

Berkeley DRM version 1.2 in VDT

Instructions and trouble shooting guide

28 May 2004

**Junmin Gu, Vijaya Natarajan
Arie Shoshani, Alex Sim
Lawrence Berkeley National Laboratory**

**srm@data.lbl.gov
<http://sdm.lbl.gov/srm>**

Introduction:

Berkeley DRM version 1.2 is distributed in VDT as a static binary form.

It consists of the following binaries:

- C++:
 - o drmServer.linux,
 - o srm-copy.linux,
 - o srm-get.linux,
 - o srm-ls.linux,
 - o srm-ping.linux,
 - o srm-status.linux,
 - o nameserv.
- Java components (for Web Service access):
 - o HRM-WSG (HRM-Web Services Gateway)
 - hrm-wsg-server
 - hrm-wsg-get
 - hrm-wsg-put
 - hrm-wsg-unpin
 - o HRM-GWS (HRM-Gateway to Web Services)
 - hrm-gws-server

Notes:

- Fonts in **green** are what you can customize.
- Fonts in **orange** are DRM command presented here.
- Fonts in **blue** are outputs from the commands.

Brief guide on using URL protocol

1. To specify a file in a MSS in general:
 - Format: srm://nameserver_host:nameserver_port/your-file-path?remoteobj=remote-corba-obj-name &msshost=your-msshost&mssport=your_mss_port
Where the order of msshost, mssport and remoteobj is interchangeable.
 - For example:
“srm://dmx01.lbl.gov:4000/nersc/users/junmin/small_testfile?remoteobj=HRMServerOnLBNL&msshost=garchive.nersc.gov”
identifies a MSS file with path /nersc/users/junmin/small_testfile from an SRM whose naming service is running at dmx01.lbl.gov:4000, and the corba object name is “HRMServerOnLBNL”, and the MSS host is “garchive.nersc.gov”.
2. To specify a file in remote gridftp server:
 - Format: gsiftp://remotehost/filepath

- For example:
"gsiftp://dmx01.lbl.gov/tmp/test.file"

Chapter 1. DRM and DataMover

Instructions:

1. Preparation before running DRM:

Making sure the following software information is in your local environment

- DRM_HOME where all the DRM software reside
 - PATH to include \$DRM_HOME
- JAVA_HOME where JDK 1.4.x resides
 - PATH/LD_LIBRARY_PATH to reflect appropriate JDK

2. Running DRM in VDT

1) Run CORBA naming service on your machine:

- "\$DRM_HOME/nameserv -Oport 4000 &"
(You can use other ports than 4000.)

2) Create your srm.rc file. You can copy the sample from the manual, please change these entries to reflect your usage:

- common*NSHost=<the hostname you ran your nameserv in step 1>
- common*NSPort=<the port for your nameserv specified in step 1>
- Since you are not running TRM, so common*TRM.. entries can be omitted
- drm*objectname=<the object name you wish DRM to register as in the naming service.>
- drm*cachehome=<the disk path you want DRM to manage>
- drm*maxCacheSize = the size of disk you want DRM to use, in MB
- coordinate the maxCacheSize with maxSizeConncurrentPinnedFiles. Both are using MBs.
- drm*localGSIFTPhostname=<hostname where your gridftp server runs for the DRM disk cache>
- drm*useTRM = false <make sure it is false >
- (optional) assign a different file path to each of drm*logfile, drm*outputFile, drm*eventLogFile.
- to simply, turn off (by not specifying):
drm*appendExistingEventLogs ,
drm*prefetchingLimit,
drm*enableFileSharing,
drm*cacheFileNamingMode.

- 3) start DRM on the machine:
- `drmServer.linux srm.rc`
- If all your entries in srm.rc are ok, you will see ‘DRM is ready’ printed by drmServer.linux.
- 4) open another window, initiate a valid grid proxy. We will test srm-* command now. Your DN needs to be in the gridmap file where DRM is running. You need to supply a resource configuration file to these commands. You can create a newfile or adding entries to the srm.rc mentioned above. The commands srm-* will look for the configuration file ‘hrm.rc’ in the running directory automatically unless you specify your configuration file via ‘-conf’ option.
- In the following example commands, if not specified, we are using ‘hrm.rc’ in our environment.
- srm-ping: in the rc file, specify these entries:


```
hrm*NSHost=<your corba naming service host>
hrm*NSPort=<your corba naming service port>
hrm*HRMName= <the same value as drm*objectname>
```

 run “`srm-ping.linux -conf <your rc file name>`”
 output should be something like this:


```
Client CONFIGURATION
DRM is alive.
Thu Apr 29 18:45:27 2004
```
 - srm-ls:

sample command (listing the content of directory /tmp from the DRM host machine):

```
srm-ls.linux -s file:/tmp
```

if you want to view the content recursively, use the “-r” option

```
srm-ls.linux -s file:/tmp -r
```
 - srm-status:

supply request id through the -r option. User id will be obtained automatically through the proxy.

