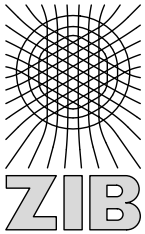
Using the LatFor Datagrid

Hinnerk Stüben



Lattice Practices '08
DESY Zeuthen, 10 October 2008

What is a (Data-) Grid?

A Grid

- integrates and coordinates resources that are not subject to centralised control
- is built from standard, open, multi-purpose protocols and interfaces that address authentication, authorisation, resource discovery and resource access
- delivers non-trivial quality of service

The Grid needs InterGrid protocols.

> Ian Foster, `http://www-fp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf` (2002)

A *Datagrid* offers storage capacity in a Grid framework.

Introduction

LDG — The LATFOR DataGrid

[`http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG.html`]

- LDG is a regional Grid for continental Europe that conforms to the International Lattice Datagrid (ILDG)
- The LDG services are operated by:
  - DESY (Hamburg and Zeuthen, Germany)
  - IN2P3 (Lyon, France)
  - JSC (Jülich, Germany)
  - ZIB (Berlin, Germany)
- Software development was sponsored by

[`http://www.d-grid.de`]
[`http://hadronphysics.infn.it`]

ILDG — The International Lattice DataGrid

[http://www.lqcd.org/ildg]

- Definition of standards for a Grid infrastructure for the lattice community
  - binary file format
  - file format schema
  - ensemble schema
  - configuration schema
  - middleware standards for interoperability in a *Grid-of-Grids*
- Aim
  - longterm storage and global sharing of gauge field configurations
- Participants
  - groups from Australia, France, Germany, Italy, Japan, UK, and US

Storage infrastructure at ZIB

Overview of this tutorial

- Motivation
- Demo: metadata catalogue, configuration downloads
- Understanding the concepts
- Software installation
- Configuration uploads

Motivation

- Rely on a well-defined data format and access mechanism
- Facilitate collaborative work
- Computer usage by ETMC:
  - generation of configurations: Barcelona, Jülich, Munich, Rome, Zeuthen
  - central storage: Edinburgh, LDG, Munich
  - measurements: Jülich, Rome, local clusters (Liverpool, Orsay, Zeuthen)
- Computer usage by QCDSF / UKQCD:
  - generation of configurations: Edinburgh, Jülich, München
  - central storage: LDG
  - measurements: Liverpool, Regensburg, Zeuthen

Downloading configurations

Demo: web interface to the metadata catalogue

[`http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/mdc.html`]

- Observe interoperability with other ILDG Grids (jqcd, usqcd)
- Click on a name → ensemble metadata human readable

  → XML representation of metadata
  → list configurations
- Click on a configuration name → configuration metadata human readable

  → XML representation of metadata

Demo: using ltools for downloading a configuration (I)

- list all ensembles: `lls --all`

```
Welcome to the Ltool-command  lls -

Showing all ensembles currently in the MDC:
===========================================

Markov-Chain-URI                                              EID Read Write
================                                              === ==== =====
mc://ldg/dik/clover_nf2/b5p29kp13632-32x64                    55  All  2
mc://ldg/dik/clover_nf2/b5p40kp13640-24x48                    56  All  2
mc://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200   53  All  6
...
```

- Explanations

`MDC`	metadata catalogue
`URI`	unified resource identifier
`EID`	ensemble ID
`Read`	read permission
`Write`	write permission (group ID)

Demo: using ltools for downloading a configuration (II)

- List configurations: `lls` *ensemble*
  Example: `lls mc://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200`

```
Welcome to the Ltool-command  lls -

The following configurations belong to
mc://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200 :

lfn://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1001
lfn://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1002
...
```

- Explanations

`mc:`	Markov chain
`lfn:`	logical filename

Demo: using ltools for downloading a configuration (III)

