



Any Data, Anytime, Anywhere

Frank Würthwein (UCSD)

Disclaimer

Giving this talk for Brian Bockelman, who is on jury duty.
Work done by Brian, with some help from Matevz, ...

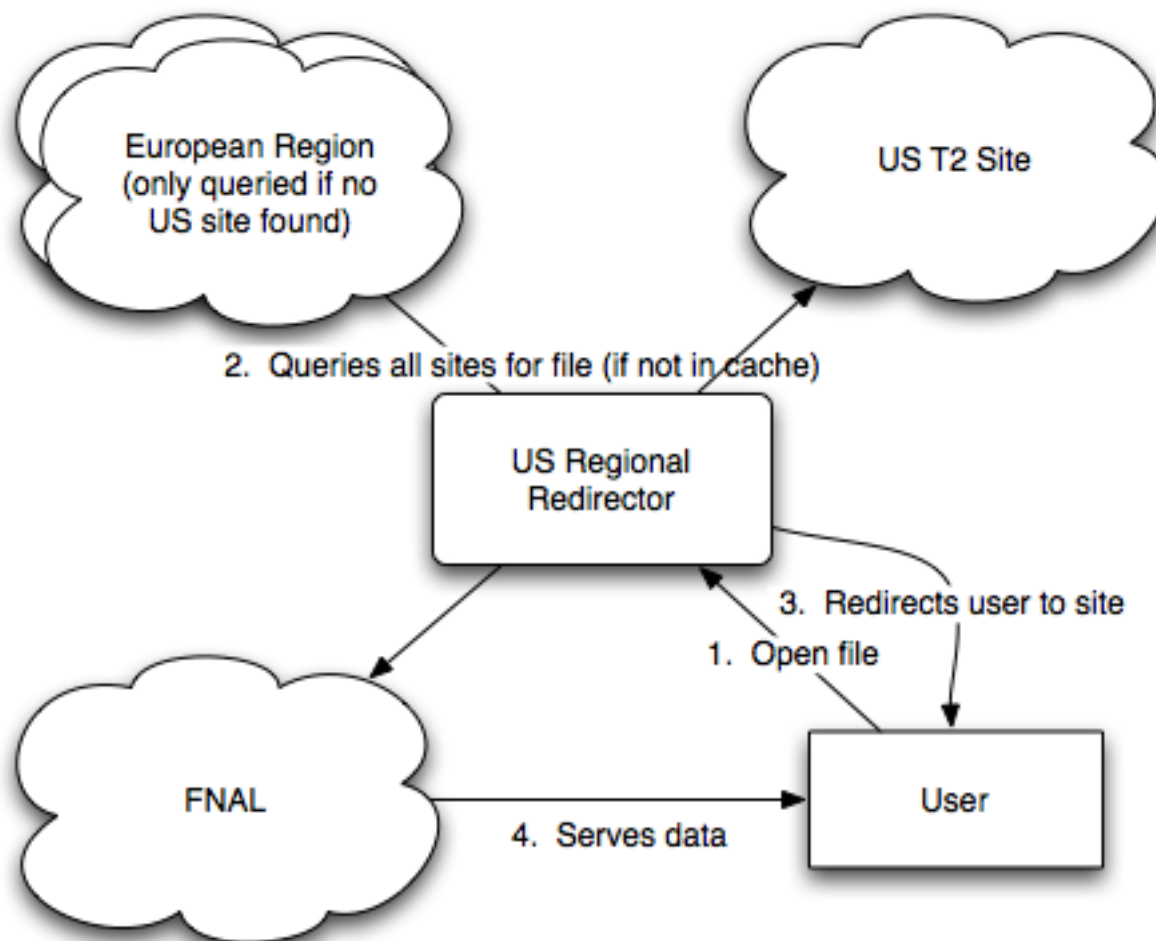
Overview

- Goals & Architecture & Use Cases
- Status
- Documentation
 - for T2
 - for T3
 - for users
- Future Directions

