XROOT and HTTP Third Party Copy

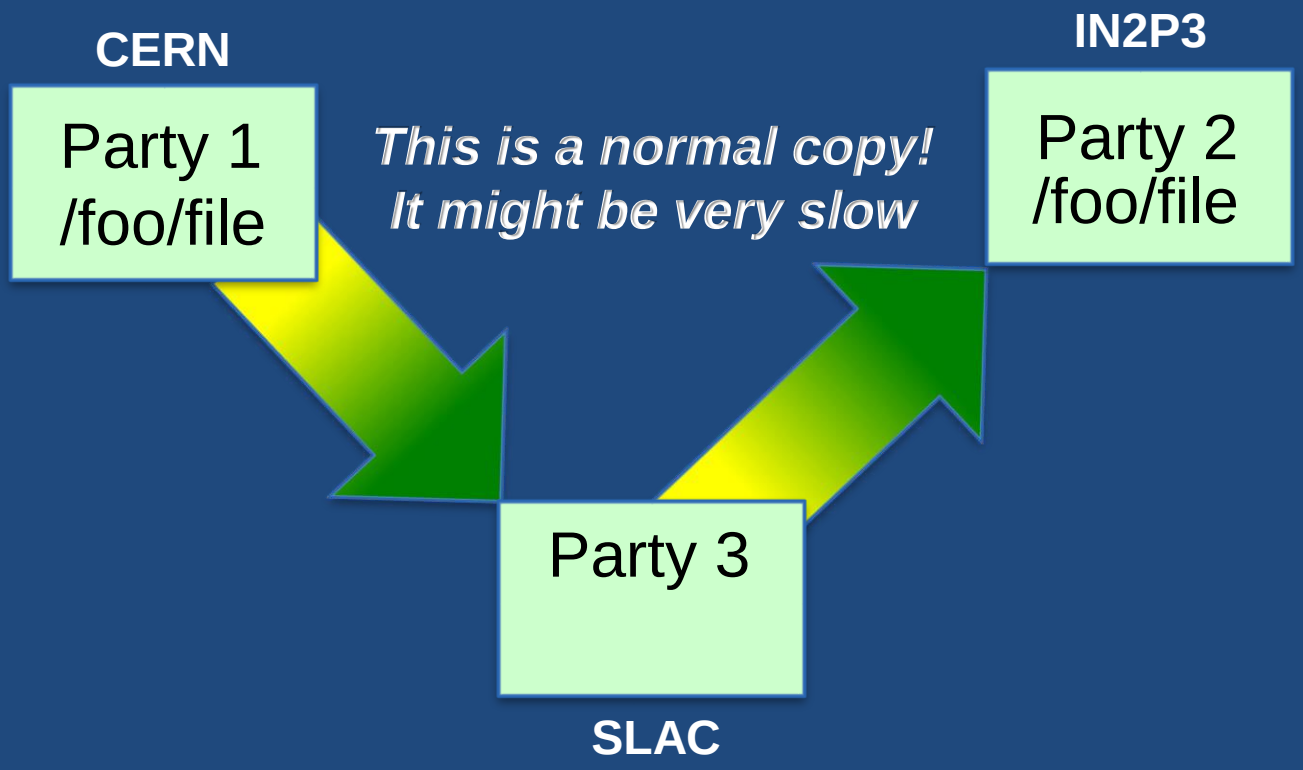
Joint WLCG & HSF Workshop

Naples, Italy

March 26-29, 2018

Andrew Hanushevsky, Alastair Dewhurst, Wei Yang

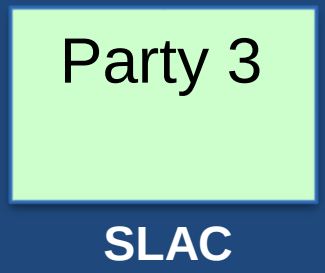
The Bane of Copying Files



xrdcp root://Party1//foo/file root://party2//foo/file

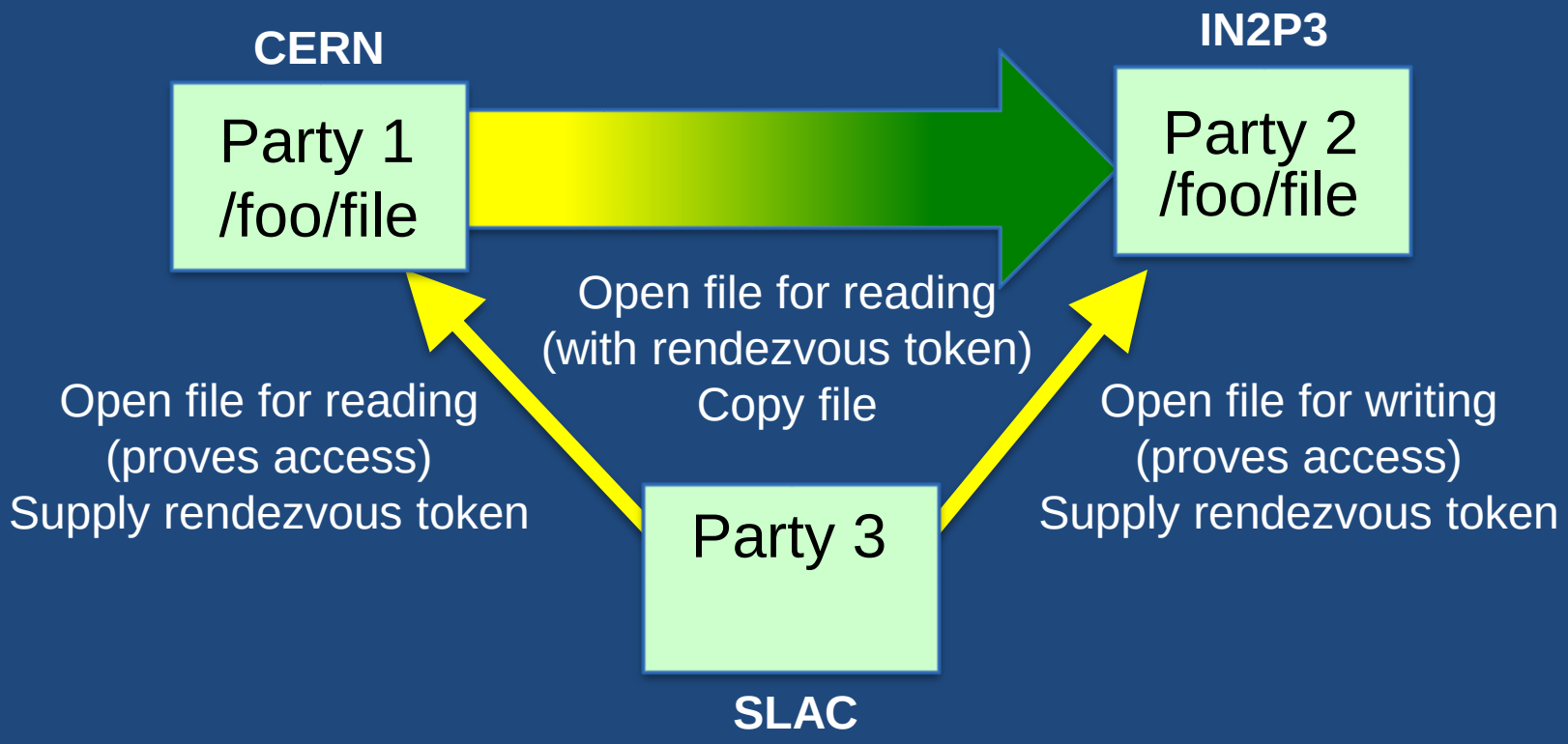
3rd Party Copy to the Rescue

This is a 3rd party copy!
It is as fast as it can be



xrdcp –tpc only root://Party1//foo/file root://party2//foo/file

XRootD 3rd Party Copy "Magic"



3rd Party Copy Security

- There is no standard for 3rd party copy
  - GridFTP & Davix use x509 forwarding
    - Dependent on x509 + identity authorization
      - Davix plans to also support sciToken authentication
  - XRootD uses a transitive closure protocol
    - Works with any kind of authentication/authoriztion
      - This also covers sciToken authrorization
  - sciToken is newly proposed HEP authorization
    - Similar to signed URL used in S3

Authentication & Authorization

- HTTP/WebDAV
  - x509+VOMS (standard does not define authentication per se), limited fine-grained authorization (but fully available with XRootD/HTTP)
- XRootD
  - x509+VOMS, grid-mapfile, or GUMS, Kerberos plus others, fine-grained capability based authorization
    - Supports VOMS roles in authorization
  - It may be possible to use voms-lcmap as well

Who Supports 3rd Party Copying?

