

Recent evolutions in the LTDP CDF project

Federico Fornari^{1,*}, *Enrico Fattibene*¹, *Alessandro Cavalli*¹, *Daniele Cesini*¹,
*Antonio Falabella*¹, *Lucia Morganti*¹, *Andrea Prosperini*¹, *Vladimir Sapunenko*¹

¹CNAF - Italian Institute for Nuclear Physics, Bologna, Italy

Abstract. In the latest years, CNAF (the national center of the Italian Institute for Nuclear Physics - INFN - dedicated to Research and Development on Information and Communication Technologies) has been working on the Long Term Data Preservation (LTDP) project for the CDF experiment, active at Fermilab from 1990 to 2011. The main aims of the project are to protect the most relevant part of the CDF RUN-2 data collected between 2001 and 2011 and already stored on tape at CNAF (4 PB), as well as to ensure the availability and the access to the analysis facility to those data over time. Lately, the CDF database, hosting information about CDF datasets such as their structure, file locations and metadata, has been imported from Fermilab to CNAF. Also, the Sequential Access via Metadata (SAM) station data handling tool for CDF data management, that allows to manage data transfers and to retrieve information from the CDF database, has been properly installed and configured at CNAF. This was a fundamental step in the perspective of a complete decommissioning of CDF services on Fermilab side. An access system has been designed and tested to submit CDF analysis jobs, using CDF software distributed via CERN Virtual Machine File System (CVMFS) and requesting delivery of CDF files stored on CNAF tapes, as well as data present only on Fermilab storage archive. Moreover, the availability and the correctness of all CDF data stored on CNAF tapes has been verified. This paper describes all these recent evolutions in detail, presenting the future plans for the LTDP project at CNAF.

1 Introduction

From the very beginning of its activities as a computing center, CNAF took part in the CDF collaboration, providing computing and storage resources for the CDF experiment [1]. Even after the end of CDF data taking activities in 2011, Fermi National Accelerator Laboratory (FNAL) continued to collaborate with CNAF in order to build and maintain a CDF Long Term Data Preservation (LTDP) facility at CNAF, with the aim to preserve the data produced by the detector for future analyses and computing activities. Technically this kind of activity is defined as “framework preservation”, implying the possibility for previously authorized scientific communities to access and use data through dedicated software services [2]. At present, CNAF hosts all the raw data and part of analysis-level Monte Carlo n-tuples belonging to the so-called CDF RUN-2 data taking phase, that lasted from 2001 to 2011. These data were copied from FNAL storage facilities to CNAF data center before 2017, with the primary purpose of keeping an extensive backup copy of the experiment’s most relevant datasets. In

*e-mail: federico.fornari@cnaif.infn.it

addition, the CDF Oracle database containing all RUN-2 metadata has been migrated from FNAL to CNAF.

2 The CDF data handling system at FNAL

The data handling system of CDF is based on the sequential access model developed originally by the D0 experiment [3]. Events are stored in ROOT [4] files, with one or more event branches, and grouped in datasets. The data volume of the CDF experiment is of multi-PByte. It consists of raw, reconstructed, n-tuples and Monte Carlo data. The data handling system at FNAL is shown in Figure 1. To store data on tape, manage and control tape resources, CDF at FNAL uses the Enstore Mass Storage System (MSS) [5]. This system was developed by Fermilab and is operated by the Storage Services Administration group of the FNAL Computing Division. The CDF software requires data to be on random access medium to access it. CDF at FNAL uses the dCache software [6] to fetch data from tape and manage the disk inventory [7].

