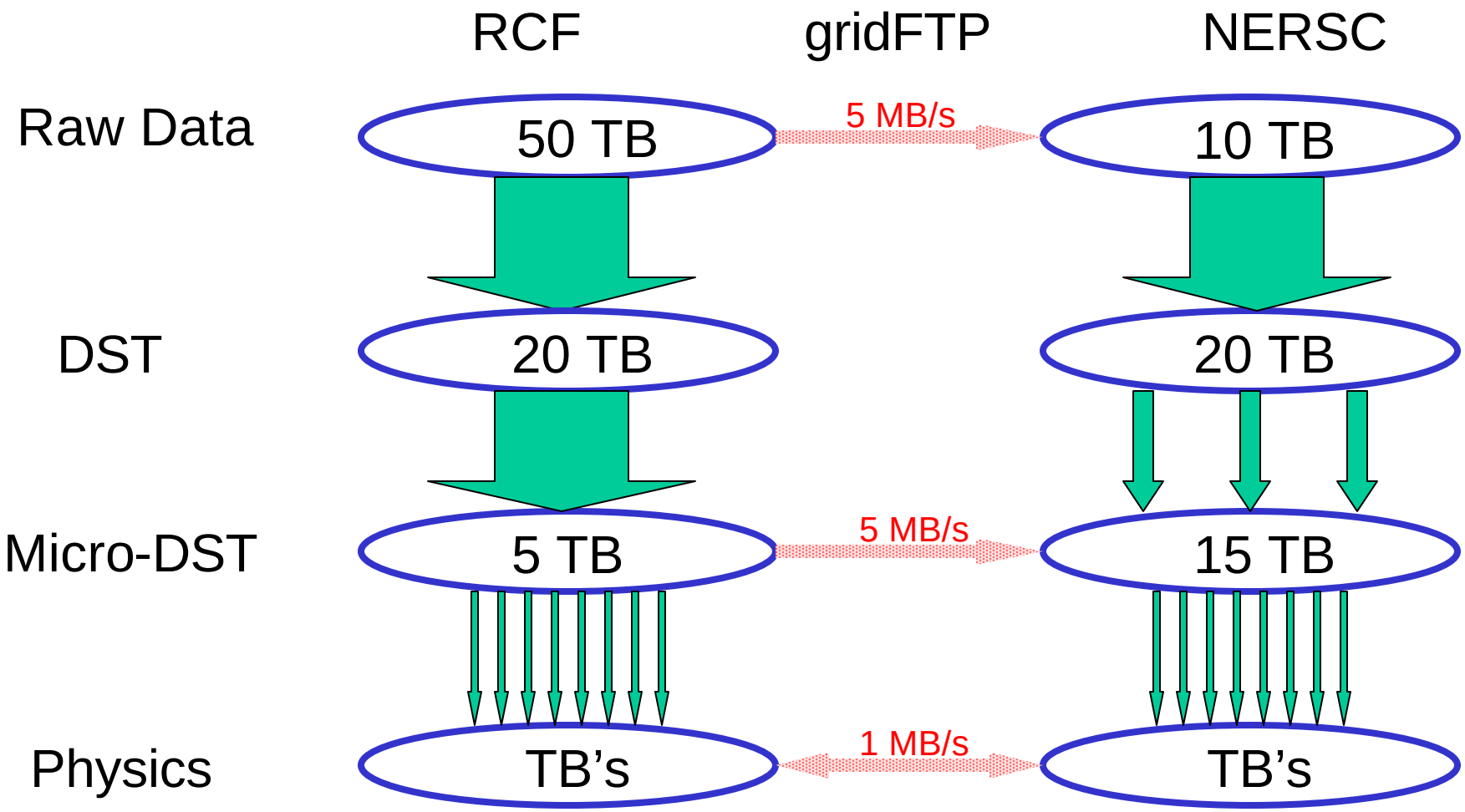
HRMs in STAR

Eric Hjort, LBNL
(STAR/PPDG Collaborations)

