

Transnational Access to Mass Storage Capacity for Computational QCD

Hinnerk Stüben* and Stefan Wollny

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Takustr. 7, 14195 Berlin, Germany

E-mail: stueben@zib.de, wollny@zib.de

Transnational Access is one part of the activities of the Hadron Physics Integrated Infrastructure Initiative (I3HP) which is funded by the European Commission in the Sixth Framework Programme. Within I3HP ZIB offers access to its mass storage system. European lattice physics collaborations can apply for mass storage capacity in order to store and share their configurations or other data (see [3]). ZIB works together with DESY (Hamburg and Zeuthen) and NIC (Jülich) on implementing a German Lattice Datagrid that conforms to the International Lattice DataGrid (ILDG). In this presentation we give an update of the implementation status and user tools.

XXIIIrd International Symposium on Lattice Field Theory

25–30 July 2005

Trinity College, Dublin, Ireland

*Speaker.

1. I3 HadronPhysics

The *HadronPhysics Integrated Infrastructure Initiative (I3HP)* is a project that originates from a joint initiative of over 2000 experimental, theoretical and computational physicists working in the field of hadron physics [1]. I3HP is funded by the European Commission in the Sixth Framework Programme. The project is structured into nine *Transnational Access Activities*, seven *Networking Activities*, and twelve *Joint Research Activities*. There are three activities that are related to lattice QCD: the Networking Activity *Computational Hadron Physics*, Transnational Access to supercomputer resources at NIC (Jülich) [2], and Transnational Access to mass storage capacity at ZIB (Berlin) [3]. NIC is one of three national German supercomputer centres, ZIB runs the supercomputer centre of the federal state of Berlin.

2. Transnational Access

The idea of Transnational Access is to give foreign researchers access to important, typically experimental facilities. It gives experimentalists the opportunity to carry out interesting experiments at facilities that they usually cannot use. The access activities related to computational QCD have a similar intention. They provide access to computational facilities. While experimentalists have to travel to the corresponding laboratories, computational facilities can be accessed via the internet.

Researchers who want to use a facility have to write a scientific proposal that is being peer reviewed. How to apply for using the mass storage system at ZIB is explained on our web page [3]. It is not necessary to write an application for just downloading configurations. For downloading a certificate is required (see section 5) and access must be permitted by the collaboration who has generated the configurations.

3. Lattice datagrids

In large scale lattice QCD projects one typically stores gauge field configurations and propagators. Gauge fields are stored permanently while propagators are stored for a limited period of time. Propagators require much more storage space than configurations. Hence they are kept at the site hosting the computer on which they were calculated. On the other hand, QCD gauge field configurations are smaller and, in the case of dynamical fermions, much more expensive to generate. This has led to the idea of sharing configurations in order to fully exploit them.

The *International Lattice DataGrid (ILDG)* [4] was started to make QCD gauge field configurations available at an international level. So far, ILDG infrastructures are being built up in Japan, UK, USA and Germany. The German ILDG Grid is called *LatFor DataGrid (LDG)* [5]. ILDG develops standards for data formats and a common middleware. The definition of metadata and a binary file format are completed (see [4]). The LatFor DataGrid conforms to the ILDG standards.

For this Transnational Access Activity we have decided to integrate the storage system at ZIB into the broader LDG and ILDG activities. Hence the LatFor DataGrid became a joint effort of DESY (Hamburg and Zeuthen), NIC (Jülich) and ZIB (Berlin).

4. Components of LDG

The main hardware components are *storage elements (SE)* that have *hierarchical mass storages systems (HSM)* attached. Each site (in Berlin, Hamburg, Jülich, and Zeuthen) operates such a storage element. The hierarchical mass storages systems are large tape libraries that work with tape robots. The total capacity of the HSM system at ZIB is about 1.2 PetaByte. Data stored on the system is very safe because there always exist two copies on different tapes.

The storage elements are small servers that run the dCache software [6] which was developed by DESY and Fermilab for storing huge amounts of data distributed among heterogeneous server nodes. From a user's perspective the distributed storage element servers provide a single virtual filesystem tree. Data may reside in the server's disk cache or might be migrated to tape. The dCache software performs data exchanges to and from the attached tape libraries automatically and invisible to the user.

Besides the storage back-ends there are user interfaces at the front-end. In order to set up a Grid infrastructure, there are several software components (middleware) needed in addition. These components are a *virtual organisation*, *Grid information services*, a *file catalogue*, an *access control service*, and a *metadata catalogue*. At the middleware level LCG-2 software is used supplemented by developments of DESY. LCG is the *Large Hadron Collider (LHC) Computing Grid* [7].