Goals

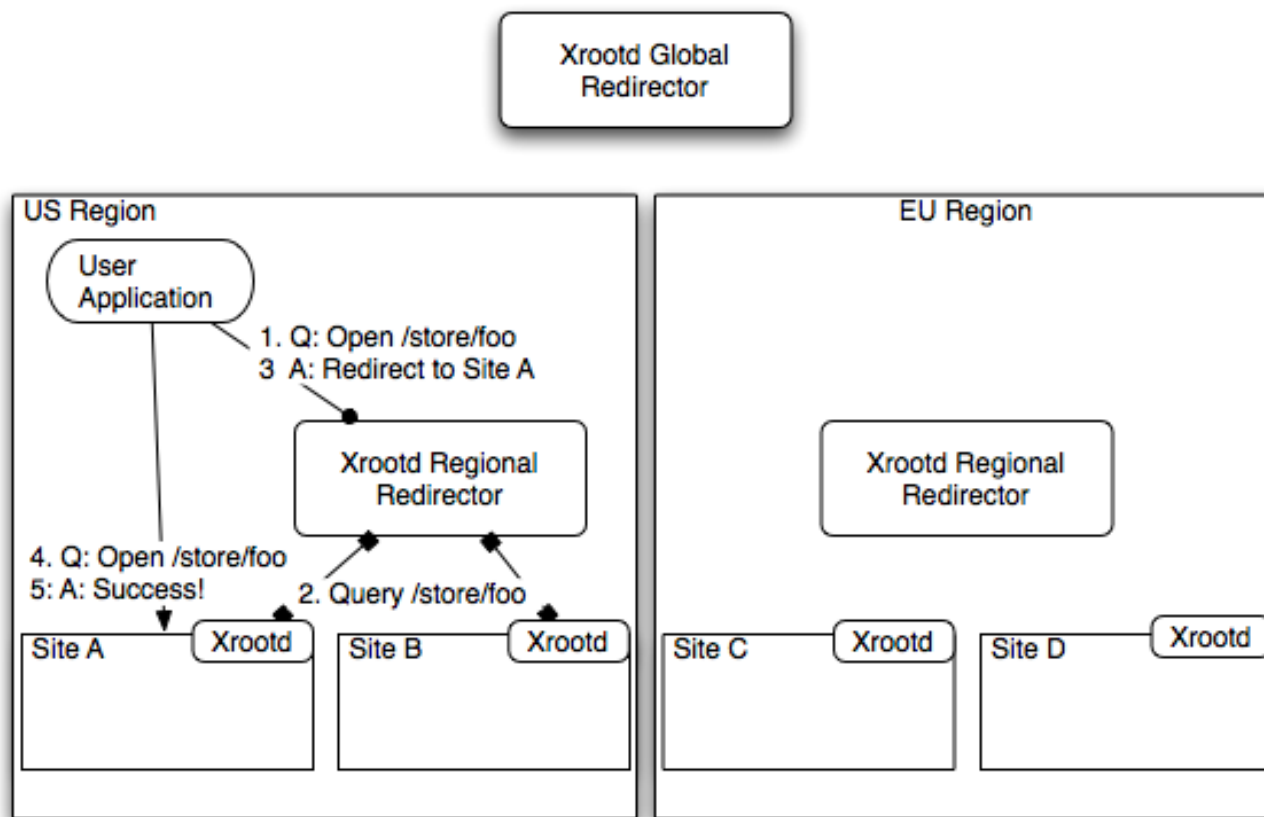
- Have any data available from anywhere at anytime.
- Focus on:
 - Physicists from their laptops/desktops
 - Debugging, fireworks access to all events
 - Failover for grid jobs when file that was supposed to be locally available is actually not locally available.
 - Site storage is down, dataset was deleted while job was waiting.
 - Allow for T3 without phedex
 - Access files at T3 that are not local
 - Serving files that are not within phedex namespace
 - Any user files, both within SEs or homedisks or ...

