

# Grid Storage Access

A thick, horizontal orange brushstroke underline that tapers at both ends, positioned below the title.

Neha Sharma

FermiGrid

Fermi National Accelerator Laboratory

Work supported by the U.S. Department of Energy under contract No. DE-AC02-07CH11359

# Course Contents

- ✓ Briefly discuss concept of Storage on Open Science Grid
- ✓ Discuss client side tools available to access the storage
- ✓ Perform Lab Exercises
- ✓ Monitoring tools

# Disclaimer



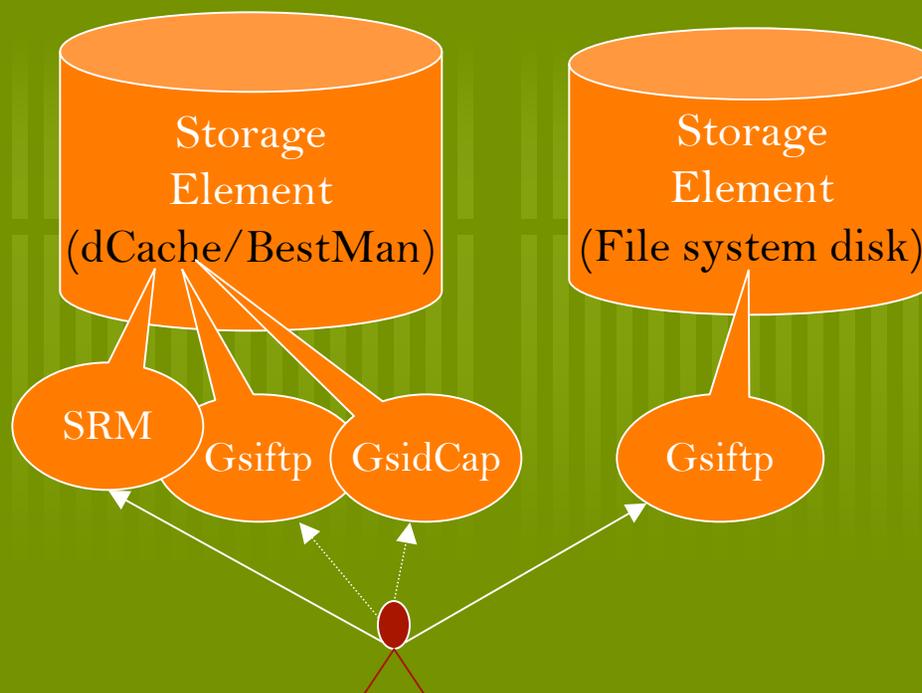
I am not the developer of any of these tools

I am a user who has more experience  
using/testing most of these tools

# Storage on Grid - High Level View

## Open Science Grid - Common, Shared, Distributed infrastructure

- Storage in OSG is represented in form of a Storage Element
- The owners of the resources provide both guaranteed and opportunistic use of managed storage.
- Go through one of the “doors”
- Each door has utility for transferring data b/w your machine & /pnfs/.. area on dCache node



# Bottom Line

- ✓ There is storage space on the Grid for you to use
  
- ✓ To use it, you need -
  - VO membership
  - Authorization
  - Valid credentials
  - Knowledge on how to use client side tools (provided by OSG Client software)

A horizontal yellow brushstroke is located in the upper left quadrant of the slide.

# Client Side Tools

# globus-url-copy

Client for requesting transfers to, from or between GridFTP servers.

```
> globus-url-copy [options] <srcUrl> <destUrl>
```

```
> globus-url-copy [options] -f <filename>
```

**<srcUrl>** may contain wildcard characters \* ? and [ ] character ranges in the filename only.

Any url specifying a directory must end with a forward slash '/'

If **<srcUrl>** is a directory, all files within that directory will be copied.

**<destUrl>** must be a directory if multiple files are being copied.

# Use cases on Grid

Case	Source	Destination
Local disk -> Remote Storage	<u>file://&lt;path&gt;</u> file:///<path>	gsiftp://<host:port>/<path>
Remote Storage -> Local disk	gsiftp://<host:port>/<path>	file://<path> file:///<path>
Remote Storage -> Remote Storage	gsiftp://<host:port>/<path>	gsiftp://<host:port>/<path>



Where:

**host** – is fully qualified host name of the node where gsiftp server is running

**port** – is the port number on which gsiftp server is listening for client requests

**path** – is the source/destination location of the data

# dccp

Client for dCap (dCache native access protocol). It provides a cp-like functionality on a PNFS filesystem (used by dCache).

```
> dccp [options] <source> <destination>
```

- ✓ <source> must be a single file while the <destination> could be a directory name as well as a file name.
- ✓ If <destination> is a directory, a new file with the same name as the source name will be created there and the contents of the source will be copied.
- ✓ If the final destination file exists in pnfs, it won't be overwritten and an error code will be returned.

# Use cases on Grid (Not used often)

Case	Source	Destination
Local disk -> Remote Storage	File location. Eg. /home/neha/testfile	gsidcap://<host:port>/ <path>
Remote Storage -> Local disk	gsidcap://<host:port>/<path>	File location. Eg./home/neha/testfile



Where:

**host** – is the fully qualified host name of the node where  
gsidcap/dcap server is running

**port** – is the port number on which gsidcap/dcap is listening  
for client requests

**path** – is the source/destination location of the data

# LAB 1



A horizontal yellow brushstroke is located in the upper left quadrant of the slide.

