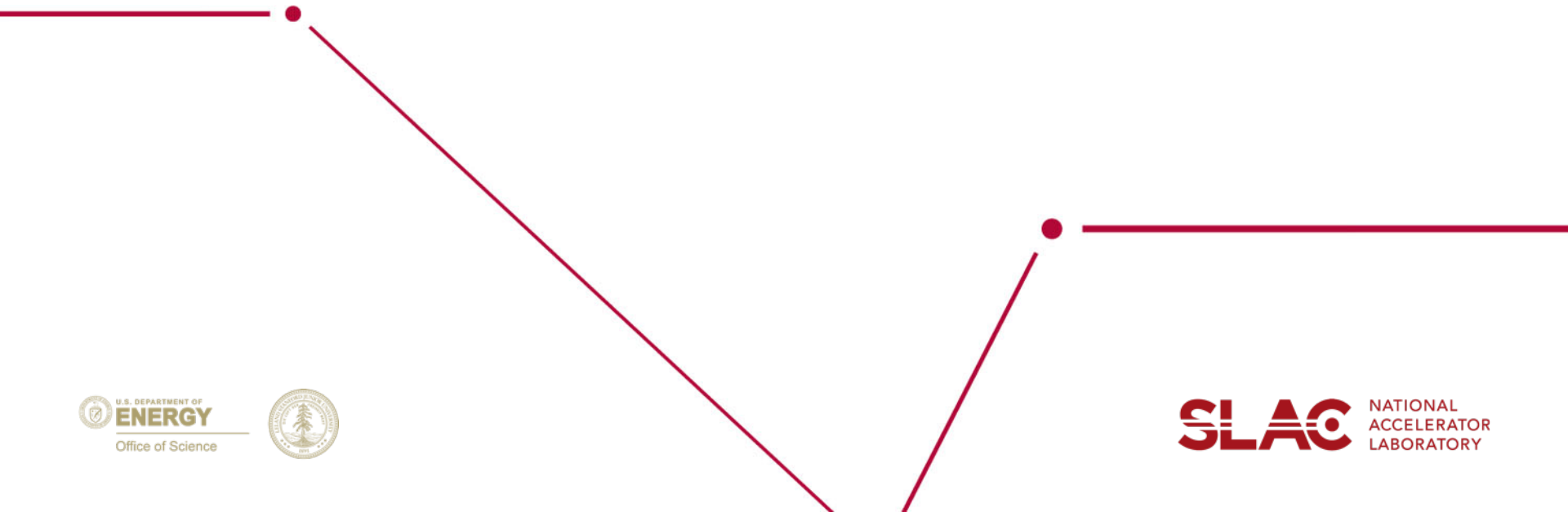
Calibration DB for LCLS2

Mikhail Dubrovin

2018-09-28



LCLS Experience with calibration stores

SLAC

- File system based calibration store

  /reg/d/psdm/XPP/xpptut15/xtc/*.xtc – data files

  calib/CsPad::CalibV1/XppGon.0:Cspad.0/pedestals/54-59.data

  HISTORY

- HDF5 calibration store

  per experiment: /reg/d/psdm/CXI/cxis0613/calib/cspad/cspad-CxiDs1.0-Cspad.0.h5

  repository: /reg/d/psdm/detector/calib/cspad/cspad-CxiDs1.0-Cspad.0.h5

- Each calibration store has API, CLI, GUI
- There are pros and cons, but they both are not DBs…
- DB makes life easier and we would like to try

Evaluation of MongoDB

SLAC

- MongoDB – open source, document based, **NoSQL** database, written in C++

- Structure of MongoDB
  - *mongod* demon runs on server (psana node(s) or your laptop…)
  - connection is provided by client using --host, --port, <authorization>
  - client supports a set of databases
    - database is a set of collections
      - collection is a storage of documents
        - document is JSON/BSON (unicode/binary) dictionary with flexible schema (NoSQL)

- Access to DBs and collections by name
- Access to documents by query or unique "_id"
- Limiting factor: **size of BSON object <16MB**

  (document can't hold double CSPAD array 2Mpixel of 8-byte words!)
  - GridFS solves "big data" issue - keeps big data in separate collections of the same DB, access it by "id_data"

Calibration DB schema on MongoDB

SLAC

- Single client contains all calibration databases (cdb). Databases are created per **experiment** and per **detector** using their prefixed names, e.g. cdb_exp1, cdb_exp2,… cdb_det1, cdb_det2,…
  - Database is a set of collections.