- Download a configuration: `lget` *`logical-file-name`*
  E.g.: `lget lfn://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1001`

```
 Welcome to the Ltool-command  lget -
 Testing grid-proxy-init:

Trying to start grid-proxy-init...
Your identity: /O=GermanGrid/OU=ZIB/CN=Hinnerk Stueben
Enter GRID pass phrase for this identity:
Creating proxy ..................................... Done
Your proxy is valid until: Sat Nov 25 00:00:27 2006
Proxy started ... grid-proxy...ok

 Trying to get binary ...

 Virtual Organisation is ildg
 Executing lcg-cp  --vo ildg \
 lfn:/grid/ildg/ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1001 \
 file:/home/bzbstueb/conf.01.1001

Checking nonzero size of downloaded File ...ok.
```


Demo: using ltools for downloading a configuration (IV)

- `grid-proxy-init` is only called if there is no valid grid-proxy:

`lget lfn://ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1002`

```
 Welcome to the Ltool-command  lget -
 Testing grid-proxy-init:
 grid-proxy...ok

 Trying to get binary ...

 Virtual Organisation is ildg
 Executing lcg-cp  --vo ildg \
 lfn:/grid/ildg/ldg/etmc/tmqcd_nf2/tlSym_b3.75_L24T48_k0.1660_mu0.0200/conf.01.1002 \
 file:/home/bzbstueb/conf.01.1002

 Checking nonzero size of downloaded File ...ok.
```

- Items to be explained:
  - `grid-proxy-init` → certificates → public key cryptography
  - virtual organisation (VO)
  - `lcg-cp` → LCG software (LCG: LHC computing grid; LHC: large hadron collider)

Binary file format

ILDG binary file format

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

- A file in ILDG binary format consists of several parts
- The parts are packaged using the LIME file format
  [`http://usqcd.jlab.org/usqcd-docs/c-lime`]
- Structure:

```
message
    record
    record
    ...
message
    record
    ...
...
```

- Each record type has a name. There have to be records of type `ildg-format`, `ildg-binary-data`, and `ildg-data-lfn`
- `ildg-format` and `ildg-binary-data` must appear in that order in the same message

Record `ildg-format`

- Example

```
<?xml version="1.0" encoding="UTF-8"?>
<ildgFormat xmlns="http://www.lqcd.org/ildg"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:schemaLocation="http://www.lqcd.org/ildg/filefmt.xsd">
<version> 1.0 </version>
<field> su3gauge </field>
<precision> 64 </precision>
<lx> 24 </lx>
<ly> 24 </ly>
<lz> 24 </lz>
<lt> 48 </lt>
</ildgFormat>
```


Record `ildg-binary-data`

- A sequence of IEEE floating point numbers in big endian format
- Precision according to `ildg-format`
- Storage sequence corresponds to the following array definitions:

  C: `double U[Lt][Lz][Ly][Lx][dim][Ncolour][Ncolour][2];`

  Fortran: `complex U(Ncolour, Ncolour, dim, Lx, Ly, Lz, Lt)`

Record `ildg-data-lfn`

- A single string with the logical filename (LFN)
- This LFN must be identical to the one given in the metadata of the configuration

Using LIME

- LIME: Lattice QCD Interchange Message Encapsulation
- The LIME package comes with a document that describes
  - the LIME format
  - the LIME API
  - some LIME utilities

  [`http://usqcd.jlab.org/usqcd-docs/c-lime/lime_1p2.pdf`]
- The utilities can be found in the directory `examples`

→ Demo: `lime_contents` utility

Security / Certificates

Digression: public / asymmetric key cryptography

- Enables secure communication over an insecure channel without prior access to a shared secret key
- A pair of keys (a *private* / *secret* key and a *public* key) is generated from a large random number
  → a private key should **always** be protected by a **strong** passphrase
- A message encrypted with a public key can only be decrypted with the corresponding private key → **secret message**
- A message encrypted with a private key can only be decrypted with the corresponding public key → **digital signature**

