The Gauge Connection - Version 3.0

Massimo Di Pierro, James Hetrick, Jim Simone, Shreyas Cholia, Carleton De Tar

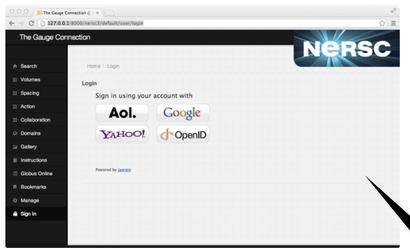
Login
Google & OpenID

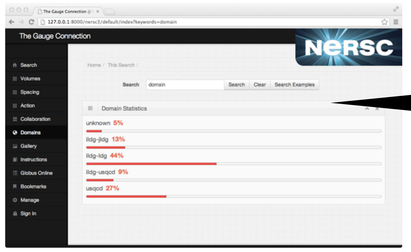
Search & Stats
by tag (size, beta, flavor, etc.)

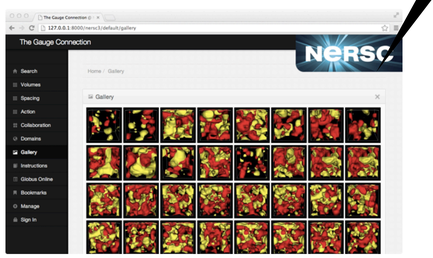
Sync ILDG
worldwide ILDG search

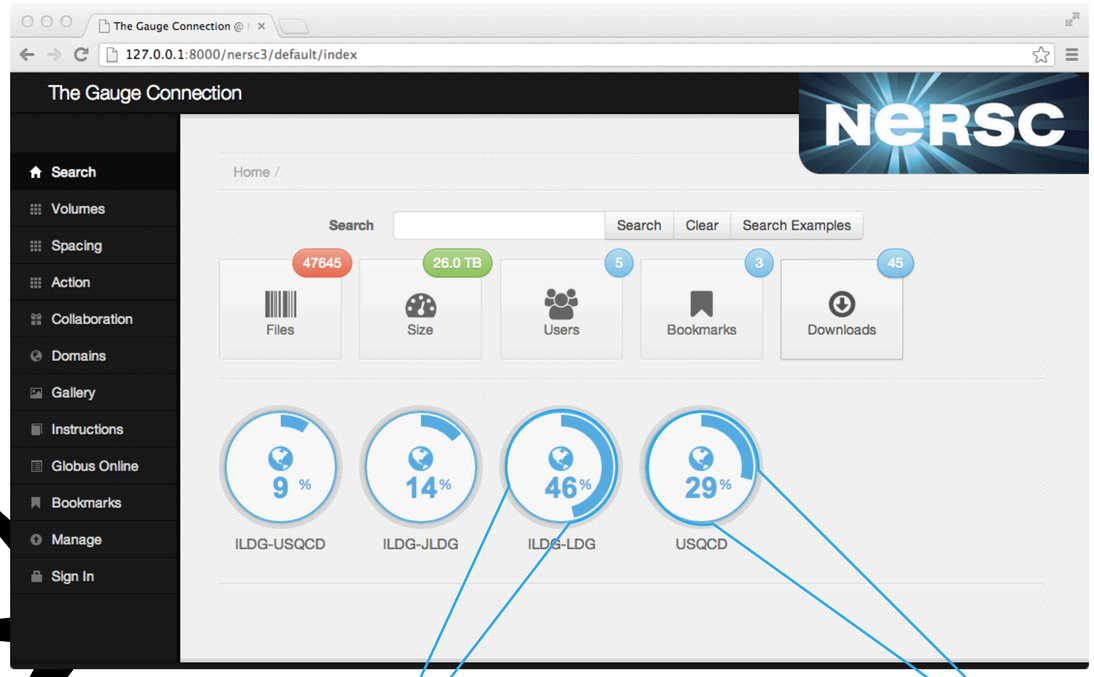
Wiki & Tagging
each ensemble has a Wiki page

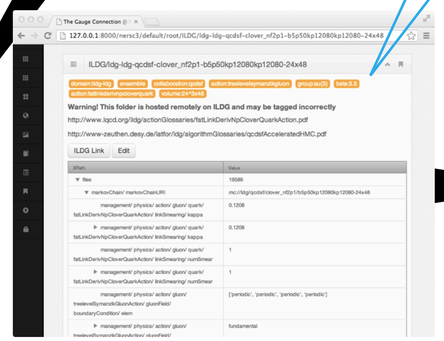
Batch Downloads
schedule batch downloads/jobs