STAR HRMs: Overview

- **STAR has computing facilities at BNL & LBL (RCF, PDSF at NERSC)**
  - **STAR computing model: Replicate all DSTs, and some raw data from RCF to PDSF**
- **First run in 2000:**
  - **Produced 5 TB raw data, a few TB's DST**
- **Second run ended Jan. 2002**
  - **Produced ~50 TB raw data, MuDst's ~ 5 TB each production**
- **Third run to begin Jan. 2003**
  - **Will need to transfer ~1 TB/week or more**
- **HRM status**
  - **In production use for STAR data transfers since 7/02**
  - **Testbed for HRMs and gridFTP**

STAR Data Processing: 2002 run



What are HRMs?

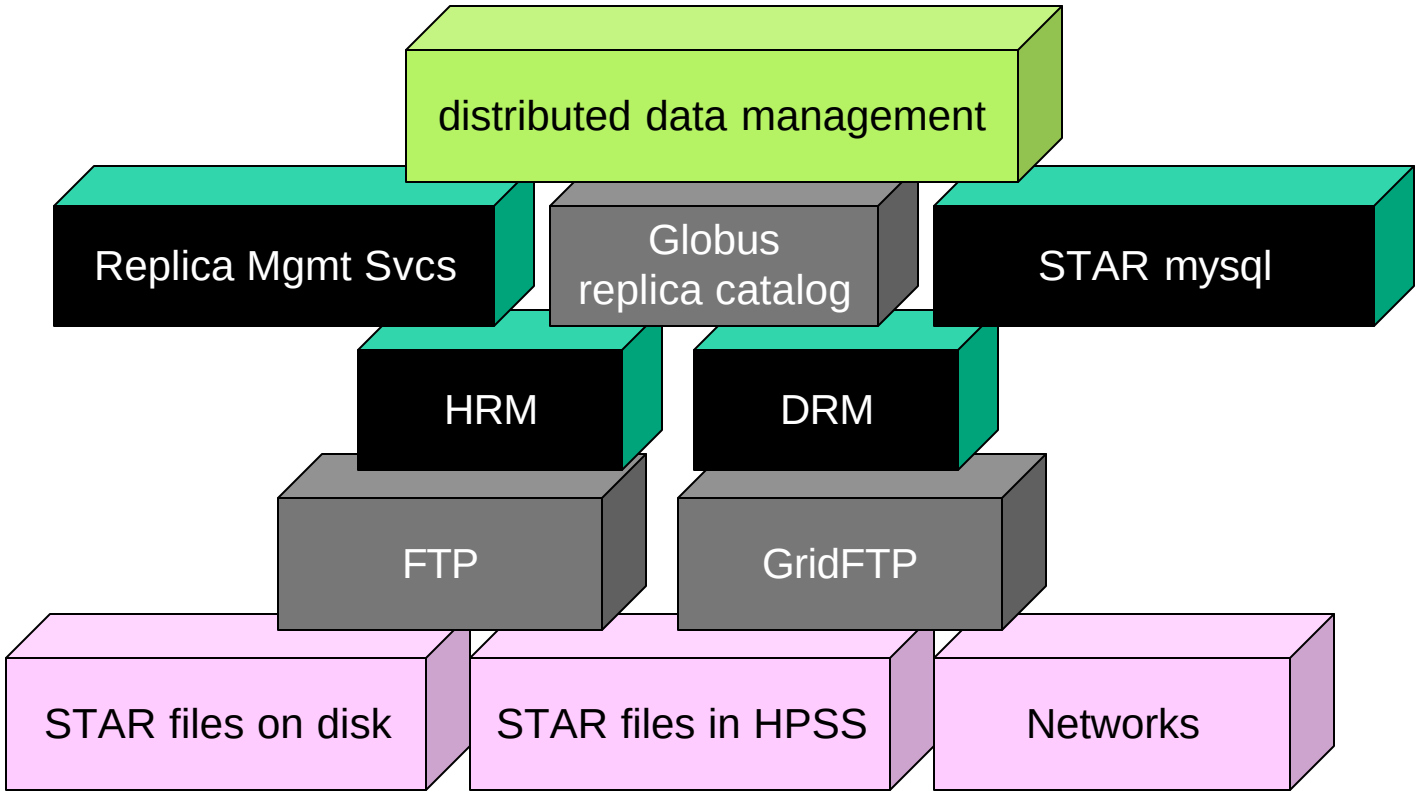
- Grid middleware developed by the Scientific Data Management Resource Group, LBNL
- Example of SRM (Storage Resource Managers)

