

Data Replication in LIGO

Kevin Flasch

for the LIGO Scientific Collaboration

University of Wisconsin-Milwaukee

- LIGO and LIGO Scientific Collaboration
- Basic Data Challenge
- Specific Problems & Challenges
- LDR
- LSCdataFind
- Successes
- Warts
- Future of LIGO
- Future of LDR

Facility dedicated to detection and use of cosmic gravitational waves

Two sites: Livingston, LA and Hanford, WA

Three interferometers: Two in Hanford, one in Livingston

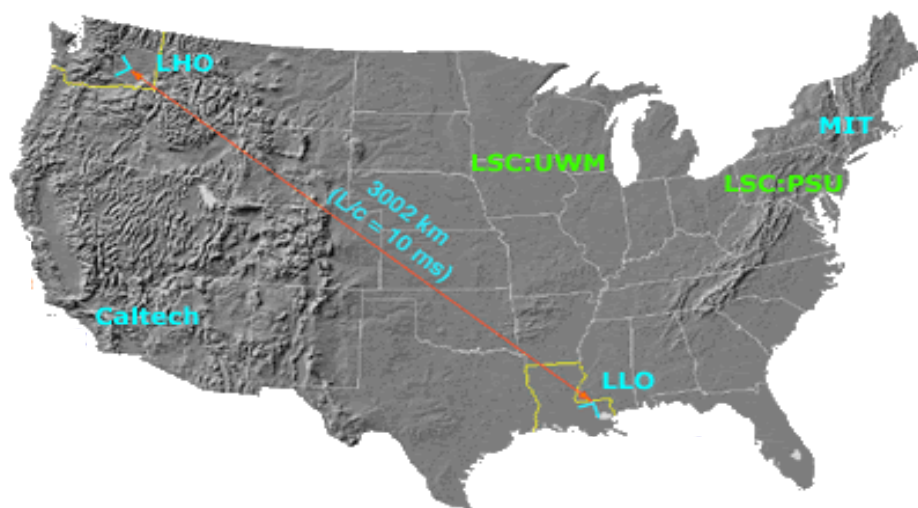
Partnership with Virgo (Italy and France) and GEO (Germany and the United Kingdom)

LIGO is supported by the NSF

4 km LIGO interferometer in Livingston, LA



The LIGO Scientific Collaboration (LSC) currently includes 428 people at 52 different institutions; data replication mainly occurs at Caltech, MIT, interferometer sites Livingston and Hanford, UWM, Penn State, Albert Einstein Institute (Germany), Cardiff (UK), Birmingham (UK)



- Basic issue is to distribute approx. one TB raw data / day to all sites
 - » Data is continually generated at both interferometer sites (LLO and LHO) during “science runs” - long periods of uninterrupted data collection ; current is S5 and has lasted over a year and a half
 - » Caltech (CIT) retrieves the data from the LHO and LLO sites and provides access to it for “Tier-2” sites (all sites besides CIT, LLO and LHO)
 - » Tier-2 sites replicate from CIT or other sites that have already transferred desired data
 - » Processed data sets (e.g., filtered or calibrated) are occasionally created at various sites. They are initially replicated from the site of origin.

Metadata Service

- » Require all data to be described in some fashion by a specific metadata schema
- » Metadata must be generated continually during a science run
- » Must be able to distribute metadata constantly and consistently to each site that needs it
- » Example of some metadata fields
 - gpsStart: 815497955 (seconds since beginning of GPS epoch)
 - gpsEnd: 815498048
 - runTag: S5
 - frameType: H1_RDS_C03_L2
 - md5: 28329c0eee60dbbde352a1ba94bca61f

- Storage of data
 - » Each site has their own in-house storage solution
 - most have some configuration of commodity hard disk drives, CIT uses SAM-QFS (disk and tape)
 - local filesystems and layout may differ as well, for example:
 - UWM uses 24 NFS-mounted storage servers
 - Cardiff stores on 100 compute nodes
 - CIT has one large filesystem with SAM-QFS
 - » Must provide a way for administrators to store incoming data on their systems in a customizable way

