Global Data Access – View from the Tier 2

Rob Gardner

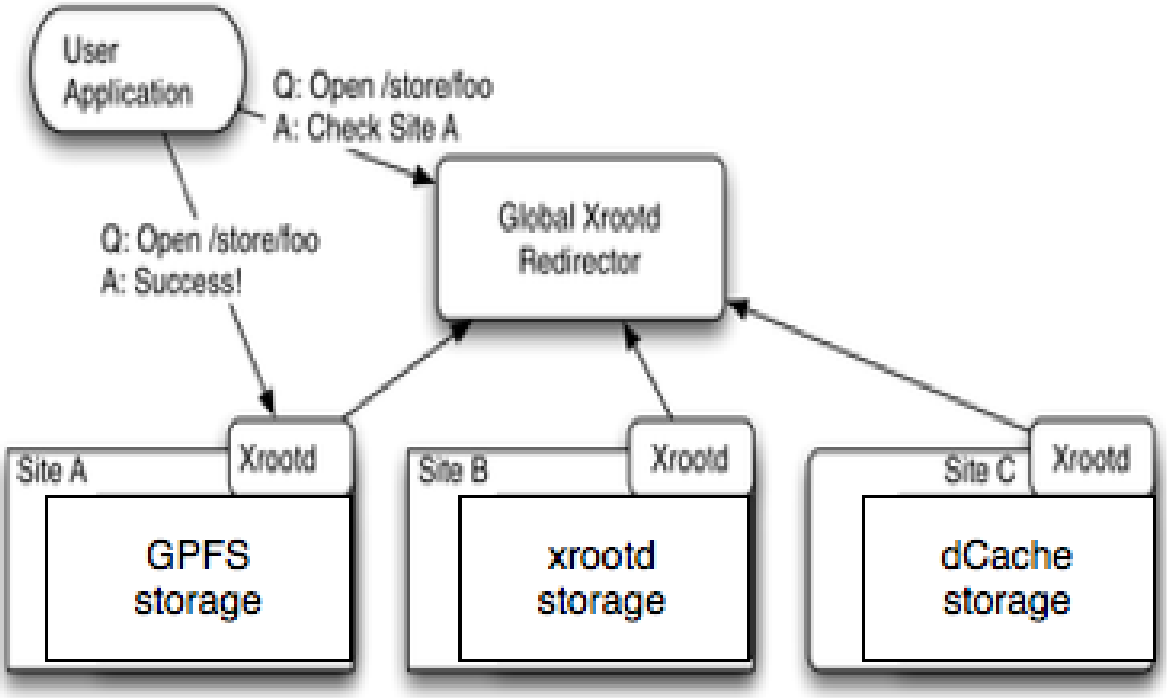
Charles Waldman

project

- We have long recognized the need for providing efficient user access to datasets at Tier 2
  - Talk: http://www.mwt2.org/~cgw/t2t/talk.html
  - Demo: http://repo.mwt2.org/t2t
- Now the reality is that we have > 5 PB in our T2 cloud
  - Typical sites have > 50K datasets, O(10M) files
  - Since gained experience with local and wide area access using both dCache and xrootd services
- Assuming a namespace convention we could start a prototype T2 access project using WLCG demonstrator findings