Certificates

- In the Grid context a certificate is a *public key certificate*
- A public key certificate is a public key digitally signed (by a trusted party) to associate the public key and a person
- The signature is by a *certificate authority (CA)*
- The German CA for LDG is *Grid Computing Centre Karlsruhe (GridKa)*
  [http://grid.fzk.de]
  [http://grid.fzk.de/cgi-bin/welcome_ca.pl]
  [http://www-grid.desy.de/certs]
- Furthers CAs: International Grid Trust Federation [http://www.gridpma.org]

Obtaining a certificate

- Generate a *certificate request*
- Contact a *registration authority (RA)* (you have to appear in person) or send a copy of your identity card an a letter with signed handwritten by a responsible person to the CA (the exact policy depends on the CA)
- Certificates are valid for a fixed period (1 year)
- The subsequent certificate can be obtained by signing a request with the old certificate
- See also
  [http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/UserDoc/newuser.html]

Generating a certificate request (for GridKa)

- Use the *Globus* command `grid-cert-request -int`
- Alternatively use *openssl*
  → follow [`http://grid.fzk.de/ca/how-to-ohne-globus.pdf`] (in German)
- Check the subject line (`OU`: organisational unit = your institute)
  - Globus: `/O=GermanGrid/OU=XXX/CN="fully qualified username"`
  - openssl: `/C=DE/O=GermanGrid/OU=XXX/CN="fully qualified username"`

  To the Grid software these are different usernames → stick to one method
- Other CAs have similar procedures
- Default filenames

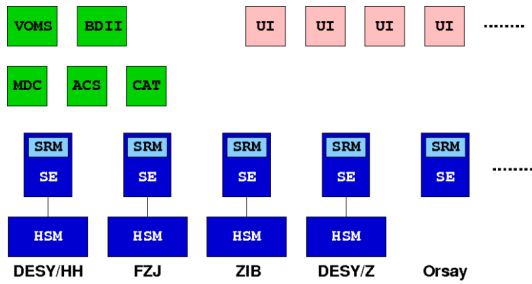
  `$HOME/.globus/usercert.pem` (certificate / public key)
  `$HOME/.globus/userkey.pem` (private key)

Virtual organisation (VO)

- In Grid computing a virtual organisation is a group of people who share a resource
- For access to LDG one has to be a member of the VO *ildg*
- The VO *ildg* is managed by DESY
- How to apply for becoming a member is described in
  [`http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/UserDoc/newuser.html`]

Software

LDG components



[hep-lat/0609012]

acronym	meaning	location of hardware/service
SE	storage element	*see figure*
SRM	storage resource manager	*see figure*
HSM	hierarchical storage manager	*see figure*
VOMS	virtual organisation management system	DESY Hamburg
BDII	Berkeley database information index	DESY Hamburg
MDC	metadata catalogue	DESY Zeuthen
ACS	access control service	DESY Zeuthen
CAT	file catalogue	DESY Hamburg
UI	user interface	the user's workstation

LDG software / middleware

- The storage elements are dCache-based [http://www.dcache.org]
- The metadata catalogue was developed at DESY Zeuthen
  [http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/Publications.html]
- The user interface is LCG-based [www.cern.ch/lcg]
- *ltools* were written at ZIB [http://www.zib.de/i3hp/ltools]
- A check-sum tool `ildg_cksum` was written at NIC/ZAM Jülich
- An RPM-based installation mechanism for the user interface, *ltools*, and `ildg_cksum` is provided by DESY Zeuthen
  [http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/UserDoc/Software.html]

LDG software installation

- No root privilege required
- Tested on
  - Fedora Core 3, Fedora Core 6
  - Scientific Linux 3, Scientific Linux 5
  - Suse Enterprise 9, 64-bit SLES 10 SP1
  - Suse 9.2, Suse 10.0, Suse 10.2, Suse 10.3
  - Ubuntu 7.10
- Instructions
  [http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/UserDoc/Software.html]
- The revocation list has to be updated regularly! (`rpm -U cert.rpm`)

Lattice tools – ltools

[http://www.zib.de/i3hp/ltools]