- Data is not distributed equally
 - » Sites must be able to pick and choose what particular data they want to replicate
 - Driven by users requests
 - » Sites must be able to tell what specific data another site has in order to replicate what it itself needs
- Users need to locate and access data
 - » Computing clusters at all sites; users may be at any one of them
 - » Users must find which sites have the data they want
 - » They must be able to locate and have their computing jobs able to locate the physical location of data at a certain site

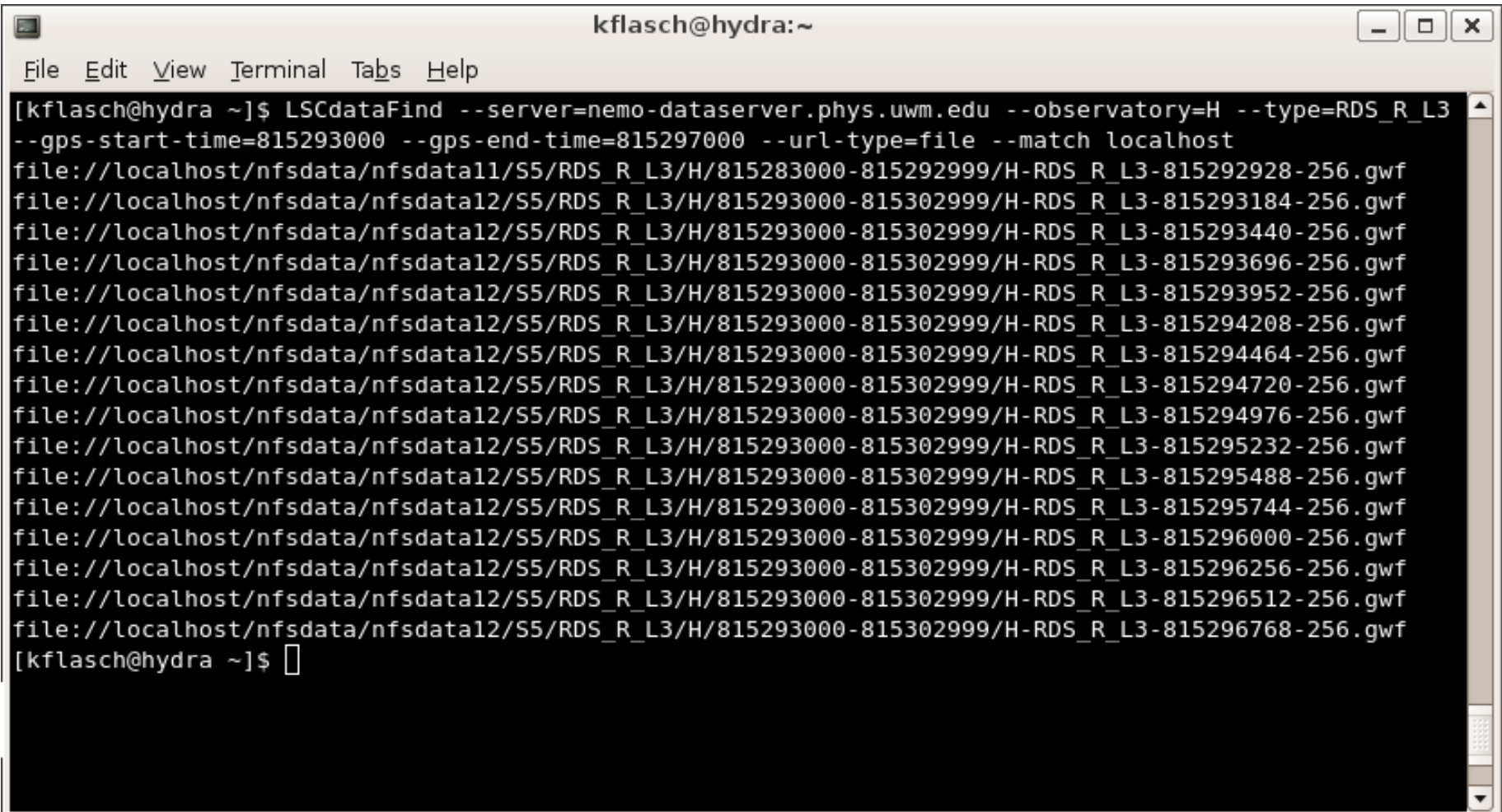
- LDR – Lightweight/LIGO Data Replicator - was created to solve these problems
 - » Lightweight: minimal code base wrapped around other services
 - » LIGO: code is based around LIGO's needs
 - » What data we have
 - custom metadata service
 - » Where data is located
 - Globus RLS
 - » Authenticated, fast data transfer
 - custom GridFTP client, standard server
 - » Ease of data transfer
 - easy for administrators to pick and choose data to replicate and data to make available



