Bulletin* (Issue No. 08–09/2007). As several people in the IT department contributed to this collaborative work and its success we would like to reproduce the article in *CNL* for a different audience.

The ALICE experiment will need a rock-solid data acquisition, selection, transfer, storage and handling system to analyse the billions of bits of data that will be generated every second. The heavy ion collisions at the LHC will generate 10 times more data per second than proton collisions. The ALICE teams have therefore been hard at it for several years designing a cutting-edge informatics system, whose reliability is regularly put to the test in the annual data challenges.

Last December groups from the collaboration and the IT department joined forces, or rather cables, in the seventh of these challenges. The teams of ALICE DAQ (data acquisition), ALICE Offline (data handling), IT/CS/IO (network) and IT/FIO (CASTOR and data storage) all took part in testing the various components of the infrastructure, from data acquisition to transfer and storage.

For the first time, the data for the challenge were generated from the ALICE experimental area. A data-handling room is being installed inside the pit and a third of the acquisition system is already in place. The optical receivers that will receive the data from the detectors in the computers here were used to generate the data, which allowed the teams to test the data-acquisition rate. "Our target is 1 GB of data per second," explained Pierre Vande Vyvre, the ALICE data acquisition project leader. This is only slightly less than the rate that will be needed when the experiment begins operation.

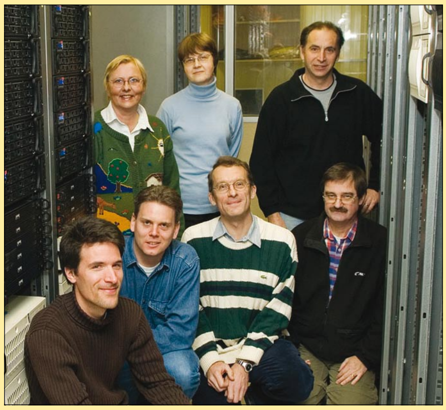
In parallel, this latest data challenge

provided the opportunity to test the buffer storage in the data acquisition room: disks are able to store up to eight hours of data independently. Normally the data flow passes through this data store but if there is a transfer problem downstream the buffer memory provides a few hours' breathing space for the necessary work to be done. In a third major innovation, the data created in the acquisition room were then transferred to the Computer Centre via the CERN optical-fibre backbone. Any disturbance of the network caused by the transfer of data at such a high speed would have been brought to light but happily none was observed.

Finally, the teams ran the four software tools responsible for operating the digital chain: DATE, the data acquisition software; ROOT, the library that allows the data to be formatted; StoreNext, a disk management software programme developed by Quantum; and the CASTOR storage management system developed at CERN.

Altogether, 1 PB (10^{15} bytes) of data was generated and stored during the 18 days of the challenge. For four days and nights running, the rate reached the requisite 1 GB per second. "The infrastructures of the ALICE data acquisition system and the CERN transfer and storage network are ready," said Pierre Vande Vyvre with satisfaction. However, this doesn't mean that the teams have no further challenges ahead of them. The final step before commissioning will be to get the entire acquisition and handling system to work with the data-handling algorithms of the various detectors.

Working in close collaboration, four teams, two from ALICE and two from IT, have assembled the puzzle that is ALICE's data acquisition, transfer, and storage system. Featured below are the four groups involved in the seventh ALICE data challenge.



The ALICE DAQ team. Front, left to right: Sylvain Chapeland, Ulrich Fuchs, Pierre Vande Vyvre, Franco Carena; back: Wisla Carena, Irina Makhlyueva, Roberto Divià. (C Soos and K Schossmaier not featured.)

The IT department's CASTOR team, responsible for data storage. From left to right: Miguel Coelho dos Santos, Jan Van Eldik, Olof Barring, Sébastien Ponce and Bernd Panzer-Steindel.

The network team (IT/CS). From left to right: Edoardo Martelli, Marc Collignon, Jean-Michel Jouanigot, Eric Sallaz.

The ALICE Offline team. Standing: Latchezar Betev and Fons Rademakers; crouching: Cvetan Cheshkov and Federico Carminati.

CASTOR is put to the test

At the other end of the digital chain, the data are stored at the CERN Computer Centre by the CASTOR storage management system. A vital centrepiece of LHC computing, CASTOR controls the storage of data on servers and the high-speed migration of data to magnetic tapes, naturally without a single bit being lost in the process. CASTOR is also involved in the transfer of data to outside institutes in the framework of the Grid. CASTOR, whose second version entered production last year, was developed entirely at CERN.