sample command:

```
srm-status.linux -r junmin@dmx.lbl.gov-40a3001e-0
```
 - srm-get:

bring a file from MSS to local disk (where client is)

sample command:

```
srm-get.linux -s gsiftp://dm.lbl.gov/tmp/anExampleToTest
-t file:/tmp/hii -b 131019086
```

Note: the caller is responsible for making sure the target path exists in the local disk. Srm-get will NOT create directories in client directory.

- **srm-copy:**
it has several capabilities, as outlined below.

To bring a file from remote location to local disk (through gsiftp):

- * source(remote url): e.g. gsiftp://dm.lbl.gov/tmp/medium.testfile, specified by “-s”
- * target: (can omit) local disk url: e.g. file:/tmp/ok.testfile specified by “-t”

If you do not provide the target url, the file is transferred into the DRM cache. The address in the cache will be printed out by srm-copy.

- * command: (bring a file from dm.lbl.gov to the default disk location)

```
srm-copy.linux -s gsiftp://dm.lbl.gov/tmp/medium.testfile -b 131019086
```

- * debug output:

```
WARNING: No entry "TARGETDIR" in the options....  
WARNING: Please provide target directory URL....  
WARNING: unless targetURL is in the input file....  
SOURCERMOD:  
gsiftp://dm.lbl.gov/tmp/medium.testfile ?&obj=DRMServerOnDmx01  
SRM-COPY: requestid=junmin@dm-4091b039-25cf50d  
SRM-COPY: FIDSET_length=1  
SRM-COPY: FID=gsiftp://dm.lbl.gov/tmp/medium.testfile  
SRM-COPY: STATUSCODE=14  
SRM-COPY: RETURN_URL=file:/home/users/junmin/medium.testfile  
SRM-COPY: STATUS_EXP=file has been transferred to the address as in  
returnUrl.  
MCOPY: gsiftp://dm.lbl.gov/tmp/medium.testfile  
FILE_TRANSFERRED_TO_CLIENT file has been transferred to the address as  
in returnUrl.  
SRM-COPY: FID=gsiftp://dm.lbl.gov/tmp/medium.testfile  
SRM-COPY: STATUSCODE=6
```

To bring a file from remote SRM to local disk:

Command:

```
srm-copy.linux -s "srm://dmx02.lbl.gov:4000/nersc/users/junmin/  
medium_test_2?msshost=garchive.nersc.gov&obj=HRMServer" –t  
file:/tmp/test5 -b 131019086 -conf drm.rc
```

Here srm-copy issues a request to DRM to contact remote SRM (located at dmx02.lbl.gov:4000/HRMServer) for a file in its HPSS . Once remote SRM has the file ready, DRM will ftp to the user specified location /tmp/test5.

Note: that “” is necessary here after –s to embrace ? and &.

To bring a file from remote to local disk through WSDL:

Need to install HRM-GSW. See below HRM-GSW manual for setups and instructions on how to communicate with other SRM Web services.

To bring a set of files from remote to local disk

Create a file, each line with entries in the following format:

Source_url<space>fileSizeInBytes<space>local_target

Where fileSizeInBytes and local_target are optional. Also, remember local_target should be of “file:” protocol.

Example of a line in such a file with source url, filesize, and local target all specified:

gsiftp://dm.lbl.gov/tmp/medium.test 131019086 file:/tmp/vdt
sample command:

srm-copy.linux -f <your_input_file>

NOTE:

- srm-copy option “-sd” is **not** supported for VDT usage.

Trouble shooting:

1. if your request got “Exceeded max allowed size”, it is possibly because:
 - a. you didn’t specify file size to srm-* , and it assumes 2GB by default and sent to DRM
 - b. in your DRM setting, the cache is less than 2GB or the entry drm*maxFileSize is less than 2GB.
If a. & b. fit your set up, you can either provide the actual size (if it is really less than your max file size) or increase the cache size to be more than 2GB.
2. If your locally compiled binary crashes immediately after starting, check your path for GCC and the GCC version that you compiled HRM and other dependencies.

Chapter 2. HRM-WSG

Introduction:

HRM-WSG is a web services gateway to DRM. It allows the interoperability from other SRM implementations.

