

Developing LHCb Grid Software: Experiences and Advances

I. Stokes-Rees, A. Soroko, C. Cioffi, University of Oxford, UK

A. Tsaregorodtsev, V. Garonne, CPPM, France

R. Graciani, University of Barcelona, Spain

M. Sanchez, University of Santiago de Compostela, Spain

M. Frank, J. Closier, CERN

G. Kuznetsov, RAL, UK

Abstract

The LHCb grid software has been used for two Physics Data Challenges, the most recent of which will have produced 90 TB of data and required over 400 processor-years of computing power. This paper discusses the group's experience with developing Grid Services, interfacing to the LCG, running LHCb experiment software on the grid, and the integration of a number of new technologies into the LHCb grid software. Our experience and utilisation of the following core technologies will be discussed: OGSi, XML-RPC, grid services, LCG middle-ware, and instant messaging.

Keywords

grid services, instant messaging, LCG, grid integration, OGSi, XML-RPC, computational grids, fault tolerance, matchmaking, pull scheduling

1 Introduction

LHCb is one of four particle physics experiments currently being developed for the Large Hadron Collider (LHC) at CERN, the European Particle Physics Laboratory. Once operational, the LHCb detector will produce data at a rate of 40 MB/s[19]. This data is then distributed around the world for 500 physicists at 100 sites to be able to carry out analysis. Before this analysis of real physics data can begin simulations are required to verify aspects of the detector design, algorithms, and theory. LHCb has worked closely with the LHC Computing Grid (LCG)[4], which is coordinating the common computing strategy for the four LHC experiments (LHCb, Atlas, Alice, and CMS). The LHCb computing model intends to utilise many aspects of LCG but will have, in addition, experiment specific components.

DIRAC (Distributed Infrastructure with Remote Agent Control) is the resulting LHCb grid computing environment. It incorporates LCG computing resources and functionality, while also, critically, allowing the integration of non-LCG resources (see figure

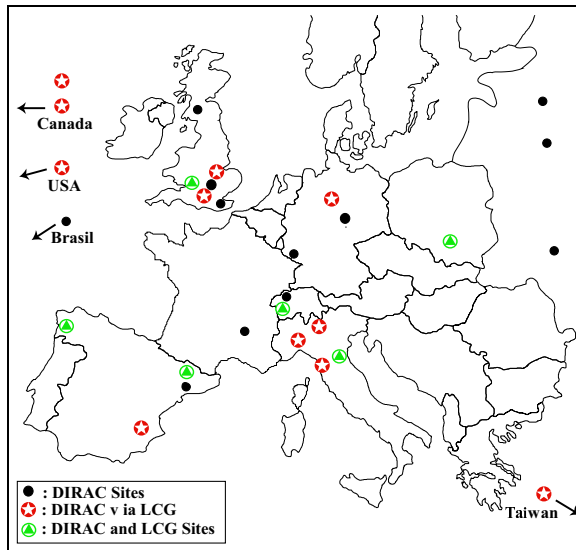


Figure 1: Sites running DIRAC. This includes a mixture of LCG sites and conventional cluster computing centres.

1). This paper recounts our experience of developing DIRAC, integrating it into the LCG grid environment, making use of existing middle-ware services and libraries, and advances from incorporating new technology such as instant messaging into our architecture.

We will start with a description of the DIRAC architecture in section 2, outline key features and advances in section 3, describe our experience developing and using grid software in section 4, and conclude in section 5 with perspectives for the future of DIRAC and computational grids.

2 Architecture

DIRAC is designed following a lightweight Agent/Service model, which emphasises a *Service Oriented Architecture*. It provides a scalable high throughput generic grid computing environment for uncoupled or loosely coupled long running computational tasks, requiring significant input data and producing large volumes of output data. The basic design objectives are to support: 100,000 queued jobs; 10,000 running jobs; and 100 sites.

The architecture is divided into four areas: Services, Agents, Resources, and User Interface. The core of the system is a set of independent, stateless, distributed *Services*. The services are meant to be administered centrally and deployed on a set of high availability machines. *Resources* refer to the distributed storage and computing resources available at remote sites, beyond the control of any central administration. An *Agent* runs on each computing resource and monitors the resource availability, requesting jobs when possible.

The *User Interface* API provides access to the Services, for control, retrieval, and monitoring of jobs and files, and has been incorporated into command line tools, GUIs, and web sites. A complete GUI interface for managing LHCb jobs has been produced by the Ganga project[1].