"Challenges like the ALICE data challenge allow us to debug the system and do the necessary fine-tuning," explained Olof Barring, who is in charge of CASTOR. "ALICE is a very important collaboration for us."

The seventh ALICE challenge involved 25 servers, each with a capacity of 5 TB. As the huge quantities of data were generated they were stored on these disks and transferred to magnetic tapes at a speed of up to 60 MB per second per tape unit. "CASTOR is operational," emphasized Olof Barring. "We now need to fine-tune it, in particular to optimize the way the servers and cassettes are filled."

Corinne Pralavorio, DSU/CO; Pierre Vande Vyvre, PH/AID; Olof Barring, IT/FIO

The deadline for submissions to the next issue of the CNL is

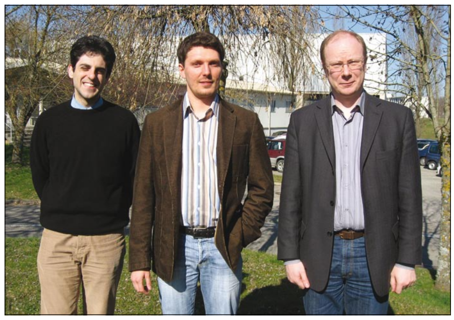
11 May

Please e-mail your contributions to cnl.editor@cern.ch

If you would like to be informed by e-mail when a new issue of CNL is available, subscribe to the mailing list cern-cnl-info. You can do this from the CERN CNL website at http://cern.ch/cnl

CERN shares Grid expertise with EU projects

The European Union is funding a number of projects related to Enabling Grids for E-sciencE (EGEE) to increase awareness of the Grid, develop applications and establish pilot infrastructures in new regions. CERN is a partner in several of these projects, all based on EGEE middleware (gLite). Associates from the different projects work within the IT department's Grid Deployment group (IT/GD), and at the end of their stay return to their home countries with an enhanced knowledge of the Grid.



The project associates in IT/GD, from left to right: Domenico Vicinanza (EUMEDGRID), Konstantin Skaburskas (Health-e-Child) and Rolandas Naujikas (BalticGrid).

Health-e-Child (www.health-e-child.org)

The Health-e-Child project is led by Siemens. It aims to build a scalable tool for clinicians who specialize in children's diseases, enabling them to share data among paediatric institutions in different countries. Sharing patients' data places security and confidentiality constraints on the software, and Grid authentication and authorization services are extended in a custom-built Grid gateway. The gateway also uses gLite data management services, including AMGA, to operate on metadata and access storage elements.

Clinicians can manipulate data, browse thumbnails of medical images and use a DICOM viewer to display full images. The software will be used to support decision-making in the fields of brain tumours, cardiology and rheumatology. The heterogeneity of imaging devices, technologies, data and protocols is a major challenge for the project.

A private Grid is being built to connect participating hospitals via the internet. Having set up the prototype infrastructure, CERN is now investigating the use of virtualization to minimize the number of servers required at each institution. Working closely with the gateway development team at CERN has proved vital for efficient knowledge transfer and rapid prototyping. The project will run for four years and has just successfully completed its first-year review.

BalticGrid (www.balticgrid.eu)

The BalticGrid project aims to extend the European Grid by integrating new partners from the Baltic States (Lithuania, Latvia and Estonia). The Baltic countries are very advanced in terms of computing infrastructure – as was witnessed by Estonia being the first country to use e-voting for parliamentary elections. Although the project is coordinated by the Royal Institute of Technology (KTH) in Stockholm, having a competent project associate at CERN ensures that there is a close collaboration with EGEE in terms of operations and infrastructure.

CERN's main contribution to BalticGrid is in SA1 (Grid operations), providing technical support for the installation, maintenance and operation of the BalticGrid production and certification infrastructures. This includes support and deployment of services such as Virtual Organization Membership Service, Service Availability Monitoring, Resource Broker, Workload Management System, LCG File Catalog, MyProxy and Information Indexes.

EUMEDGRID (www.eumedgrid.eu)

This project, which has received €1.6 m in EU funding, seeks to create national Grid initiatives in the Mediterranean area. EUMEDGRID has already established a human network and Grid presence in most countries in the Mediterranean basin, including Palestine and Israel. It hopes to add Syria to the list soon, showing that the Grid can extend scientific cooperation despite difficult political conditions.

