Datagrids for Lattice QCD

O. Büchner[a], M. Ernst[b], K. Jansen[c], Th. Lippert[a], D. Melkumyan[d], B. Orth[a], D. Pleiter[c], H. Stüben[e], P. Wegner[d], S. Wollny[e]

[a]Zentralinstitut für Angewandte Mathematik ZAM, 52425 Jülich, Germany

[b]Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

[c]John von Neumann-Institut für Computing NIC/DESY, 15738 Zeuthen, Germany

[d]Deutsches Elektronen-Synchrotron DESY, 15738 Zeuthen, Germany

[e]Konrad-Zuse-Institut für Informationstechnik ZIB, 14195 Berlin, Germany

As the need for computing resources to carry out numerical simulations of QCD formulated on a lattice has increased significantly, efficient use of the generated data has become a major concern. To improve on this, groups plan to share their configurations on a worldwide level within the International Lattice DataGrid (ILDG). Doing so requires standardized description of the configurations, standards on binary file formats and common middleware interfaces. We describe the requirements and problems and discuss solutions. Furthermore, an overview is given on the implementation of the LatFor DataGrid [1], a France/German/Italian grid which will be one of the regional grids within the ILDG grid-of-grids concept.

1. INTRODUCTION

Numerical simulations of theories describing the interaction of elementary particles have become an important approach for understanding the fundamental forces in nature. In particular, investigation of the strong interaction requires (due to its non-perturbative nature) computer simulations. The theory which is supposed to describe the strong interactions is Quantum Chromodynamics (QCD). In order to perform numerical calculations this theory is formulated on a discrete space-time lattice (lattice QCD). Results from such computations are key input for the interpretation of data obtained from experiments performed at large and costly, existing and planned particle accelerators, like HERA, CEBAF, PEP II, Tevatron or LHC.

Monte-Carlo techniques are used to generate Markov chains which consist of configurations of the fields mediating the strong interactions. In state-of-the-art simulations tens of Teraflops computing power are needed to generate such *ensembles* of so-called gauge field *configurations.*[1] Due to the large amount of high-performance computing resources being used it is of obvious importance to share these configurations among a large number of scientists by building an International Lattice DataGrid (ILDG). The main goal of ILDG is to provide means for long-term storage and global sharing of data being produced by lattice QCD. It eventually aims for providing semantic access to worldwide distributed data.

The ILDG will be implemented as a grid-of-grids with standardized interfaces to query metadata catalogues and to exchange data, i.e. configurations. ILDG currently focuses on the exchange of gauge field configurations.[2] In order to exchange such configurations it is necessary to address three kinds of tasks. Firstly, common standards were needed for metadata documents which describe these configurations. Furthermore, standards for conversion between different (binary) file formats used to store configurations had to be defined. Finally, the necessary common middleware standards and interfaces to achieve interop-

[1]See, e.g., [2] for more information on the computational requirements of lattice QCD simulations.

[2]For the future also storage of other, more storage intensive data objects will be considered.

erability among different Grid middleware suites are being developed [3].

In this paper we concentrate on the definition of the metadata and the implementation of one of the ILDG regional grids, which will we used by groups active in research on lattice QCD in France, Germany and Italy.

2. METADATA

The metadata which describes the data files containing gauge field configurations should provide a variety of information. On the one hand it should describe the data itself. Most relevant is obviously the information about the *physics* parameters or to be more specific: information on the used lattice action as well as the simulation parameters. *Data management* information providing, e.g., checksums allow the user of the data to check its integrity.

On the other hand the metadata should also provide information on the provenance of the data. This allows the user of the data to assess its quality as, e.g., the results of simulations may depend on the used *algorithms* and in particular the chosen parameters. A typical example for the latter is the convergence criterium of iterative solvers which are used in all simulations of lattice QCD. Information on the *source code* and the used *machines* are important for the provider of the data for back-tracking possible errors. While the probability for this to become necessary is small due to the large efforts of all people being involved in simulations of lattice QCD, the possibility of data being produced using programs containing bugs or machines generating wrong data can nevertheless not be excluded.