The general separation between Services and Agents is that Services are *stateless* and *reactive*, whereas Agents are *stateful* and *proactive*. The Services can be distributed across several machines, or run from a single “server”. This allows easy replication for redundancy and load-balancing.

2.1 Job Management Services

Jobs are described using the text based ClassAd Job Description Language (JDL) designed by the Condor project for use with the Condor Matchmaking scheduling system[21]. The JDL file is submitted to the Job Receiver Service which registers the job in the Job Database and notifies the Optimiser Service. The Optimiser Service sorts jobs into different job queues and dynamically re-prioritises queue ordering. Agents monitor resource availability. When they detect “free slots”, they submit a job request to the Matchmaker Service, which interrogates the various Job Queues and returns a suitable job, based on the resource’s profile. These components are illustrated in figure 2.

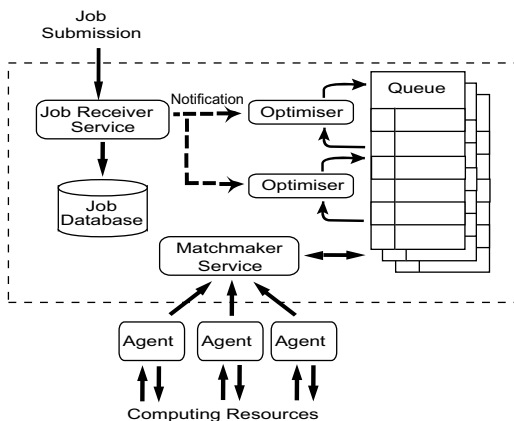


Figure 2: Core DIRAC Services

2.2 Data Management Services

The DIRAC Data Management Services provide fault tolerant transfers, replication, registration, and

meta-data access for files both at DIRAC computing centres and long term mass storage sites.

A *Storage Element* (SE) is an abstracted interface to Internet-accessible storage. It is defined entirely by a host, a protocol, and a path. This definition is stored in the Configuration Service (see section 2.3), and can be used by any Agent, Job, Service or User, either for retrieving or uploading files. Protocols currently supported by the SE include: gridftp, bbftp, sftp, ftp, http, rfiio or local disk access. The SE access API is similar to the Replica Manager interface of the EDG project.[15]

The *File Catalogue Service* provides a simple interface for locating physical files from aliases and universal file identifiers. This has made it possible to utilise two independent File Catalogues, one from the already existing LHCb Bookkeeping Database, and another using the AliEn File Catalogue from the Alice experiment[2]. In the recent LHCb Data Challenge they were both filled with replica information in order to provide redundancy to this vital component of the data management system, and to allow performance comparisons to be made.

Within a running job, all outgoing data transfers are registered as *Transfer Requests* in a transfer database local to each Agent. The requests contain all the necessary instructions to move a set of files in between the local storage and any of the SEs defined in the DIRAC system. Different replication, retry, and fail-over mechanisms exist to maximise the possibility of successfully transferring the data (see section 3.4). This system disconnects the data transfer from the job execution in a manner similar to that done by Condor Stork[14]. It recognises that data placement is a significant, but under appreciated, part of computational grid infrastructure.

2.3 Configuration Service

It is necessary for Services, Agents, Jobs, and Users to be able to “find” each other, and to learn about the properties of the other components. This is a common issue in all service oriented architectures. Services also need their own configuration mechanism. It was felt that the existing approaches, such as LDAP, UDDI[12], MDS[10], and R-GMA[24, 6], were powerful, yet complex, and required significant infrastructure to utilise.

As such, and in keeping with the principles of simplicity and lightweight implementation, a network-enabled categorised name/value pair system was implemented, which overloads the Python ConfigParser API and the Microsoft Windows INI file format. Components which use the Configuration Service do so via a *Local Configuration Service* (LCS). This retrieves information from a local file, from a remote service, or via a combination of the two.

This system proved to be very robust, and the multi-threaded XML-RPC server was able to handle high volumes of concurrent requests, as illustrated in figure 3. The simplicity of the data format and the access API allowed site administrators to easily edit their local

configurations, and for multiple Configuration Service implementations to be developed.