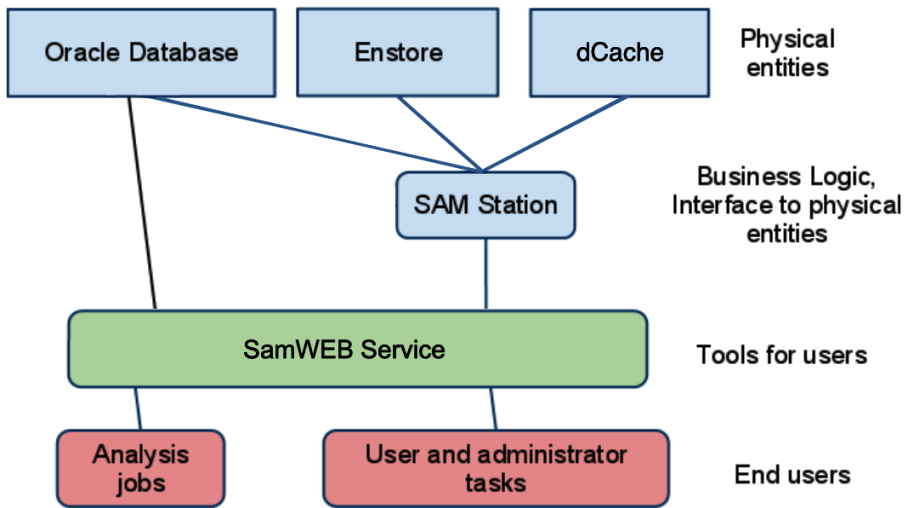


Figure 1. Schematic overview of the CDF data handling system at FNAL. The various components are described in the following sections.

2.1 SAM station

The SAM station is a set of code products which interact with an Oracle database in order to store and give access to metadata about CDF data files, such as their locations and contents. A product-management package called Unix Product Support (UPS) [8], developed at Fermilab, provides access to software products commonly used by Fermilab experiments. UPS can handle multiple databases and multiple versions of a software, thus being used to setup the SAM station. In order to analyze RUN-2 data, CDF collaboration has decided to use the basic SAM functionalities as explained in the following. The original form of the SAM products includes a command-line and a C++ interface called `sam_cpp_api`. The latter is used when a CDF executable uses SAM functionality through the `diskcache_i` software, which is a UPS product that delivers files to AspectC++ (AC++) jobs. Both the command-line and the C++

interface use an obsolete and proprietary network communication system called CORBA [9]. Consequently, SAM was updated to SamWEB, which replaces `sam_cpp_api` and the previous command-line interface with an `http` interface [10]. CDF RUN-2 software uses `dCache` and `diskcache_i` with SamWEB. There are also two products which wrap the `http` calls: `samweb` client and `ifdh` client.

2.2 SamWEB

SamWEB is a `http`-based server application developed with Apache, and can be accessed with a web browser or via `curl` in a shell. `https` connections are mandatory in order to perform CDF files delivery to AC++ jobs, but require X.509 certificates. In order to access files from the CDF `dCache` servers, users must have their CILogon [11] certificate's Distinguished Name (DN) added to the server access list by the Scientific Data Storage and Access group. `http` requests from users can be sent directly to the server via `curl` or via a wrapper. The wrapper methods are denoted as `samweb` and `ifdh`. For both these clients, Open Science Grid CERN Virtual Machine File System [12] (OSG CVMFS), hosting both the Fermilab and CDF software repositories, is required. Authentication is enforced for important commands, so a Kerberos certificate in its X.509 format is needed.

`ifdh` stands for Intensity Frontier Data Handling client, and is used by CDF jobs to manage SAM projects. A SAM project represents the pool of files available for delivery, and the collection of processes pulling files from that pool. The project is typically started by specifying the dataset definition name. Then, the SAM snapshot, i.e. a list of files from a query at some time, is created. Internal to SAM, a project is the unix process called `pmaster` running on the station node. `pmaster` responds to requests from processes and coordinates the delivery of files specific to the project.

`ifdh` is setup to perform copy of files through `dCap/dccp` (a POSIX like interface for accessing `dCache` using a proprietary data transfer protocol, which can emulate POSIX access across the LAN or WAN) from CDF Enstore via PNFS (Parallel Network File System) and must be used on systems where PNFS is not mounted. In addition, an X.509 certificate is required, and the user's DN must be present in the `dCache` access file.

3 Moving CDF data handling system to CNAF

Starting from the end of February 2019, Fermilab has decided to give up maintaining the CDF data handling system and turned off all its SAM station and SamWEB instances. As a matter of fact, these applications need to run on a Scientific Linux 6 operating system, which has been made end-of-life starting from November 20th 2020. Fermilab will continue supporting Scientific Linux 6 and Scientific Linux 7 based on Red Hat Enterprise Linux 6 and Red Hat Enterprise Linux 7, respectively, but moving forward it is switching over to CentOS. Exactly the same policy is adopted by CNAF. In order to maintain the CDF data handling system in a LTDP perspective, CNAF has chosen Singularity [13], a virtualization software that allows to run the CDF system in Scientific Linux 6 containers on a CentOS 7 host machine.