One of the "gridified" applications demonstrated by CERN during the first EU review was a compute-intensive application developed by Italian and Tunisian scientists for modelling the intrusion of saltwater in coastal aquifers. This is useful for evaluating which pumping rates are sustainable without inducing contamination by saltwater. The project was a specific example of interdisciplinary collaboration between north and south Mediterranean countries.

CERN is not contributing hardware to EUMEDGRID but provides knowledge, expertise and support. At the end of the project it is hoped that sustainable Grid resource centres will operate in all of the participating countries.

John Shade, IT/GD

Editor's note

The article "CERN gLite community uses virtualization", published in the last issue of *CNL* (Vol. 42, issue 1, January–March 2007), describes some results based on several tools, including Xen and SmartDomains.

The following article expands on these two tools. Xen is a virtualization tool for supporting virtual servers, and SmartDomains is a virtual domain manager for Xen. SmartDomains has been developed by CERN openlab in collaboration with HP Labs to automate the deployment and management of virtual machines (VMs).

The vGrid tool described in the previous article is a web interface that makes SmartDomains easy to use in the context described in that article.

The CNL editors

SmartDomains deploys Xen VMs

What is Xen? Xen is a virtualization tool that enables a server to support multiple virtual servers. With Xen, users can share a server, isolate applications and boot different operating systems, saving time and money.

"This old box of mine is not really what I need, but for many reasons I cannot switch to a better OS, and by the way I do not need one but five machines, some with SLC to run gLite, some with Debian or Fedora; a small one that is always up for my personal web server, and two big ones that have a lot of memory and Geant4 libraries that would work overnights and weekends." Is this you? What would you say if booting or updating a pool of virtual machines (VMs) was as simple as pressing a button?

Xen is not pure magic – the software, from XenSource and Cambridge University, has to be installed. This involves downloading a distribution that works, installing the files under /boot, updating /etc/grub.conf, and rebooting on Xen kernel. Xen is used more and more at CERN so help on this point is not hard to find.

Is it worth it? With Xen there is no need to buy more boxes, but still, if it takes until lunchtime to find the images, remember the configuration that worked, create logical volumes then launch several machines, there is little time left to actually use them before saving and compressing the images or cleaning everything. As soon as #!/bin/sh comes to mind it may be worth considering a switch to SmartDomains.

So what is SmartDomains? This is a virtual domain manager for Xen that is powered by SmartFrog, an open source configuration management system developed at HP Labs. SmartDomains enables users to write configurations of VM pools and their workflows. It automatically prepares file systems and storage back-ends, boots VMs and checks their liveness, and finally cleans physical hosts back to their initial states, potentially saving and tarring images.

VMs can be launched in a whole computing centre without extra work. The description of the distributed virtual pool is kept in a local file that is used to launch the set-up with a simple start command on localhost. Finally everything can be cleaned in the same way. SmartDomains is based on a peer-to-peer layer that distributes the management work seamlessly.

With SmartDomains there is a language to learn, so those who are ruthless with computers may want to try a few manual VM boots before switching. For the others, starting with SmartDomains default configurations will simplify things.

SmartDomains was developed at CERN openlab by Xavier Gréhant (openlab fellow) and Olivier Pernet (openlab summer student 2006). Several people at CERN have experience in SmartDomains: we are collaborating with the IT department's Grid Deployment group (IT/GD), which is using it in the context of gLite certification. Andreas Unterkircher is driving this use of the application. The description language used in SmartDomains enables template descriptions to be kept and extended. gLite testers edit their own templates via a web interface (vGrid), which was created by Omer Khalid to simplify the use of SmartDomains.

SmartDomains builds on SmartFrog's elaborate description language and its mechanism for deploying distributed Java components. As part of the openlab activities we are investigating applications for SmartFrog at CERN, especially in virtual resource management.

Since SmartFrog is an extensible component-based framework, SmartDomains is also extensible. With its complete base components for Xen it seems to have the potential to become a fully featured virtual resource management system; and because SmartDomains shifts the complexity of run-time resource management to the descriptions, it seems to be particularly adapted to launching batch execution environments. We are keen to find groups who are interested in developing use cases in these areas.

Support for virtualization

Within the IT department, Grid virtualization is organized as a stack of components that are developed or supported by the openlab team and the IT/GD group.

- Xen: support and expertise is provided by Havard Bjerke;
- SmartFrog: support is provided by Xavier Gréhant;
- SmartDomains: development and support is provided by Xavier Gréhant;
- vGrid: development and support is provided by Omer Khalid.
- Virtual machine converter: Dimitar Shiyachki contributed to its development.