Global Xrootd Federation



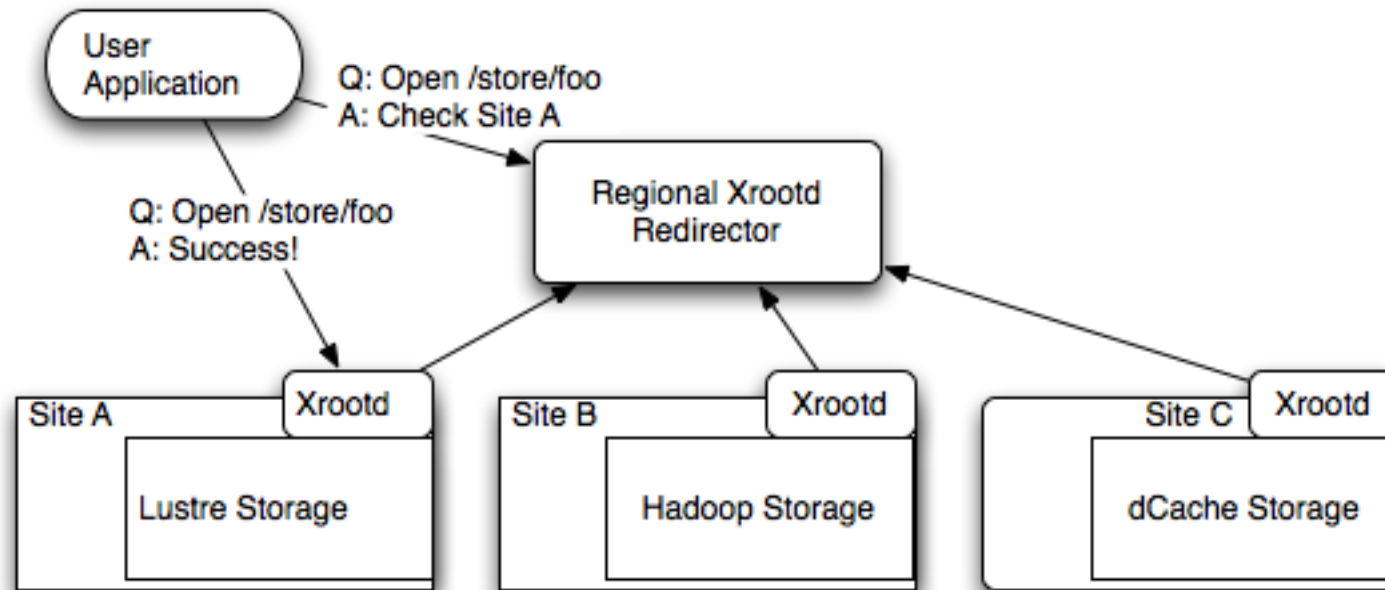
Long Term Goal

Architecture (I)



Deploying the global system in terms of federated regional systems

Regional Architecture (II)



Interactively, always access redirector.

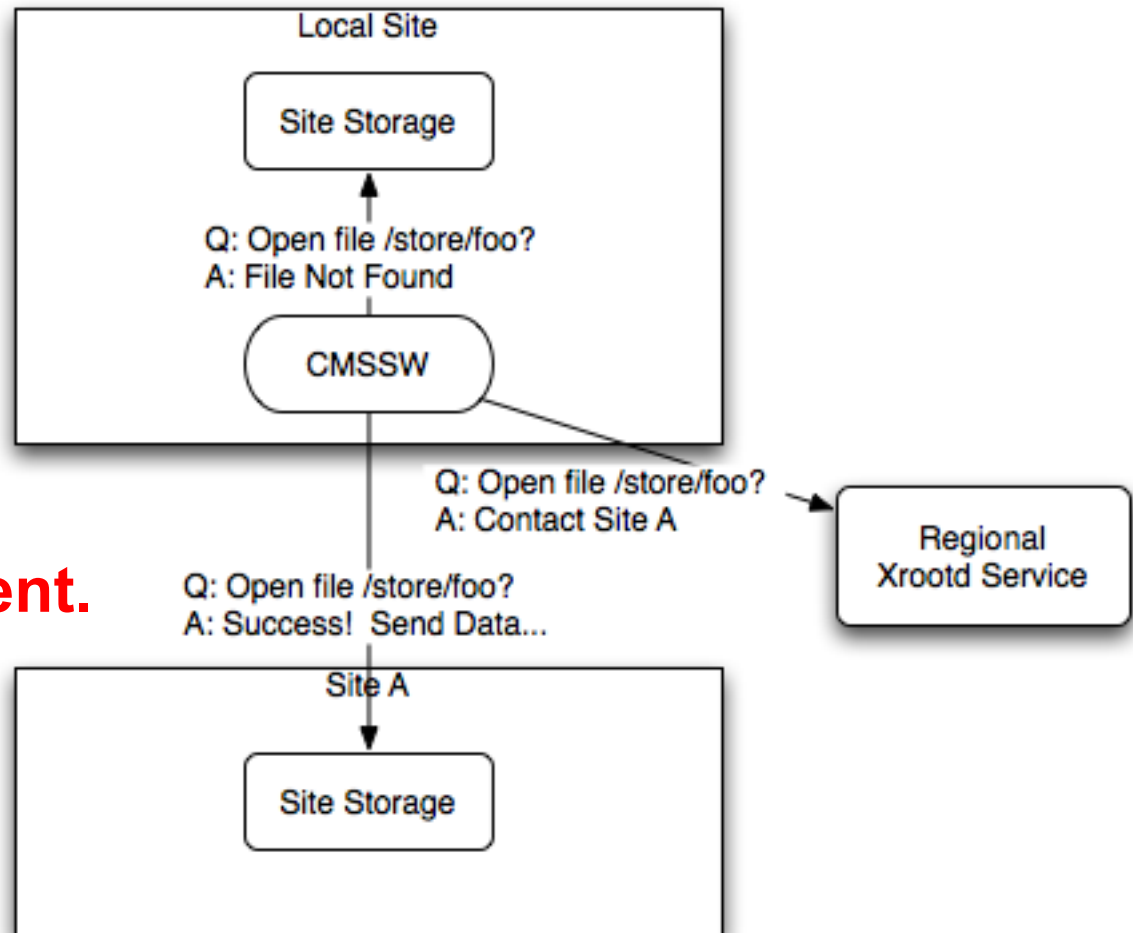
Same namespace, irrespective of where the data is.

