XRootD and FTS Workshop @ JSI

Monday, 27 March 2023 - Friday, 31 March 2023

Jozef Stefan Institute

Book of Abstracts

Contents

XRootd presentations / **1**

Data-Aware Scheduling for Opportunistic Resources (with XRootD and HTCondor)

Author: Robin Hofsaess[1]

Co-authors: Gunter Quast [1]; Manuel Giffels [1]; Matthias Jochen Schnepf ; Max Fischer

[1] *KIT - Karlsruhe Institute of Technology (DE)*

Corresponding Author: robin.hofsaess@cern.ch

In the talk, I want to present our ideas for a data-aware scheduling mechanism for our opportunistic resources attached to GridKa, the T1 center in Germany.
Opportunistic resources are non permanent computing sites (partly with cache storages) distributed in Germany that provide resources for the HEP community from time to time.
We are planning to implement a hash-based distribution of datasets to different resources, inspired by Ceph/CRUSH, in combination with HTCondor scheduling and XRootD caching.
This will enable us to to schedule jobs to the according site without the need of a separate data management.

XRootd presentations / **2**

Kubernetes and XrootD

Author: Fabio Andrijauskas[1]

[1] *Univ. of California San Diego (US)*

Corresponding Author: fabio.andrijauskas@ucsd.edu

Bioscience, material sciences, physics, and other research fields require several tools to achieve new results, discoveries, and innovations. All these research fields require computation power. The Open Science Grid (OSG) provides ways to access the computation power from different sites for several research fields. Besides the processing power, it is essential to access the data for all simulations, calculations, and other kinds of processing. To provide data access to all jobs on the OSG, the Open Science Data Federation (OSDF) have ways to create the required data access. The primary way to provide data on OSDF is the XrootD on a Kubernetes infrastructure on the National Research Platform. This work aims to show if there is any overhead using XrootD in a Kubernetes environment. To test this, we set an XrootD origin on bare metal and an XrootD origin using Kubernetes on the same host and request files using files size 500MB, 1GB, and 10GB. The results show a 2% larger performance on the transfer rate using bare metal than Kubernetes XrootD origin. In conclusion, there is no statistical difference between XrootD running on Kubernetes or bare metal.

10 minutes presentation

XRootd presentations / **3**

Experience deploying xCache for CMS in Spain

Author: Carlos Perez Dengra[1]

Co-author: Jose Flix Molina [2]

[1] *PIC-CIEMAT*

[2] *CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES)*

Corresponding Author: carlos.perez.dengra@cern.ch

Over the last few years, the PIC Tier-1 and CIEMAT Tier-2 sites in Spain have been exploring XCache as a content delivery network service for CMS data in the region. This service aligns with the WLCG data management strategy towards HL-LHC. The caching mechanism allows data to be located closer to compute nodes, which has the potential to improve CPU efficiency for jobs, especially for frequently accessed data. Additionally, since many CMS jobs read data from remote sites using XRootD redirectors, there is significant room for improvement using this technology. We have successfully deployed XCache services at both the PIC and CIEMAT sites, and have configured them to cache popular CMS data based on ad-hoc data access popularity studies. Additional previous verification process revealed that there is no significant degradation in CPU efficiency for non I/O intensive tasks reading data from either site in the region, despite the distance between the two sites being 600km with 9ms latency. Hence, a single cache scenario for the region has been studied, with the cache placed at the PIC Tier-1 and serving data to both sites. This presentation aims to highlight our deployment experience and the benefits we have seen from using XCache in the region, as well as potential future use cases in the context of our data management strategy.

FTS - Site Reports / **4**

Update on FTS at RAL

Author: Rose Cooper[None]

Corresponding Author: rose.cooper@stfc.ac.uk

The File Transfer Service (FTS3) is a data movement service developed at CERN, designed to move the majority of the LHC's data across the WLCG infrastructure. Currently, the Rutherford Appleton Laboratory (RAL) Tier 1 runs two production instances of FTS, serving WLCG users (lcgfts3), and the EGI community (fts3egi). During this talk, we are going to present the status of these production instances at the RAL Tier 1 site, as well as changes and developments planned for FTS at RAL over the next year.
The first of the planned changes is in relation to RAL's involvement with the Square Kilometre Array (SKA) experiment, and the UK SKA Regional Centre (UKSRC). Here we are engaged with helping and designing their networking and data transfer requirements, which will begin with the deployment of a SKA FTS instance, so we can begin the testing on their requirements. The second change is the planned integration of token authentication/authorization methods, which aims to improve accessibility to the service to both our existing and new users'communities. Testing is currently underway on integrating our EGI instance with EGI Check-in, and we intend for the SKA instance to integrate with INDIGO IAM once it is deployed.