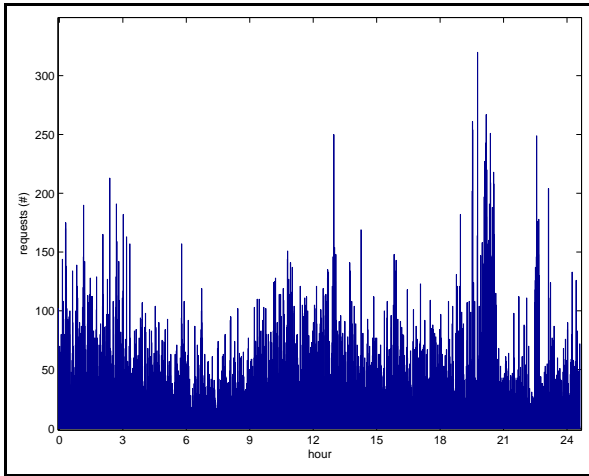


Figure 3: XML-RPC requests per minute to the Configuration Service over a 24 hour period

2.4 Monitoring and Accounting Services

The *Job Monitoring Service* provides an interface for Agents and Jobs to update job state and for other Services or Users to query job state. This only retains information for jobs which are active in the system. Jobs which have completed or failed are eventually cleared to the *Job Accounting Service*. There are three access modes to the Monitoring Services: an API, a web-interface, and command line tools.

2.5 Agent

The Agent handles DIRAC jobs at a computing site by submitting them to the Local Resource Management System (LRMS). The current system supports PBS, LSF, BQS, Sun Grid Engine, Condor, Globus, EDG/LCG, Fork, and In-Process resource types. The job JDL is inspected by the Agent to handle staging and registering of input or output files. The Agent monitors the progress of the job and sends status updates to the Monitoring Service.

The design consists of a set of pluggable Agent modules. The modules are executed in sequence in a continuous loop. Typically a site runs several agents each having its own set of modules, for example job management modules or data management modules. This feature makes the DIRAC Agent very flexible, since new functionality can be added easily, and sites can choose which modules they wish to have running. Further details are discussed in section 3.2.

Agents can operate in a *cycle-scavenging* mode at the cluster level, where they only request and execute jobs when the local resources are under-utilised. This idea comes from global computing models, such as SETI@Home, BOINC, and distributed.net [23, 3, 7], which perform cycle-scavenging on home PCs.

3 Key Features and Advances

This section discusses four aspects which have been key to the success of DIRAC: the pull scheduling paradigm, lightweight modular agents, the use of instant messaging, and mechanisms to provide fault tolerance.

3.1 Pull Scheduling

DIRAC emphasises *high throughput* rather than *high performance*. This idea was first presented by the Condor project[5], from which DIRAC borrows heavily in terms of philosophy for designing generic distributed computational systems[18]. It advocates immediately using computing resources as they become available, rather than attempting global optimisations of all jobs over all resources. In the Condor approach, which we will call the *pull* paradigm, computing resources request computing tasks by announcing their availability. In contrast a *push* paradigm has a scheduler which monitors the state of all queues and assigns jobs to queues as it wishes.

For *push* scheduling to work, all the information concerning the system needs to be made available at one place and at one time. In a large, federated, grid environment this is often impractical as information may be unavailable, incorrect, or out of date. Even if it is available, job allocation complexity grows quadratically with the number of jobs and resources, where every possible allocation combination must be evaluated to select an optimal schedule. While there are efficient heuristic approaches that in practise approach an optimal solution, such algorithms generally require complete and up to date information regarding system state, and are typically designed to operate on homogeneous computing resources with $10^2 - 10^3$ queued jobs.

As a result of this, *push* scheduling in a grid environment has proven to be problematic. By contrast, the DIRAC Central Services simply maintain queues of prioritised jobs (see section 2.1) and allocate the highest priority job which matches the resource's profile. The Condor Matchmaking libraries facilitate dynamic resource definitions, as opposed to the traditional batch system which contains queues consisting of static characteristics[21].

The previously difficult task of determining where free computing resources exist is now distributed to the remote Agents (see section 2.5) which have an up to date view of their sites state. Since jobs are grouped into queues based on common requirements, the worst case is that each Agent job request will be compared against each queue once, where the number of queues is much less than the total number of queued jobs.

Both long matching time and the risk of job starvation can be avoided through the use of an appropriate Optimiser to move “best fit”, “starving”, or “high-priority” jobs to the front of the appropriate queue. As

reported elsewhere[9], this allows a mixture of standard and custom scheduling algorithms.