The relevance of the physics information has already been stressed. When designing the metadata documents this part requires particular care. Actions used for simulating lattice QCD can often be parameterized in different ways. Queries for particular physics parameters should however return all references to data objects which were generated using effectively the same parameters. The data description therefore has to be *unique*. It is impossible to anticipate which actions will be used in future simulations. As long-term data procurement is one of the primary goals of ILDG the metadata documents have to be organized in an *extensible* way. This means that the standards for the metadata documents will be allowed to change such that older documents continue to conform to the modified standards without being modified themself. Finally, *simplicity* and *generality* have been identified as requirements for the design of metadata documents. The design should for instance be general enough to allow the description of data objects other than gauge configurations, like fermion propagators, particle correlations functions or operator eigenvalue spectra.

Todays best choice for implementing the outlined requirements for metadata design is XML. XML is extensible by design and, since XML is verbose, is making metadata documents both human readable and easy to parse by computers. Finally, standards on the structure and contents of XML documents can be enforced by using XML schemata.

As an example let us consider the most simple action, the so-called Wilson gauge action which is defined by the following equation:

$$S_G = \frac{\beta}{18V} \sum_{x,\mu<\nu} \mathrm{ReTr}(1 - U^{\Box}_{x,\mu\nu}),$$

where β is the gauge coupling, V the number of sites on the lattice and $U^{\Box}_{x,\mu\nu}$ the product of gauge fields along a plaquette. The following example shows a small piece of a metadata document for a gauge configuration that was generated using this action:

```
<plaquetteGluonAction>
    ...
    <couplings>
        <elem>
            <beta>5.2</beta>
        </elem>
    </couplings>
</plaquetteGluonAction>
```

Today a large number of different actions are being used, some of them with a larger number of parameters, resulting in more complicated data structures. Since configurations are grouped into ensembles which share important properties,

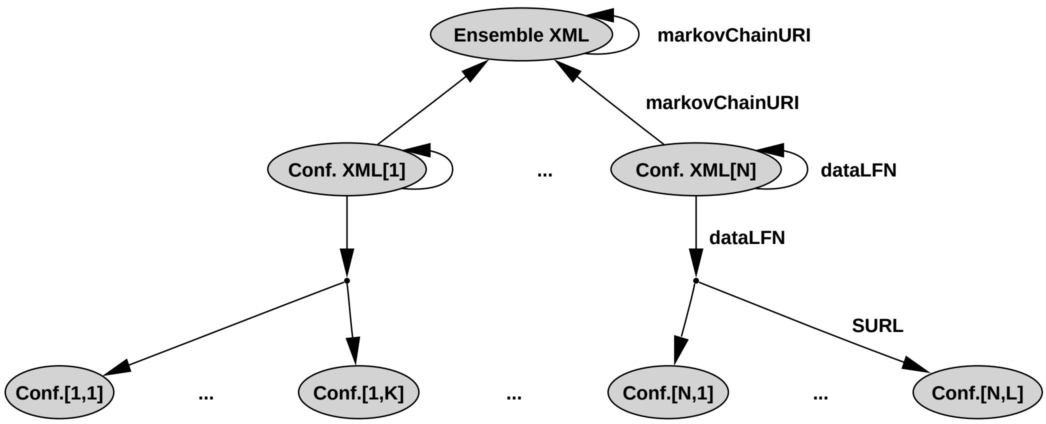

Figure 1. Linking metadata documents and data files with data files being replicated within the Grid.

splitting the metadata information into two documents allows to avoid replication of these data structures. Two schemata were therefore developed by the ILDG metadata working groups [4] and have been adopted by the research community [5].

The lattice datagrids will have to store three kind of objects: the ensemble and configuration XML documents as well as the binary data files. The first two will be stored in a Medata Catalogue (MDC), for the latter various Storage Elements (SE) providing access to mass storage systems will be used. The different objects are linked by the so-called `ensembleURI`, an unique identifier for each ensemble, and the `dataLFN`, a Logical File Name (LFN) which is assigned to each binary data file uploaded to the datagrid. Note that the mapping from configuration XML to ensemble XML document is a many-to-one relation, while for each configuration XML document a unique LFN is assigned. Within the grid several replicas of the binary data file might exist. A file catalogue takes care of mapping the LFN to a physical file locator, i.e. a SURL. The links between the various data objects is schematically shown in Fig. 1.