- LDR runs at each site as a few separate daemons
 - » LDRMaster : monitors other daemons
 - » LDRSchedule: finds and schedules files for transfer
 - » LDRTransfer: supervises transfer and storage of files
 - » LDRMetadataServer: serves local metadata to other sites
 - » LDRMetadataUpdate: updates local metadata database
- Relies on a few other important pieces: MySQL, Globus RLS (Replica Location Service), Globus GridFTP Server, pyGlobus (python port of Globus Toolkit)

- Each site fulfills certain roles
 - » some publish new data, some provide data, some replicate data (or any combination)
 - » new data is published into metadata catalog and RLS for other sites to replicate
- Local storage
 - » each site has its own storage solution
 - » administrator modifies a “local storage module” to govern how incoming data will be stored and recorded
 - functions like `newHoldingFile()`, `enterFile()`, `newFileCallback()`, `failedTransferCallback()`

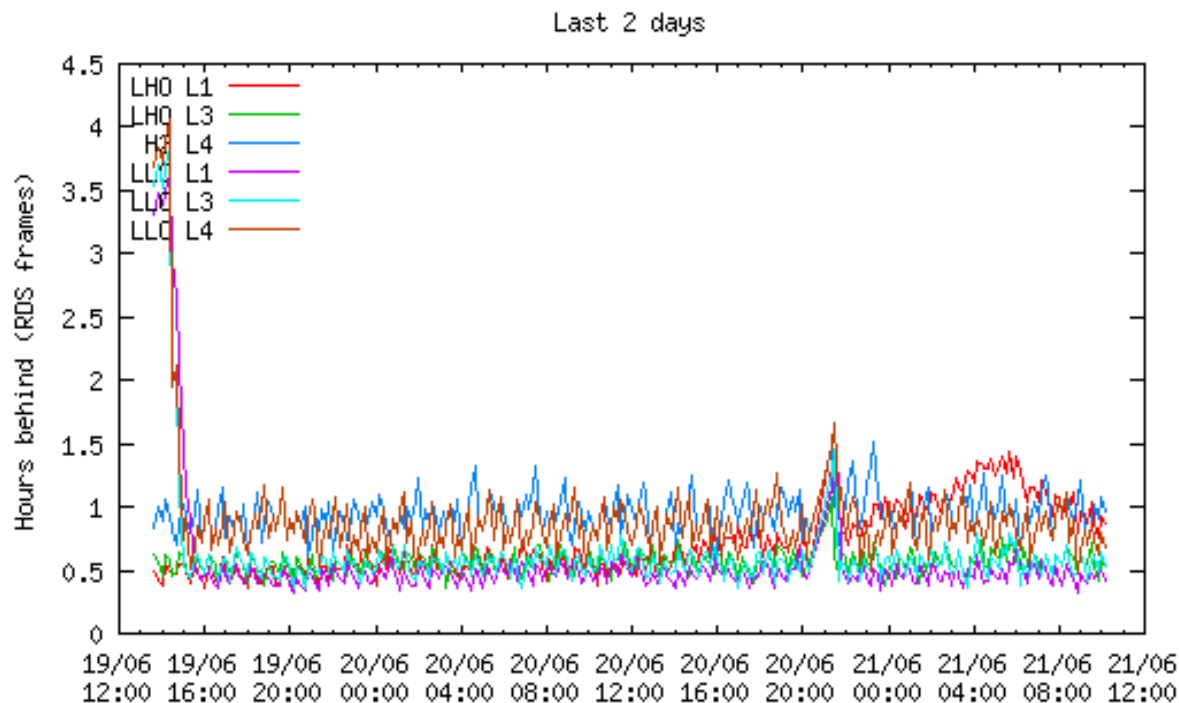
- Needed a way for users to easily find available data
- Work already done for LDR itself to find data to replicate to other sites, so a user tool was based on the LDR backend: LSCdataFind
- Uses a local RLS and Metadata service to allow users to specify characteristics about data they want (metadata fields like gpsStart, for example) and receive usable physical locations