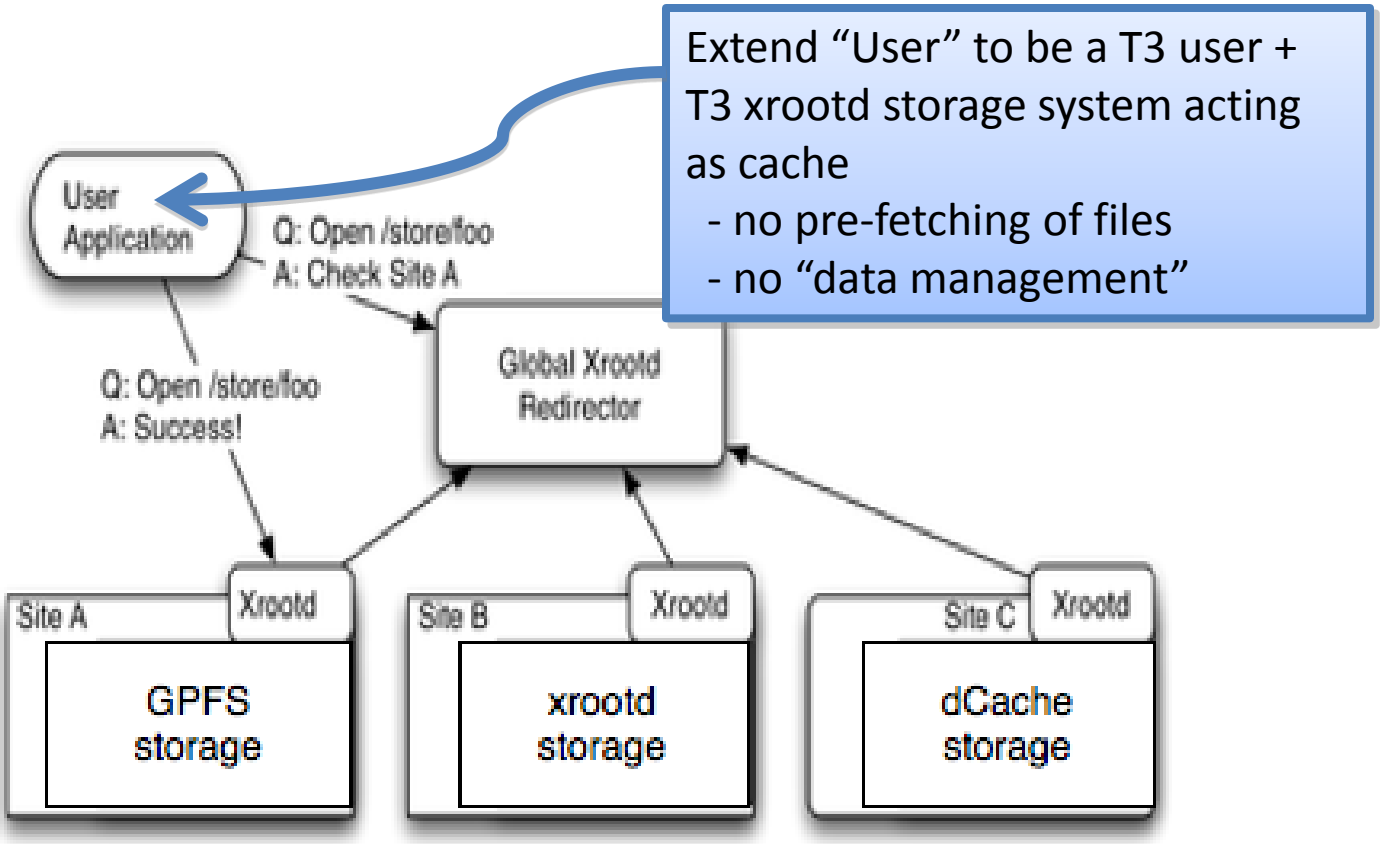
diagram

(courtesy Brian Bockelman)



diagram

(courtesy Brian Bockelman)



(some) questions

- On T2 with many data servers what additional services may be required?
  - An xrootd on each dCache pool, eg?
- What local caching strategy is best on the client side – block or file?
  - And associated additional services, (eg. frm, squid)

First steps

- Try out namespace convention & dq2 client
- Setup needed T2 site level xrootd federation services
- Register with SLAC global redirector – functional tests
- Create namespace-to-site path translation method
  - Test with xrdcp
- Test local direct read access:
  - ROOT client and xrootd storage
  - ROOT client and dCache, GPFS storage
- Same, remote client access. Study latency issues
- Next: investigate caching on client (user) side