  Disk Resource Managers (DRMs)

  Tape Resource Managers (TRMs)

  Hierarchical Resource Managers (HRM = TRM+DRM)
- API and CLI are supported
- For STAR the HRMs make the 3-step transfer HPSS->disk->disk->HPSS look like a simple 1-step copy.

Components of STAR site-to-site data distribution



DGRA
Application

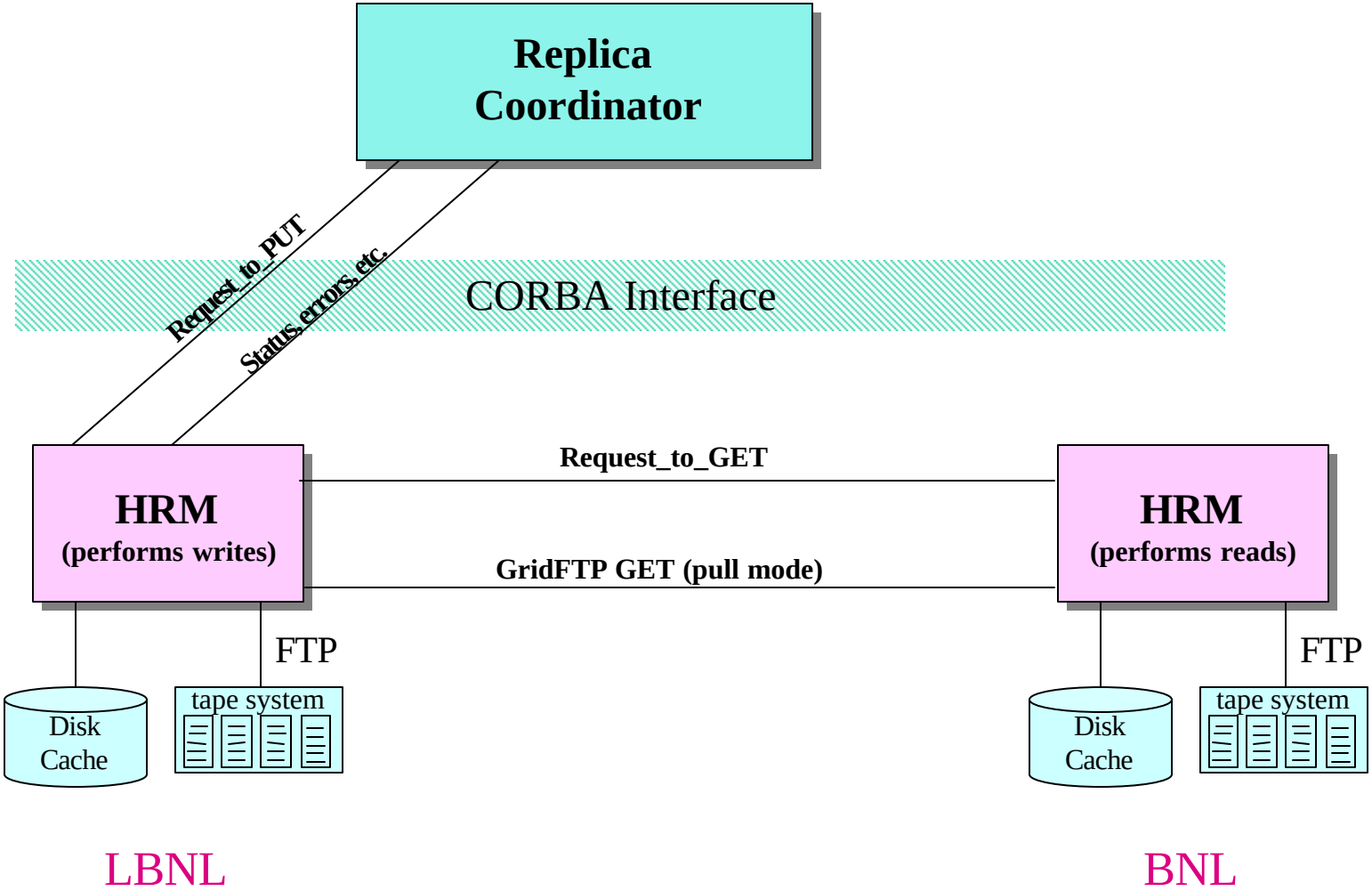
Collective

Resource

Connectivity

Fabric

File Replication: BNL to LBNL



Starting HRM servers

Instructions for running HRM's at RCF and PDSF
Eric Hjort
10/22/02

General:
HRM working directory at both locations is ~hjort/hrm
Configuration file is hrm.rc.

To get things running at RCF:
Open two windows on stargrid02.rcf.bnl.gov.
cd to the HRM working directory in each.
In first window, clean the cache disk: perl cleancache.pl
In first window, start nameserver if not already running: nameserv -OAport XXXX &
In first window, start TRM (use starpftp account): trm.linux
In second window, start DRM: drmServer.linux hrm.rc -OAport XXXX

To get things running at PDSF:
Open four windows on pdsfgrid2.nersc.gov
cd to the HRM working directory in each.
Load necessary modules in each: source setup
In first window, clean the cache disk: perl cleancache.pl
In first window, start nameserver if not already running: nameserv -OAport XXXX &
In first window, start TRM (use starofl account): trm.linux
In second window, start DRM: drmServer.linux hrm.rc
In third window, start FMT (optional): fmt.linux

To transfer files from RCF to PDSF:
In fourth window get your proxy: grid-proxy -init
Execute the hrm-mcopy command contained in the .cmd file of your choice.
(For this first try use starsink_2001_2237034.cmd)
To monitor with FMT look at http://pdsfgrid2.nersc.gov:6099/fmt_applet_eric.html
…or open yet another window and do "perl rate.pl" in the HRM directory.