A screenshot of a terminal window titled 'kflasch@hydra:~'. The window has a menu bar with 'File', 'Edit', 'View', 'Terminal', 'Tabs', and 'Help'. The terminal shows the execution of the 'LSCdataFind' command with various options. The output lists 15 files, each with a full path to a .gwf file on the local host. The files are named with a combination of directory paths, time ranges, and file identifiers. The terminal ends with a prompt for the user to enter another command.

```
kflasch@hydra:~  
File Edit View Terminal Tabs Help  
[kflasch@hydra ~]$ LSCdataFind --server=nemo-dataserver.phys.uwm.edu --observatory=H --type=RDS_R_L3  
--gps-start-time=815293000 --gps-end-time=815297000 --url-type=file --match localhost  
file://localhost/nfsdata/nfsdata11/S5/RDS_R_L3/H/815283000-815292999/H-RDS_R_L3-815292928-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815293184-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815293440-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815293696-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815293952-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815294208-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815294464-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815294720-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815294976-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815295232-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815295488-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815295744-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815296000-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815296256-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815296512-256.gwf  
file://localhost/nfsdata/nfsdata12/S5/RDS_R_L3/H/815293000-815302999/H-RDS_R_L3-815296768-256.gwf  
[kflasch@hydra ~]$
```

- Replicated over 770 TB of raw and processed S5 data so far
- Reliable (good enough) transfer rates (10-15 MB/s CIT -> UWM)
- Usable tool (LSCdataFind) for users to locate data at sites
- Small core development team
- Involved community
- Dependable, in production software!

Plot of time delay of transfer of data from interferometer sites to CIT for further Tier-2 replication



- No 24/7 reliability
 - » Issues coping with sites going down
 - » Unintelligent backend doesn't determine best/other places to go
- Had issues with RLS reliability (problems addressed thanks to the RLS team!)
- Not very user/administrator friendly
 - » Relies on learning much new terminology and software and support from the LSC community
 - » Interface is clumsy and obfuscated

- Next data run S6 is slated to begin in June of 2009
 - » LDR must be able to scale to amount of data it will need to track and replicate
- Enhanced and Advanced LIGO
 - » Enhanced LIGO (S6) will increase the sensitivity of the interferometers
 - » Advanced LIGO will greatly increase the sensitivity and therefore replication and storage requirements for all new data
 - » Advanced LIGO will also likely involve increased demand for greater turnaround in specific data replication

- Move Metadata daemons to WSRF-compliant services, probably built on Globus Java WS core
- Integrate Lots Of Small Files / pipelined GridFTP
 - » We replicate many big files, but increasingly more small files such as user processed ones ; pipelining will help us maintain good transfer rates
- Improve monitoring by leveraging Globus MDS 4
- Investigate integrating Globus RFT and Globus DRS
- Focus on stability and scaling...

- Metadata
 - » about 17,800,000 files tracked at CIT currently; We have managed to continue scaling our metadata services to this point
 - » Starting to feel strain and will need to cope with scaling much higher for S6
- Data transfer
 - » Current data rates are acceptable and will continue to be
 - » No worries about scaling with GridFTP; only limitation is network
- User demands
 - » Currently, we are able to handle user requests for data location
 - » Expect more users, more queries and faster expected response time

- Current Development Team
 - » Stuart Anderson, Gerald Davies, Kevin Flasch, Filippo Grimaldi, Steffen Grunewald, Ben Johnson, Scott Koranda, Dan Kozak, Greg Mendel, Brian Moe, Murali Ramsunder, David Stops, Igor Yakushin
- Alumni
 - » Bruce Allen, Paul Armor, Keith Bayer, Patrick Brady, Junwei Cao, Mike Foster, Tom Kobiialka, Adam Mercer
- More information
 - » LIGO: <http://www.ligo.caltech.edu/>
 - » UWM LSC: <http://www.lsc-group.phys.uwm.edu/>
 - » LDR: <http://www.lsc-group.phys.uwm.edu/LDR/>