3. MIDDLEWARE IMPLEMENTATION

In this section we will describe the implementation of the Lattice Forum (LatFor, [6]) DataGrid which will become part of the ILDG grid-of-grids concept. The aim of this regional grid is to deploy a datagrid infrastructure for research on lattice QCD in France, Germany and Italy. This research is done by groups at large national laboratories as well as by smaller groups located at universities with quite different local computing facilities being available.

For accessing the datagrid User Interfaces (UI) need to be installed at all sites. Furthermore, Storage Elements (SE) are planned at those sites where significant compute resources have been installed. Mass storage will be provided for permanent storage and medium sized storage for keeping temporary replicas. Finally, a set of central information services are needed. A schematic view on the LatFor DataGrid architecture is shown in Fig. 2.

The infrastructure installed so far is based on the LCG-2 compatible grid infrastructure deployed at DESY [7]. The software for Linux-based UIs is provided as a tar-ball which consists of the Globus-2 tools for authentication and the LCG data management client utilities.

Until this conference four storage elements have been installed at DESY (Hamburg), DESY (Zeuthen), ZAM (Jülich) and ZIB (Berlin). The last three sites provide significant compute resources for lattice QCD simulations on massively parallel computers. All storage elements are

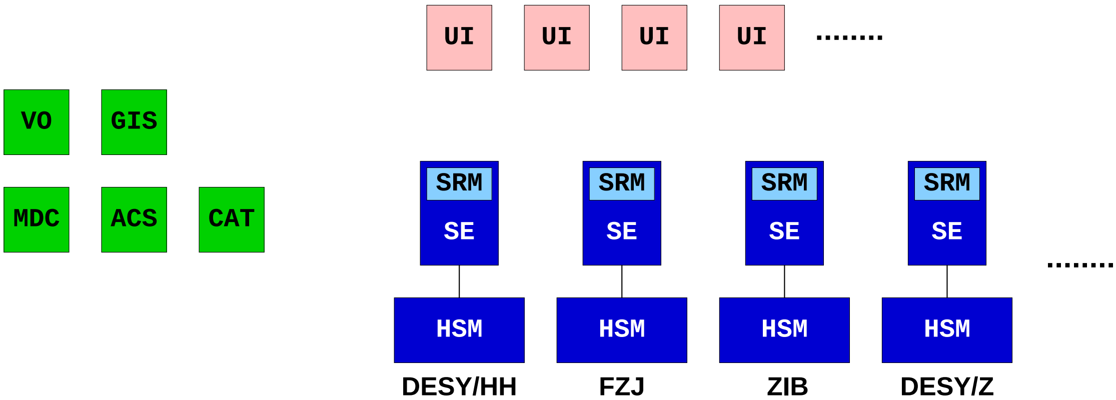

Figure 2. Schematic overview on the LatFor DataGrid architecture. It consists of the following components: User Interfaces (top right), Storage Elements (bottom right) and central services (left). Details are described in the text.

based on *dCache*, a software jointly developed by DESY and FNAL [8]. dCache provides access to the local mass storage back-ends via a SRM interface.

The central information services consist of the usual grid information services (GIS), a virtual organization (VO) *"ildg"*, a file catalogue (CAT), the metadata catalogue (MDC) and an access control service (ACS). Currently the EDG file catalogue is used but it will be replaced by the LCG file catalogue LFC, soon. It is expected that this catalogue will fulfill all needs of the LatFor DataGrid.

The owner of data sets are supposed to have full control on the access rights. These rights can be set for each ensemble and will be enforced for all data objects which belong to this particular ensemble. Access rights are retrieved and modified via the metadata catalogue, which will be described in more detail in the next section. The access control service (ACS) is responsible for retrieving the access control information from the metadata catalogue and performing the required ACL set operations.

4. METADATA CATALOGUE

For the metadata catalogue of the LatFor DataGrid the following requirements have been identified:

- For being ILDG conform it should be able to store extensible XML documents.
- The front-end has to be implemented as a web service which provides a set of services defined by ILDG. Through these services it will be possible to query all ILDG metadata catalogues in the same way.
- As a back-end a relational database is desirable as these databases are still more mature and more commonly used.
- The core of the metadata catalogue should also be usable for other research communities adopting a similar approach.

