Sharing lattices throughout the world: an ILDG status report

Carleton DeTar[*]

Physics Department, University of Utah, Salt Lake City, UT 84112, USA

E-mail: `detar@physics.utah.edu`

The international lattice data grid, a system for sharing gauge configuration files throughout the world, is operational for the most part. I give a status report, give some pointers on accessing lattice files, and highlight some of the available gauge configurations.

---

[*]Speaker.

1. Introduction

The rapid expansion of internet services in the 1990's made it easy to publish gauge configuration files for general public access. The US Gauge Connection at NERSC was established in 1997 for this purpose[1]. It was simple and effective, but it was not the ideal solution for broad international use. The development of a truly international lattice data grid (ILDG) began some years later (2002) as a result of discussions organized by the UKQCD [2]. Its primary purpose is to facilitate the advancement of our field by making it easy for scientific collaborations to share gauge configuration files with other collaborations throughout the world. To reach this goal required first the development of regional or national grids for intraregional file sharing and storage. This effort typically exploited existing grid infrastructure and tools. The ILDG was then built upon the regional grids.

As a consequence of this history, the ILDG is a consortium of autonomous regional grids [3]. Current participants are centered in Australia (CSSM), Europe (Austria, Switzerland, Germany, France, Italy)(LDG), Japan (JLDG), the United Kingdom (QCDGrid or DiGS), and the United States. Additional participants are, of course, always welcome. Cooperation involves the standardization of several components: binary file formats, metadata markup schema (QCDml), user authentication methods, and replica-catalog and file-transfer interoperability. These standards are achieved through the efforts of two international working groups, namely, the Metadata and the Middleware Working Groups. An ILDG board with international representation supervises the work of these groups. The board chair rotates on an annual basis. To foster coordination, for the past five years the ILDG has held semiannual internet conferences through VRVS (Virtual Room Videoconferencing System) with worldwide participation.

Funding for this effort has been left to countries hosting the regional grids. For the most part funding has been minimal if not completely absent. The manpower, grid software, and necessary file storage capacity have been donated and/or borrowed from other funded projects. Nonetheless, over the course of five years, the ILDG has developed into a resource that is ready for use.

This is a brief status report. Rather than dwelling on the technical details of grid operation, I will try to take the perspective of a user.

2. Access policies

Lattice files in a grid repository may have access restrictions. These are set by the contributing collaboration. Thus it is always a good idea to check with the collaboration that created the files if you find there are restrictions. The authors are often comfortable releasing recent lattices selectively for noncompeting projects, or they might be willing to collaborate if there are overlapping interests. There is a spectrum of such policies ranging from permissive to more restrictive:

- All lattices are available as soon as they are generated (some collaborations).

- Recent lattices are available after negotiation (most collaborations).

- All lattices are available approximately six months after the first publication (universal).

Of course, in a publication based on another collaboration's lattices it is proper to cite an appropriate journal reference, where possible, and to acknowledge the collaboration that produced them.

It is very encouraging to note a historical trend toward increasing openness.

3. Files

The binary lattice files adhere to a prescribed format [4] called LIME [5]. (The name is a take-off on the name for the conventional MIME standard for e-mail, in this case meaning "lattice interchange message encapsulation"). The standard gives collaborations considerable leeway in adding their own variations. Essentially, the file consists of a series of records (messages), only three of which have required content: the main one, of course, is the binary payload. A second required message specifies the precision and space-time dimensions. A third contains the unique logical file name (LFN), for example,

```
lfn://USQCD/MILC/asqtad/2_plus_1_flavor/2064f21b676m007m050/
    series_1/l2064f21b676m007m050b.1530
```

The first part of the LFN (`USQCD` in this case) is a standardized prefix identifying the collaboration, and the rest is left to the discretion of the collaboration.

The logical file name (LFN) is used by the catalogs and file transfer utilities to identify the file. The mechanism works as follows: a user requests the file by naming its LFN. The retrieval software consults a replica catalog (RC), which associates the LFN with a universal resource locator (URL). The URL locates the file on the internet.

Each binary file is accompanied by a metadata file containing the archiving history, checksums and plaquette values, etc. Each ensemble of gauge configurations also has a metadata file containing information about the gauge action and the parameters used to generate it. Both types of metadata files are XML and must conform to a standard schema, collectively called QCDml. The Metadata Working Group sets all of these standards [6]. Please see Tomoteru Yoshie's poster contribution [7].

4. Mode of access

For a detailed account of file discovery and retrieval, please see Dirk Pleiter's poster contribution [8].

4.1 Grid certificates

Before you can download lattices you must obtain an X.509 grid certificate from an appropriate certificate authority (CA). The ILDG accepts certificates from any member of the International Trust Federation [9]. Once you have a certificate, you must register it with the ILDG "virtual organization" [10].

4.2 Discovery

The obvious next step in accessing a set of lattices is to identify the ensemble(s) of interest. This is the purpose of the metadata catalog (MDC).