From root prompt access LFN “/store/foo” as:

```
TFile::Open("root://xrootd.unl.edu//store/foo");
```

Architecture (III)

Fallback access is the first step for deployment.



T3 Use Case

- Princeton lumi group asked how to make a W/Z skim at LPC and then transfer the skim to Princeton T3.
 - You can go through StoreResults
 - Or you can simply configure Princeton T3 to access LPC based skim via xrootd.
 - Make LPC a local SE to Princeton for CRAB to recognize it.
 - Adjust TFC at Princeton to direct part of namespace to LPC ...
 - ... or rely on fallback upon error.

Files outside PhEDEx

- You want to share your ntuples with your colleagues, but they are not in DBS, nor PhEDEx.
 - Export /store/user via xrootd and tell them via email.
- But your files are sitting in your home area on a login pool at your T3.
 - Export that home area via xrootd and tell them via email.
- But you don't want to pollute the global namespace with your home area.
 - Install a local xrootd redirector and export the home area only to that local redirector, and not the global redirector.

Status

Status

- We have two US redirectors
 - Production & itb
- We have successful scale tests for hadoop sites.
 - Lustre should behave the same.
- We have not so successful scale tests for sites that use xrootd door in dcache.
 - Problem is with libdcap, not the sites
 - FNAL scales well because it does NOT use dcache doors for xrootd export.
- We have Nagios based service monitoring
 - Up/down, email alerts
- We have MonALISA based IO monitoring
- We have documentation
- We are ready for the 1st step towards production use.
 - (see slide 19 for proposal)

Deployment Status

- Production
 - Nebraska
 - Caltech
 - UCSD
 - T3_FNALLPC
- Integration
 - Purdue
 - Wisconsin

Note about FNAL:

- Only data that is on disk can be accessed.
- All data at LPC is on disk.
- Therefore only LPC is “supported” as of today.
- We will be working out more extensive policy before data taking starts.

