

DQ2 CLIENTS NEWS

Angelos Molfetas on behalf of the DDM team

Introduction

- New features in the upcoming DQ2 clients release
 - Pool File Catalog (PFC) generation
 - Global Physical Dataset Path (PDP)
- PDP particularly relevant to US Tier 3 sites
- Changes will appear in upcoming stable release (0.1.36)

Pool File Catalog Generation

- Old CLI (dq2-ls -P -G):
 - Hard to maintain code
 - Cluttered other dq2-ls code
- New CLI (dq2-PFC-gen):
 - Conversation (SURL -> TURLS) methods:
 - regular expressions in ToA (default)
 - lcg-cp
 - lcg-getturls
 - Validation option (regular expressions vs lcg-getturls)

Examples (dq2-PFC-gen)

- To generate PFC using ToA regular expressions:
 - `dq2-PFC-gen -L PIC_HOTDISK -t cond09_mc.000012.gen.COND`
- To generate PFC using LCG-GT:
 - `dq2-PFC-gen -L PIC_HOTDISK -t -l cond09_mc.000012.gen.COND`
- To generate PFC using LCG_GETURLs:
 - `dq2-PFC-gen -L PIC_HOTDISK -t -b cond09_mc.000012.gen.COND`
- To validate Regular Expressions vs LCG-GETURLs:
 - `dq2-PFC-gen -L PIC_HOTDISK -t -V cond09_mc.000012.gen.COND`
- Further examples please see:
 - <https://svnweb.cern.ch/trac/dq2/wiki/Using%20dq2-PFC-gen>

Global Physical Dataset Path (PDP)

- Background:
 - Requested for Tier-3 sites
 - Store datasets in a commonly accessible area
- Support in the dq2 clients:
 - Download to PDP using dq2-get
 - List datasets and files in PDP using dq2-ls
 - Base path and prefix can be specified
- Implications:
 - Lesser reliance on Site Services for Tier-3
 - Lesser reliance on LFC for Tier-3

PDP example:

DSN: mc10_7TeV.

105563.Pythia_Gee_003_800.merge.AOD.e574_s933_s946_r1652_r1700_tid192840_00

PDP: /grid/atlas/dq2/mc10_7TeV/AOD/e574_s933_s946_r1652_r1700/mc10_7TeV.

105563.Pythia_Gee_003_800.merge.AOD.e574_s933_s946_r1652_r1700_tid192840_00/

base path and prefix: *{bpath}/{prefix}/mc10_7TeV/AOD/e574_s933_s946_r1652_r1700/mc10_7TeV.*

105563.Pythia_Gee_003_800.merge.AOD.e574_s933_s946_r1652_r1700_tid192840_00/

dq2 clients PDP examples

- Download to PDP area:
 - `dq2-get -g mc10_7TeV.105563.Pythia_Gee_003_800.merge.AOD.e574_s933_s946_r1652_r1700_tid192840_00`
- Download to PDP area with modified base path and prefix:
 - `dq2-get -g --bpath=/tmp/angelos/ --prefix=temp_datasets mc10_7TeV.105563.Pythia_Gee_003_800.merge.AOD.e574_s933_s946_r1652_r1700_tid192840_00`
- List mc* datasets in PDP area:
 - `dq2-ls -g mc*`
- List mc* datasets in PDP area with modified base path:
 - `dq2-ls -g --bpath=/tmp/angelos mc*`
- Further examples, please see:
 - <https://svnweb.cern.ch/trac/dq2/wiki/Using%20physical%20dataset%20paths%20in%20the%20dq2%20clients>

PDP consistency

- Consistency between PDP and CC:
- Site Services maintains versioning consistency for non-frozen datasets
- Dq2-ls checks consistency between central catalogs and PDP
 - Files and checksums
 - Inconsistencies can be listed by dq2-ls by:
 - Consistency mode
 - Debug mode
 - User driven
- Dq2-get establishes the difference between PDP and CC and downloads missing files

PDP consistency example 1

```
sudo mkdir /grid/  
sudo chown angelos /grid/  
dq2-put -s test user.angelos.`uuidgen`  
rm /grid/atlas/dq2/user/angelos/4b50d35d-0801-4669-bee3-34fa846df896/user.angelos.  
4b50d35d-0801-4669-bee3-34fa846df896/c3124767-0057-4132-abd5-6a1832315184  
dq2-ls -g -V user.angelos.4b50d35d-0801-4669-bee3-34fa846df896  
Consistency check log:  
WARNING: file on catalogs c3124767-0057-4132-abd5-6a1832315184 (dataset  
user.angelos.4b50d35d-0801-4669-bee3-34fa846df896) does not exist on disk
```