Each regional grid maintains a catalog of the files in its repository and publishes the MDC for those files. The regional grids then republish the contents of the world's catalogs through their web portals. This process is in various stages of development. Eventually all the regional grids will catalog the world's ensembles and be globally searchable.

A web-based global catalog is available in Australia (CSSM)[11], Europe (LDG)[12], Japan (JLDG)[13], and the US (JLab)[14]. Web-based search engines for searching the MDC's are also available in Australia and Japan. The UK provides a Java-based browser application to perform data discovery [15]. The United States currently has not populated its own MDC, but lattices in the older NERSC format continue to be available by means of a web interface through the Gauge Connection [1], and Brookhaven National Laboratory [16] has recently posted lattices from the RBC collaboration.

4.3 Downloading files

The final step involves the file transfer itself, which can proceed through various protocols directly (http or gridftp) or through the Storage Resource Manager (SRM v2) service framework. The latter allows client and server to negotiate the transfer protocol and endpoint that are actually used.

Client software is supposed to support SRM. The actual transfer protocol depends on the regional grid that dispenses the files. Australia, Japan and UK publish their data via a gridftp service. The JLDG additionally has a convenient web interface that provides the necessary download script [13]. The LDG dispenses files via SRM and provides a set of convenient tools called LTOOLS [17] to allow integrated access to their metadata, file catalogs, and storage elements. The UK uses its own set of free grid tools [15]. Finally, in the US, NERSC and BNL currently use http, but produce files only in the old NERSC format. An initial storage element is in development at FNAL, where there are now many files in ILDG format ready for SRM access and cataloging.

The middleware working group is working to alleviate this complex situation. Recently packages have been added to the LDG LTOOLS framework (which is relatively straightforward to install) that provide an SRM client (srmcp) and an initial version of an integrated client (ildg-client/ildg-get) that allows retrieval of metadata documents and files through http or SRM. This software, as well as the crucial catalogues such as the metadata and file catalogs are in the process of maturation, and the process should become better integrated and more reliable in the future.

5. What ensembles are or will soon be available?

I give a brief survey of some important and highly useful lattice ensembles now or soon to be available to the world community. As might be expected, they reflect a healthy diversity of approaches to lattice physics. I have not attempted an exhaustive list – rather I have included what I feel might be particularly interesting. Nor have I attempted to list all the important lattice parameters, such as lattice size and quark mass. The details are available in the meta data catalogs. There is such an impressive variety to choose from, surely there is something for every taste.

N_f	Action	Collaboration	a(fm)	cfgs(approx)	Availability
0	tpLW	Adelaide	0.128	400	Available
2	FLIC/tpLW	Adelaide	0.125	$400 \times$ couple	In production

Table 1: Selected ensembles available through the Australian CSSM. The abbreviation "tpLW" stands for the tadpole Lüscher-Weiß action. For configuration number, the notation $400 \times n$ implies n ensembles of approximately 400 configurations each, typically with different choices of quark masses.

N_f	Action	Collaboration	a(fm)	cfgs(approx)	Availability
2	NP-clover/Wilson plaq	QCDSF	0.07/0.11	5000×16	Negotiable
2	tmQCD/tlSym	ETMC	0.07/0.11	3500×15	Negotiable
$2+1$	stout/tpLW	QCDSF	0.08	1000×16	In progress
$2+1+1$	Stouted tmQCD/tlSym	ETMC	0.07/0.09	2500×15	In progress

Table 2: Selected ensembles available through the European LDG. Notation is the same as in previous tables. The abbreviation "tlSym" stands for the tree-level Symanzik gauge action. The abbreviation "tmQCD" stands for the twisted mass formulation.

5.1 Australia

The Adelaide group has only recently gained access to computing facilities that make it possible to begin serious unquenched calculations using the fat link irrelevant clover (FLIC) action. Production has begun as indicated in Table 1.

5.2 Europe

Europe (Table 2) has a couple of vigorous dynamical fermion efforts underway. The QCDSF collaboration has been generating some very large two-flavor ensembles based on the nonperturbative clover action and is following up with $2+1$ flavors of "stout" fermions. The ETMC has also been generating large two-flavor QCD ensembles with twisted mass fermions and plans to follow up with a full complement of flavors using a stouted version of the action.

5.3 Japan

In Japan (Table 3) the CP-PACS and JLQCD collaborations have access to very substantial computational resources, which have been devoted to two ambitious lattice generation projects using clover fermions and overlap fermions. The CP-PACS and JLQCD collaborations have recently released a significant set of full-QCD ensembles based on the Wilson-clover fermion action with the Iwasaki gauge action.

5.4 UKQCD and RBC Collaborations

Using QCDOC computers, the UKQCD collaboration has teamed up with the US RBC collaboration to generate a series of lattice ensembles based on the domain wall fermion formulation (Table 4). A large set has been released for general public access. A few of the most recent ensembles still have restricted access. The UKQCD has also been generating some full QCD lattices using the Asqtad formulation.