Nagios State Monitoring

Xrootd Redirectors (xrootd-redirectors)

Host	Services	Actions
xrootd-itb.unl.edu	Xrootd Redirector Heartbeat Purdue Xrootd Redirector Heartbeat UCSD Xrootd Redirector Heartbeat Wisconsin	
xrootd.unl.edu	Xrootd Redirector Heartbeat Caltech Xrootd Redirector Heartbeat FNAL Xrootd Redirector Heartbeat Nebraska Xrootd Redirector Heartbeat UCSD	

Xrootd Redirectors (xrootd-redirectors)

Host	Status	Services	Actions
xrootd-itb.unl.edu	UP	2 OK 1 CRITICAL	
xrootd.unl.edu	UP	4 OK	

Xrootd Servers Caltech (xrootd-servers-caltech)

Host	Status	Services	Actions
cithep160.ultraight.org	UP	1 OK	
cithep172.ultraight.org	UP	1 OK	
cithep230.ultraight.org	UP	1 OK	
cithep251.ultraight.org	UP	1 OK	
gridftp-16-23.ultraight.org	UP	1 OK	

Xrootd Servers FNAL (xrootd-servers-fnal)

Host	Status	Services	Actions
cmsrv32.fnal.gov	UP	1 OK	

Xrootd Servers Purdue (xrootd-servers-purdue)

Host	Status	Services	Actions
cmsdbs.rcac.purdue.edu	UP	1 OK	
crabserver.rcac.purdue.edu	UP	1 OK	

Xrootd Servers UCSD (xrootd-servers-ucsd)

Host	Status	Services	Actions
uaf-3.t2.ucsd.edu	UP	1 OK	
uaf-4.t2.ucsd.edu	UP	1 OK	
uaf-5.t2.ucsd.edu	UP	1 OK	
uaf-6.t2.ucsd.edu	UP	1 OK	

Xrootd Servers UNL (xrootd-servers-unl)

Host	Status	Services	Actions
red-gridftp1	UP	8 OK	
red-gridftp10	UP	8 OK	
red-gridftp11	UP	8 OK	
red-gridftp12	UP	8 OK	
red-gridftp2	UP	8 OK	
red-gridftp3	UP	8 OK	
red-gridftp4	UP	8 OK	
red-gridftp5	UP	8 OK	
red-gridftp6	UP	8 OK	
red-gridftp7	UP	8 OK	
red-gridftp8	UP	8 OK	
red-gridftp9	UP	8 OK	

For detailed documentation:

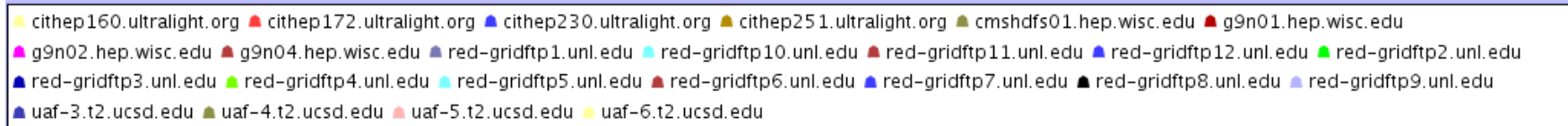
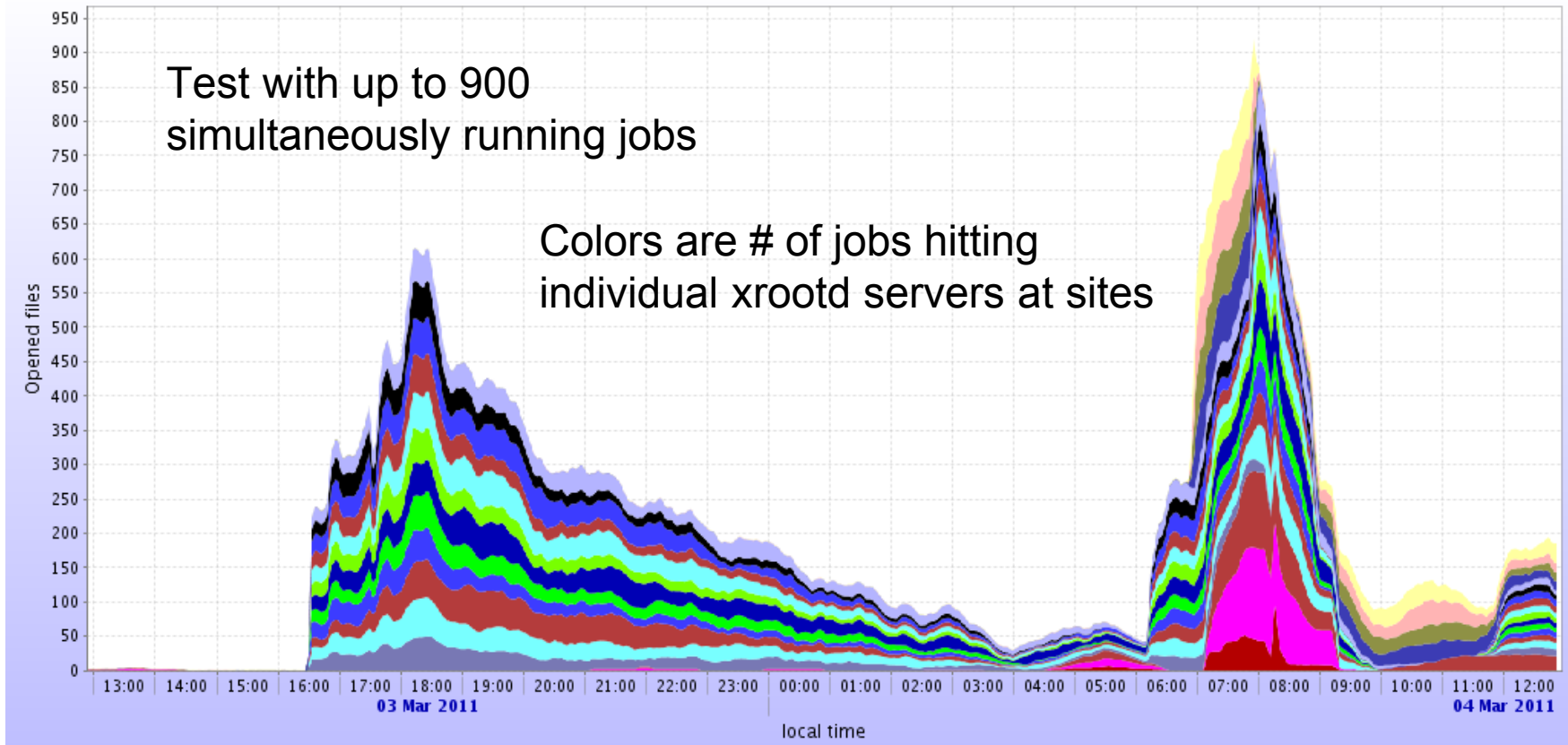
<https://twiki.cern.ch/twiki/bin/view/Main/XrootdMonitoring>