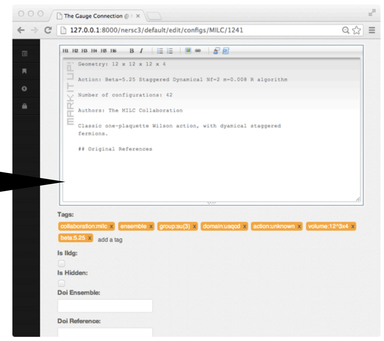


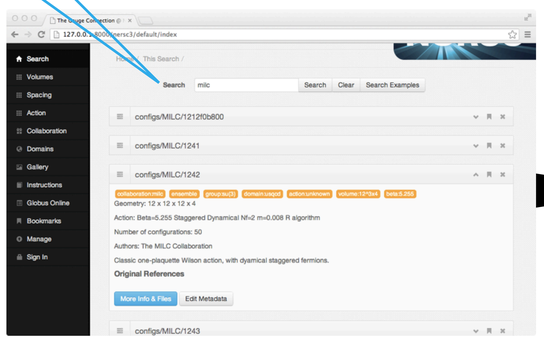


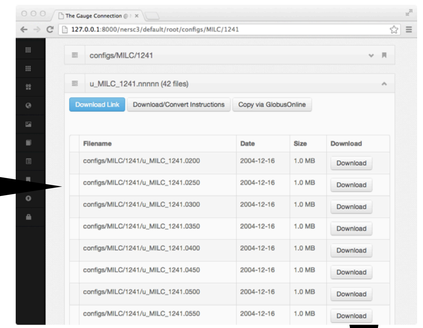












Abstract and Motivations

The "Gauge Connection" is a data archive serving the Lattice QCD community which has been operated by the National Energy Research Scientific Computing (NERSC) center since 1998. It is now being updated to provide a number of modern features and to offer a simple path to publishing gauge field ensembles for those who prefer it.

The archive stores over 16TBytes of gauge ensembles locally. The new interface also collects information about all the ILDG ensembles made available on third party ILDG servers and provides a single place to search them all.

Ensembles can be searched by name or by tags. In the ILDG case the tags are generated from the metadata. Tags can also be added manually. Each ensemble (local or remote) is associated to a wiki page. The content of the wiki page is automatically populated with known information such as links to the papers referencing the action and the ensemble (DOI and arXiv).

Local ensembles are stored and are delivered in the NERSC format but they can be converted to ILDG or FermiQCD format by the download script itself.

The system has an access control mechanism with four user roles: administrators (can manage every aspect), editors (can edit the wiki pages linked to the data), registered users (can download data and comment), anonymous visitors (can browse the wiki pages and query the data by tags).

It allows bookmarking of data ensembles, it remembers recent searches, and it logs all downloads in order to keep track of usage and most popular ensembles.

About NERSC

The NERSC data is stored on the High Performance Storage System (HPSS), a tape storage system developed by a collaboration between five Department of Energy laboratories and IBM with significant contribution form Universities.

HPSS stores more than 2.5Peta Bytes of data, including 16 Tera Bytes of Lattice QCD data. HPSS has an FTP interface, a GridFTP interface, a web interface and a Globus Online Interface.

The "Gauge Connection" data archive has been operated by the National Energy Research Scientific Computing (NERSC) center since 1998 and it is now being updated to provide a number of modern features. With over 16TBytes of data the NERSC archive is one of the largest public repositories of lattice ensembles.

Features

- Social Media Authentication and Access Control
- Keeps copy of ILDG metadata about remote ensembles
- Each ensemble is associated with a Wiki page and tags
- Ensembles are searchable by name and tags
- Tags include: domain, collaboration, beta, spacing, mass, action, group
- Allows download of individual files as well as bulk downloads
- Tree browser of ILDG metadata
- Generates statistical charts about data
- Keeps detailed logs about usage and downloads

Building Blocks

The system is based on web2py, a framework for rapid development of secure database driven web applications. It is written in Python and supports many standard databases including Sqlite, MySQL, PostgreSQL, Oracle, MSSQL, Informix, DB2, Sybase, Firebird, and Google Bit Table. The SQL is generated automatically.

The system has a module Model-View-Controller design which separates the data representation, data presentation, and workflow.

The system allows download individual files using the web interface as well as bulk downloads using qcdutils and Globus Online (work in progress).

The system exposes a RESTful web services APIs based on the JSON protocol which can be used to access the data programmatically.

```
$ qcdutils.py http://qcd.nersc.gov/nersc/api/files/demo
http://qcd.nersc.org/nersc/api/files/demo
target folder: demo
total files to download: 1
downloading demo.nersc
demo.nersc 100% |###############| Time: 00:00:00 654.52 K/s
completed download: 1/1

$ qcdutils.py --covert ildg --float demo/demo.nersc
converting: demo/demo.nersc -> demo/demo.nersc.ildg
  (precision: f, size: 4x8x8x8)
100% |##################################|

$ qcdutils.py demo/qcdutils.catalog.db
demo.nersc created on 2011-06-17T13:42:30.876812
   [14e7cf9106bfcc16388aeac285ccdad9]
demo.nersc.ildg created on 2011-06-17T13:42:34.424604
```