USING GRID TECHNOLOGIES FOR LATTICE QCD*

M. Ernst, A. Gellrich (DESY, 22607 Hamburg, Germany)
A. Haupt, D. Melkumyan, D. Pleiter†, P. Wegner (DESY, 15738 Zeuthen, Germany)
K. Jansen (NIC/DESY, 15738 Zeuthen, Germany)
O. Büchner, Th. Lippert, B. Orth (NIC/ZAM, 52425 Jülich, Germany)
H. Stüben, S. Wollny (ZIB, 14195 Berlin, Germany)

Abstract

We present efforts to put an international datagrid infrastructure into operation to store and share data generated by compute intensive simulations of Quantum Chromodynamics (QCD) formulated on a lattice. This international datagrid is planned to be a grid-of-grids. In this contribution we will concentrate on the implementation of one of the regional grids used by research groups in Europe, the LatFor DataGrid (LDG). We report on our metadata catalogue and discuss our experience with the used technologies for XML-Java binding and Java object contents persistence. Finally, we will give an overview on the LDG infrastructure and middleware, which is based on LCG-2 middleware components.

INTRODUCTION

Numerical simulations of lattice Quantum Chromodynamics (QCD) require a huge amount of computational resources. To carry-out such calculations highly scalable capability computers (like apeNEXT [1], BlueGene/L [2], or QCDOC [3]) providing TFlops of compute power are required.

Grid technologies can help to improve exploitation of these precious resources, e.g. by sharing the produced data on a global level. The International Lattice DataGrid (ILDG) [4, 5] has been founded in 2001 to define the required standards needed for a grid infrastructure to be used for research on lattice QCD. This infrastructure should provide the means for longterm storage and global sharing of so-called gauge configurations. Groups from Australia, France, Germany, Italy, Japan, UK and the US joined this effort. Two working groups on Metadata and Middleware have been established to define the standards.

As the development of grid technologies is currently mainly driven by large experiments, it is important to notice that the research community of lattice QCD is typically organised in small groups. Many of these groups work at universities and do not have the resources for setting-up a complex local infrastructure and have to rely on the locally available (and often heterogeneous) computing facilities.

REQUIREMENTS

Sharing simulation results requires semantic access to worldwide distributed data. This can only be achieved if an unique description of the metadata can be enforced. Today's standard solution for such a task is the use of XML for metadata documents which have to conform to an XML schema. Defining such a schema turned out to require significant efforts, mainly because discretisation of QCD on a 4-dimensional space-time lattice is not unique. As a result many different actions are used in lattice QCD simulations and a flexible and, in particular, an extensible markup schema had to be implemented.

Data sharing furthermore requires standards on the binary file format. The adoption of such a standard has been achieved and the community is currently in the process of adopting this standard in their software and data production environment. Finally, a definition of common middleware interfaces is required, which is however still under discussion.

In the following we will concentrate on the implementation of the LatFor[1] DataGrid (LDG) [6, 7], a LCG-2 based infrastructure currently used by lattice QCD research groups in France, Germany, Italy and Spain.

METADATA CATALOGUE

The metadata catalogue for the LatFor DataGrid (LDG) was supposed to fulfil the following requirements:

- Its functionality should allow to load, store and query XML documents which conform to an extensible schema.
- A web-service conforming to ILDG specifications should be used as front-end.
- Primarily for easier maintenance we opted for a standard relational database as back-end.
- The used technologies should be general enough to make the catalogue potentially useful also for other research communities.

For the implementation we adopted the strategy of mapping XML documents on Java object contents and Java object contents on relational tables.

* Talk given by D. Pleiter at CHEP06, Mumbai (India).
† dirk.pleiter@desy.de

[1] Lattice Forum (LatFor) is a joint initiative of German lattice physicists to coordinate the physics program and share resources.

There exist several software solutions for XML-Java binding. For us two selection criteria were most relevant: performance and coverage of the XML schema specification. We eventually chose JAXB, which has become an open source project [8]. Using an XML binder we can use the ILDG metadata schemata to generate a set of annotated Java classes.[2] These classes are the basis for software which transforms XML documents into Java object contents and vice versa.

For Java object contents persistence we chose for another open source project, Hibernate [9]. This software provides the required functionality of automated object-relational mapping as well as loading and storing object contents. Hibernate supports various SQL databases, we actually use MySQL.

The front-end interface to the metadata catalogue consists of a few functions to query and download metadata documents, which are in the process of being standardised by ILDG. Furthermore, a large set of functions have been implemented for uploading documents and manipulating tables used for access control. For the web-service implementation Apache Tomcat and Axis are used. Conforming to ILDG specifications read operations are fully open, i.e. no authentication and authorisation mechanisms are required. To secure write access we mandated use of the Globus Security Infrastructure (GSI). So far, we used COGkit to implement the required authentication and authorisation mechanisms, but we are in the process of re-implementing this part using gLite trustmanager.

