Datagrids for Lattice QCD

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

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

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

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

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

[e]Konrad Zuse Institut 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 discus solutions. Furthermore, an overview is given on the implementation of the LatFor DataGrid [1], 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, SLAC 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-up 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. 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 converting between different (binary) file formats used to store configurations had to

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