

# **XRootD Status and Plans**

# XRootD Development Team (2023/2024)

- Server / OFS / OSS
  - Andrew B. Hanushevsky
  - David Smith
- Client / XrdEc / Python
  - David Smith
  - Guilherme Amadio
- CMake / Packaging / Testing / CI
  - Guilherme Amadio
- ▶ TLS
  - Andrew B. Hanushevsky
  - David Smith
  - Guilherme Amadio
- XCache
  - Alja Mrak Tadel
  - Matevž Tadel

- HTTP Protocol Plugin
  - Cedric Caffy
- HTTP Third-Party Copy (TPC) Plugin
  - Brian Bockelmann
  - Cedric Caffy
  - Elvin A. Sindrilaru
- GSI Authentication Plugin
  - Gerardo Ganis
  - Guilherme Amadio
- XrdOssCsi
  - David Smith
- SciTokens Plugin
  - Brian Bockelmann
  - Derek Weitzel
- EPEL / Debian Packaging
  - Mattias Ellert



- Server
  - Make maxfd configurable (default is 256k)
  - Use SHA-256 by default for signatures and message digests
  - Switch to a fixed set of DH parameters (compatibility with OpenSSL 1.0.2)
  - Allow specfication of minimum and maximum creation modes
  - Better detection of private IPV6 addresses (check also for unique local address)
  - Include token information in the monitoring stream (subject, user, vorg, role, groups)
- XCache
  - New function for file eviction
  - Allow origin to be a locally mounted directory (e.g. XCache for Ceph/Lustre)
- Client
  - New subcommand for xrdfs cache to allow for cache evictions
  - Do not enforce TLS when --notlsok option is used in combination with root:// URL
  - Increase default number of parallel event loops to 10 (affects XCache)

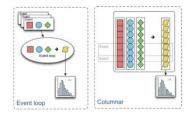


I/O performance studies of analysis workloads on production and dedicated resources at CERN

A. Sciabà, J. Blomer, P. Canal, D. Duellmann, E. Guiraud, A. Naumann, V.E. Padulano, B. PanzerSteindel, A.J. Peters, M. Schulz, D. Smith

### **Analysis Grand Challenge ttbar analysis**

- Simplified analysis from CMS used as technical demonstrator in IRIS-HEP
  - Input dataset 3.6 TB, 2300 ROOT files, 1.5 GB/file consisting of CMS 2015 Open Data
- Columnar analysis paradigm
  - · Distributed using a map-reduce concept
- Original <u>Coffea</u> implementation
  - ROOT-less, parallelism via Python futures or Dask
- RDataFrame port (<u>talk</u>)
  - ROOT-based, parallelism via implicit multithreading, Dask and other

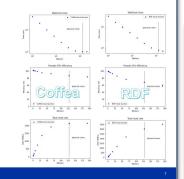


- Measure performance and scalability
  - Local parallelism on client node
  - Data read from local node vs. directly from EOS via xrootd vs. via an XCache instance
  - · NOT a comparison between Coffea and RDF
    - Simply, different workloads with different behaviors



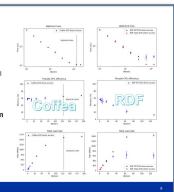
#### Local access

- · Scalability is excellent
  - Some bottleneck appears for high numbers of workers
  - Overcommitting does not help for Coffea, but it increases throughput for RDF
  - Indication that Coffea is more CPUconstrained
- · The CPU efficiency comparably high
  - · I/O not a strong bottleneck
- Local, fast SSD storage is always going to work well
- · Aggregate read rates up to 3 GB/s



#### **Direct access to EOS**

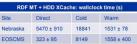
- Scalability still good when parallelism is via multiprocess
- RDF implicit multithreading does not perform well with xrootd and many threads
- · RDF via Dask is multiprocess, scales better
- CPU efficiency practically constant around 60% with multiprocess parallelism
- I/O time is not negligible anymore but no bottlenecks
- · Two EOS instances tested
- EOSCMS by default, but CERNBOX produces similar (slightly worse) results