A virtual organisation (VO) is an organisational unit in a Grid infrastructure. The VO representing the LatFor DataGrid is called *ildg*. Grid information services handle e.g. authentication. The file catalogue maps logical filenames to physical locations and manages replica of files. The access control service (ACS) handles access permissions. (In LDG not all data is necessarily public. The ACS allows to store public and private data in the same environment.) The metadata catalogue holds the metadata and makes it possible to list metadata and perform search operations.

5. Important concepts

Three important concepts are (1) authentication via certificates, (2) logical filenames and physical locations of files, and (3) data formats.

In a Grid context users are authenticated by presenting a certificate. The mechanism is similar to public key authentication of the secure shell. Like in the secure shell, users generate a pair of keys. In both cases the private key is (or should be) protected by a passphrase. In the secure shell authentication is enabled by putting the public key onto the target computer. In a Grid the public key is signed by a *certificate authority (CA)*. In order to obtain the signature one typically has to apply for it in written form and send a copy of one's identity card or passport. By this means the user becomes known to the Grid. CAs for this Grid are the LCG Regional Centres (see [7]) which provide an international infrastructure.

When retrieving data (gauge field configurations) from the Grid the user usually specifies a *logical filename (LFN)*. The Grid middleware translates the LFN into a physical location. The LFN has to be unique within a virtual organisation. By convention, this is ensured by beginning any LFN with the name of a collaboration, eg:

```
lfn:qcdfs/b5p40kp13610-24x48/bqcd.561.1.1.00125.tar
```

Replicas of files can exist in the disk caches of several storage elements. The middleware is supposed to find the best copy. In general, the user does not have to take care on which storage element a file can be found, although the physical location can be used as well, eg:

```
srm://dcache.zib.de/pnfs/zib.de/data/ildg/qcdsf/\
b5p40kp13610-24x48/bqcd.561.1.1.00125.tar
```

However, on uploading files it is recommended to specify a physical location. One then has complete control on where data is stored. Otherwise the system would pick a storage element.

Data formats were defined by ILDG. There are conventions for formats of metadata [8] and binary data [9]. On uploading binary data correct metadata have to be supplied. In future all configurations should be stored in the format described in [9].

6. User tools

The authors have written a set of easy to use command line tools. The motivation was to simplify the corresponding LCG commands by using natural defaults and combining sequences of commands and to prevent erroneous use of LCG commands which could lead to inconsistencies in the metadata and file catalogue. The commands are:

<code>lget</code>	get a local copy of a configuration (or its metadata)
<code>lput</code>	put a configuration on the Grid
<code>lls</code>	list all logical file names of an ensemble
<code>linit</code>	initialise an ensemble in the metadata catalogue
<code>lupdate</code>	update metadata
<code>lvalidate</code>	check conformance of metadata to QCDml

The tools are called *lattice tools* (*ltools*) and can be downloaded from [10].

7. Usage

At the time of writing the system is mainly used by the QCDSF collaboration for storing configurations from full simulations of $O(a)$ improved QCD with $N_f = 2$ Wilson fermions. Soon the configurations of the SESAM/T χ L/GRAL projects (dynamical $N_f = 2$ standard Wilson fermions) will be stored. The χ LF collaboration plans to store their configurations whose generation is about to start (dynamical $N_f = 2$ Wilson twisted mass fermions).

8. Summary

In this Transnational Access Activity a hard- and software infrastructure was set up that is tailored to storing configurations from simulations of QCD and that is well integrated into the International Lattice DataGrid activities of the Computational Hadron Physics community. Some collaborations have started using it. The system is open to more users.

9. Acknowledgement

We acknowledge the support from the European Community-Research Infrastructure Activity under FP6 “Structuring the European Research Area” programme (HadronPhysics, contract number RII3-CT-2004-506078).

References

- [1] <http://www.infn.it/eu/i3hp/>
- [2] <http://www.fz-juelich.de/nic/i3hp-nic-ta/>
- [3] <http://www.zib.de/i3hp/>
- [4] <http://www.lqcd.org/ildg/>
- [5] <http://www-zeuthen.desy.de/latfor/ldg/>
- [6] <http://www.dcache.org/>
- [7] <http://lcg.web.cern.ch/lcg/>
- [8] http://www.ph.ed.ac.uk/ukqcd/community/the_grid/QCDm1.1/
- [9] <http://www-zeuthen.desy.de/~pleiter/ildg/#filefmt>
- [10] <http://www.zib.de/i3hp/ltools/>