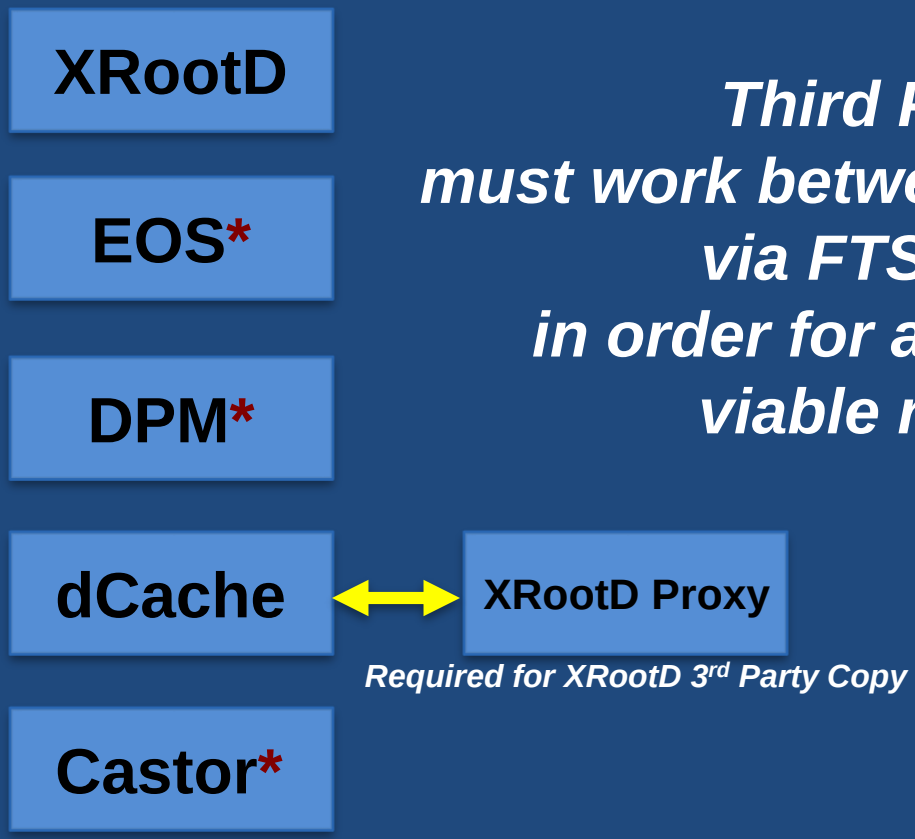
- GridFTP via globus-url-copy
  - Since ~2005
    - Globus defined protocol
- HTTP via davix-cp
  - Since 2016
    - dCache/DPM non-standard HTTP extensions
- XRootD via xrdcp
  - Since 2010
    - XRootD Collaboration defined protocol

If GridFTP Works, Why This Talk?

- Globus has ended support for Open Source Globus Toolkit, including GridFTP
  - Replaced by a $$$ subscription model
    - Globus Connect
- But…
  - WLCG and Grid Community Forum promises support until ~2021.
    - Support cost ⬆ whenever OpenSSL ABI changes
  - Inevitable that a free replacement must be found
    - Currently only HTTP/WebDAV & XRootD support 3PC
      - Must evaluate these as practical replacements

ATLAS Storage Ecosystem

Other LHC sites are in a similar situation



Third Party Copy
must work between all combinations
via FTS and Rucio
in order for a protocol to be a
viable replacement

Required for XRootD 3rd Party Copy

*****Uses XRootD core (i.e. XRootD Inside!)*

SLAC NATIONAL ACCELERATOR LABORATORY

XRootD 3rd Party Copy Evaluation

D – Direct invocation F – Using FTS R – Using Rucio

X – Fully functional X – Issues found X – Broken

	Castor	dCache	DPM	EOS	XRootD
Castor				F R	
dCache					
DPM					
EOS	F R				D F R
XRootD					D F R