# Srmclient commands

## Some useful things to know ...

- ✓ SRM has 2 implementations - V1.1 and V2.2
- ✓ There is difference in what works with V1.1 and V2.2
- ✓ User (with a certain set of grid credentials) is authorized to read/write data from/to a designated storage location.
- ✓ This location is called “**root path**”

Example - On Fermilab Public dCache, the root path for users of “fermilab” VO is `/pnfs/fnal.gov/usr/fermigrid/volatile/fermilab`

# What works with which Version?



dCache  
version

SRM version

Supported  
srmclients

1.7.0-\*

1.1

srmcp, srm-advisory-delete,  
srm-get-metadata, srmping,  
srm-get-request-metadata

1.8.0-\*

2.2 & 1.1

All srmclients

srm-advisory-delete is  
deprecated in V2

# srmcp

Perform data transfers to/from or between SRM servers

```
> srmcp [options] <srcUrl> <destUrl>
```

```
> srmcp [options] -copyjobfile <file>
```

**file** – is a file containing <srcUrl> <destUrl> mappings

At least one of the srcUrl or destUrl **MUST BE** an srmUrl - i.e of form  
`srm://<host:port>/[webservicepath?SFN=]/<path>`

# Use cases on Grid

Case	Source	Destination
Local disk-> Remote Storage	file:/// <path>	srm:// <host:port> / [webservice path?SFN=] / <path>
Remote Storage -> Local disk	srm:// <host:port> / [webservice path?SFN=] / <path>	file:/// <path>
Remote Storage1 -> Remote Storage 2	srm:// <host:port> / [webservice path?SFN=] / <path> gsiftp:// <host:port> <path>	gsiftp:// <host:port> <path> srm:// <host:port> / [webservice path?SFN=] / <path>

**Host** - hostname of the srm/gsiftp server node

**Port** - port on which srm/gsiftp server is listening for client requests

**Path** - source/destination location of data

# srmls, srmmv, srmrm

**srmls** - List contents of a directory

```
> srmls [options] <srmUrl>
```

**srmmv** - Move file/directory from one location to another

```
> srmmv [options] <srcUrl <destUrl>
```

```
> srmrm [options] <srmUrl>
```

srcUrl/srmUrl/destUrl is of form - srm://<host:port>/[webservicepath?SFN=]/<path>

# srmmkdir, srmrmdir

**srmmkdir** – Create a new directory

```
> srmmkdir [options] <srmUrl>
```

New dir inherits permission mask from parent directory

**srmrmdir** – Remove an existing directory

```
> srmrmdir [options] <srmUrl>
```

Can remove directories recursively as long as they are all empty

srmUrl is of form - srm://<host:port>/[webservicepath?SFN=]/<path>

# Space Reservation



Space reservation is a promise by the storage system to make certain amount of storage space of certain type available for usage for a specified period of time.

Available in V2.2

# High level view

Step 1 - User makes a Space Reservation Request



Step 2 - If approved, SRM Server gives back a Space Token



Step 3 - User uses this Space Token to copy files in the reserved area.



# srm-reserve-space

Reserve space in advance for the upcoming requests

```
> srm-reserve-space [options] a1=v1 a2=v2 a3=v3 a4=v4 a5=v5 <srmUrl>
```

<srmUrl> - srm://<host:port>

a[i]

retention\_policy

access\_latency

lifetime

desired\_size

guaranteed\_size

v[i]

REPLICA, CUSTODIAL

ONLINE, NEARLINE

(in seconds)

(in bytes)

(in bytes)

**Note** : This function will only work if Space Reservation is enabled, configured correctly and your VO is authorized to reserve space.

# srmcp using space token

```
> srmcp [options] -space_token=?? <srcUrl> <destUrl>
```

```
> srmcp [options] -space_token=?? -copyjobfile <file>
```

**file** – is a file containing <srcUrl> <destUrl> mappings

Atleast one of the srcUrl or destUrl **MUST BE** an srmUrl -  
i.e of form srm://<host:port>/[webservicepath?SFN=]/<path>

# srm-release-space

Release an occupied space.

```
> srm-release-space [options] -space_token=?? <srmUrl>
```

```
<srmUrl> - srm://<host:port>
```

# LAB 2



# Monitoring tools

- ✓ srm-get-request-status (V1.1)
- ✓ SRM monitoring tool called “srmwatch”

# Fermilab Public dCache

- ✓ dCache version - 1.7.0-39. Only V1.1 srmclients are supported
- ✓ Port numbers
  - ✓ Kerberized dCap - 24725, 24736
  - ✓ Plain dCap - 24136, 24125
  - ✓ GSI dCap - 24536, 24525
  - ✓ SRM - 8443
  - ✓ GSI Ftp - 2811

## Request from developers

*We would like to get feedback on your experience with various srmclients and improvements you would like to see in future releases.*

*Thank you.*

Please fill in the “srmclients feedback” form before you leave.