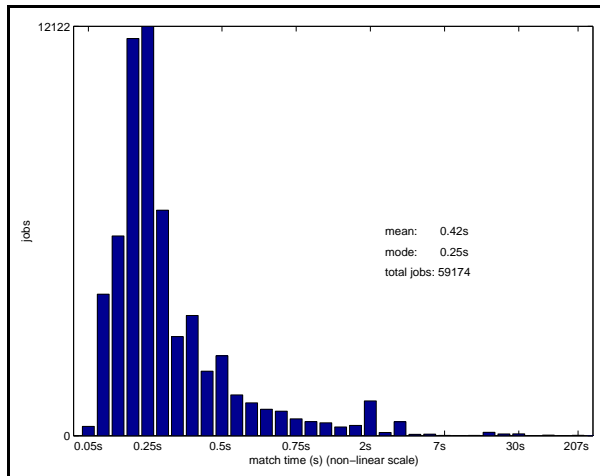


Figure 4: Match time distribution for 60,000 jobs during DC04

Figure 4 shows the match times for jobs during LHCb DC04. 97% of the time this operation takes less than one second even with tens of thousands of queued jobs, thousands of running jobs, and dozens of Sites requesting jobs concurrently.

3.2 Lightweight Modular Agents

By providing simple abstractions of *Computing Elements* (CE) and *Storage Elements* (SE), and exposing simple APIs to the Core Services, it was possible to implement lightweight Agents (see section 2.5) which can be installed and run entirely in user-space on any Computing Resource. This allows the rapid utilisation of heterogeneous systems in a federated manner — the most general objective of computational grids. Local site administrators simply install a one megabyte self-contained package with all the necessary software for the DIRAC Agent.

The configuration allows local policies on queue usage to be applied, and selection of specific Agent modules to run. This modularity gives administrators great flexibility and control, and makes it easy to write custom modules.

The only pre-requisite is a recent version of the Python interpreter and outbound Internet connectivity, in order to contact the DIRAC Services. This allows the agent to run under virtually any computing and network environment, including behind firewalls and private networks utilising Network Address Translation (NAT) to reach the Internet. Installation entirely in regular user-space mitigates the security risks present in software which requires “root” access and system wide installation.

3.3 Instant Messaging for Grid Services

DIRAC has incorporated an instant messaging framework into all the components: Services, Agents,

Jobs, and User Interface. This provides reliable, asynchronous, lightweight, and high speed messaging between components. Public demand for instant messaging has led to highly optimised packages which utilise well defined standards, and are proven to support thousands to tens-of-thousands of simultaneous users. While these have primarily been for person-to-person communication, it is clear that machine-to-machine and person-to-machine applications are possible, and it is in these areas DIRAC has demonstrated a novel application of the technology.

While the DIRAC Services expose their APIs via XML-RPC, due to the simplicity, maturity, and robustness of this protocol, the need to expose a monitoring and control channel to the transient Agents and Jobs led to the use of instant messaging. No *a priori* information is available about where or when an Agent or Job will run, and local networks often will not allow Agents or Jobs to start an XML-RPC server that is generally accessible. This suggests a client-initiated dynamic and asynchronous communications framework is required.

Extensible Messaging and Presence Protocol (XMPP), now an IETF Internet Draft[13], is currently used in DIRAC. This has grown out of the open-source, non-proprietary, XML based Jabber instant messaging standard. XMPP provides standard instant messaging functionality, such as one-to-one messaging, group messaging (“chat”), and broadcast messaging. An RPC-like mechanism exists called Information/Query, (IQ) which can be used to expose an API to any XMPP entity. The *roster* mechanism facilitates automatic, real-time monitoring of XMPP entities via their *presence*.

The DIRAC Services use XMPP in places where fault tolerant, asynchronous messaging is important. For example, the Job Receiver Service uses XMPP to notify the Optimiser Service when it receives a new job. When the Optimiser gets this message, it will then sort the new job into the appropriate queues. The IQ functionality has the potential to allow users to retrieve real-time information about running jobs, something which is critical for interactive tasks, or for job steering. It also greatly facilitates debugging and possible recovery of stuck jobs.

XMPP is specifically designed to have extremely lightweight clients, and gracefully handles dynamic availability of entities, buffering all messages until an entity is available to retrieve them. By matching the XMPP IQ functionality to standard XMPP messages, it is possible for users with a standard XMPP client to locate and communicate with Agents, Jobs and Services from anywhere. This has already been put to good use in DC04 for controlling and monitoring the state of Agents.