- *ltool* commands hide details of the underlying grid infrastructure

command	task
lls	list configurations or ensembles
lget	download configuration or (configuration or ensemble) metadata
lput	upload configuration and its metadata (performs various checks)
linit	initialise new ensemble
lupdate	update (replace) metadata
lvalidate	check conformance of metadata (configuration or ensemble) to XML schema
ladm	access control related tasks
lid	list various IDs (ensembles, projects, groups, managers, admins, users)

Metadata

Metadata

- Document data
- Enable semantic search
- Concentrate on the description of SU(3) configurations
- Standardise description of data
  - uniqueness (a specific action should always have the same description)
  - extensibility (for new actions)
  - generality (allow inclusion of other data sets, eg propagators)
- Use XML and XML schemata (XML: extensible markup language)

QCDml – XML schemata for marking up SU(3) configurations

- There are metadata for ensembles and configurations
- To learn the details of the XML schemata see
  - the ILDG web page [http://www.lqcd.org/ildg]
  - tutorial by C. Maynard and D. Pleiter at Lattice 2004 [hep-lat/0409055]
  - the metadata catalogue [http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/mdc.html]
- Metadata can be generated
  - with XML tools
  - by using templates

Sample template for configuration metadata (I)

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<gaugeConfiguration xmlns="http://www.lqcd.org/ildg/QCDml/config1.3">
    <management>
        <crcCheckSum>#crc_check_sum#</crcCheckSum>
        <archiveHistory>
            <elem>
                <revisionAction>generate</revisionAction>
                <participant>
                    <name>#participant_name#</name>
                    <institution>#participant_institution#</institution>
                </participant>
                <date>#today#</date>
            </elem>
        </archiveHistory>
    </management>
```


Sample template for configuration metadata (II)

```
    <implementation>
        <machine>
            <name>#machine_name#</name>
            <institution>#machine_institution#</institution>
            <machineType>#machine_type#</machineType>
        </machine>
        <code>
            <name>#code_name#</name>
            <version>#code_version#</version>
            <date>#code_date#</date>
        </code>
    </implementation>
```

Sample template for configuration metadata (III)

```
    <algorithm>
        <parameters>
            <name>stepSize</name>
            <value>#para_step_size#</value>
        </parameters>
        <parameters>
            <name>numberSteps</name>
            <value>#para_number_steps#</value>
        </parameters>
        <parameters>
            <name>solverResiduum</name>
            <value>#para_solver_residuum#</value>
        </parameters>
        <parameters>
            <name>rho</name>
            <value>#para_rho#</value>
        </parameters>
    </algorithm>
    <precision>#precision#</precision>
```


Sample template for configuration metadata (IV)

```
    <markovStep>
        <markovChainURI>#markov_chain_uri#</markovChainURI>
        <series>#markov_series#</series>
        <update>#markov_update#</update>
        <avePlaquette>#average_plaquette#</avePlaquette>
        <dataLFN>#data_lfn#</dataLFN>
    </markovStep>
</gaugeConfiguration>
```

- The `markovChainURI` of configuration and ensemble must match!

Sample template for configuration metadata (V)

placeholder	replacement
#crc_check_sum#	559813009
#participant_name#	Hinnerk Stueben
#participant_institution#	ZIB
#today#	2006-11-26T10:47:38+01:00
#machine_name#	jubl
#machine_institution#	NIC/ZAM Juelich
#machine_type#	IBM Blue Gene/L
#code_name#	BQCD
#code_version#	3.3.0
#code_date#	2006-08-28T14:43:00+02:00
#para_step_size#	0.01
#para_number_steps#	100
#para_solver_residuum#	1e-14
#para_rho#	0.06
#precision#	double
#markov_chain_uri#	mc://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48
#markov_series#	536
#markov_update#	2485
#average_plaquette#	0.5439676525
#data_lfn#	lfn://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime

Naming conventions

- `markovChainURI`:

  `mc://ldg/`*collaboration*`/`*project*`/`*ensemble*

  Example:

  `mc://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48`

- `dataLFN`:

  `lfn://ldg/`*collaboration*`/`*project*`/`*ensemble*`/`*configuration*

  Example:

  `lfn://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime`

- *collaboration*, *project*, and *ensemble* must match!

Storage location

- The sequence

  collaboration/*project*/*ensemble*/*configuration*

  is also used as the storage location (it represents a directory structure).

  Example:

  `srm://dcache.zib.de/pnfs/zib.de/data/ildg/`
  `qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime`

- This is handled by `lput`.

Ensemble metadata example (I)

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<markovChain xmlns="http://www.lqcd.org/ildg/QCDml/ensemble1.3">
    <markovChainURI>mc://ldg/qcdsf/clover_nf2/b5p40kp13625-24x48</markovChainURI>
    <management>
        <collaboration>qcdsf</collaboration>
        <projectName>clover_nf2</projectName>
        <archiveHistory>
            <elem>
                <revisionAction>add</revisionAction>
                <participant>
                    <name>Hinnerk Stueben</name>
                    <institution>ZIB</institution>
                </participant>
                <date>2006-10-30T17:26:36+01:00</date>
            </elem>
        </archiveHistory>
    </management>
```