XRootd presentations / **5**

Porting of XRootD to Windows as a part of EOS-wnc

Author: Gregor Molan[1]

[1] *Comtrade 360's AI Lab*

Corresponding Author: gregor.molan@cern.ch

XRootD provides fast, low latency, and scalable data access. It also provides a hierarchical organization of a filesystem-like namespace organized as a directory. As part of CERN EOS, XRootD assures another possibility for a fast connection for data transfer between the client and the EOS FST.

This is the presentation of Comtrade's work at the CERN's project of productization of EOS, and it is the presentation of XRootD porting to Windows as a part of the EOS client porting from Linux to

Windows. All functionalities of the EOS client ported on Windows should ultimately be the same as those on Linux. XRootD is a part of the EOS client implementation on Linux and the first approach is to port the XRootD to provide EOS implementation on Windows. To make the best use of all the advantages and possibilities of Windows, the transfer of XRootD to Windows is designed to support the functionalities of XRootD and not to transfer the original code from Linux to Windows.

XRootD implementation on Linux is technically investigated as a group of components to port EOS client functionalities from Linux to Windows adequately. The list of external libraries is presented for each of these components. Presented is the list of the majority of Linux libraries used in XRootD, where there are Windows alternatives. If the porting of the XRootD to Windows is limited to essential functionalities, the most important is the port of the xrdcp binary to Windows. Except for networking and security, appropriate libraries for Windows are available for all other functionalities.

According to the determined missing Windows libraries for network and security, network and security should be either implemented on Windows as part of xrdcp or we should provide a Windows version of these libraries. Within a collaboration between CERN openlab and Comtrade, Comtrade invested and provided a port of XRootD and xrdcp binary with no encoded connection (security). Based on Comtrade's estimation, the investment needed for porting missing XRootD libraries to Windows is out of the scope of Comtrade internal investments for XRootD. To complete this implementation, an appropriate outside investment is needed. The final result will be the complete port of the XRootD to Windows. Finally, porting XRootD to the Windows platform would bring additional possibilities for using Windows for particle physics experiments.

XRootd presentations / **6**

XRootD in the UK: ECHO at RAL-LCG2 and developments at Tier-2 sites

Author: James William Walder[1]

[1] *Science and Technology Facilities Council STFC (GB)*

Corresponding Author: james.william.walder@cern.ch

ECHO is the Ceph-backed erasure-coded object store, deployed at the Tier-1 facility RAL-LCG2. It' s frontend access to data is provided via XRootD - using the XrdCeph plugin via the libradosstriper library of Ceph, with a current usable capacity in excess of 40PB.
This talk will cover the work and experiences of optimising for, and operating in, Run-3 of the LHC, and the developments towards future Data Challenges and needs of HL-LHC running.
In addition, a summary of the XRootD activities in the UK is presented, including the ongoing migrations of a number of Tier-2 sites from DPM to a CephFS+XRootD storage solution.

FTS - Communities and Collaborations / **7**

FTS Community Talk: CMS

Author: Katy Ellis[1]

[1] *Science and Technology Facilities Council STFC (GB)*

Corresponding Author: katy.ellis@stfc.ac.uk

This presentation will describe the usage of FTS by the CMS experiment at the Large Hadron Collider during the start of Run-3. I will describe the particular features recently developed for, and employed by CMS for our unique user case as well as current challenges and efforts to optimise performance

on the boundary between FTS and Rucio. I will also discuss the future transfer requirements of CMS.

XRootd presentations / **8**

Open Science Data Federation - OSDF

Author: Fabio Andrijauskas[1]

[1] *Univ. of California San Diego (US)*

Corresponding Author: fabio.andrijauskas@cern.ch

All research fields require tools to be successful. A crucial tool today is the computer. The Open Science Grid (OSG) provides ways to access computational power from different sites. Open science data federation (OSDF) provides data access to the OSG pool using several software stacks. OSDF has received upgrades related to storage space, monitoring checks, monitoring stream collection, and new caches. New monitoring systems provide a way to detect a problem before the user; a new cache can provide more data to the users, new origins create more storage available, and new monitoring streams enable a sophisticated debug model. All these improvements create a new way to provide data to OSG and others. The OSDF is receiving many investments and will create more ways to provide scientific data.