The two main outstanding issues for the use of Instant Messaging are the security and authentication implications of a “tunnelled” control channel into remote computing sites, and the scalability to tens of thousands of XMPP entities communicating across the

same instant messaging network. For the first, a group at Lawrence Berkley Laboratory (USA), have developed an XMPP server which accepts Globus x509 GSI certificates for authentication. This is a positive step and we will be working with them to investigate how this can be used for end-to-end security and authentication. For the second, many high performance commercial XMPP servers are available, however the freely available open source servers still demonstrate some robustness and scalability issues which have created problems when XMPP is used heavily, for example by thousands of jobs broadcasting status updates (see section 4.2).

3.4 Fault Tolerance

In a distributed computing environment it is impossible to assume that the network, remote storage, and remote services will constantly be available. The result is that any remote operation may fail in one of three ways: failure to connect to the remote resource; stall during the remote operation; or, exit with a failure.

These failures often are not permanent, so a retry at a later time or to an alternate equivalent resource may be successful and allow the parent operation to complete, albeit with a delay incurred due to the retry. In order to cope with these failure modes the following mechanisms were used in DIRAC:

Retry Many commands retry with a time delay in order to overcome any network outages, service request saturation, or service failure and restart.

Duplication Numerous services have a duplicate backup service available at all times.

Fail-Over When contacting critical services, after the retry limit is reached, a request to an alternate service is attempted.

Caching In the Local Configuration Service, the remotely fetched data can be cached locally for future retrieval.

Watchdog Monitors components to ensure continuous availability and restart on failures.

All Services and Agents are run under the *runit* watchdog[20]. It ensures that the component will be restarted if it fails, or if the machine reboots. It also has advanced process management features which limit memory consumption and file handles, so one service cannot incapacitate an entire system. Automatic time-stamping and rotation of log files facilitates debugging, and components can be paused, restarted, or temporarily disabled. Furthermore, none of this requires root access.

4 Implementation and Operational Experience

DIRAC has been developed over the past two years by a core team of two to four developers, with extensive

input, contributions, testing, and deployment feedback from the LHCb Data Management Group, and computing centre administrators. It has aimed to bridge the computing requirements of LHCb with the capabilities available at the collaborating computing centres, and to provide a basis for evaluating grid computing approaches, particularly the functionality offered by the LCG environment.

4.1 OGSA and OGSi

The framework proposed by OGSA and specified in detail by OGSi was seen as a very positive step toward increased interoperability between grid software components. The DIRAC team invested several months of intensive work in developing DIRAC Services as Java GT3 components, however this was abandoned shortly before Globus and IBM jointly announced their intent to discontinue OGSi and in its place proposed WSRF (Web Services Resource Framework)[11].

In principal, the DIRAC team supports the idea of dynamic, stateful, transient Grid Services, as compared to Web Services which are static and stateless. A common standard for security, lifetime, service data, and publish/subscribe event notifications are all required, however in the end we found OGSi was unworkable for the following reasons:

heavyweight and complex impossible to develop lightweight clients, difficult to run as a regular user, significant infrastructure required for deploying Grid Service container

not standards compatible unable to leverage existing Web Services tools

poor documentation for installation, maintenance, debugging, development

poor implementation many bugs in GT3 and constant exceptions being thrown

Together these made it difficult to develop, debug, deploy, and maintain OGSi Grid Services. Similar experiences were recorded by others[22]. We also investigated using a pure Python implementation of OGSi, pyGridWare, prepared by Lawrence Berkley Laboratory (USA), but this was not sufficiently complete to be usable. While it is understood that the more recent versions of GT3 have corrected many of the early technical problems, the combined facts that OGSi no longer had a future, and the complexity of development under GT3 forced development of DIRAC to return to Python and XML-RPC.

4.2 Data Challenge 2004

The DIRAC system has been used for the LHCb Data Challenge 2004 (DC04), held from May to July 2004. DC04 had three goals: to validate the LHCb distributed computing model based on the combined use

of LCG and conventional computing centres; to verify LHCb physics software; and, to generate simulation data for analysis. 150,000 jobs were run, consuming over 400 processor-years of CPU power, and producing 90 terabytes of data. This data was redistributed across the centres for both organised (i.e. planned and predictable) and chaotic analysis of the results.

The system operated smoothly with a sustained level of over 2000 running jobs, and 600 gigabytes of data generated and replicated daily. Figure 1 shows the participating sites, and figure 5 shows a snapshot of the running job distribution. Once installed, the DIRAC Agents ran autonomously and restarted after failures or reboots. A central team watched the monitoring system and alerted site administrators when problems were detected with the local resources.

The twenty participating sites varied enormously in size, from 20 CPU clusters shared heavily with other users to large 500+ dedicated CPU clusters. A mailing list and weekly phone conferences allowed the DIRAC software developers, site administrators, and data challenge managers to discuss progress and solve problems.