    Collections are created per detector and for GridFS **big data**, e.g.
    - collections in experiment DB: detA, detB,…detZ, **fs.chanks, fs.files**
    - collections in detector DB: detN, **fs.chanks, fs.files**
  - Collection is a storage of documents.
    - Document contains metadata about calibration and reference to **big data**, e.g.

      {"_id":"5b6cdde71ead144f115319be","experiment":"cxid9114",
      "run":116,"run_end":"end","detector":"cspad_0001","ctype":"pedestals",
      "time_sec":1402940673, "version":"v0", "id_data":"5b6cdde71ead144f11531974",
      "data_type":"ndarray","data_dtype":"float32","data_size":"2296960",
      "data_ndim":"2", "data_shape":"(5920, 388)", …}

Top level calibration constants access API

SLAC

Access in LCLS2 network

```
from psana.pscalib.calib.MDBUtils import calib_constants
det = ‘cspad_1234’
data, doc = calib_constants(det, exp='cxi12345', ctype=’pedestals’,
                run=56, time_sec=None, vers=None)
```

Any combination of parameters which uniquely defines latest available in DB constants is allowed.

WWWeb service interface

- in python

```
from psana.pscalib.calib.MDBWebUtils import calib_constants
data, doc = calib_constants(det, exp='cxi12345', ctype=’pedestals',
                        run=56, time_sec=None, vers=None)
```

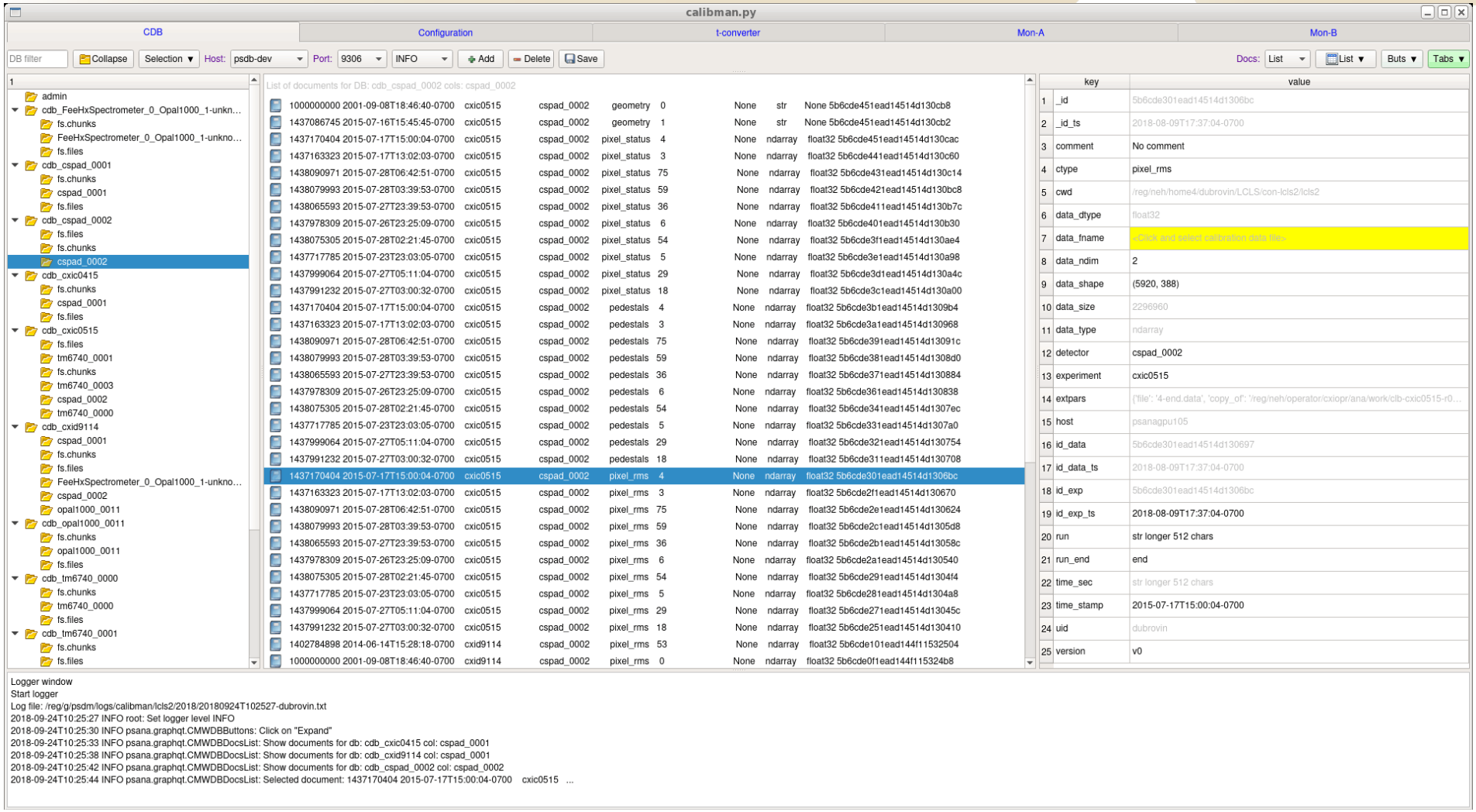
- similar www interface is available in C++ for DRP