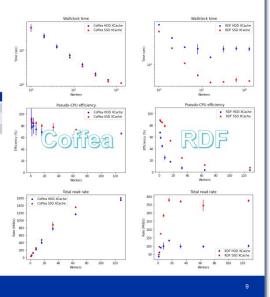
#### **HDD/SDD-based XCache**

 Compared performance of direct access to Nebraska and CERN, cold cache and warm cache

Coffea + HDD XCache: wallclock time (s)				RDF MT + HD	
Site	Direct	Cold	Warm	Site	Direct
Nebraska	442 ± 16	608	133 ± 6	Nebraska	5470 :
EOSCMS	139 ± 8	325	137 ± 3	EOSCMS	323 ±



- Performance results
- · A cold cache is slower than direct access!
  - · Due to sparse file access and network latency
- Multiprocess scales very well
- · HDD XCache almost as good as SSD XCache
- RDF multithreaded scales very poorly "out of the box"
- All connections multiplexed into one ⇒ bottleneck!
- · SDD XCache helps a lot, but scalability is still broken

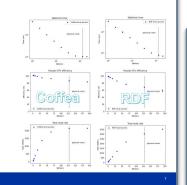






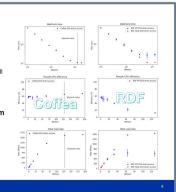
#### Local access

- · Scalability is excellent
  - Some bottleneck appears for high numbers of workers
  - Overcommitting does not help for Coffea, but it increases throughput for RDF
  - Indication that Coffea is more CPUconstrained
- · The CPU efficiency comparably high
  - · I/O not a strong bottleneck
- Local, fast SSD storage is always going to work well
- · Aggregate read rates up to 3 GB/s



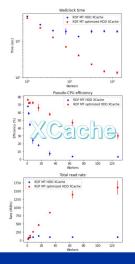
#### **Direct access to EOS**

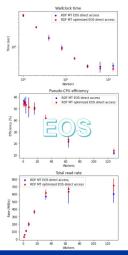
- Scalability still good when parallelism is via multiprocess
- RDF implicit multithreading does not perform well with xrootd and many threads
- RDF via Dask is multiprocess, scales better
- CPU efficiency practically constant around 60% with multiprocess parallelism
- I/O time is not negligible anymore but no bottlenecks
- · Two EOS instances tested
- EOSCMS by default, but CERNBOX produces similar (slightly worse) results



### RDF + Xrootd performance optimization

- Scalability with ROOT multithreading and XCache can be improved
  - XRD\_PARALLELEVTLOOP=10 on the client largely improves Xrootd performance
  - Prevent the connection multiplexing by adding different client names to the file names root://client1@eoscms.cern.ch//eos/myfile.root
- Enormous impact when reading from XCache
  - Obvious as it is a single server and multiplexing a big bottleneck
- Effect negligible when reading from EOS
  - Files already spread over hundreds of disk servers, multiplexing irrelevant