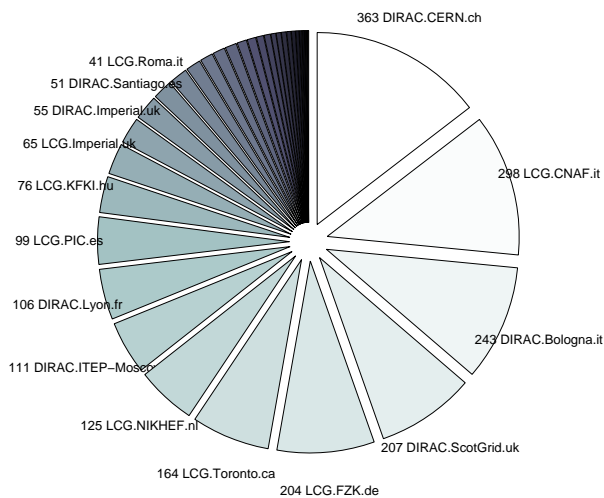


Figure 5: Representative snapshot of running jobs per site during DC04. Notice a mixture of “standard” sites (prefixed DIRAC), and sites accessed via LCG (prefixed LCG). Numbers indicate running jobs at that site.

Another 20 sites were accessed via LCG, and are discussed in section 4.3. In total, these 40 sites provided more than 3000 worker nodes. At the typical level of 5000 globally queued jobs, the Matchmaker Service responded to Agent job requests in an average of 0.42 seconds (see figure 4). More than 40,000 jobs were completed in the month of May with an average duration of 23 hours, running on average at 93% load, the remaining 7% being I/O operations. Each job produced on average 400 megabytes, which was replicated to several sites for redundancy and to facilitate later data analysis. To date the system has produced, stored and transferred 20 terabytes of data.

At the time of writing, there have been four major outages in the DIRAC core services availability which resulted in jobs failing, jobs stalling, or sites failing to

get new jobs:

Deleted Database One of the high availability core servers, which is monitored 24 hours a day by CERN IT staff, reached 90% of full on the local hard drive. This was due to a large and very actively used database on the server. The IT staff intervened by stopping MySQL and deleting part of the database, resulting in the loss of all queued jobs.

Distributed Denial of Service Early efforts to incorporate instant messaging into all aspects of DIRAC resulted in very effective distributed denial of service attack on the server hosting DIRAC and the instant messaging hub. Thousands of jobs were simultaneously sending status information, and in many cases were (unnecessarily) sharing this information with each other, resulting in an extremely high, and unmanageable, message volume which compromised the performance of other services running on the same server.

Network Failure CERN experienced a site wide network failure for approximately one day due to efforts required to isolate an internally compromised machine. All services were unavailable during this time, and it was proposed that a fail-over system be prepared at an external site. This was not completed due to the infrequency of extended total network failure at CERN and the effort required to configure and manage a second DIRAC system.

MySQL Connection Limit MySQL has a default limit of 100 simultaneous database connections. The multi-threaded XML-RPC Services take their DB connection handles from a connection pool local to each Service. If the pool is emptied, the Service creates more connection handles. At a point of high load, due to the ramp up of LCG sites in DC04, one Service repeatedly emptied its connection pool and claimed all 100 available connections, thereby blocking all other Services from communication with the DB. Until the MySQL connection limit was increased and the Services set a pool limit, the DIRAC Services were effectively unavailable. This occurred over the weekend and resulted in a day of lost job matching, although running jobs continued.

Only a few significant bugs were identified in the DIRAC software. These appeared early in the data challenge and were quickly resolved. They generally centred around service scalability and availability, requiring the implementation of operation buffering, timeouts, and fail-over mechanisms on both the client and server sides (see section 3.4). Once these early bugs were resolved, “standard” computing sites observed stable performance. Experience with LCG is discussed in section 4.3.

The small size of DIRAC, buffering of transfer requests, use of a local job database, and independence

from the local batch system, all meant that it was possible to stop the Agent, even while jobs are still running on the site, perform a software update, and recommence the Agent without losing existing jobs or transfers. This greatly facilitated rapid resolution of bugs, and was even extended to a prototype Update Agent module which would perform automatic Agent software updates.