Command Line Interface

SLAC

```
cdb
cdb -h
cdb print
cdb print -e cxix25615
cdb print -d camera_0_cxids1_0
cdb convert -e xcs01116 <authorizatin> # converts LCLS to LCLS2 calibrations
cdb add -e cxi12345 -d camera_0_cxids1_0 -c pedestals -r 123 -f my.txt
cdb get -e cxix25615 -d cxids1_0_cspad_0 -c pedestals -s 1520977960 -p -f peds.txt
cdb get -e xcsh8215 -d xcsendstation_0_cspad_0 -c pedestals -r 100 -f my.txt
cdb get -d xcsendstation_0_cspad_0 -c pedestals -r 100 -f my.txt
cdb deldoc -e cxix25615 -d cxids1_0_cspad_0 -c pedestals -r 125
cdb deldoc -e cxix25615 -d cxids1_0_cspad_0 -c pedestals -s 1520977960
cdb delcol -e cxix25615 -d cxids1_0_cspad_0
cdb delcol -d cxids1_0_cspad_0
cdb deldb -e cxix25615
cdb deldb -d cxids1_0_cspad_0
cdb deldb --dbname cdb_cxids1_0_cspad_0
cdb delall
cdb export --dbname cxix25615
cdb import --dbname cxix25615 --iofname cdb-...arc
```

GUI for Calibration DB maintenance



Summary

SLAC

- Calibration DB for LCLS2 is designed on MongoDB
- Easy access to constants is provided by python API, CLI, and GUI
- World wide web access (read-only) is provided by python and C++ API

- Implementation is nearly completed, work on tests and examples

SLAC

Additional slides

API: access to calibration constants

SLAC

Low level API

```
import psana.pscalib.calib.MDBUtils as mu

client, expname, detname, db_exp, db_det, fs_exp, fs_det, col_exp, col_det =\
mu.connect(host='psanaphi105', port=27017, experiment='cxi12345',
                detector='cspad_0002', **kwargs)

doc = mu.find_doc(col_det, query={<dict-of-parameters-to-find-document>})
data = mu.get_data_for_doc(fs, doc)

… long list of methods
```

CLI: add constants

SLAC

Command:

```
cdb add –e exp12 –d det34 -c pedestals -r 567 –v V123  -f cons.txt …
cdb add –e exp12 –d det34 -c pedestals -s 1500000000 –v V123 -f cons.txt …
cdb add –e exp12 –d det34 -c pedestals –t 2018-03-26T15:00:01-0700 -f cons.txt …
```

- adds/creates two DBs: cdb_exp12 and cdb_det34
- with collections: fs.chunks, fs.files, det34
- data from cons.txt is saved in collections fs.* with unique “id”
- document is created with all necessary metadata and “id” for data and saved in collections named as det34

Duplication of data and documents for experiment and detector DBs allows to export DBs independently

