

Production Operations Procedure: Data access issues

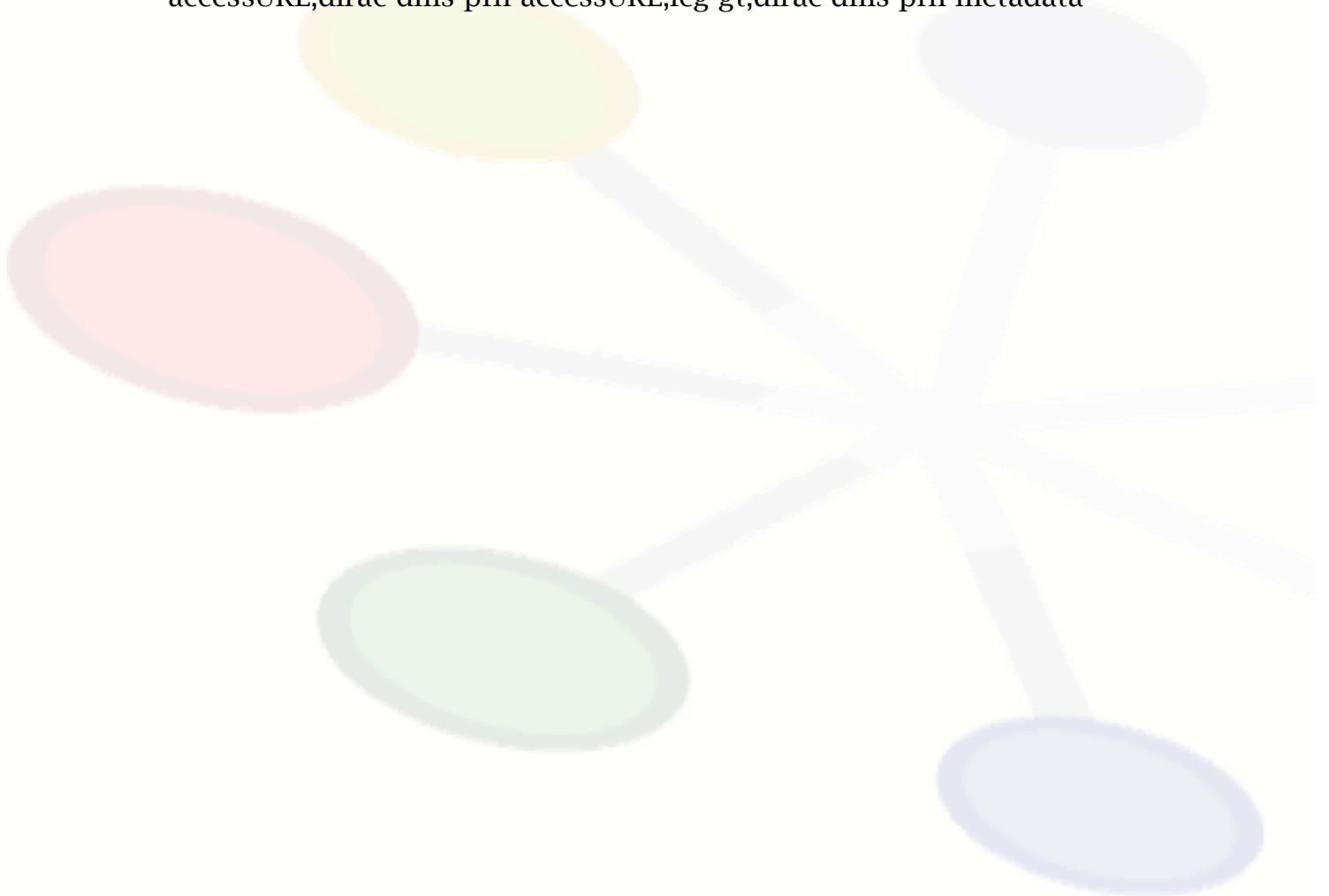
Date: 01/09/08

Author: Nick Brook

Version: 1.1

Description: Check the status of a Grid file if there are issues with data access

Tools To Use: `lcg-ls`; `dirac-dms-lfn-replicas`; `lcg-lr`; `dirac-dms-lfn-accessURL`; `dirac-dms-pfn-accessURL`; `lcg-gt`; `dirac-dms-pfn-metadata`



Sometimes there will be issues associated with data access during a production job. It is possible to make a set of preliminary checks on the problematic file using DIRAC tools or lcg utils. (Note: Currently only lcg-ls (of the lcg utils) is usable in the DIRAC environment.) Common access protocols used by LHCb are: rfiio (at RAL), root (at CERN and CNAF), gsidcap (at IN2P3, NL-T1, GridKa and PIC.) Often it is worth wrapping the lcg commands with “time” to measure the response time; the DIRAC commands usually have verbose reporting progress as they go on.