Web service clients can send calls to HRM-WSG. HRM-WSG component will then forward the calls to DRM, and gets the result back to the user.

How to configure and run HRM-WSG

Requirements:

1. Java version - j2sdk1.4.x, please don't use j2sdk 1.5.0 beta.
2. Please make sure you have a valid proxy file.
3. Please run DRM or HRM before you start HRM-WSG.
4. Edit the \$DRM_HOME/config/hrm_wsg.rc file according to your setup.
The most important ones are as follows :
 - common*NSHost= *<the nameserver host your DRM is connected to>*
 - common*NSPort= *<the nameserver port your DRM is connected to>*
 - drm*objectname= *<your DRM object name, please look into DRM's hrm.rc file>*
 - wsg*CorbaPort= *<CORBA NS port>* e.g. 8090
 - wsg*RecachingMode=*<either true or false>*, default false.
 true => OVERWRITE mode is yes.
 false => OVERWRITE mode is no.
 - wsg*MSSHHostName = *<YOUR MSS HOST NAME>*
 e.g. garchive.nersc.gov
 - wsg*servesHRM = false
 - wsg*PutModeType=*<pull or push>*, default is push mode.
 HRM-WSG runs either on pull or push mode
 - if it runs on *push* mode, after a hrm-wsg-get call, DRM returns TURL, then the user should gsiftp the file by himself and then he can release the file by using hrm-wsg-unpin.
 - if it runs on *pull* mode, DRM will do the gsiftp and pull the file to the TURL location. user just needs to release the file by using hrm-wsg-unpin.
5. The file \$DRM_HOME/config/ogsa_wsg.rc is used to launch the web service for SRM. It contains the following parameters. Please edit the parameters according to your setup.
*wsg*HTTPHostName=<host_name>*
*wsg*HTTPHostIP=<host_ip_address>*

```

wsg*ProxyFile=<proxy_file_location>
wsg*GridMapFile=<grid_map_file_location>
wsg*HTTPPort=<web_server_port> e.g. 8080
wsg*HTTPGPort=<web_server_secure_port> e.g. 8443
wsg*ServiceHandle=<service_handle>
e.g. /ogsa/services/HRM_WSG
wsg*InvocationId=<invocation_instance_id> e.g. test

```

How to run HRM-WSG:

0. Edit the rc files according to the instruction above.
1. Please have JAVA_HOME and DRM_HOME environment variables in the shell.
 e.g. setenv \$JAVA_HOME /usr/local/j2sdk1.4.2_04
 e.g. setenv \$DRM_HOME /software/drm-1.2
 <parent directory where the binaries exists>
2. After that just type \$DRM_HOME/hrm-wsg-run.sh from the command prompt.
 - This step will launch the HRM-WSG server in the port that you Specified in the ogsa_wsg.rc file with the service handle and Invocation Id given in the ogsa_wsg.rc file.
 - Once the server is successfully launched, it deploys the required Web services (including ogsa grid related mandatory services) automatically and creates the instance (InvocationId) that you have given in the ogsa_wsg.rc file.
 It writes a WSDL file in this directory, usually named SERVICE_HANDLE+”_”+INVOCATION_ID.wsdl.
 - The server writes the debug log files in the \$DRM_HOME/ogs/server_*.log location.

Trouble shooting

1. If you receive GSSEception :
 Reason : May be your proxy is not valid
2. If you receive IOException : Chain Null
 - Reason :
 Server may be already running and the process is not exited properly, or port is not released.
 - Possible solution:
 Please kill the previous process and start again.
3. If you receive Corba error :
 - Reason :
 HRM-WSG did not bind to the correct DRM object.
 - Possible solution:
 Please make sure you have the correct *common*NSHost*, *common*NSPort*, *drm*objectname* in the config/hrm_wsg.rc file.

Chapter 3. Client programs to HRM-WSG

Section 1: How to run HRM-WSG client:

Requirements

1. JDK 1.4.x, not tested with java 1.5.x
2. Make sure you have a valid proxy.
If –proxyfile =<proxyfile path> is not provided in the \$DRM_HOME/properties/hrm-wsg-* .properties file, default user proxy is used. Otherwise the proxy file is read from the path and used for the authentication.
3. Please set JAVA_HOME and DRM_HOME environment variables in the shell.
e.g. setenv \$JAVA_HOME /usr/local/j2sdk1.4.2_04
e.g. setenv \$DRM_HOME <parent directory where the binaries exists>
4. DRM is up and running.
5. HRM-WSG is running.

Optional property setup for HRM-WSG clients.

