

XRootD Roadmap To Start The Second Decade

Root Workshop

SaaS-Fee

March 11-14, 2013

Andrew Hanushevsky, SLAC

<http://xrootd.org>

Truly A Decade!

- # 1997 – Objectivity, Inc. collaboration
 - Design & Development to scale Objectivity/DB
 - First attempt to use commercial DB for Physics data
 - Successful but problematical
- # 2001 – BaBar decides to use root framework vs Objectivity
 - Collaboration with INFN, Padova & SLAC created
 - Design & develop high performance data access system
 - Work based on what we learned with Objectivity
- # 2003 – First deployment of **XRootD** system at SLAC
- # 2013 – Wide deployment with several implementations
 - ALICE, ATLAS, CMS, EXO, Fermi, LSST; among others
 - Protocol also available in dCache, DPM, and EOS

Diversity Hastens Innovation

- # Deployment has become diverse
 - Feature set is determined by experiments
 - Challenge to make it all cohesive
 - Yet backward compatible
 - We have largely succeeded
 - Even BaBar's old code can use the new servers!
 - Acceleration of significant additions
 - 3.3.0 (latest) saw 28 new features
 - Deceleration of bug reports
 - Only 8 relatively minor bug reports open

Significant Additions In 3.3.0

- # Beta introduction of new client
- # New “f” stream monitoring
- # User-friendly xrdcp CLI
- # Third party copy
- # Integrated checksums at redirector level
- # Plug-in version checking
- # EPEL Conformance

Looking Forward

- # New client
- # Proxy readv pass-through
- # IPV6
- # Configurable I/O throttling
- # Internal bridge protocol
- # New protocol plug-in
 - Plus *much* more that we don't have time for

New Client

Complete rewrite

- Emphasis on scalability & performance
 - Better use of threads and fully asynchronous

Will be the platform for future features

- Current client only patched for major issues
 - Support will continue for at least 2 years
- New client switch in 2013-14 timeframe

See the next talk!

Proxy readv Pass-Through

- # Current proxy server unrolls readv's
 - This may severely impact performance
- # Solved by integrating readv interfaces
 - Allows proxy to pass through readv's
 - Impact: some plug-ins need to be recompiled
- # Code supplied by Brian Bockelman
 - Expect this in 4.0

IPV6

- # We have decided to do a clean transition
 - More maintainable than a simple “patch”
 - Based on a new IP agnostic framework
 - And it has major side-effects
 - All network sensitive plug-ins need to change
 - Fortunately, the changes are apparent and minimal
- # Expect this in 4.x
 - Alpha code is available now
 - The xrdipv6 git branch

Configurable I/O Throttling

- # Support for federating marginal sites
 - As well as reducing I/O impact otherwise
- # Implemented as a server plug-in
 - Existing code path not effected w/o plug-in
- # Developed at UNL
 - Brian Bockelman for CMS
- # Will be made generally available
 - Expect this in 4.x

Internal Bridge Protocol

- # Makes it possible to map protocols
 - Protocol **x** to **xroot** protocol and execute
 - **xroot** protocol responses back via protocol **x**
- # All done within the xroot server
 - Allows plugging in other protocols
 - Leverages existing & future **XRootD** features
 - E.g. Monitoring, sessions, security, etc
- # Expect this in 4.x

An HTTP Plug-in

- # Yes, **XRootD** will respond to two protocols
 - Native xroot and http on the same port
 - Possible via the new bridge protocol
 - As an aside, **XRootD** already is multi-protocol
 - It also supports proof protocol but not via a bridge
- # A WLCG collaborative project
 - Headed by Fabrizio Furano

HTTP To Replace Xroot?

- # This simply augments service possibilities
 - Ad hoc downloads (e.g. wget)
 - Possible SRM replacement interface (e.g. webdav)
 - FTP replacement for data distribution
 - Though xrdcp would still be faster
 - And offer more functionality (e.g. multi-source copying)
- # Replacement for xroot protocol?
 - We don't think so for many reasons

HTTP Is Not The Holy Grail??!

- # We don't think HTTP is high performance
 - Neither does Google
 - “HTTP was not particularly designed for latency ... a FIFO queue ... redundant headers ... header sizes ... latency can be significant”
 - <http://www.chromium.org/spdy/spdy-whitepaper>
 - Nor does Microsoft
 - “HTTP does not allow for out- of-order or interleaved responses ... users are often disappointed in the perceived performance of websites”
 - <http://tools.ietf.org/html/draft-montenegro-httpbis-speed-mobility-01>

Can HTTP Be Fixed?

- # Google is proposing SPDY
 - <http://en.wikipedia.org/wiki/SPDY>
- # Microsoft is proposing S+M (Microsoft Speed + Mobility)
 - [http://en.wikipedia.org/wiki/Microsoft SM](http://en.wikipedia.org/wiki/Microsoft_SM)
- # Proposals going through IETF & W3C
 - One or the other will likely become HTTP 2.0
 - Each reduces network bandwidth & latency
 - Neither really address CPU utilization issues
 - Hey, servers are dirt cheap especially in the cloud!

Bottom Line

- # HTTP is a serial session-less ASCII text protocol with inherent performance issues
 - Leading to significantly increased bandwidth, CPU, and latency thus limiting scalability
- # This is not a high performance protocol
 - It was never meant to be!
 - It is a highly enhanced yet minimalistic FTP
 - So, no current network based full-featured file access protocol associated with *performance* uses HTTP

But HTTP Is Everywhere!

Naively, yes.

- But any serious use of HTTP to construct a robust application requires custom clients negating the “everywhere” concept
 - Just ask Amazon, Facebook and Twitter

The old adage applies ...

- A standard knife can substitute for a scalpel
 - But would you want your surgeon to use it?

So, we will use it for what it was designed

XRootD Collaboration

- # Mutually interested institutions contributing to development & maintenance
 - SLAC (founder)
 - CERN (2010)
 - Duke (summer 2011)
 - JINR (fall 2011)
 - UCSD (winter 2011)
 - UNL (expected spring 2013)

Conclusion

- # **XRootD** is under active development
 - Always looking for new ideas
 - Feel free to suggest them
 - Be a contributor – it's open source
 - Available via a git repo and soon to be on github!
 - Consider joining the **XRootD** collaboration
 - It costs no money to join
- # See more at <http://xrootd.org/>

Acknowledgements

Current Software Contributors

- ATLAS: Doug Benjamin, Patrick McGuigan,
- CERN: Lukasz Janyst, Andreas Peters, Justin Salmon
- Fermi: Tony Johnson
- JINR: Danila Oleynik, Artem Petrosyan
- Root: Gerri Ganis, Bertrand Bellenet, Fons Rademakers
- SLAC: Andrew Hanushevsky, Wilko Kroeger, Daniel Wang, Wei Yang
- UCSD: Matevz Tadel
- UNL: Brian Bockelman
- WLCG: Fabrizio Furano, David Smith

US Department of Energy

- Contract DE-AC02-76SF00515 with Stanford University