Further information

For an online tutorial on SmartDomains, see http://smartdomains.sourceforge.net. To download SmartDomains, visit http://sourceforge.net/projects/smartdomains. For more information, contact xavier.grehant@cern.ch.

A presentation on the gLite test-bed can be found at twiki.cern.ch/twiki/bin/view/Virtualization/GDVirtualization.

See also http://cern.ch/openlab for news about openlab's activities.

Xavier Grehant, IT/DI, openlab

Parallel evolution: DILIGENT and ETICS

DILIGENT and ETICS, two European Commission-funded projects that are leaders in the European Grid scene, recently achieved milestone releases of their software, thanks largely to their collaboration with each other.

ETICS (E-infrastructure for Testing Integration and Configuration of Software) offers Grid-enabled automated building and testing of software. Just over a year after the project launched [in January 2006], its final release candidate was made public. The service, however, has been in use for some time by several projects, including DILIGENT (a DIgital Library Infrastructure on Grid Enabled Technologies).

"We started to use ETICS when it was three months old for our build and deployment testing activities," explained DILIGENT's Pedro Andrade. "At the moment we are the only project also exploring some of the ETICS test functionality."

DILIGENT offers digital library management services that enable the dynamic creation and deployment of digital libraries.

"The way they are using the system gave us many requirements for the testing system," explained Alberto Di Meglio, the ETICS project manager. "For example, consider the case where you want to execute a test and have the system build or deploy components on a given node before the test starts. A number of specific

The ETICS interface and DILIGENT tools.

requirements for this case were written for us by DILIGENT."

Adopting software still in development for mission-critical tasks is a risky decision, but as Andrade explains, the collaboration has been successful.

"DILIGENT is a complex system, with software developed by seven partners and integrating 242 different software components, and ETICS also has a learning curve. It took some time but with the help of the ETICS team this was possible, and I think now, with our release alpha, we can say that the effort invested in the beginning has completely paid off."

The close collaboration between these projects is also helped by a common partner: Italy's Engineering Ingegneria Informatica SpA. Paolo Fabriani works on both projects for Engineering, writing the ETICS web application and managing the building of the DILIGENT software.

"The collaboration has been very good," said Fabriani, "DILIGENT had some constraints and advanced requirements but ETICS was able to take these into account. Engineering is also interested in using ETICS for some pilot projects within the company, and we're also looking at ETICS for some EU projects starting later this year."

DILIGENT is planning the beta release of its software this summer, which will add a greater number of services to the basic package of the alpha release.

Meanwhile, ETICS has some 15 projects registered to use their service, including national efforts such as Grid-Ireland and the Italian EGRID. They are also involved with the US Grid community. The University of Wisconsin-Madison, a project partner, is proposing the ETICS service to the major US Grid projects like Open Science Grid and TeraGrid. ETICS is in the process of making plans for its second year, moving toward a 1.0 release of its service this spring.

Owen Appleton, EGEE contributing editor to International Science Grid This Week (iSGTW)

This article was published in full in *iSGTW* on 7 March (see www.isgtw.org/?pid=1000321). For further information see the DILIGENT website at www.diligentproject.org, and the ETICS website at http://etics.web.cern.ch/etics.

Workshop prepares for Large Hadron Collider

Processing data for the Large Hadron Collider (LHC), the next-generation particle accelerator under construction at CERN, Switzerland, is one of the driving forces for the development of Grid technology.

The task of processing the 15 PB of data the LHC is predicted to generate each year falls to the Worldwide LHC Computing Grid (WLCG) collaboration, which recently met at CERN for a week-long workshop in preparation for the start-up of the LHC later this year.

"WLCG doesn't have its own Grid," explained WLCG's Jamie Shiers. "It uses Grids from several different projects. We use two major Grid infrastructures: EGEE (Enabling Grids for E-sciencE) in Europe, and OSG (Open Science Grid) in the United States, as well as many individual research centres. That we can get more than 200 sites to work together is really impressive."

Shiers coordinates the WLCG service and runs the WLCG "service challenges". These challenges stress the capacity and reliability of WLCG's service to ensure it will be able to cope with the needs of the LHC experiments.