40 minutes presentation

XRootd presentations / **9**

XCache

Author: Matevz Tadel[1]

[1] *Univ. of California San Diego (US)*

Corresponding Author: matevz.tadel@cern.ch

XCache overview
developments in 5.x
plans
30 minutes

XRootd presentations / **10**

XRootD monitoring discussion

Author: Matevz Tadel[1]

[1] *Univ. of California San Diego (US)*

Corresponding Author: matevz.tadel@cern.ch

30 minutes
10 minutes introduction
20 minutes discussion

XRootd presentations / 11

LHCOPN/LHCONE Status and Updates

Authors: Edoardo Martelli[1]; Marian Babik[1]

[1] *CERN*

Corresponding Authors: edoardo.martelli@cern.ch, marian.babik@cern.ch

In this talk we'll give an update on the LHCOPN/LHCONE networks, current activities, challenges and recent updates. We will also focus on the various R&D projects that are currently on-going and could impact XRootD and FTS. Finally, we will also cover our plans for mini-challenges and major milestones in anticipation of the DC24.

XRootd presentations / 12

Welcome

Corresponding Author: jona.javorsek@ijs.si

Welcome and logistics

FTS - State of Affairs / 13

Welcome

Corresponding Author: jona.javorsek@ijs.si

Welcome talk: workshop logistics, Ljubljana survival guide and an introductory word from the head of the institute

XRootd presentations / 14

XRootD Features

Author: Andrew Bohdan Hanushevsky[1]

[1] *SLAC National Accelerator Laboratory (US)*

Corresponding Author: andrew.bohdan.hanushevsky@cern.ch

We will review the new XRootD features added since the last workshop.

15

Kingfisher: Storage Management for Data Federations

Author: Brian Bockelman[1]

[1] *Morgridge Institute for Research*

Corresponding Author: brian.bockelman@cern.ch

A cornerstone of translating the raw capacity of a distributed system into an effective source of shared computing power is the methodical management of all the resources. While one commonly thinks of managing processing resources - CPUs, GPUs, memory - there's surprisingly little attention paid to the management of storage resources. Questions abound: How much storage should be set aside? When can it be reclaimed? How should it be reclaimed? How can it be subdivided?

The Kingfisher project, just beginning, is planning to explore different storage management techniques through the use of its LotMan library which tracks the space usage and local policies for storage. The first intended application for LotMan is the XCache configuration of XRootD and its interaction with HTCondor.

This presentation will cover the basic concepts being developed for LotMan, the use cases we hope to tackle through the year, and be an opportunity for the XRootD developers to discuss on how to best integrate a policy engine into XCache.

XRootd presentations / **16**

To the OSPool and Beyond: The guts of the OSDF client

Author: Brian Bockelman[1]

[1] *Morgridge Institute for Research*

Corresponding Author: brian.bockelman@cern.ch

The Open Science Data Federation (OSDF) delivers petabytes of data each month to workflows running on the OSPool. To do so, one requires a reliable set of client tools. This presentation will take a look "under the hood" of the current OSDF client tooling, covering:

- Discovery of nearby cache instances.
- Acquisition of credentials for transfer, automated or otherwise.
- Experiences maintaining the client in Go.
- Integration with the HTCondor Software Suite.
- Monitoring and telemetry of performance.

Finally, we'll cover the how we plan to make the client more usable, especially in applications beyond the OSPool, over the coming year.

FTS - Communities and Collaborations / **17**

The journey of a file in Rucio

Author: Radu Carpa[1]

[1] *CERN*

Corresponding Author: radu.carpa@cern.ch

This talk focuses on the Rucio data management framework and its interaction with FTS.

XRootd presentations / **18**

RNTuple: ROOT's Event Data I/O for HL-LHC

Author: Jakob Blomer[1]

[1] *CERN*

Corresponding Author: jakob.blomer@cern.ch

This talk provides an introduction to RNTuple, ROOT's designated TTree successor. RNTuple is active R&D, available in the ROOT::Experimental namespace. Benchmarks using common analysis tasks and experiment AODs suggest a 3x - 5x better single-core performance and 10%-20 smaller files compared to TTree. The talk will specifically focus on RNTuple's I/O scheduling and optimization opportunities for remote reading with XRootD.

XRootd presentations / **19**

Experience with XCache in Virtual Placement

Author: Ilija Vukotic[1]

[1] *University of Chicago (US)*

Corresponding Author: ilija.vukotic@cern.ch

Virtual Placement is a way to approximate a CDN-like network for the ATLAS experiment. XCache is an important component in a Virtual Placement mechanism and is expected to substantially improve performance and reliability, while simultaneously decreasing bandwidth needed. I will explain how we configure it, deploy and use it, share experience in more than one year of running it.