Issues

FTS was unable to verify checksum for Castor & EOS transfers

Surprise! It's Actually Not Dismal

- Many/most large sites are behind a firewall
  - Require a proxy to access site storage anyway
- XRootD Proxy 3rd Party Copy works
- XRootD proxy is LHC storage interoperable
  - Castor, dCache, DPM, and EOS
- The proxy solves 3rd Party Copy issues
  - GridFTP also provides proxy capability
    - So, this is a 1-for-1 replacement

HTTP 3rd Party Copy Evaluation

Brian's talk on HTTP storage ecosystem provides a lot more details

D – Direct invocation F – Using FTS R – Using Rucio

X – Fully functional X – Issues found X – Broken

	Castor	dCache	DPM	EOS	XRootD
Castor	D F	D F	D F	D F	D F
dCache	D F	D F	D F	D F	D F
DPM	D F	D F	D F	D F	D F
EOS	D F	D F	D F	D F	D F
XRootD	D F	D F	D F	D F	D F

<u>Issues</u>

Castor and EOS do not support WevDAV.

No known XRootD endpoints have enabled https.

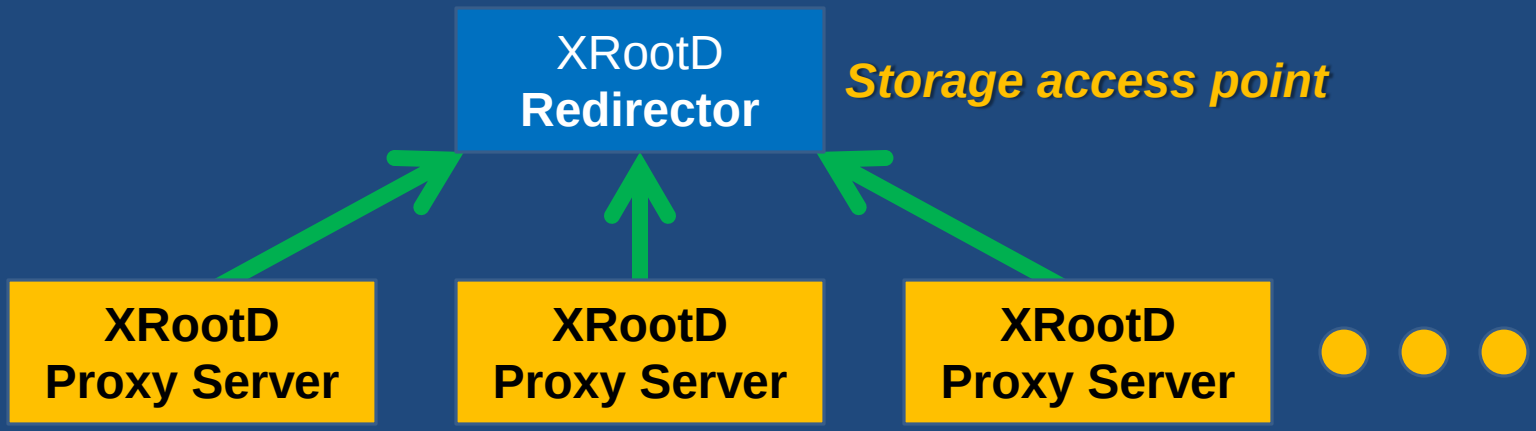
Rucio replication not tested.

XRootD Deployment Scenario

- Castor, DPM, EOS, Vanilla XRootD ready
  - Need to add ofs.tpc directive to config file
    - Should this be made the default?
- dCache requires an XRootD proxy front
- All need x509 proxy cert with prod attribute
  - Maybe add x509 forwarding, if we must
- Set AGIS/RUCIO ATLAS site settings
  - read_wan, write_wan, delete_wan

XRootD Scaling

- XRootD supports multiple DTN's
  - Implemented as a cluster of proxy servers
    - Logically a single access point
      - See BNL setup as an example



Conclusion

- We do know that GridFTP can be replaced
  - For us, HTTP & XRootD are widely deployed
    - All ATLAS and most LHC sites usually have both
- More evaluation is needed
  - To make sure all SE combinations work
- Free data management looks quite doable
  - We are starting to plan now
    - Expect a transition period as things fall into place