"This workshop was the first to address the full WLCG team," he said. The sites that have agreed to contribute to LHC data analysis in the WLCG collaboration are arranged in a number of "tiers". CERN will serve as the Tier-0 site, which will collect and distribute data to 12 Tier-1 sites. Some 150 Tier-2 sites will contribute to processing the data. This was the first workshop to address all of these tiers, allowing all the partners in the WLCG collaboration to discuss the challenges that face them in getting ready to start taking data.

WILL VENTERS

Around 300 represesentatives attended the WLCG workshop, from 86 sites worldwide.

"This workshop helps us to prepare for the dress rehearsals this summer, which will be end-to-end tests of all components." These rehearsals will test the full chain of computing, from the trigger systems used by the LHC experiments to record collisions through to distributing the data and performing analysis around the globe.

Organizers designed the workshop to be highly interactive. "In previous workshops it was mostly just presentations," said Shiers. "This time we wanted a more motivational event. Jos Engelen, CERN's chief scientific officer, gave a talk on LHC physics, not just the Higgs boson. We also organized a visit to the ATLAS detector. Many people had been told how big it was, but it's a different scale from what most people imagine." For many members of the collaboration this meeting was their last chance to see the detectors before the experiments close this summer in preparation for the start-up of the LHC.

"There will be a pilot run of the LHC starting in November this year," said Shiers. "As far as the WLCG service is concerned, we won't be perfectly ready but we'll be ready enough. We all need the pilot run, as there is a limit as to what we can do in testing. There will be some pain but I don't foresee major problems, it will certainly be a busy year."

Owen Appleton, EGEE contributing editor to iSGTW

This article was published in *iSGTW* on 7 February.

WLCG improves India's access to LHC data

By typing in a simple command on Monday 12 February Gavin McCance initiated the first file transfer between CERN and India on the Worldwide LHC Computing Grid (WLCG).

Connection via the WLCG will give Indian researchers the same level of access to Large Hadron Collider particle physics experiments as researchers in European or North American countries.

"It's a good feeling to be involved in a collaboration like this," said McCance, WLCG/CERN.

The file transfer, containing 1 GB of test data, reached India in less then seven minutes and represents the first drop of an imminent deluge on the WLCG transfer service.

Once the CERN Large Hadron Collider is generating particle physics data, the WLCG, which runs on EGEE and Open Science Grid infrastructures, will be needed to manage and process the files. Since the amount of data generated will be larger than one institution could handle, the task of storing and processing the data will be divided among many sites. These sites are organized in a "tier" system.

India will be the home of two of the Tier-2 sites. One site in Mumbai, the Tata Institute of Fundamental Research (TIFR), will be the main site within India for analysing data from the CMS experiment. The other site, the Variable Energy Cyclotron Centre (VECC) in Calcutta, will be responsible for the calibration and operation of one of the subdetectors in the ALICE experiment. The capacity of the network, though modest now, can be improved quickly.

Collaborators in the CERN–India transfer. From left, standing: Andrea Sciabà (CERN), Gavin McCance (CERN), Patricia Méndez Lorenzo (CERN), Vinod Boppanna (BARC, India). Sitting: Digamber Sonvane (BARC), India Vikas Singhal (VECC).

"It's 45 Mbits now, which by our standards is almost negligible," said Jamie Shiers, coordinator for the WLCG. "But it's a start, and it could soon be increased to 1 Gbit and even beyond."

The test transfer went beautifully. "We were getting basically the same rate as transfer to comparable sites in the US," said Shiers. "Compared to what has been available before now, this is an improvement of light-years."

This new connection highlights the growing ability of distant members within international collaborations to play a vital role in basic operations.

In the past, scientists visiting research centres like CERN would be essentially out of touch once they returned to India. Files are now sent to India around the clock at an increasing rate of transfer. Approximately 1 GB of test data reaches Calcutta every three minutes.

"We've collaborated for a long time with India," said Shiers. "But in the last few years this has become increasingly productive – thanks in large part to technology."

Danielle Venton, iSGTW editor

This article was published in *iSGTW* on 21 February. For more information, visit the VECC website at www.veccal.ernet.in and the TIFR website at www.tifr.res.in.

How to access CERN websites and avoid certificate warning

Many users appear to be uncertain about the best way to publish or access a CERN website that has been created using CERN IS Web Services. Taking the example of the CERN site *Teachers*, should the url be http(s)://teachers.web.cern.ch/teachers or simply http(s)://cern.ch/teachers?