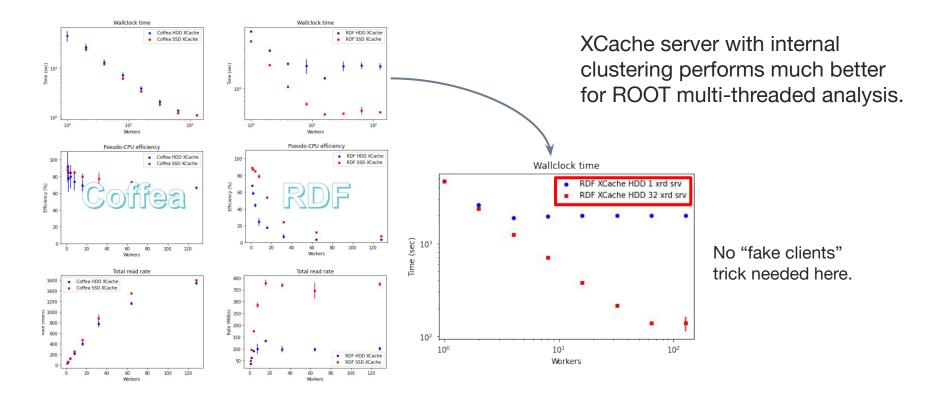






10







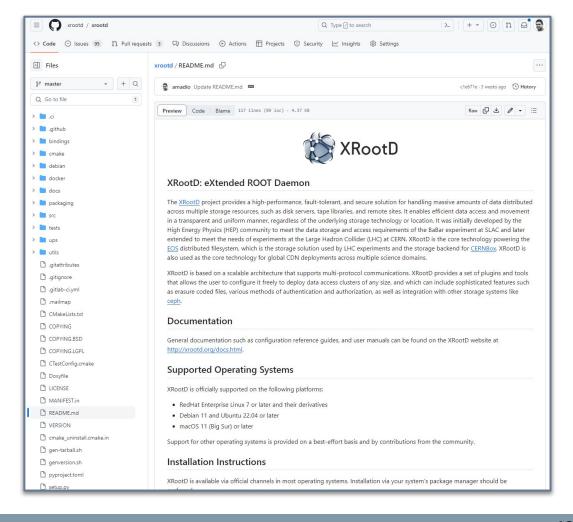
- ► XRootD 5.6.0 (cont.)
  - Build System / C++
    - Modernization of CMake build system (requires CMake 3.16+)
    - Support building against musl libc (e.g. on Alpine Linux, Void Linux)
    - Make codebase C++17 ready (migration to C++17 as baseline in XRootD 5.7.0)
    - XrdCeph git submodule merged back into the main repository
    - Build option ENABLE\_CRYPTO removed. OpenSSL always required with XRootD 5.x
  - Testing and Continuous Integration
    - Updated continuous integration system based on GitHub Actions
    - New xrd-docker to automate building/running container-based tests
    - New test.cmake script added to automate configure/build/test cycle
  - RPM/DEB Packaging
    - Rewrite of both DEB/RPM packaging to simplify building official repository
  - Python bindings
    - Better support for Python 3.x, rewrite of build system in accordance with PEP517