HRM Configuration file

```
##########################################################################
# HRM.RC: configuration file for the HPSS Resource Manager
# Jan. 2001
##########################################################################
common*NSHost=pdsfgrid2.nersc.gov
common*NSPort=XXXX
common*HRMName=HRMServerPDSF
common*TRMName=TRMServerPDSF
common*TRMHost=pdsfgrid2.nersc.gov
common*TRMRefFile=/export/data/star/hrmdata/hrm-var/trm.ref
common*VerboseConfig=true
##########################################################################
#Client
hrm*HRMLogFile=/export/data/star/hrmdata/hrm-var/out.client.log
hrm*EnableLogging=true
##########################################################################
# DRM options
# will be changed to share some info
drm*objectname=HRMServerPDSF
drm*cachehome=/export/data/star/hrmdata/hrm-cache
drm*maxCacheSize=200GB
drm*maxConcurrentFTP=6
drm*concurrencyLevel=30
drm*localGSIFTPhostname=pdsfgrid2.nersc.gov
drm*logfile=/export/data/star/hrmdata/hrm-var/log.cache.hrm
drm*useTRM=true
drm*NumFTPStreams=1
drm*FTPBufferSize=1000000
#drm*FTPBlockSize=500000
drm*maxSizeConcurrentPinnedFiles=30GB
drm*maxNumConcurrentPinnedFiles=500
drm*pinDuration=2
#drm*outputFile=hrm.output
drm*showDebugMessages=true
##########################################################################
# Make sure you have "w" permission in the directories files are put.
# for HSI, path to the hsi is needed. If the path is in the env variable,
# just putting nothing will be enough. For example, hrm*HSI=
# If HSI does not exist on local system, remove the entry,
# and it'll not be used. If the entry is listed, HSI will be used.
trm*HSI=/usr/local/bin/hsi
trm*PFTP=/usr/local/bin/pftp
trm*DiskPath=/export/data/star/hrmdata/hrm-cache
trm*HPSSHostName=archive.nersc.gov
# HPSSHSIHostName is needed for -h and -p
# if not provided, HPSSHostName will be used
trm*HPSSHSIHostName=archive.nersc.gov:XXXX
# user name and password does not have to be set
# if those two entries are not there, it'll ask during the intial setup
#trm*HPSSUserName=HPSS_user_name
#trm*HPSSUserPassword=HPSS_user_passwd
trm*SupportedProtocols=gsiftp;file
trm*PFTPMaximumAllowed=2
trm*TRMListFilesStaged=/export/data/star/hrmdata/hrm-var/out.listfilesstaged.log
trm*TRMLogFile=/export/data/star/hrmdata/hrm-var/out.trm.log
trm*EnableLogging=true
trm*EnableRecovery=false
trm*RecoveryPath=/export/data/star/hrmdata/hrm-var/trmcache
##########################################################################
hrm*PFTP=/usr/local/bin/pftp
hrm*HRMLogFile=/export/data/star/hrmdata/hrm-var/out.client.log
hrm*EnableLogging=true
hrm*HPSSHostName=archive.nersc.gov
#hrm*HPSSUserName=
#hrm*HPSSUserPassword=
hrm*PFTPMaximumAllowed=2
##########################################################################
fmt*RMIHost=pdsfgrid2.nersc.gov
fmt*RMIPort=XXXX
fmt*RMIServerName=HRMFMTServerPDSF
fmt*WEBPort=XXXX
fmt*FMTPort=XXXX
fmt*URLCodePath=/usr/local/pkg/HRM/hrm
### Give the log file size in KB, 1000000 = 1 MB
fmt*LOGFileSize=1000000
fmt*LOGFileCount=3
fmt*NumFilesPerThread=10
fmt*FrequencySecForMonitor=60
fmt*FMTLogLoc=/export/data/star/hrmdata/hrm-var
##########################################################################
```

Prepare for transfer

Do grid-proxy-init:

```
[hjort@pdsfgrid2 ~/hrm]$ grid-proxy-init
Your identity: /O=Grid/O=Lawrence Berkeley National Laboratory/OU=PDSF/CN=Eric Hjort/Email=ELHjort@lbl.gov
Enter GRID pass phrase for this identity:
Creating proxy ..................................................... Done
Your proxy is valid until Sat Jan 25 02:17:08 2003
```

Run daqman.pl at PDSF:

-uses globusrun to execute the script daqFileLister.pl at RCF which queries file catalog:

```
my @query = `globusrun -s -r "$remoteHost" –f $rslDir/$dataSet.rsl`;
```

- returns list of files and sizes (GASS)
- formats file list and command file for hrm-mcopy

File transfer example

Starsink_2003_4022058.cmd:

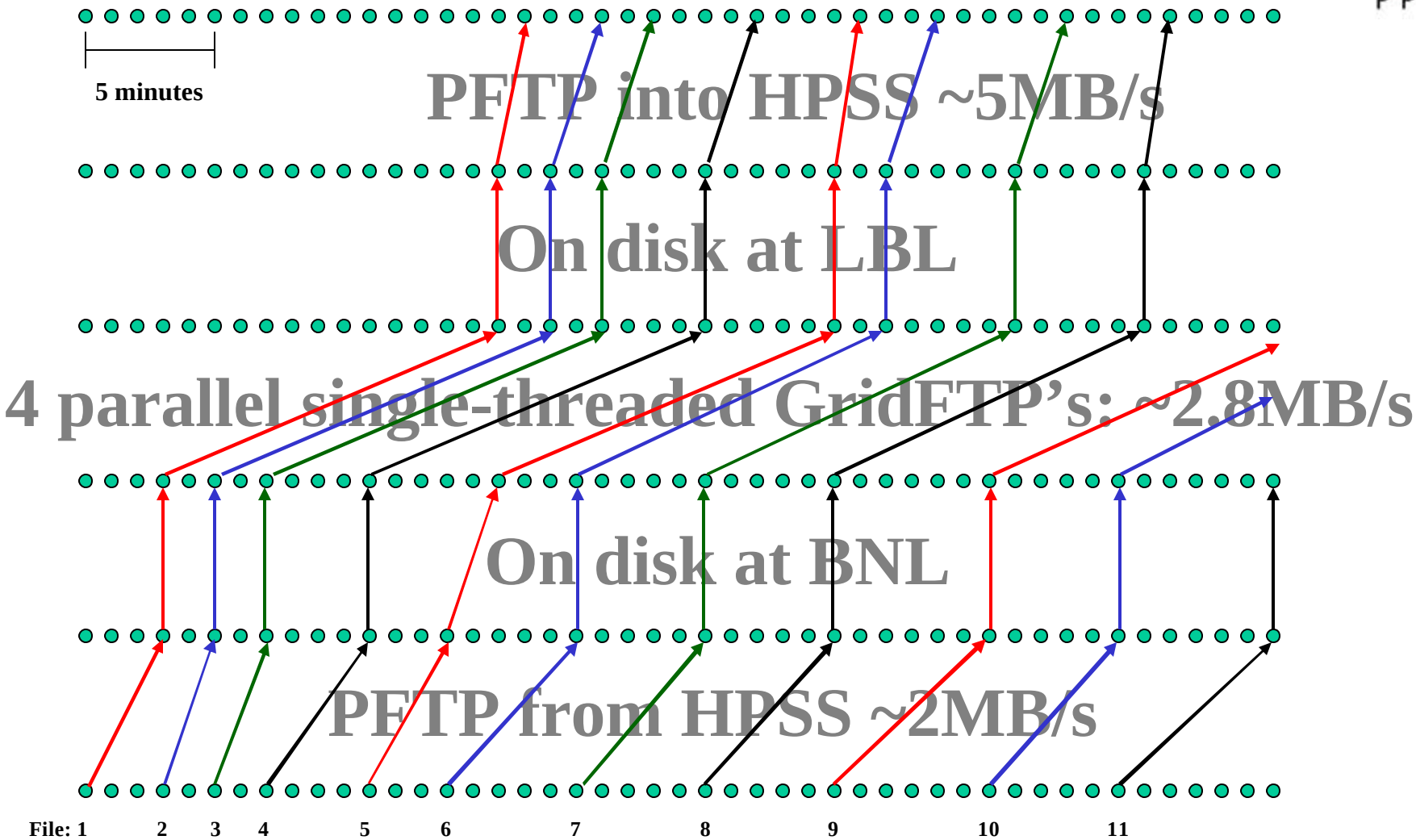
```
hrm-mcopy.linux -d -f starsink_2003_4022058.txt -c hrm.rc
-t gsiftp://archive.nersc.gov/nersc/projects/starofl/raw/daq/2003/022
-l starsink_2003_4022058.log
```

Starsink_2003_4022058.txt:

```
hrm://stargrid02.rcf.bnl.gov:XXXX/DRMServerOnHRM/hpss.rcf.bnl.gov:XXXX/h
ome/starsink/raw/daq/2003/022/st_physics_4022058_raw_0040010.daq 524886016
hrm://stargrid02.rcf.bnl.gov:XXXX/DRMServerOnHRM/hpss.rcf.bnl.gov:XXXX/h
ome/starsink/raw/daq/2003/022/st_physics_4022058_raw_0040011.daq 525139968
```

Other CLI commands: hrm-copy, hrm-get, hrm-mget hrm-put, hrm-mput, hrm-ls, hrm-status

File transfer schematic



Files are ~525MB each.