LHCb access their files using the Storage Resource Manager, SRM, at the Tier-1 centres. To control the data pools used at the sites the data is associated with a space token as given in the table below.

<u>Type of data</u>	<u>SRM spacetoken</u>
RAW	LHCb_RAW
RDST	LHCb_RDST
DST	LHCb_M-DST or LHCb_DST

The following is a list of useful LCG and DIRAC commands:

1. Which versions of lcg util & GFAL are being used?

```
lcg-ls --version
```

Example:

```
(DIRAC3-Production)[~/DIRAC3]$ lcg-ls --version
lcg_util-1.6.13
GFAL-client-1.10.15
```

2. What are the physical replicas of the logical file? Note the DIRAC command also return the DIRAC SE name.

```
dirac-dms-lfn-replicas <LFN>
lcg-lr lfn:/grid/<LFN>
```

DIRAC example:

```
(DIRAC3-Production)[brook@lxplus253 ~/DIRAC3]$ dirac-dms-lfn-replicas
/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0000069794.raw

2008-09-01 12:00:37 UTC dirac-dms-lfn-replicas/DiracAPI INFO: Replica Lookup
Time: 0.28 seconds
{'Failed': {},
 'Successful': {'/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0000069794.raw':
 {'CERN-RAW': 'srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0
000069794.raw'}}}
```

LCG example:

```
~ > lcg-lr
lfn:/grid//lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0000069794.raw

srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0
000069794.raw
```

3. Check whether a TURL can be generated at the site for data access. Note the DIRAC tools report the version of GFAL & lcg utils being used. The DIRAC SE name can be found following the procedure outlined above.

```
dirac-dms-lfn-accessURL <lfn> <SE>
dirac-dms-pfn-accessURL <pfn> <SE>
lcg-gt <pfn> <protocol> --st <space token>
```

LFN Example:

```
(DIRAC3-Production)[brook@lxplus253 ~/DIRAC3]$ dirac-dms-lfn-accessURL
/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0000069794.raw CERN-RAW

2008-09-01 11:08:10 UTC dirac-dms-lfn-accessURL INFO:
ReplicaManager.getReplicaAccessUrl: Attempting to get access urls for 1
replicas.
2008-09-01 11:08:10 UTC dirac-dms-lfn-accessURL INFO:
ReplicaManager.getReplicaAccessUrl: Resolving replicas for supplied LFNs.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO:
ReplicaManager.__getPhysicalFileAccessUrl: Attempting to get access urls for 1
files.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO: Using lcg_util from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/lcg_util.py
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO: The version of
lcg_utils is 1.6.13
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO: Using gfalthr from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/gfalthr.py
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO: The version of gfalthr
is 1.10.15
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO: StorageElement.isValid:
Determining whether the StorageElement CERN-RAW is valid for use.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO:
StorageElement.isLocalSE: Determining whether CERN-RAW is a local SE.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO:
StorageElement.getAccessUrl: Generating protocol PFNs for SRM2.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL INFO:
StorageElement.getAccessUrl: Attempting to get access urls for 1 physical
files.
2008-09-01 11:08:11 UTC dirac-dms-lfn-accessURL ERROR:
SRM2Storage.__gfal_turlsfromsurls: Failed to perform gfal_turlsfromsurls:
[SE][StatusOfGetRequest] http://srm-lhcb.cern.ch:8443/srm/managerv2: Can not
handle specified protocol Unknown error 18446744073709551615
2008-09-01 11:08:12 UTC dirac-dms-lfn-accessURL INFO:
StorageElement.getAccessUrl: Generating protocol PFNs for RFIO.
2008-09-01 11:08:12 UTC dirac-dms-lfn-accessURL INFO:
StorageElement.getAccessUrl: Attempting to get access urls for 1 physical
files.
{'Failed': {},
 'Successful': {'/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0000069794.raw':
{'RFIO':
'castor://castorlhcb:9002/?svcClass=lhcbraw&castorVersion=2&path=/castor/cern.
ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0000069794.raw'}}}}
```