Ensemble metadata example (II)

```
<physics>
    <size>
        <elem>
            <name>X</name>
            <length>24</length>
        </elem>
        <elem>
            <name>Y</name>
            <length>24</length>
        </elem>
        <elem>
            <name>Z</name>
            <length>24</length>
        </elem>
        <elem>
            <name>T</name>
            <length>48</length>
        </elem>
    </size>
```


Ensemble metadata example (III)

```
    <action>
        <gluon>
            <plaquetteGluonAction>
                <glossary>http://www.lqcd.org/ildg/actionGlossaries/plaquetteGluonAct
                <gluonField>
                    <gaugeGroup>SU(3)</gaugeGroup>
                    <representation>fundamental</representation>
                    <boundaryCondition>
                        <elem>periodic</elem>
                        <elem>periodic</elem>
                        <elem>periodic</elem>
                        <elem>periodic</elem>
                    </boundaryCondition>
                </gluonField>
                    <beta>5.40</beta>
            </plaquetteGluonAction>
        </gluon>
```

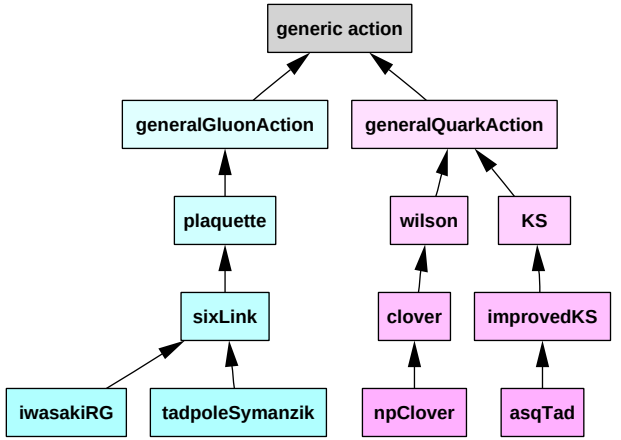
Ensemble metadata example (IV)

```
            <quark>
                <npCloverQuarkAction>
                    <glossary>http://www.lqcd.org/ildg/actionGlossaries/npCloverQuarkActi
                    <quarkField>
                        <normalisation>sqrt2kappa</normalisation>
                        <boundaryCondition>
                            <elem>periodic</elem>
                            <elem>periodic</elem>
                            <elem>periodic</elem>
                            <elem>antiperiodic</elem>
                        </boundaryCondition>
                    </quarkField>
                        <numberOfFlavours>2</numberOfFlavours>
                        <kappa>0.13625</kappa>
                        <cSW>1.8228</cSW>
                </npCloverQuarkAction>
            </quark>
        </action>
    </physics>
```