1. parameters for srm_get are given in \$DRM_HOME/properties/params_wsg-get.properties.
2. parameters for srm_put are given in \$DRM_HOME/properties/params_hrm-wsg-put.properties
3. parameters for srm_unpin are given in \$DRM_HOME/properties/params_hrm-wsg-unpin.properties.
4. To overwrite the parameters contains in the \$DRM_HOME/properties/params_hrm-wsg-* .properties file, from the command line,
use the following options.
-host=<webservice hostname>
-port=<webservice portname>
-path=<webservice pathname>
-s=<source_file_name>
-sd=<path_to_input_file_name_contains_all_the_source_file_name>
5. Please see the \$DRM_HOME/samples/gws/ directory for some sample input files.

How to run HRM-WSG Client programs:

The following three programs take parameters from properties file or from command line. If command line is used, parameters from properties file are overwritten.

hrm-wsg-get

hrm-wsg-get –help
shows usage

hrm-wsg-get [-s <single_source> | -sd <multiple sources (path_to_input_file)>
supports request_to_get (for both single source and multiple sources)
e.g. hrm-wsg-get –host dmx.lbl.gov –port 8443
–path ogsa/services/HRM_WSG/test
–s srm://garchive.nersc.gov/nersc/vijayaln/medium.testfile
e.g. hrm-wsg-get –sd samples/wsg/test_get –host dmx.lbl.gov

hrm-wsg-get –status <request_id> --- Show status of the given request id
e.g. hrm-wsg-get –status 1 –host dmx.lbl.gov

hrm-wsg-get –est

[–s <single_source> | –sd <multiple sources (path_to_input_file)>
Show estGetTime of request_to_get
e.g. hrm-wsg-get –est –host dmx.lbl.gov
–s srm://garchive.nersc.gov/nersc/vijayaln/write/medium.testfile

hrm-wsg-put

hrm-wsg-put –help ---- shows usage

hrm-wsg-put [-s <single_source> -t <target_destination> -size <file_size> |
–sd <multiple sources (path_to_input_file)>
supports request_to_put (for both single source and multiple sources)
e.g. hrm-wsg-put
–s gsiftp://srm.lbl.gov/%2Fhome/data/junmin/testfiles/medium.testfile
–t srm://garchive.nersc.gov/nersc/vijayaln/medium.testfile2
–size 131019086
e.g. hrm-wsg-put
–sd samples/wsg/test_put

hrm-wsg-put –copy

[–s <single_source> -t <target_destination> -size <file_size> |
–sd <multiple sources (path_to_input_file)>
supports srm copy (for both single source and multiple sources)
e.g. hrm-wsg-put –copy –sd samples/wsg/test_put

hrm-wsg-put –est

[–s <single_source> -t <target_destination> -size <file_size> |
–sd <multiple sources (path_to_input_file)>
Show estPutTime of request_to_put
e.g. hrm-wsg-put –est –sd samples/wsg/test_put

hrm-wsg-unpin

hrm-wsg-unpin –help ----- show usage

hrm-wsg-unpin [-TURL <turl> | -sd <multiple turls (path_to_input_file)>
-unpin <request_id>

supports unpin (release) the given turl(s) for the request id.

e.g. hrm-wsg-unpin –host dmx.lbl.gov –unpin 3 –TURL
gsiftp://dmx.lbl.gov/data/test/89021_11795_1434477648

e.g. hrm-wsg-unpin –host dmx.lbl.gov –unpin 3 –sd samples/wsg/test_unpin

hrm-wsg-unpin –TURL <turl>

supports advisory delete the given turl(s).

e.g. hrm-wsg-unpin –TURL

gsiftp://dmx.lbl.gov/data/test/1085689021_11795_1434477648

Section 2: How to run Fermi SRM client with HRM-WSG server:

Please get Fermi SRM client tar file and unzip it.

Note: We do not provide or distribute Fermi SRM client.

Steps:

0. Please read \${FERMI_CLIENT}/srm/README
1. cd to \${FERMI_CLIENT}/srm directory.
2. setenv SRM_PATH \${FERMI_CLIENT}/srm directory.
3. Fermi SRM client uses \${HOME}/.srmconfig/config.xml.
4. Please copy \$DRM_HOME/hrm-wsg-fnal_templates/config.xml to \${HOME}/.srmconfig/config.xml

Please edit the values in the \${HOME}/.srmconfig/config.xml file.

The important parameters that needs to be edited are :

- webservices path=<the webservices handle and Invocation Id>,
e.g. /ogsa/services/HRM_WSG/test
- mapfile location=<the map file location>,
e.g. \${FERMI_CLIENT}/srm/conf/hrm-wsg-axis.map
- proxyfile location=<your proxy file location>
e.g. /tmp/x509.....