XRootd presentations / **20**

Getting the most out of XCache

Author: Ilija Vukotic[1]

[1] *University of Chicago (US)*

Corresponding Author: ilija.vukotic@cern.ch

XCache grew to be a quite stable, performant and function rich caching server for the HEP community.
I will propose a few developments that could help its adoption, simplify and optimize its operation in large distributed systems.

XRootd presentations / **21**

What's up with the XRootD client

Corresponding Author: michal.simon@cern.ch

XRootd presentations / **22**

XRootD Release Schedule and Future Plans

Corresponding Author: amadio@cern.ch

- Current release procedure/automation
- Discussion on development workflow
- Plans for 5.6 and 6.0 releases later this year
- Python bindings (drop Python2 for good, packaging work)

XRootd presentations / **23**

Evolution of XRootD Testing and CI Infrastructure

Corresponding Author: amadio@cern.ch

- Recent CI developments (+Alpine, +Alma, -Ubuntu 18)
- Supported platforms and compilers
- Full (or almost full) migration from GitLab CI to GitHub Actions
- Test coverage and static analysis
- Plans for improving the docker-based tests, running them in CI

XRootd presentations / **24**

OU XRootD Site Report

Corresponding Author: hs@nhn.ou.edu

XRootd presentations / **25**

XRootD usage at GSI

Corresponding Author: s.fleischer@gsi.de

XRootd presentations / **26**

Analysis of data usage at BNL

Corresponding Author: hito@bnl.gov

XRootd presentations / **27**

XCache Developments & Plans

Corresponding Author: matevz.tadel@cern.ch

XRootd presentations / **28**

Getting the most out of XCache

Corresponding Author: ilija.vukotic@cern.ch

XRootd presentations / **29**

Experience with XCache in Virtual Placement

Corresponding Author: ilija.vukotic@cern.ch

XRootd presentations / **30**

Kingfisher: Storage Management for Data Federations

Corresponding Author: brian.bockelman@cern.ch

XRootd presentations / **31**

XRootD pgRead & pgWrite

Corresponding Author: andrew.bohdan.hanushevsky@cern.ch

XRootd presentations / **32**

XrdEc: the whole story

Corresponding Author: michal.simon@cern.ch

XRootd presentations / **33**

XRootD Plugins

Corresponding Author: andrew.bohdan.hanushevsky@cern.ch

XRootd presentations / **34**

XCache Developments & Plans

Corresponding Author: matevz.tadel@cern.ch

XRootd presentations / **35**

XRootD Plugins

Corresponding Author: andrew.bohdan.hanushevsky@cern.ch

XRootd presentations / **36**

A Brief History of the dCache Xroot Implementation (virtual)

Corresponding Author: arossi@fnal.gov

Workshop wrap-up / **37**

Don’t be a Stranger, Please! :-)

Corresponding Author: michal.simon@cern.ch

Workshop wrap-up / **38**

Many Thanks & Future Outlook

Corresponding Author: andrew.bohdan.hanushevsky@cern.ch

XRootd presentations / **39**

Outlook on EOS, the CERN storage solution for LHC Run3 and beyond

Corresponding Author: amadio@cern.ch

FTS - Site Reports / **40**

FTS3 at FNAL (virtual)

Authors: Lorena Lobato Pardavila[1]; Lorena Lobato Pardavila[None]

[1] *Fermi National Accelerator Lab. (US)*

Corresponding Authors: lorena.lobato@cern.ch, llobato@fnal.gov

- Outline
- Introduction
- Configurations
  - CMS configuration –physical server
  - Public configuration –containers
- Differences
- Advantages & disadvantages of each configuration
- Summary

FTS - State of Affairs / **42**

FTS3: State of affairs

Author: Mihai Patrascoiu[1]

Co-authors: Joao Pedro Lopes ; Steven Murray [1]

[1] *CERN*

Corresponding Author: mihai.patrascoiu@cern.ch

Last year in review, showcasing the evolution of the FTS project, as well as touching on what's new in the FTS world, community engagement and the future direction.

FTS - State of Affairs / **43**

Tape, REST API and more

Author: Joao Pedro Lopes[None]

Co-author: Mihai Patrascoiu [1]

[1] *CERN*

Corresponding Author: joao.pedro.batista.lopes@cern.ch

This talk will present recent QoS improvements, go into the details of the Tape REST API and how it is implemented in FTS & Gfal2, showcase Gfal2 tape interaction over HTTP and finally, look at what's upcoming in the tape world, such as Archive Metadata and Tape REST API evolution.