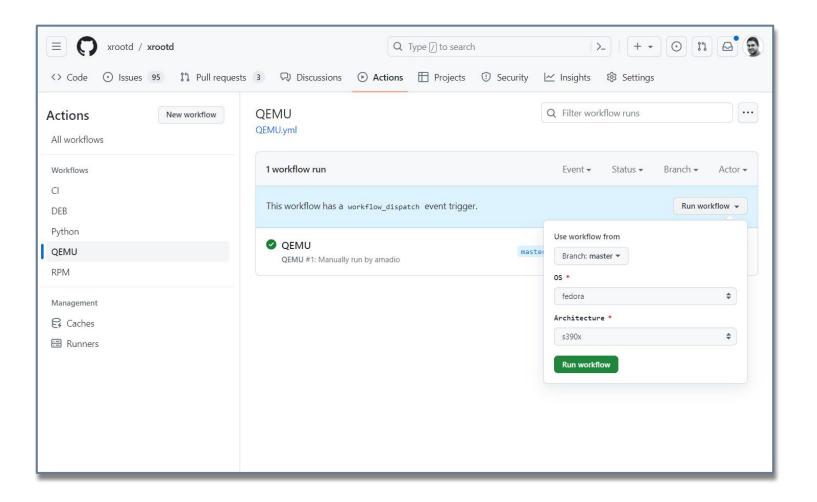


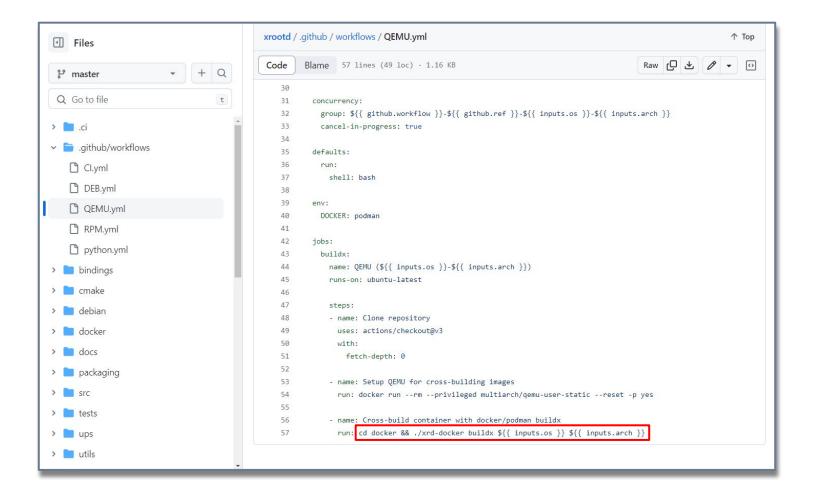
- ► XRootD 5.6.0 (cont.)
  - XrdHttp
    - Include Accept-Ranges in HEAD response
    - Allow XRootD to return trailers indicating failure
    - Report cache object age for caching proxy mode
    - Return 405 instead of 500 error code on deletion of non-empty directory
    - Return 404 instead of 500 error code on GET request on non-existent file
  - Plugins
    - GSI: Add option to display DN when it differs from entity name
    - ZTN: Allow to point to a token file using cgi ?xrd.ztn=tokenfile
    - ZTN: Allow option -tokenlib none to disable token validation (used by EOS)
  - Miscellaneous
    - Updated docs: README.md, INSTALL.md, TESTING.md, CONTRIBUTING.md
    - Sandboxing / hardening settings added to systemd service units (commented out)
    - Make output of xrdcrc32c tool consistent with xrdadler32



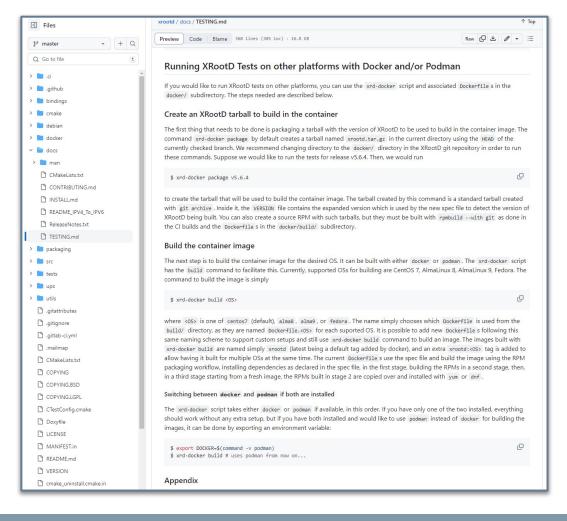
- New README in Markdown
- GitHub Actions
  - Continuous Integration
  - RPM / DEB Packages
  - Python wheels
  - QEMU cross-platform
- CDash Dashboard