PFN Example:

```
(DIRAC3-Production)[brook@lxplus253 ~/DIRAC3]$ dirac-dms-pfn-accessURL
srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0
000069794.raw CERN-RAW

2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO:
ReplicaManager.__getPhysicalFileAccessUrl: Attempting to get access urls for 1
files.
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO: Using lcg_util from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/lcg_util.py
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO: The version of
lcg_utils is 1.6.13
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO: Using gfalthr from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/gfalthr.py
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO: The version of gfalthr
is 1.10.15
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO: StorageElement.isValid:
```

```

Determining whether the StorageElement CERN-RAW is valid for use.
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO:
StorageElement.isLocalSE: Determining whether CERN-RAW is a local SE.
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO:
StorageElement.getAccessUrl: Generating protocol PFNs for SRM2.
2008-09-01 11:04:28 UTC dirac-dms-pfn-accessURL INFO:
StorageElement.getAccessUrl: Attempting to get access urls for 1 physical
files.
2008-09-01 11:04:29 UTC dirac-dms-pfn-accessURL ERROR:
SRM2Storage.__gfal_turlsfromsurls: Failed to perform gfal_turlsfromsurls:
[SE][StatusOfGetRequest] httpg://srm-lhcb.cern.ch:8443/srm/managerv2: Can not
handle specified protocol Unknown error 18446744073709551615
2008-09-01 11:04:29 UTC dirac-dms-pfn-accessURL INFO:
StorageElement.getAccessUrl: Generating protocol PFNs for RFIO.
2008-09-01 11:04:29 UTC dirac-dms-pfn-accessURL INFO:
StorageElement.getAccessUrl: Attempting to get access urls for 1 physical
files.
{'Failed': {},
 'Successful': {'srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0
000069794.raw': {'RFIO':
'castor://castorlhcb:9002/?svcClass=lhcbraw&castorVersion=2&path=/castor/cern.
ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0000069794.raw'}}}

```

LCG utils

```

~ > lcg-gt srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0
000069794.raw --st LHcb_RAW root

castor://castorlhcb.cern.ch:9002?svcClass=lhcbraw&castorVersion=2&path=/castor
/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0000069794.raw
3903832

```

4. What is the status of the file on the site? NEARLINE indicate a file is on an active tape whilst ONLINE indicates that the file has been staged to the MSS online cache. For data being accessed from an application the file should have already been staged ie it should be ONLINE.

```

dirac-dms-pfn-metadata <PFN> <SE>
lcg-ls -l <PFN>

```

DIRAC Example:

```

(DIRAC3-Production)[brook@lxplus253 ~/DIRAC3]$ dirac-dms-pfn-metadata
srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0
000069794.raw CERN-RAW

2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO:
ReplicaManager.__getPhysicalFileMetadata: Attempting to get metadata for 1
files.
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO: Using lcg_util from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/lcg_util.py
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO: The version of lcg_utils
is 1.6.13
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO: Using gfalthr from:
/afs/cern.ch/lhcb/software/DEV/DIRAC/DIRAC3_v0r4p0/Linux_x86_64_glibc-
2.3.4/lib/python2.4/site-packages/gfalthr.py
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO: The version of gfalthr
is 1.10.15
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO: StorageElement.isValid:
Determining whether the StorageElement CERN-RAW is valid for use.
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO:
StorageElement.isLocalSE: Determining whether CERN-RAW is a local SE.
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO:
StorageElement.getFileMetadata: Attempting to get metadata for 1 physical
files.
2008-09-01 12:11:54 UTC dirac-dms-pfn-metadata INFO:
StorageElement.getFileMetadata: No pfn's generated for protocol RFIO.
{'Failed': {},
 'Successful': {'srm://srm-
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHcb/CCRC/29581/029581_0
000069794.raw': {'Cached': 1,

```

```
'Directory': False,  
'File': True,  
'Migrated': 1,  
'Permissions': 511,  
'Size': 1511609407L}}
```