Teachers.web.cern.ch is a DNS alias of the server that hosts the *Teachers* website on the CERN web servers. When accessing the secure form of the url with the format https://teachers.web.cern.ch/teachers/, users will see a warning message about a "security certificate" and then a window that says "There is a problem with this website's security certificate". However, it is still possible to continue by selecting "Continue to this website (not recommended)", and users will finally access the site.

The certificate that the web server uses for https connections is issued to the real name of the server, webh09.cern.ch. Owing to the way an https connection is established, a server can only have one such certificate. This is a limitation of the technology. Therefore, the url https://teachers.web.cern.ch/teachers/ will always display the warning that you see. Unfortunately nothing can be done about that. Previous versions of Internet Explorer and other browsers also displayed this warning, but it was only a pop-up message that could easily be ignored rather than an intermediate page.

The best way to avoid this problem when accessing a CERN website is to use the format https://cern.ch/teachers (or https://www.cern.ch/teachers). This will redirect you to the correct hostname without a security certificate warning.

We recommend that users publish CERN urls in the format http://cern.ch/web-site-name for publicly visible pages, or https://cern.ch/web-site-name for secure connections to, for example, password-protected areas of sites.

Nicole Crémel, IT/UDS; Andreas Wagner, IT/IS

Automatic screen lock is enforced on NICE PCs

As of April automatic screen lock, a password-protected screensaver, will be enabled for all CERN users on NICE PCs. This is being done to increase security.

After a certain number of minutes of inactivity, the screen will be locked and the screensaver will run even if the user had turned off this option before the new setting was introduced.

Users who want to change the setting should go to the WinServices account status webpage (at https://www.cern.ch/WinServices/Services/UserAccounts/AccountStatus.aspx). The change will be effective after the next log-on. Please note that the time out value is not enforced and users can still change it directly on their PC.

If you want to learn more about this setting – for instance, how to set the screensaver time out – go to www.cern.ch/WinServices/Help/?kbid=900190.

The NICE team

A humorous take on the arrival of Windows Vista

When space is available in *CNL* we like to add a note of humour, so, with the permission of globecartoon, we are reproducing this announcement for Windows Vista from the Chappatte website (see www.globecartoon.com).

Windows Vista as viewed by Chapatte.

Patrick Chappatte (known simply as Chappatte in his cartoons) is a Lebanese–Swiss cartoonist who draws for *Le Temps*, *Neue Zürcher Zeitung* (the Sunday edition) and the *International Herald Tribune*. He has also worked as an illustrator for the *New York Times* and as a cartoonist for *Newsweek*.

You can learn about Chappatte's new book, *Globalized*, at www.globecartoon.com/book. This book contains Chappatte's best cartoons from the *International Herald Tribune*, 2005–2007. *Globalized* "will take you on a ride over an over-heating planet, through a clash of civilizations, into a quagmire in Iraq and past a lot of Bush". The book costs €14.95.

Calendar

April

23–27 **HEPiX 2007 (spring meeting)**
Hamburg, Germany
http://hepix2007.desy.de

24–27 **HealthGrid 2007**
Geneva, Switzerland
http://geneva2007.healthGrid.org

May

7–11 **OGF20 and EGEE User Forum**
Manchester, UK
www.ogf.org; www.eu-egee.org/uf2

14–17 **CCGrid 2007**
Rio de Janeiro, Brazil
http://ccGrid07.lncc.br

21–22 **DEISA Symposium 2007**
Munich, Germany
www.deisa.org/news_events/future_events.php

21–24 **Terena Networking Conference 2007**
Copenhagen, Denmark
http://tnc2007.terena.org

25–29 **HPC 2007 and SCS Spring Simulation Multiconference**
Norfolk, Virginia, US
http://hosting.cs.vt.eduv/hpc2007

26–30, **IEEE International Parallel and Distributed Processing Symposium (IPDPS)**
Long Beach, California, US
www.ipdps.org

27–30 **ICCS 2007**
Beijing, China
www.gup.uni-linz.ac.at/iccs

June

4–8 **TeraGrid '07**
Madison, Wisconsin, US
www.union.wisc.edu/teraGrid07

5–7 **Grid Asia 2007**
Singapore
www.ngp.org.sg/Gridasia/2007

September

2–7 **CHEP'07**
Victoria, British Columbia, Canada
www.chep2007.com

October

1–5 **EGEE'07**
Budapest, Hungary
www.eu-egee.org/egee07

November

5–9 **HEPiX 2007 (autumn meeting)**
St Louis, Missouri, US
https://www.hepix.org