Ensemble metadata example (V)

```
    <algorithm>
        <name>qcdsfAcceleratedHMC</name>
        <glossary>http://www-zeuthen.desy.de/latfor/ldg/algorithmGlossaries/qcdsfAccelera
        <reference>Phys.Lett. B564 (2003) 235-240</reference>
        <exact>true</exact>
        <parameters>
            <name>timeScaleRatio</name>
            <value>3</value>
        </parameters>
    </algorithm>
</markovChain>
```


Hierarchy of actions in the ensemble metadata schema



[hep-lat/0409055]

Uploading configurations

Uploading a configuration (I)

- Upload a configuration: `lput` *`configuration-metadata`*
  Example: `lput qcdsf.536.02485.xml`

```
Welcome to the Ltool-command  lput -

Testing grid-proxy-init
grid-proxy...ok

Virtual Organisation is ildg

Storage Element is dcache.zib.de
Checking that URI and LFN are built according to convention:
OK

Checking that ensemble mc://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48 exists in MDC:
OK

Checking that configuration does not exist in MDC:
OK

Checking that LFN does not already exist:
OK
```


Uploading a configuration (II)

```
All checks ok, will now start to upload metadata and binary...

Starting to upload metadata to MDC ...

...MDC Operation ok.

Starting to upload binary ...

Executing lcg-cr  --vo ildg -l \
lfn:/grid/ildg/ldg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime \
file:/home/bzbstueb/conf/qcdsf.536.02485.lime \
-d srm://dcache.zib.de/pnfs/zib.de/data/ildg/qcdsf/\
                        clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime
guid:ae514f13-dd42-4943-a284-3fcb0e4d4d4a

Data successfully copied to Storage Element !
LFN is: lfn://ldg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.536.02485.lime
```

- `guid`: globally unique identifier

Uploading a configuration (III)

- Define defaults in `$HOME/.ltools`

```
## Default virtual organisation (for up- and downloads)
default_vo ildg

## Only needed for lputting data :
## default UPLOAD Storage element. If nothing is given here, by uploading
## the --se <SE> option hast to be specified by uploading a file.

default_upload_se dcache.zib.de
```

- A template for `.ltools` can be found in the installation directory `$LROOT/opt/ltools-?.?/template.ltools`
- The storage elements are listed in `$LROOT/etc/ltools.rsc`

Uploading a configuration (IV)

- `lput` also handles failures
  Example: `lput qcdsf.537.00170.xml`

```
Welcome to the Ltool-command  lput -

normal messages, all checks ok

Starting to upload binary ...

Executing lcg-cr  --vo ildg -l \
lfn:/grid/ildg/ldg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.537.00170.lime \
file:/home/bzbstueb/conf/qcdsf.537.00170.lime \
-d srm://dcache.zib.de/pnfs/zib.de/data/ildg/qcdsf/clover_nf2/b5p25kp13600-24x48/qcdsf.53
lcg_cr: File exists

 Fatal Error...lcg-cr command did NOT succeed !
Error-message is : Bad file descriptor
Rolling back : Trying to restore former MDC-status:

 ...MDC was rolled back to former consistent status.
```

→ try again!

Initialising an ensemble

- Initialising an ensemble: `linit` *`project-ID ensemble-metadata`*
- Listing project-IDs: `ladm --show-projects`

Concluding remarks

Next release of ltools

- The next release of *ltools* will be available soon
- It will interoperate with the other ILDG Grids
- `lls --ildg` will list ensembles from all ILDG Grids
- `lget` will work transparently for all ILDG Grids

Help / Support / Further reading

- Hint: one can use `grid-proxy-init -valid ...` to extend the validity of the proxy certificate
- All *ltool* commands understand the option `--help`
- Problems can be reported to `ldg-adm@desy.de`
  Please append the output from `{lget|lput} --debug` *`file`* to your e-mail
- Further reading
  - Dirk Pleiter's talk at HackLatt 2008
    [`http://ukqcd.epcc.ed.ac.uk/training/2008/talks/ILDG.pdf`]
  - Papers and Presentations
    [`http://globe-meta.ifh.de:8080/lenya/hpc/live/LDG/Publications.html`]