Data Management presented the greatest overall challenge. A number of sites experienced significant data transfer delays or failures, resulting in transfer backlogs. Large sites would quickly fill their queues with hundreds of jobs, producing 40 Gb of data and all finishing at approximately the same time, therefore saturating the site's outbound bandwidth or the target server's inbound bandwidth. Although DIRAC supports a wide range of transfer protocols (see section 2.2), difficulties in using every one of these were encountered. In particular we note the lack of a simple user-level installation of a grid-ftp client as a major stumbling block in its acceptance. From a global view the system has difficulty in identifying fatally failed transfers (i.e. those that will not be retried) and transfers which are outstanding but queued. There are plans to use the XMPP interface to Agents to resolve this. The Transfer Request mechanism performed well and eventually flushed data backlogs.

4.3 Integration with LCG

LCG is required to make possible the storage and processing of the vast quantities of data produced by the LHC experiments. It will bring together hundreds of computing centres around the world and provide an aggregated computing power equivalent to over 70,000 of today's fastest processors. One of the broad objectives for DIRAC is to provide a smooth transition from cluster based to grid based computing for the LHCb experiment and to integrate LHCb computing with the LCG resources.

The following are the primary challenges LHCb has faced in trying to use LCG for DC04 and during integration with DIRAC:

Computing power and queue normalisation

Job submissions to LCG include a job duration. This only makes sense in some sort of normalised time units (for CPU-bound jobs), however there was no good mechanism to do this. Some sites internally normalised their queue times, resulting in LHCb jobs running with less time than expected. Some sites overload the processors or use Hyper-Threading, similarly throwing off the expected execution time. This resulted in thousands of jobs being terminated by queue time limits.

Job working space LHCb jobs required 500 megabytes to 1.5 gigabytes of working space. NFS mounted job working directories resulted in

overwhelming NFS. Other sites did not provide sufficient space for LHCb jobs to complete.

Availability of output files after failure

LCG jobs run in a "sandbox" which is erased when a job completes, is cancelled, or aborts. Output files are often required to diagnose the cause of failures, but these are not available if the job was cancelled or aborted.

Security certificates The x509 GSI certificates do not allow sharing of job details with others within the VO. This makes collaborative work difficult. There were also bugs found in the handling of proxy certificates which would result in the use of almost-expired certificates and the subsequent failure of jobs.

Operating on large volumes of jobs None of the LCG tools are responsive or easy to use with large numbers of jobs. During DC04 hundreds or thousands of jobs would be submitted to LCG each day. Monitoring and managing these jobs was extremely difficult.

Poor job scheduling It was difficult to evenly spread DC04 jobs across the available LCG resources due to problems with the Ranking algorithm and Estimated Response Time values. At times all jobs submitted to LCG would end up queued at a single overloaded site.

Lack of API documentation The LCG command line tools are largely written in Python, the same language as DIRAC. It would be preferable for DIRAC to use the same APIs as utilised in these tools, however the lack of useful documentation made this difficult.

5 DIRAC Future Developments

The service oriented architecture of DIRAC proved that the flexibility offered by this approach allows faster development of an integrated distributed system. The *pull* paradigm Agent/Service model has scaled well with a large and varying set of computing resources, therefore we see the future evolution of DIRAC along the lines of the services based architecture proposed by the ARDA working group at CERN[17] and broadly followed by the EGEE middle-ware development group[16]. This should allow DIRAC to be integrated seamlessly into the ARDA compliant third party services, possibly filling functionality gaps, or providing alternative service implementations. The use of two different File Catalogues in the DIRAC system is a good example of leveraging the developments of other projects, and being able to "swap" services, provided they implement a standard interface.

DIRAC currently operates in a trusted environment, and therefore has had only a minimal emphasis on security issues. A more comprehensive strategy is required for managing authentication and authorisation

of Agents, Users, Jobs, and Services. It is hoped that a TLS based mechanism can be put in place with encrypted and authenticated XML-RPC calls using some combination of the GridSite project[8], and the Clarens Grid Enabled Web Services Framework, from the CERN CMS project.

While the *pull* model works well for parameter sweep tasks, such as the physics simulations conducted during DC04, it remains to be seen if individual analysis tasks, which are more chaotic by nature, and require good response time guarantees, will operate effectively. A new class of Optimiser is planned which will allocate time-critical jobs to high priority global queues in order that they be run in a timely fashion.

Expanded use of the XMPP instant messaging framework should allow both Jobs and Agents to expose a Service interface, via the XMPP IQ mechanisms. This has great promise for user interactivity, and real-time monitoring and control of Agents and Jobs.