PoS(LATTICE 2007)009

N_f	Action	Collaboration	a(fm)	cfgs(approx)	Availability
2	Clover/Iwasaki	CP-PACS	0.20 0.15 0.10	1000×4 1000×4 800×4	Available
2	Clover/Plaquette	JLQCD	0.10	1200×5	In preparation
2+1	Clover/Iwasaki	CP-PACS+ JLQCD	0.12 0.10 0.07	800×5×2 800×5×2 600×5×2	Available (new)
2	Overlap/Iwasaki	JLQCD	0.12	O(500)×6	After 1st spectrum paper

Table 3: Selected ensembles available through the Japanese JLDG. Notation is the same as in previous tables.

N_f	Action	Collaboration	a(fm)	cfgs(approx)	Availability
2+1	DWF/Iwasaki	UKQCD/RBC	0.12	800x4	Available
2+1	DWF/DBW2	UKQCD/RBC	0.12	300-1000	Available
2	DWF/DBW2	UKQCD/RBC	0.12	1000x3	Available
2+1	Asqtad/tpLW	UKQCD	0.12	3000	Available
2+1	DWF/Iwasaki	UKQCD/RBC	0.09	??	Restricted
2+1	DWF/Iwasaki	UKQCD/RBC	0.09	?? × 2	In production

Table 4: Selected ensembles available through the UKQCD and RBC collaborations. Notation is the same as in previous tables. The question mark indicates the information was unknown at the time of this writing.

5.5 United States

In the US (Table 5) over the past few years the MILC collaboration has been generating and publishing an extensive set of full QCD ensembles based on the Asqtad formulation. The most recent additions are at a lattice spacing of 0.06 fm. Further work is underway to extend and add to the existing ensembles at 0.09 and 0.12 fm. All lattices are open to general access as soon as they are created. The LHPC collaboration has embarked on a major effort to generate dynamical clover ensembles on anisotropic lattices. The US RBC production of domain wall lattices was included in the previous section.

Acknowledgments

References

[1] The Gauge Connection: `http://qcd.nersc.gov/`.

N_f	Action	Collaboration	a(fm)	cfgs(approx)	Availability
$2+1$	Asqtad/tpLW	MILC	0.15	600 x 4	Available
$2+1$	Asqtad/tpLW	MILC	0.12	600 x 5	Available
$2+1$	Asqtad/tpLW	MILC	0.09	500×2	Available
				500×1	New, available
$2+1$	Asqtad/tpLW	MILC	0.06	500×2	In production Available
$2+1$	Aniso clover/tpLW	LHPC	0.125	500×4	Planned
$2+1$	Aniso clover/tpLW	LHPC	0.10	$?? \times 4$	Planned
$2+1$	Aniso clover/tpLW	LHPC	0.08	$?? \times 4$	Planned

Table 5: Selected ensembles available through the NERSC gauge connection or by contacting the collaborations directly. Notation is the same as in previous tables.

[2] C. T. H. Davies, A. C. Irving, R. D. Kenway and C. M. Maynard [UKQCD collaboration], Nucl. Phys. Proc. Suppl. **119** (2003) 225 [arXiv:hep-lat/0209121]. ; A. C. Irving, R. D. Kenway, C. M. Maynard and T. Yoshie, Nucl. Phys. Proc. Suppl. **129** (2004) 159 [arXiv:hep-lat/0309029].

[3] A. Ukawa, Nucl. Phys. Proc. Suppl. **140** (2005) 207 [arXiv:hep-lat/0409084]. ; K. Jansen, PoS(LAT2006)013 (2006) [arXiv:hep-lat/0609012].

[4] ILDG file format:
http://www-zeuthen.desy.de/~pleiter/ildg/ildg-file-format-1.1.pdf.

[5] LIME software and documentation: http://usqcd.jlab.org/usqcd-docs/c-lime/.

[6] QCDml markup schema and examples: http://www.ccs.tsukuba.ac.jp/ILDG/;
http://www.lqcd.org/ildg/.

[7] T. Yoshie, *Marking up lattice QCD configurations and ensembles*, these proceedings (2007).

[8] D. Pleiter, *Towards an Interoperable International Lattice Datagrid*, these proceedings. (2007).

[9] International Trust Federation: http://www.gridpma.org.

[10] Registration site for the ILDG virtual organization:
https://grid-voms.desy.de:8443/vo/ildg/vomrs/.

[11] CSSM portal: http://cssm.sasr.edu.au/ildg/.

[12] LDG portal: http://www-zeuthen.desy.de/latfor/.

[13] JLDG portal: http://www.jldg.org/lqa/index.html.

[14] JLab MDC portal: http://usqcd.jlab.org/mdc-web-client/index.jsp.

[15] DiGS software: http://forge.nesc.ac.uk/projects/qcdgrid/.

[16] Domain wall fermion lattices from the RBC collaboration: https://qcdlattices.bnl.gov/.

[17] LTOOLS software for the LDG: http://www-zeuthen.desy.de/latfor/ldg/doc/.