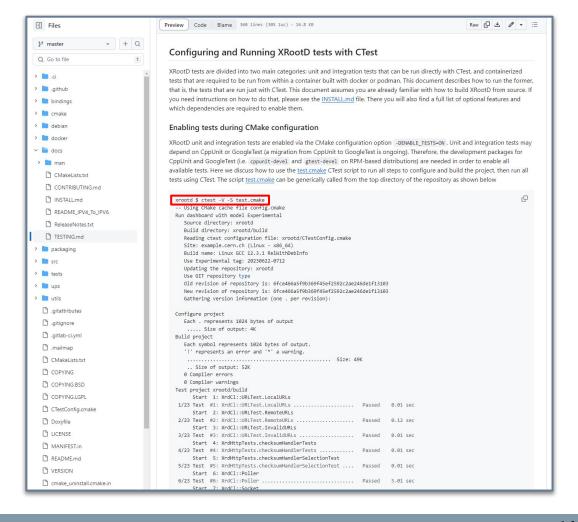




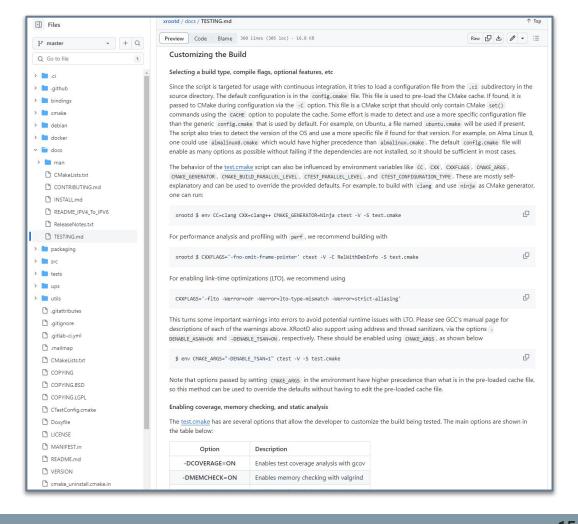
- New README in Markdown
- GitHub Actions
  - Continuous Integration
  - RPM / DEB Packages
  - Python wheels
  - QEMU cross-platform
- CTest config/build/test script
- CDash Dashboard



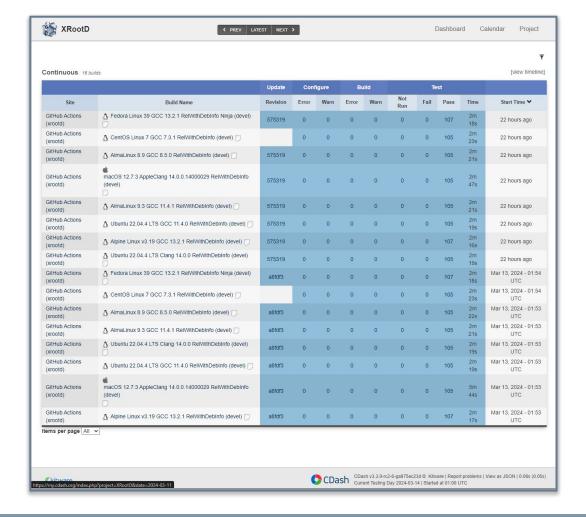
- New README in Markdown
- GitHub Actions
  - Continuous Integration
  - RPM / DEB Packages
  - Python wheels
  - QEMU cross-platform
- CTest config/build/test script
- CDash Dashboard



- New README in Markdown
- GitHub Actions
  - Continuous Integration
  - RPM / DEB Packages
  - Python wheels
  - QEMU cross-platform
- CTest config/build/test script
- CDash Dashboard



- New README in Markdown
- GitHub Actions
  - Continuous Integration
  - RPM / DEB Packages
  - Python wheels
  - QEMU cross-platform
- CTest config/build/test script
- CDash Dashboard



## Next Release: XRootD 5.7.0

- Move to C++17 standard as baseline
- Remove using namespace std; from all headers and source files
- Allow changing C++ standard via CMAKE\_CXX\_STANDARD
- Update min/default RSA key bits to 2048
- Full migration of testing infrastructure to GoogleTest (CppUnit tests removed)
- New only-if-cached directive for Xcache
- New tpc.route option to force connecting back via same IP on HTTP-TPC
- Disallow renegotiation for TLSv1.2 and earlier

## **Future Plans**

- Continue to extend testing coverage
  - HTTP, TPC, XCache, authentication methods (host, krb5, gsi, tokens, sss, unix)
- Resolve at least ~20% of outstanding enhancements per year
- Support for alternative protocols (e.g. SFTP, an EGI request)
- Native Oauth2 plugin
  - Enable using CERN SSO tokens with XRootD/EOS
- SSH Authentication
  - SSH keys/certificates as possible alternative to X509 certificates
- Better, multi-protocol integration with FTS
  - Reduced interface needed for file transfers
- Range cloning of files to allow updating of EC files
- Plugin for handling RUCIO datasets within ZIP archives