MonALISA IO monitoring

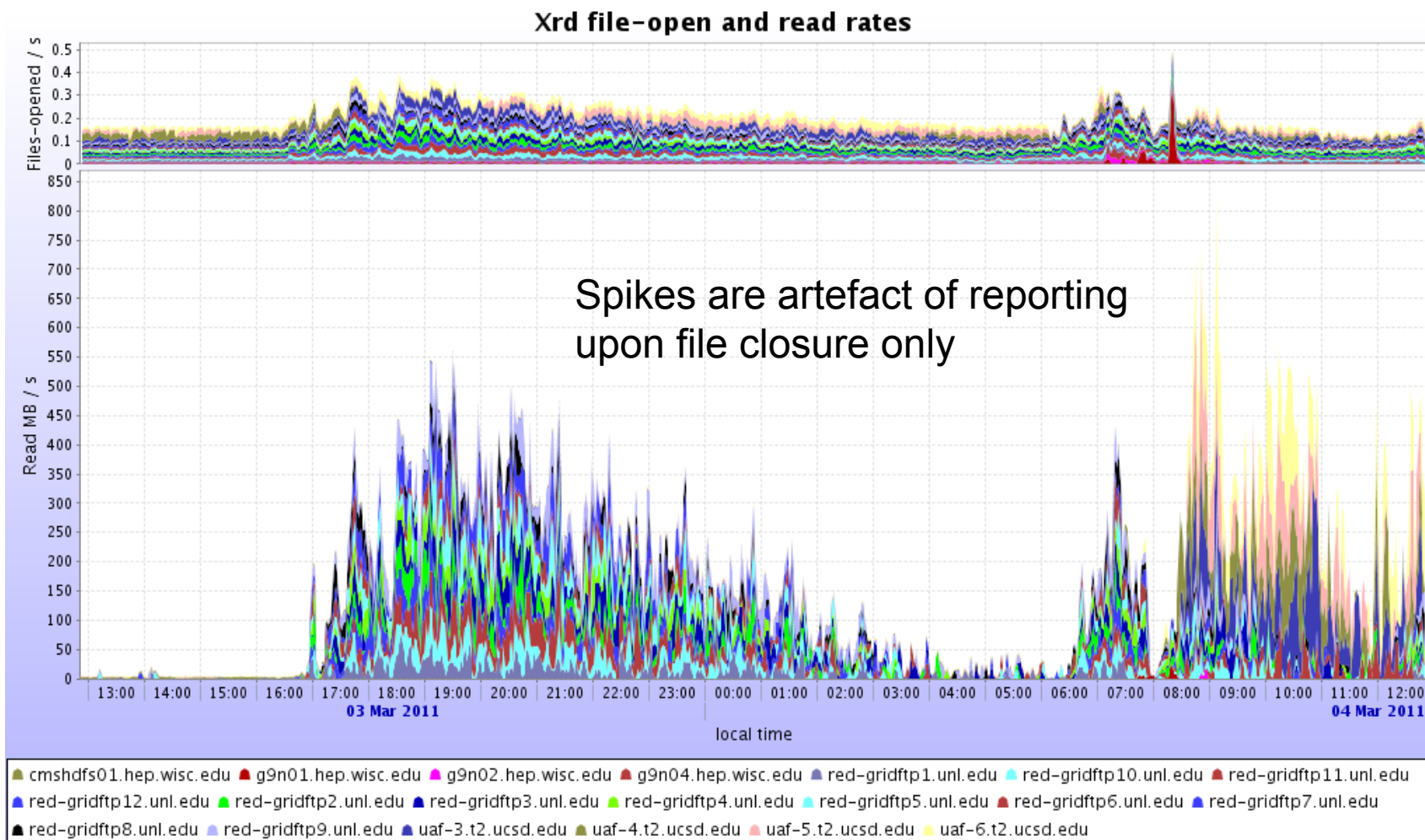
- An initial deployment exists here:
 - <http://xrootd.t2.ucsd.edu/display>
- The next few slides show some of the information we are starting to collect.
- More to come in the future.