FTS - State of Affairs / **44**

FTS & Tokens

Author: Shubhangi Misra[None]

Co-author: Mihai Patrascoiu [1]

[1] *CERN*

Corresponding Authors: shubhangi.misra@cern.ch, mihai.patrascoiu@cern.ch

This talk will describe the future strategy of tokens in FTS, as well as implementation milestones to fully integrated tokens into the FTS landscape.

FTS - Site Reports / **45**

FTS3 @ CERN

Author: Steven Murray[1]

Co-authors: Joao Pedro Lopes ; Mihai Patrascoiu [1]

[1] *CERN*

Corresponding Author: steven.murray@cern.ch

The FTS3 @ CERN site report, presenting the number of instances, volume of data served each year, database setup and various operation tips and tricks discovered throughout the years.

FTS - Communities and Collaborations / **46**

FTS3: Cloud Storage transfers

Author: Mihai Patrascoiu[1]

Co-author: Eraldo Silva Junior [2]

[1] *CERN*

[2] *CBPF - Brazilian Center for Physics Research (BR)*

Corresponding Author: mihai.patrascoiu@cern.ch

This presentation will show all that's needed to get your FTS instance configured to serve cloud storage transfers. The final part of the presentation will show our plan to simplify this process and make things easier to configure and more intuitive over all.

FTS - Monitoring / **47**

FTS3@CERN: Service Health Monitoring

Author: Mihai Patrascoiu[1]

Co-authors: Joao Pedro Lopes ; Steven Murray [1]

[1] *CERN*

Corresponding Author: mihai.patrascoiu@cern.ch

This talk will show an overview of the health and alarm metrics used at the FTS3@CERN deployment. The full lifecycle will be presented, from the software changes and scripts needed, to logging extraction via FluentBit and ultimately to the Grafana display.

FTS - Monitoring / **48**

FTS3: The Monitoring Zoo

Author: Joao Pedro Lopes[None]

Co-author: Mihai Patrascoiu [1]

[1] *CERN*

Corresponding Author: joao.pedro.batista.lopes@cern.ch

The word "monitoring" is used everywhere in the FTS world. This talk wants to dive into the different types of monitoring present in the FTS world and explain what each of them means.

FTS - Communities and Collaborations / **49**

FTS Community Talk: ATLAS (virtual)

Author: Mario Lassnig[1]

[1] *CERN*

Corresponding Author: mario.lassnig@cern.ch

The ATLAS view on data management and FTS involvement

FTS - Communities and Collaborations / **50**

FTS Community Talk: LHCb (virtual)

Author: Ben Couturier[1]

[1] *CERN*

Corresponding Author: ben.couturier@cern.ch

FTS - Communities and Collaborations / **51**

EGI Data Transfer Activities (virtual)

Author: Andrea Manzi[None]

Corresponding Author: andrea.manzi@egi.eu

The talk will focus on the activities in EGI related to Data transfer and orchestration, in particular focusing on integration with EGI Check-in AAI in the context of the EGI-ACE project and the new EOSC Data transfer service in EOSC Future project. An overview of the new EGI lead project inter-Twin will be also given and the role FTS has there in the infrastructure supporting Scientific Digital Twins.

FTS - Site Reports / **52**

FTS3: BNL Deployment

Author: Hironori Ito[1]

[1] *Brookhaven National Laboratory (US)*

Corresponding Author: hito@bnl.gov

An overview of the FTS3 deployment at BNL

XRootd presentations / **53**

A Prometheus XRootD exporter based on mpxstats (virtual)

Corresponding Author: j.knedlik@gsi.de

FTS - Monitoring / **54**

FTS-Noted Project (virtual)

Author: Edoardo Martelli[1]

Co-author: Maria Del Carmen Misa Moreira [1]

[1] *CERN*

Corresponding Authors: maria.del.carmen.misa.moreira@cern.ch, edoardo.martelli@cern.ch

An overview of the FTS-Noted project, aimed at shaping traffic through dynamic network switches.

FTS - Communities and Collaborations / 55

FTS-Alto project (virtual)

Author: Y. Richard Yang[None]

Corresponding Author: yang.r.yang@yale.edu

An overview of the FTS-Alto project in collaboration with Dr. Richard Yang and his research group (Yale University)