# **XRootD Packaging and Distribution**

- New official RPM repositories, include Alma 9, Fedora https://xrootd.slac.stanford.edu/dload.html#official-rpm-repositories
- RPM and DEB packaging scripts available in official GitHub repository
- XRootD is available via official channels in most distributions already
  - Alma, Arch, CentOS, Debian, Fedora, Gentoo, Manjaro, Raspbian, Rocky, Ubuntu, etc
- XRootD is also available on repositories that work across other OSs
  - Homebrew, macPorts, Nix, Spack, conda, etc
- Installation via official channels strongly encouraged





- XRootD 5.5.2
  - Enable ZTN authentication with macaroons-based tokens
  - Extend number of parallel copies from 4 to 128 (--parallel option to xrdcp)
- XRootD 5.5.3
  - Support user-provided script for computing checkums
- XRootD 5.5.4
  - ZTN plugin enabled by default
  - Fixes for authentication failures across daylight savings change
  - Support certificates with dates in year 2049 or later
- XRootD 5.5.5
  - Enable XrdClProxy plugin to work with pgRead
  - Fix creation of zip archives with many entries
  - Fix for mixing of reused file handles coming from external table (seen on EOS AMS)



### XRootD 5.6.1

- Use kernel provided uuid on macOS
- Set RPATH that works for binaries and libraries on macOS

- HTTP: Fix chunked PUT creating empty files
- SciTokens: Update maximum header size and line length in INI files
- Fix template for default ZTN token location
- Change the thread-id returned to OpenSSL 1.0.x to improve performance
- Insert CRLs containing critical extensions at the end of the bundle
- XrdClHttp: Add pgWrite support to the HTTP client plugin
- Export readv comma separated limits via XRD\_READV\_LIMITS environment variable
- Python: Allow build customization via environment variables (e.g. CXX, CXXFLAGS)
- Fix promotion of **root://** URLs to use TLS encryption (bug introduced in 5.6.0)



### XRootD 5.6.3

- Export project version in XRootDConfig.cmake module
- Create environment file within xrd.adminpath
- Return an error if xrdfs rm fails to delete any file
- Initial packet marking support in HTTP TPC

### XRootD 5.6.4

- Use full certificate chain for verification
- Migrate tests to GoogleTest and run without containers
- Add integrity check for headers and fix header dependency issues
- Fixes on SPARC architecture and GNU/Hurd (external contributions)
- Fix crash on pss.origin directive without specifying a port (uses protocol default)

- Support GCC 14
- Export project version in XRootDConfig.cmake module



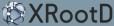
### XRootD 5.6.6

- Use full certificate chain for verification
- Migrate tests to GoogleTest and run without containers
- Add integrity check for headers and fix header dependency issues
- Fixes on SPARC architecture and GNU/Hurd (external contributions)
- Fix crash on pss.origin directive without specifying a port (uses protocol default)

### XRootD 5.6.7

- Fix crash at teardown when using copies with multiple streams
- Fix TPC initialization to take into account control stream (was always using 2 streams before)

- Only claim to be TLS capable if TLS initialization succeeds (--notlsok no longer needed)
- Create CDash dashboard for XRootD and enable submissions in test.cmake
- Fix build on FreeBSD



- XRootD 5.6.9
  - Python
    - Fix iteration over a file with Python3
    - Fix crash with raw strings in prepare call
  - HTTP TPC
    - Fix 500 server response code if X-Number-Of-Streams > 100
  - XrdSciTokens
    - Add stat permissions to create, modify and write operations
    - Allow creation of parent directories if necessary
    - Fix bug when scope includes basepath or /