Scaling Test

Xrd number of connections / files

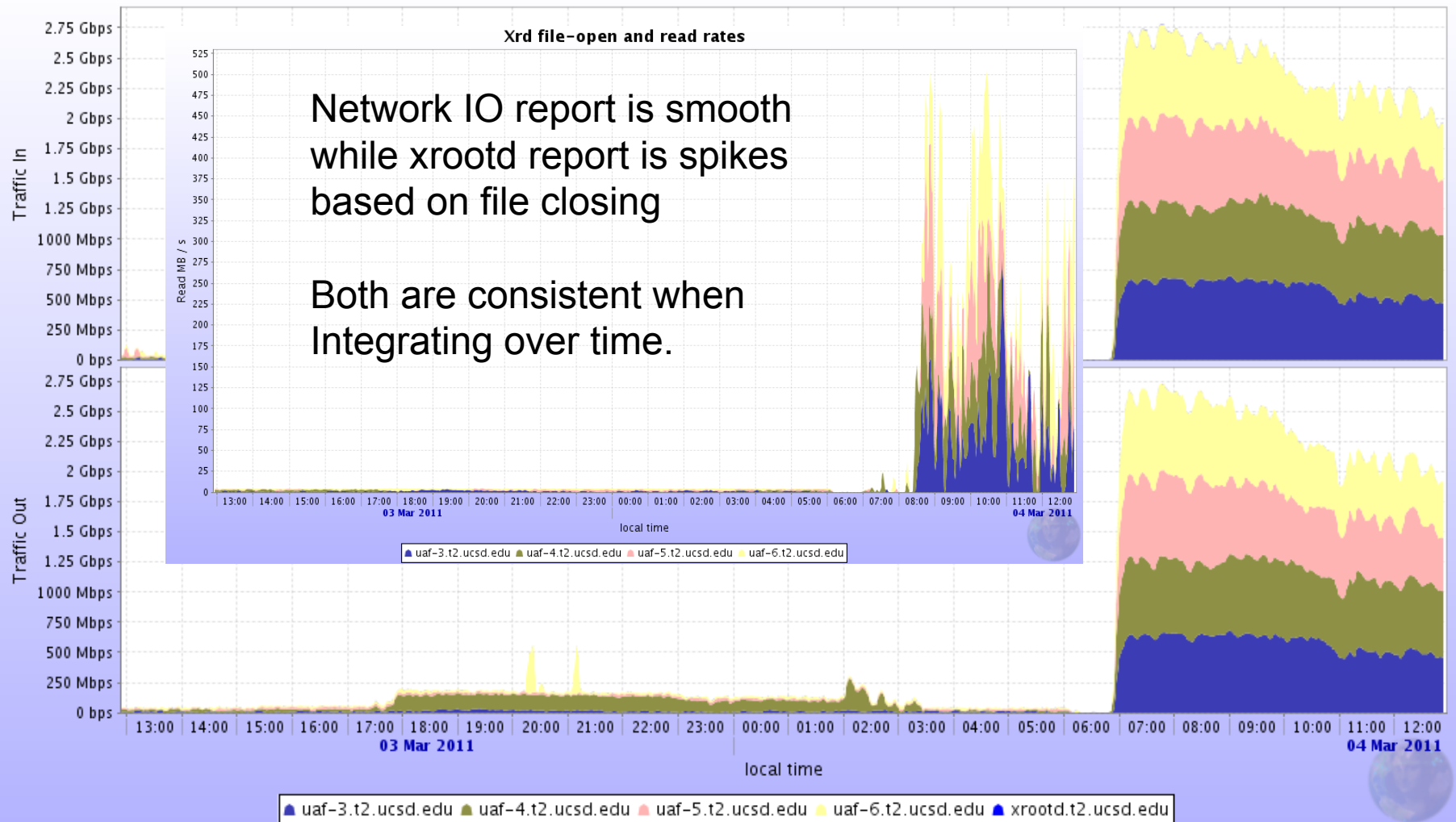


Rates as accounted by xrootd



Example UCSD only

Network traffic on CMS::UCSD::MLSensor



Conclusions

We are ready for:

- failover ops at all T2, T3 sites.

Expect by Summer:

- Interactive use for debugging.
- T3 “without” phedex.

<https://twiki.cern.ch/twiki/bin/view/Main/XrootdUscmsTimeline>

Action Items for Sites

- All US T2 and T3 should change their TFC to include fallback to the US redirector.
- Florida should add xrootd on Lustre and go through the path towards production.
- MIT should add xrootd on dcache into itb.
- Interested in a couple “betatester” T3s that make their CE local to the T2,LPC that are in production xrootd list.

Documentation

Documentation Overview

<https://twiki.cern.ch/twiki/bin/view/Main/CmsXrootdArchitecture>

Documentation

For Users

We have the following user documentation available also:

- [Xrootd Client Usage](#) - How to utilize the current demonstrator infrastructure.

For Admins

The following documentation is aimed at the sysadmins of CMS sites:

- How to integrate Xrootd into your site. Select the appropriate one for your SE technology.
 - [HdfsXrootdInstall](#) - Joining a HDFS site to the global Xrootd cluster.
 - [DcacheXrootd](#) - Joining a dCache site to the global Xrootd cluster.
 - [PosixXrootd](#) - Joining a POSIX site to the global Xrootd cluster.
 - [T3Xrootd](#) - Configuration of Xrootd for USCMS T3 sites.
- [Checklist for production sites](#). Requirements for a site to reach (and maintain) production status.
 - [Changes for the TFC](#). Required changes for TFC at sites to support Xrootd monitoring.
- [Xrootd Monitoring Tests](#). An overview of the monitoring tests performed.

Future Directions

Two Additional Major Thrusts

- Load leveling across US T2s
 - In general, the load across the T2s is not uniform. Part of the reason for that is dataset availability.
 - Imagine a high and low watermark inside glideinWMS such that jobs run at any site if the destination site is not keeping up.
- Caching
 - At T3s to improve performance.
 - At any site, and for the general OSG user community.

Summary & Conclusions

- We have several sites in production.
- We are ready to start with failover use from all sites to the US redirector.
- We are ready for a couple T3s to define the production xrootd sites as “local SEs”
- This is just the beginning !!!
 - A lot more work needed, and more interesting uses to come.
 - For now, the goal is to gain some small scale experience through routine ops.