PDP consistency example 2

```
touch /grid/atlas/dq2/user/angelos/4b50d35d-0801-4669-bee3-34fa846df896/user.angelos.4b50d35d-0801-4669-bee3-34fa846df896/non_exist_fn1
```

```
touch /grid/atlas/dq2/user/angelos/4b50d35d-0801-4669-bee3-34fa846df896/user.angelos.4b50d35d-0801-4669-bee3-34fa846df896/non_exist_fn1
```

```
dq2-ls -g -V user.angelos.4b50d35d-0801-4669-bee3-34fa846df896
```

Consistency check log:

WARNING: file on disk non_exist_fn1 (dataset user.angelos.4b50d35d-0801-4669-bee3-34fa846df896) does not exist in the central catalogs

WARNING: file on disk non_exist_fn2 (dataset user.angelos.4b50d35d-0801-4669-bee3-34fa846df896) does not exist in the central catalogs

```
rm /grid/atlas/dq2/user/angelos/4b50d35d-0801-4669-bee3-34fa846df896/user.angelos.4b50d35d-0801-4669-bee3-34fa846df896/non_exist_fn2
```

```
rm /grid/atlas/dq2/user/angelos/4b50d35d-0801-4669-bee3-34fa846df896/user.angelos.4b50d35d-0801-4669-bee3-34fa846df896/non_exist_fn1
```

```
dq2-ls -g -V user.angelos.4b50d35d-0801-4669-bee3-34fa846df896
```

Consistency check log:

No problems detected

PDP consistency example 3

```
dq2-get -g user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8
dd if=/dev/urandom of=/grid/atlas/dq2/user/angelos/42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8/user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8/new_file_aamm1
bs=10000 count=1
dd if=/dev/urandom of=/grid/atlas/dq2/user/angelos/42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8/user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8/new_file_aamm2
bs=10000 count=1
dd if=/dev/urandom of=/grid/atlas/dq2/user/angelos/42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8/user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8/new_file_aamm3
bs=10000 count=1
```

```
dq2-ls -V -g -f user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8
Consistency check log:
WARNING: file on disk new_file_aamm1 (dataset user.angelos.42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8) does not exist in the central catalogs
WARNING: file on disk new_file_aamm3 (dataset user.angelos.42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8) does not exist in the central catalogs
WARNING: file on disk new_file_aamm2 (dataset user.angelos.42b1e86b-9d7c-4bf3-8985-
a4ebbb1d4bc8) does not exist in the central catalogs
```

PDP consistency example 3 (cont)

```
dq2-put -s /grid/atlas/dq2/user/angelos/42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8/  
user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8/ user.angelos.  
42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8  
Dataset user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8 exists already, confirm to  
create new version 2 and add files (y/n)? y  
Gathering list of files... done  
new_file_aamm1: pre-checking file  
new_file_aamm3: pre-checking file  
new_file_aamm2: pre-checking file
```

```
dq2-ls -g -f user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8  
...  
new_file_aamm1  
new_file_aamm3  
fd898802-5742-4a5e-ba65-3bc66a2345b7  
new_file_aamm2  
...  
dq2-ls -g -V user.angelos.42b1e86b-9d7c-4bf3-8985-a4ebbb1d4bc8  
Consistency check log:  
No problems detected
```

Miscellaneous

- Reminder: dq2-put2 is in the clients
- Dq2-get download priority:
 1. DQ2_LOCAL_SITE_ID,
 2. DQ2_LOCAL_SITE_ID's site
 3. tier 1 in close site
 4. any close site
 5. random site
- Support for STOMP protocol in the client for traces submitted to ActiveMQ server (see Donal Zang's presentation)
- Support planned for DQ2 Accounting CLI/Client API transparently using Oracle or HBASE back end (please see Mario Lassnig's presentation)
- Support for Python 2.6
- Dropping support for python 2.3 and python 2.4

Summary

- New CLI for PFC generation
- Support for PDP in dq2 clients
- User driven support for consistency verification in dq2-ls
- Additional Information:
 - Change log: <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DQ2Clients>
 - Roadmap: <https://svnweb.cern.ch/trac/dq2/roadmap>
 - Latest Release: https://svnweb.cern.ch/trac/dq2/wiki/0.1.36_rc3
 - Relevant wikis: <https://svnweb.cern.ch/trac/dq2/wiki>