

Direct HTTP I/O from ROOT, using Davix

- Rod Walker, LMU, 28th April 2014

- Motivation
- Relation to FAX
- DAVIX
- Integration status
- Comment & conclusions

Motivation

- HEP protocol zoo is reduced
 - dcap, rfio possible to reduce to just xrootd
 - still HEP specific protocol
- Good to use standard protocols
 - must support vector reads, for TTreeCache
 - NFS4.1
 - not fully accepted/supported yet
 - client needs root setup, not available on all OS's
 - HTTP/WebDav
 - supported by major storage used by ATLAS, dcache, dpm, eos, storm
 - many clients, no admin required – long available TWebFile in Root

Xrootd FAX(AAA?) vs WebDav

	Xrootd FAX	https
Global File Name	some.redirector.ch//atlas/rucio/user/hito:user.hito.xrootd.mwt2-1M	https://voatlasrucio-redirect-prod-01.cern.ch/redirect/data11_7TeV/AOD.512720._000001.pool.root.1
Redirection to TURL	regional,hierarchical	central, redundant, horizontal scaling
File lookup	P2P ask all storage	Rucio Db query
Replica selection	Choose from BW mesh	Geoip/metalink(all)
Client	Root,xrdcp	Root, curl,aria2c,firefox*,...
Server	Dcache,dpm,xrootd,eos	SRM+apache, S3, ..
Proxy server	xrootd	Nothing caches https objects. Reverse proxy?

* - Browser has non-voms p12 cert. Depends on SRM config(needs gridmap)

Davix

- Root long supported http urls in TFile
 - not really https. No redirection, metalinks, failover
- DAVIX should address this
- Added to 5.34-15, but not yet built in afs/cvmfs
- Private build showed some problems
 - davix bugs, dCache http responses
 - needs some iteration

DAVIX(2)

- No new performance tests yet
 - previous tests(with http) showed comparable performance to dcap, xrootd
- Ruco redirect using geoip works
 - needs Rucio registered file – migration started
 - also providing metalink (all replicas)
 - DAVIX will handle metalinks too

Integration steps

- Goal is to use https io for analysis jobs
 - at least retire rfio/dcap in favour of https/xrootd
- Pilot aria2c mover builds redirect url (gLFN)
 - Root sees urls as input files
 - aria2c copy to scratch non-root files
- Root built with davix(with fixes) in cvmfs
 - with no need for special .rootrc
 - old root versions will never work
 - cannot use by default on a queue
- HC comparisons to FAX
- Prefer quicker dev/test/deploy cycle to find problems

First things last

- Is io over WAN even worthwhile?
 - few hundred T2 analysis jobs read at few GB/s from local storage
 - could easily fill pipe from WAN
 - source storage(s) now serves local WNs plus remote WNs
 - cluster/storage/networks planning then meaningless
 - standard heavy use, or massive failover not useful
 - what does that leave
 - casual interactive from desktop
 - throttled large scale use, e.g. T3 no storage
 - light failover, eg. access to 100th file fails
 - data/job co-location still rules ok
- still worthwhile but not to be oversold

Conclusion

- Rucio redirect service looks good
 - no scaling tests yet
 - awaits root deployment, dcache fix
- DAVIX in Root, but not built by default
 - expts agree addition to LCG externals
 - Davix improvements are then independent (lib)
- Hammercloud tests to be done when ready
- Massive WAN TFile io not wise
 - standard protocol access and gLFN is a good thing
 - eg. dq2-get outputs becomes browser click or standard parallel client