For the metadata catalogue implementation several software tools for binding Java to XML schemata have been evaluated. *JAXB* from Sun was finally found to provide best coverage of the XML schema specification. This technology allows to transform a given XML schema to a set of Java classes and annotate the generate classes with special tags called xdoclets. These classes can then be used in an application that transforms XML documents into Java content objects and vice versa. As a drawback of this approach it should be noted that not all features of the XML schema specification are supported yet. In the prototype implementation of the LatFor metadata catalogue therefore modified schemata have

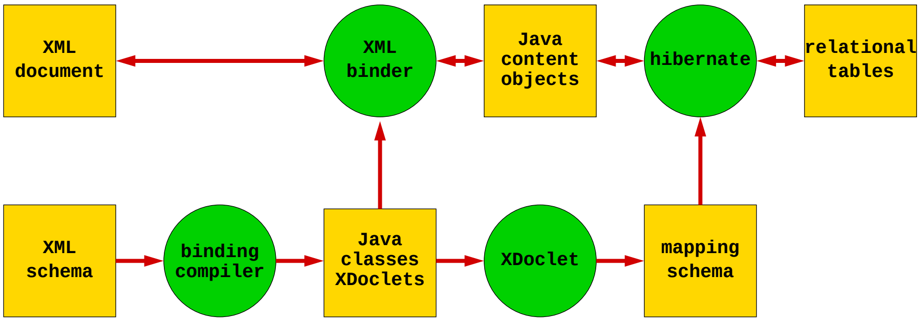

Figure 3. Transforming XML documents conforming to a given schema into relational tables.

been used. They were chosen to be equivalent to the original schemata in the sense that any document which conforms to one of the original schemata also conforms to the modified schema.

For interfacing with the database back-end the open source software *hibernate* is used. This allows to store Java objects in relational databases. It provides an object oriented view of existing relational data using an object-relational mapping schema. This schema can be automatically generated using XDoclet taking the already generated xdoclets on input. Each of the Java classes generated by the XML binder is mapped onto relational tables. Various SQL databases can be used, the LatFor prototype metadata catalogue has however only been tested using MySQL.

The full transformation of a XML document into relational tables (and vice versa) is schematically shown in Fig. 3.

The metadata catalogue is queried using XPath expressions. It returns a pointer to all documents for which matching with this expression gave a non-empty result. For these kind of queries the chosen approach has a significant performance advantage as typically only simple elements are queried. The catalogue translates these kind of XPath expressions directly into SQL queries, which are much faster to execute.

Access to the metadata catalogue is realized by web-services implemented using Tomcat and Axis. For all services which result in a modification of the catalogue contents authentication is mandatory. Authentication is done via the Globus Security Infrastructure (GSI) for which one of the Commodity Grid (CoG) Kits is used.

5. CONCLUSIONS

We presented an international initiative for setting-up a Grid infrastructure which provides semantic access to worldwide distributed data obtained in simulations of lattice QCD. For research groups in France, Germany and Italy a pilot Grid infrastructure based on LCG-2 middleware components and dCache has been put into operation. This regional Grid will become part of the International Lattice DataGrid, which is planned to be a grid-of-grids. For the implementation of a metadata catalogue new software technologies for XML-Java binding and object-relational mapping have been explored and successfully used for a prototype implementation.

REFERENCES

1. `http://www-zeuthen.desy.de/latfor/ldg`
2. H. Simma et al., these proceedings.
3. B. Joo and W. Watson, Nucl. Phys. Proc. Suppl. **140** (2005) 209 [arXiv:hep-lat/0409165].
4. A. Ukawa, Nucl. Phys. Proc. Suppl. **140** (2005) 207 [arXiv:hep-lat/0409084]; see also: `http://www.lqcd.org/ildg`.
5. C. M. Maynard and D. Pleiter, Nucl. Phys. Proc. Suppl. **140** (2005) 213 [arXiv:hep-lat/0409055].
6. `http://www-zeuthen.desy.de/latfor`
7. A. Gellrich et al., these proceedings.
8. `http://www.dcache.org`