XPath expressions have to be used to query the metadata catalogue. The metadata catalogue returns a list of identifiers for all matching documents. An XML document matches when applying the XPath expression results in a non-empty set of nodes. A returned document identifier can by used to download the document itself. As typically queries are restricted to simple elements, we implemented an algorithm for mapping those queries directly onto SQL queries. For instance, the XPath query expression `/gaugeConfiguration/markovStep[markovChainURI = `*uri*`]/dataLFN` is mapped to the SQL query expression `select dataLFN from MarkovStepType where markovChainURI = "`*uri*`"`.

With this optimisation we find a very good performance for typical queries even if the catalogue contains $> 10,000$ documents. However, performance is an issue as materialisation of XML documents is rather expensive. On our currently used hardware[3] we measured 0.02 and 0.04 seconds for loading and storing an XML document, respectively. As we expect the catalogue to store $O(100,000)$ documents, the time for dumping or restoring all XML documents will be of $O(0.5)$ and $O(1)$ hour, respectively.

The LDG metadata catalogue is maintained at DESY in Zeuthen and since beginning of 2006 regularly used by production jobs, which proves its stability.

[2] We actually use a slightly simplified version of the ILDG schemata.
[3] A 2.8 GHz Dual-Xeon server with 2 GB of main memory.

MIDDLEWARE

The middleware of the LatFor DataGrid (LDG) is based on the LCG-2 software. The infrastructure relies on a number of central services maintained at DESY in Hamburg. This includes a VOMS service for the VO "ildg", a Berkeley Database Information Index (BDII) service and a file catalogue.[4]

Until now four dCache-based [10] Storage Elements (SE) have been made available to the VO "ildg" at DESY, ZAM Jülich and ZIB Berlin. These sites do provide the biggest share of compute resources currently available for lattice QCD calculations in Germany and are expected to be able to contribute the required storage for $O(100.000)$ gauge configurations, which corresponds to $O(10-100)$ TBytes of data.

On the client site only a few LCG software components for authentication and data management are required. To simplify access to the metadata catalogue and the storage elements a set of user and project management tools have been implemented, which hide details of the underlying grid infrastructure (for further details, see [11]). The LCG-based User Interface (UI) has been compiled for a number of Linux flavours, including Scientific Linux, SUSE Linux, SUSE Linux Enterprise Server and Debian. For the other components Java and Perl has been used to keep the software portable. An RPM-based installation mechanism is provided for the UI and the other client software components. Also CA certificates and Certificate Revocation Lists (CRL) are regularly distributed as RPMs.

CONCLUSIONS

The use of grid technologies becomes increasingly relevant for research on lattice QCD. In the near future this will mainly focus on setting-up datagrid infrastructures for sharing data generated on high performance capability computers.

When implementing a metadata catalogue for the LatFor DataGrid we explored the use of XML-Java binding and Java object contents persistence technologies. Based on available software solutions we implemented a metadata catalogue which meets our requirements. It supports an extensible XML schema, although not all XML schema specifications are fully supported, yet. The front-end is flexible and could easily be adapted to ILDG interface specifications. By using an SQL server as back-end we can rely on a standard, well-supported technology. Furthermore, this allowed us to improve performance of XPath queries for simple elements, which are mapped to fast SQL queries. However, we have to face performance issues for operations which require materialisation of XML documents.

LCG and LCG-compliant software has been used to setup an infrastructure for groups performing research on lattice QCD in France, Germany, Italy and Spain, which is

[4] Currently we are still using the EDG file catalogue, but plan to move to the LCG File Catalogue (LFC).

also accessed by users in Cyprus and UK. Storage elements have been installed at several HPC sites and the client software has been made available for different flavours of Linux. The LatFor DataGrid may therefore be considered as an example for a successful application of the LCG software stack beyond the community of experimental high energy physics.

ACKNOWLEDGEMENTS

This work has partially been funded by the HEP-Grid initiative [12] and the European Community-Research Infrastructure Activity under FP6 "Structuring the European Research Area" programme (HadronPhysics, contract number RII3-CT-2004-506078).

REFERENCES

[1] F. Belletti *et al.* [APE collaboration], "apeNEXT: Experiences from Initial Operation," these proceedings.

[2] IBM Journal of Research and Development, Vol. 49, No. 2/3, 2005.

[3] P. A. Boyle *et al.*, "QCDOC: Project status and first results," J. Phys. Conf. Ser. **16** (2005) 129.

[4] `http://www.qcd.org/ildg`

[5] A. Ukawa, "Status of International Lattice Data Grid: An overview," Nucl. Phys. Proc. Suppl. **140** (2005) 207 [arXiv:hep-lat/0409084].

[6] `http://www-zeuthen.desy.de/latfor/ldg`

[7] O. Büchner *et al.*, "Datagrids for lattice QCD," Nucl. Inst. and Methods in Physics Research, Vol. 559, Issue 1 (2006) 57-61.

[8] `https://jaxb.dev.java.net`

[9] `http://www.hibernate.org`

[10] `http://www.dcache.org`

[11] H. Stüben and S. Wollny, "Using the mass storage system at ZIB within I3HP," Nucl. Phys. Proc. Suppl. **153** (2006) 300.

[12] P. Malzacher, "The German HEP-Grid initiative," these proceedings.