LCG example

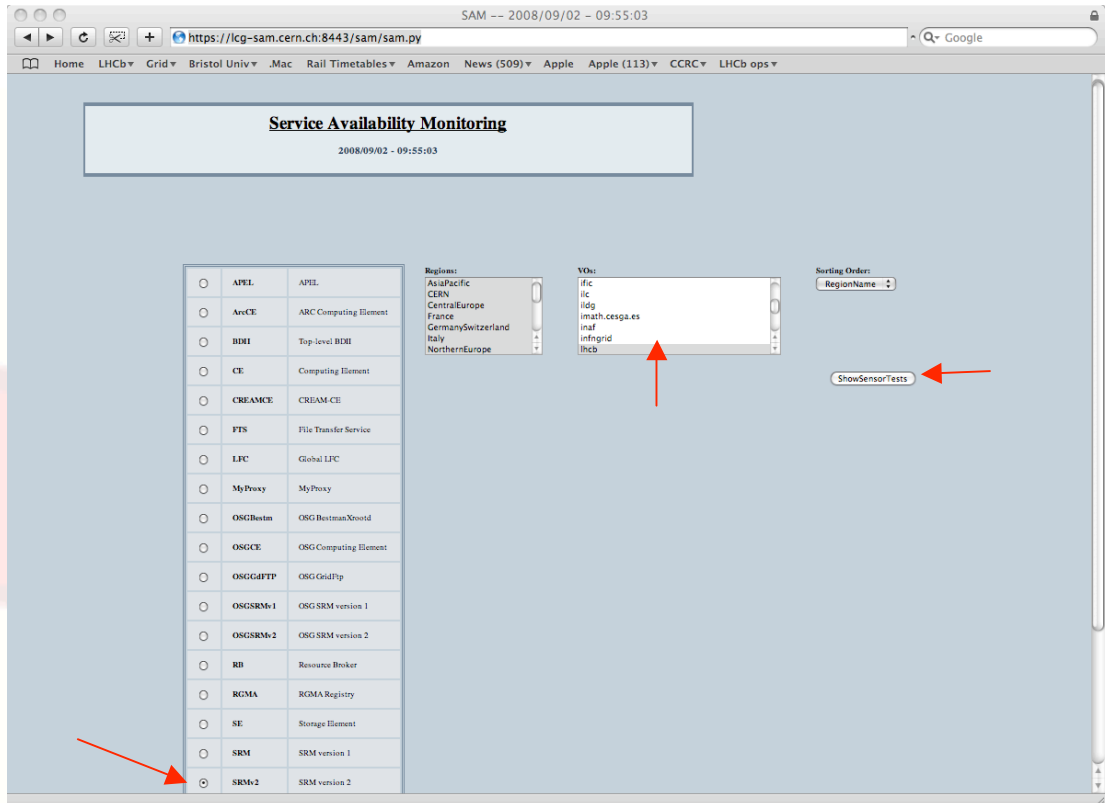
```
~ > lcg-ls -l srm://srm-  
lhcb.cern.ch/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581/029581_0  
000069794.raw  
  
-rwxrwxrwx 1 2 2 1511609407 ONLINE_AND_NEARLINE  
/castor/cern.ch/grid/lhcb/data/CCRC08/RAW/LHCb/CCRC/29581_000069794.ra  
w
```



Another useful page to look at that gives you the status of the SRM SE at each of the Tier-1 sites is the SAM web page

<https://lcg-sam.cern.ch:8443/sam/sam.py>

Select the “SRMv2” from the left hand table and “lhcb” from the VO list and press “ShowSensor Tests” button as indicated below



The output page should look similar to this (but hopefully with all sites passing!)

No	SiteName	NodeName	Status	Inch
CERN				
1	CERN-PH03	cern-03b01a01a0	OK	OK OK OK OK
France				
2	FRAN01	cern-01a01a01a0	ERROR	OK error error error
Germany/Switzerland				
3	DE-SW13	cern-01a01a01a0	OK	OK OK OK OK OK
Italy				
4	IT-NA13	cern-01a01a01a0	OK	OK OK OK OK OK
Northern Europe				
5	SE-SM13	cern-01a01a01a0	ERROR	OK error error error error
South Western Europe				
6	ES	cern-01a01a01a0	ERROR	OK error error error error
UK				
7	UK-LX13	cern-01a01a01a0	OK	OK OK OK OK OK