CLI: cdb print -e xcsm9816

```
cdb print -e xcsm9816
2018-03-20T15:25:30 MDB_CLI INFO: MongoDB client host:psanaphi105 port:27017

dbnames ['admin', 'cdb_xcsendstation_0_cspad2x2_0', 'cdb_xcsendstation_0_epix100a_1', 'cdb_xcsendstation_0_epix100a_2', 'cdb_xcsendstation_0_epix100a_3', 'cdb_xcsendstation_0_epix100a_4', 'cdb_xcsm9816', 'config', 'local',
'mydbname']

 DB cdb_xcsm9816 contains 7 collections: ['xcsendstation_0_epix100a_1', 'fs.chunks', 'xcsendstation_0_epix100a_3', 'fs.files', 'xcsendstation_0_epix100a_2',
'xcsendstation_0_epix100a_4', 'xcsendstation_0_cspad2x2_0']

 COL xcsendstation_0_epix100a_1 contains 66 docs
 Details for collection xcsendstation_0_epix100a_1 66 documents:
 20 document keys:
   run_end       experiment     extpars       _id          version
   data_dtype    uid            cwd           comment      time_sec
   data_size     ctype          run           data_ndim    id_data
   host          data_type      data_shape    detector     time_stamp
doc#  time_sec            time_stamp  experiment            detector          ctype  run              ts_data  data_type data_dtype
   0 1476442278 2016-10-14T03:51:18-0700   xcsm9816 xcsendstation_0_epix100a_1    pixel_mask  20     2018-03-20T11:42:46-0700   ndarray   float32
   1 1476676936 2016-10-16T21:02:16-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   340    2018-03-20T11:42:47-0700   ndarray   float64
   2 1476418323 2016-10-13T21:12:03-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   7      2018-03-20T11:42:47-0700   ndarray   float64
   3 1476521140 2016-10-15T01:45:40-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   66     2018-03-20T11:42:48-0700   ndarray   float64
   4 1476444527 2016-10-14T04:28:47-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   30     2018-03-20T11:42:49-0700   ndarray   float64
   5 1476609833 2016-10-16T02:23:53-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   240    2018-03-20T11:42:49-0700   ndarray   float64
   6 1476705090 2016-10-17T04:51:30-0700   xcsm9816 xcsendstation_0_epix100a_1    pedestals   439    2018-03-20T11:42:50-0700   ndarray   float64
...

 COL xcsendstation_0_cspad2x2_0 contains 69 docs
 Details for collection xcsendstation_0_cspad2x2_0 69 documents:
 20 document keys:
   run_end       experiment     extpars       _id          version
   data_dtype    uid            cwd           comment      time_sec
   data_size     ctype          run           data_ndim    id_data
   host          data_type      data_shape    detector     time_stamp
doc#  time_sec            time_stamp  experiment            detector          ctype  run              ts_data  data_type data_dtype
   0 1476542126 2016-10-15T07:35:26-0700   xcsm9816 xcsendstation_0_cspad2x2_0    pixel_mask  101    2018-03-20T11:42:24-0700   ndarray   float32
   1 1476457527 2016-10-14T08:05:27-0700   xcsm9816 xcsendstation_0_cspad2x2_0    pedestals   49     2018-03-20T11:42:25-0700   ndarray   float64
   2 1476417759 2016-10-13T21:02:39-0700   xcsm9816 xcsendstation_0_cspad2x2_0    pedestals   6      2018-03-20T11:42:25-0700   ndarray   float64
   3 1476676936 2016-10-16T21:02:16-0700   xcsm9816 xcsendstation_0_cspad2x2_0    pedestals   340    2018-03-20T11:42:26-0700   ndarray   float64
   4 1476418323 2016-10-13T21:12:03-0700   xcsm9816 xcsendstation_0_cspad2x2_0    pedestals   7      2018-03-20T11:42:26-0700   ndarray   float64
```

CLI: cdb convert -e cxif5315

```
cdb convert -e cxif5315
Scan: /reg/d/psdm/CXI/cxif5315/calib
  converted   2 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/pixel_bkgd
  converted   8 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/pixel_rms
  converted   1 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/common_mode
  converted   8 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/pedestals
  converted   8 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/pixel_status
  converted   1 files from: /reg/d/psdm/CXI/cxif5315/calib/CsPad::CalibV1/CxiDs2.0:Cspad.0/geometry
```

- command convert - converts old LCLS1 calibration data to LCLS2 DB
- converted all xcs, xpp, mec, cxi calibrations to server running on scratch disk
- a few issues were found/fixed in old calib data
  - a few zero-size files
  - files with wrong names
  - fies with incorrect data
  - non-text files in /reg/d/psdm/XCS/xcsm9816/calib/Xtcav::CalibV1/XrayTransportDiagnostic.0:Opal1000.0/pedestals/
- access to calibration info works fine.

Document for metadata

{"_id":"5b6cdde71ead144f115319be","experiment":"cxid9114","run":
116,"run_end":"end","detector":"cspad_0001","ctype":"pedestals","time_sec":
1402940673,"time_stamp":"2014-06-16T10:44:33-0700","version":"v0","comment":"No
comment","extpars":{"file":"116-end.data","copy_of":"./work/clb-cxid9114-r0116-peds-ave-
CxiDs1.0:Cspad.0.txt","exp":"cxid9114","run":"0116","comment":"calibrun-
dark","user":"koglin","host":"psanacs059","cptime":"2014-06-16T10:47:27","zone":"PDT"},"
uid":"dubrovin","host":"psanagpu105","cwd":"/reg/neh/home4/dubrovin/LCLS/con-lcls2/
lcls2","id_data":"5b6cdde71ead144f11531974","data_type":"ndarray","data_dtype":"float32
","data_size":"2296960","data_ndim":"2","data_shape":"(5920, 388)"}