5. Please use hrm-wsg-fnal_templates/hrm-wsg-axis.map file.
 - Copy the hrm-wsg-fnal_templates/hrm-wsg-axis.map file to the \${FERMI_CLIENT}/srm/conf directory.

- Edit the \${HOME}/.srmconfig/config.xml file to point the correct map file path.
6. Some sample commands :
- **srmcp**
srm://dmx.lbl.gov:8443/nersc/users/vijayaln/medium_test?msshost=garchive.nersc.gov srm://dmx02.lbl.gov:8443/tmp/test.file
Will contact the target web service dmx02.lbl.gov to copy the file from SRM running on dmx.lbl.gov to the local machine dmx02.lbl.gov to the file location /tmp/test.file
 - **srmcp**
gsiftp://dmx17.lbl.gov:8443/garchive.nersc.gov/nersc/users/vijayaln/medium_test srm://dmx17.lbl.gov:8443/tmp/test.file
Get the file using gsiftp to dmx17.lbl.gov /tmp/test.file

Chapter 4. HRM-GWS

Introduction

HRM-GWS is a gateway component for DRM to other SRM web services implementations.

Requirements

1. JDK sdk 1.4 or higher, not tested with java sdk 1.5.x.
2. HRM-GWS must be registered with the same naming service host where DRM is registered. For example, if DRM is registered with dmx.lbl.gov:3000, then HRM-GWS should also be registered on the same naming service host.
3. The objectname for the HRM-GWS should be in this format:
YOUR_DRM_OBJEC_NAME + “WSDLBroker”

For example : if you name your DRM as DRMServer, then your HRM-GWS OBJECT_NAME should be DRMServerWSDLBroker.

4. Please edit the hrm.rc file

gws*objectname=<HRM-GWS object name>

e.g. DRMServerWSDLBroker (see 3 for details)

gws*WsdlPath=<WSDL service handle you want to talk to>

e.g. /ogsaservices/HRM_WSG/test

gws*UserKey=<your userkey.pem file path>

gws*UserTrustedCertificates=<your .globus/certificates dir path>

gws*UserCert=<your usercert.pem file path>

gws*debug=<true or false>

gws*log4j=<log4j properties file path>

e.g. logs/log4j_gws.properties

gws*MapFile=<map file location>

map file is used only when GWS contacts the glue type web service,
for axis type web service map file is not used.

How to run HRM-GWS

1. Please start DRM, HRM-WSG on site_1
2. Please set JAVA_HOME and DRM_HOME environment variables in the shell.
e.g. setenv \$JAVA_HOME /usr/local/j2sdk1.4.2_04
e.g. setenv \$DRM_HOME <parent directory where the binaries exists>
3. Type \$DRM_HOME/hrm-gws-run.sh to run the HRM-GWS server.
This step will start the HRM-GWS server. The logs are written in

- The \$DRM_HOME/logs/hrm-gws-* .log location.
4. (Optionall) Another set of DRM, HRM-WSG server can be run on another site site_2 to test the web services communications.
 5. HRM-GWS can be tested with the sample client hrm-wsg-client-run.sh
Please edit the \$DRM_HOME/properties/params_hrm-gws-client.properties file. For example in this case, it shoule be:


```
hostname=<the hostname of the webservice running on site_1>
portname=<the portname of the webservice running on site_1>
pathname=<web service handle running at site_1>
e.g. /ogsa/services/HRM-WSG/test
```
 6. HRM-GWS Client program supports single source file and set-of sources


```
source=<single_source_file_name_can_be_entered>
sourceDir=<path_to_input_file_name_contains_all_the_source_file_name>
```

 Please see the \$DRM_HOME/samples/gws directory for some input file samples.
 7. The parameters in the \$DRM_HOME/properties/params_gws.properties file can be overwritten by the command line options.


```
-host <webservice hostname>
-port <webservice portna me>
-path <webservice handle name>
-s <source_file_name>
-sd <path_to_input_file_name_contains_all_the_source_file_name>
```
 8. Some sample SourceUrl's are given.
 Simple srm: type urls:
`srm://garchive.nersc.gov/nersc/users/vijayaln/medium_test`
 Wsdl type srm urls are used to forward the calls to another HRM-WSG running on another site:
`srm://dmx.lbl.gov:8443/nersc/users/vijayaln/medium_test?msshost=garchive.nersc.gov`
 gsiftp:// type url's:
`gsiftp://srm.lbl.gov/%2Fhome/data/testfiles/medium_test`