Transfer summary

```
-----------------------------------------------------
ProductionMinBias/FullField/P02gc/2001
  Total: Files = 2219 Size = 72224427667
  Done: Files = 2219 Size = 72224427667
-----------------------------------------------------
ProductionMinBias/FullField/P02gd/2001
  Total: Files = 5230 Size = 170838965222
  Done: Files = 5205 Size = 169899231310
-----------------------------------------------------
ProductionMinBias/ReversedFullField/P02gc/2001
  Total: Files = 2481 Size = 78236912931
  Done: Files = 2475 Size = 78236446184
-----------------------------------------------------
ProductionMinBias/ReversedFullField/P02gd/2001
  Total: Files = 9968 Size = 280379132628
  Done: Files = 1042 Size = 30676838504
-----------------------------------------------------
MinBiasVertex/ReversedFullField/P02ge/2001
  Total: Files = 11255 Size = 272904825041
  Done: Files = 11255 Size = 272904825041
-----------------------------------------------------
Central/FullField/P02gc/2001
  Total: Files = 356 Size = 14315337753
  Done: Files = 0 Size = 0
-----------------------------------------------------
productionCentral/FullField/P02gc/2001
  Total: Files = 1835 Size = 82546770872
  Done: Files = 499 Size = 22297630054
-----------------------------------------------------
productionCentral/FullField/P02gd/2001
  Total: Files = 3332 Size = 135130583115
  Done: Files = 3037 Size = 123185490802
-----------------------------------------------------
productionCentral/FullField/P02ge/2001
  Total: Files = 16405 Size = 647215181998
  Done: Files = 9206 Size = 351395522801
-----------------------------------------------------
productionCentral/ReversedFullField/P02gc/2001
  Total: Files = 1111 Size = 48512747722
  Done: Files = 1108 Size = 48377975087
-----------------------------------------------------
productionCentral/ReversedFullField/P02gd/2001
  Total: Files = 0 Size = 0
  Done: Files = 0 Size = 0
-----------------------------------------------------
productionCentral/ReversedFullField/P02ge/2001
  Total: Files = 23348 Size = 985288790151
  Done: Files = 19193 Size = 809058765945
-----------------------------------------------------
productionCentral600/FullField/P02gc/2001
  Total: Files = 3681 Size = 141749835842
  Done: Files = 0 Size = 0
-----------------------------------------------------
productionCentral600/ReversedFullField/P02gc/2001
  Total: Files = 3430 Size = 130355176779
  Done: Files = 0 Size = 0
-----------------------------------------------------
productionCentral1200/FullField/P02gc/2001
  Total: Files = 1127 Size = 44852010396
  Done: Files = 0 Size = 0
-----------------------------------------------------
productionCentral1200/ReversedFullField/P02gc/2001
  Total: Files = 497 Size = 19707009843
  Done: Files = 0 Size = 0
-----------------------------------------------------
Summary:
  Total Files = 86275 Size = 3124257707960
  Total Done: Files = 55239 Size = 1978257153395
```

An example MuDst's being transferred. 2 TB and 55k files transferred, 1.1 TB and 30k files to go when this summary was made.

This transfer was completed in 8/02.

Ongoing/future work

- File catalog at PDSF
- HRM/File catalog interface
- Transfers PDSF -> RCF
- Bandwidth improvement
- Globus RLS (Replica Location Service)
- Increased automation