Furthermore, with this Service interface to Agents, a peer-to-peer network of directly interacting Agents is envisioned. This would reduce, and possibly even eventually eliminate, the reliance on the Central Services, as Agents could dynamically load-balance by taking extra jobs from overloaded sites.

DIRAC will follow with interest the anticipated release of Web Service-Resource Framework (WSRF), the successor to OGSF, with the hopes that the ease of use, language neutrality, tool availability, and complexity level will make it suitable for future versions of DIRAC. Increased integration with LCG and more detailed performance analysis of the DIRAC architecture is planned for the coming year. Finally, it is hoped that computational grids can benefit from the maturing technologies in peer-to-peer file sharing, instant messaging, and global computing which benefit from economies of scale. We believe a convergence and integration of these areas will provide solutions to many outstanding issues in grid computing.

Acknowledgements

We gratefully acknowledge the involvement of the LHCb Collaboration Data Management Group, and the managers of the LHCb production sites. Help from the LCG Experiment Support Group, particularly Roberto Santinelli and Flavia Donno, was invaluable for integrating DIRAC with LCG.

References

- [1] D. Adams, P. Charpentier, U. Egede, K. Harrison, R. Jones, J. Martyniak, P. Mato, J. Moscicki, A. Soroko, and C. Tan. The ganga user interface for physics analysis on distributed resources. In *Computing in High Energy Physics (CHEP 04)*, September 2004.
- [2] AliEN. <http://www.alien.cern.ch>.
- [3] BOINC. <http://boinc.berkeley.edu>.
- [4] CERN. The LHC Computing Grid Project. <http://lcg.web.cern.ch/LCG/>.
- [5] Condor. <http://www.cs.wisc.edu/condor/>.
- [6] A. W. Cooke et al. Relational Grid Monitoring Architecture (R-GMA). 2003.
- [7] distributed.net. <http://www.distributed.net>.
- [8] A. T. Doyle, S. L. Lloyd, and A. McNab. Gridsite, gacl and slashgrid: Giving grid security to web and file applications. In *Proceedings of UK e-Science All Hands Conference 2002*, Sept 2002.
- [9] D. G. Feitelson and A. M. Weil. Utilization and Predictability in Scheduling the IBM SP2 with Backfilling. In *12th International Parallel Processing Symposium*, pages 542–546, 1998.
- [10] S. Fitzgerald. Grid Information Services for Distributed Resource Sharing. In *Proceedings of the 10th IEEE International Symposium on High Performance Distributed Computing (HPDC-10'01)*, page 181. IEEE Computer Society, 2001.
- [11] Globus and IBM. <http://www.globus.org/wsrp/#announcement>, January 2004.
- [12] W. Hoschek. The Web Service Discovery Architecture. In *Proceedings of the 2002 ACM/IEEE conference on Supercomputing*, pages 1–15. IEEE Computer Society Press, 2002.
- [13] IETF. Extensible Messaging and Presence Protocol. <http://www.ietf.org/html.charters/xmpp-charter.html/>.
- [14] T. Kosar and M. Livny. Stork: Making data placement a first class citizen in the grid. In *24th IEEE International Conference on Distributed Computing Systems (ICDCS2004)*, Tokyo, Japan, March 2004.
- [15] P. Kunszt, E. Laure, H. Stockinger, and K. Stockinger. Advanced Replica Management with Reptor. 5th International Conference on Parallel Processing and Applied Mathematics, Sept 2003.
- [16] E. Laure. EGEE Middleware Architecture. In *EDMS 476451*. CERN, June 2004.
- [17] LHC. Architectural Roadmap Towards Distributed Analysis - Final Report. Technical report, CERN, November 2003.
- [18] M. Livny, J. Basney, R. Raman, and T. Tannenbaum. Mechanisms for High Throughput Computing, 1997.
- [19] N. Neufeld. The LHCb Online System. *Nuclear Physics Proceedings Supplement*, 120:105–108, 2003.
- [20] G. Pape. runit Service Supervision Toolkit. <http://smarden.org/runit/>.
- [21] R. Raman, M. Livny, and M. H. Solomon. Match-making: Distributed resource management for high throughput computing. In *HPDC*, pages 140–, 1998.
- [22] G. Rixon. <http://wiki.astrogrid.org/bin/view/Astrogrid/GlobusToolkit3Problems>, December 2003.
- [23] SETI@Home. <http://setiathome.ssl.berkeley.edu/>.
- [24] B. Tierney, R. Aydt, D. Gunter, W. Smith, V. Taylor, R. Wolsky, and M. Swany. A Grid Monitoring Architecture, Jan 2002.