CNAF instantiated a couple of CentOS 7 servers and deployed the CDF services (SAM station and SamWEB) using Scientific Linux 6 Singularity containers. On the first machine, a SAM station named `cdf-cnaf-ltdp`, together with its SamWEB `http` server, has been installed and configured to manage CDF files delivery from CNAF tape system [14], based on IBM Spectrum Protect (formerly Tivoli Storage Manager - TSM) and IBM Spectrum Scale (formerly General Parallel File System - GPFS). On the second machine, a SAM station named `cdf-fnal-ltdp` together with its SamWEB `http` server has been installed and configured to manage CDF files delivery from FNAL Enstore/`dCache`. This is because some CDF users might

need to run jobs using “production” files present only on FNAL tape system. CDF raw data files are reconstructed by a CDF “production” executable and the resulting outputs are called “production” files. The CDF data copied at CNAF (about 4 PBytes) only consists of raw data files and condensed production output n-tuples. Moreover, the reason for the implementation of two separate SAM stations lies in the fact that the services must be started with an unchangeable configuration parameter indicating which location prefix (CNAF’s or FNAL’s) to retrieve for the files.

The workflow to access CDF data through the CNAF data handling system is shown in Figure 2. Once users have completed the registration procedure at CNAF, they are able to access the CDF services setup for data analysis. As a matter of fact, if they only need to analyze CDF files that are already at CNAF, they can interact with the SamWEB server interfacing with *cdf-cnaf-ltdp* SAM station. The CDF job will setup CDF software from OSG CVMFS, which is properly mounted on the machine, and then it will use *ifdh* to copy files locally through *cp* command from CNAF tapes via GPFS. *ifdh* interacts with SamWEB server, interfacing with the SAM station, which in turn sends a query to the Oracle database and gets the proper file locations (the ones starting with CNAF storage mount point). These are passed to *ifdh*, which is then able to construct and execute the specific *cp* command. An X.509 certificate is required as well, but in this case it is not the CILogon one for Fermilab, since CILogon Base is not present among the recognized Grid Certification Authorities on servers at CNAF. Each CDF user accessing the machine hosting the *cdf-cnaf-ltdp* SAM station is then provided with a personal X.509 certificate (with standard one year duration) issued by a custom Grid Certification Authority, named CDF Custom CA, created on purpose with Globus Toolkit [15] and installed on the machine. This means that no Kerberos client is required.

In case of data present only at FNAL, the *cdf-fnal-ltdp* SAM station is used as interface to the FNAL storage infrastructure. In this case a Kerberos client is required, in order to initialize a X.509 CILogon certificate and authenticate against FNAL dCache servers. When files are

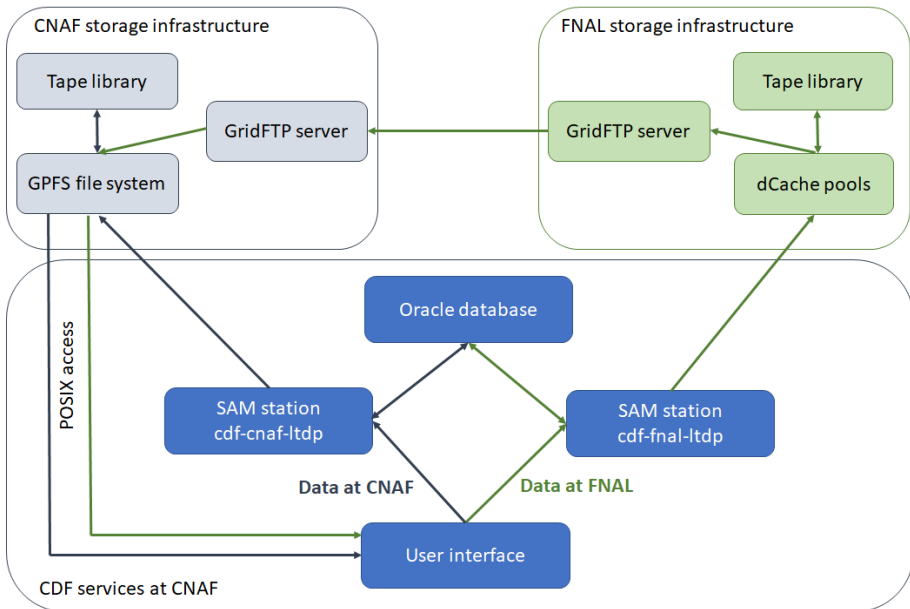


Figure 2. Flowchart showing CDF files delivery procedure from both CNAF and FNAL tape systems.

read from FNAL tape system, they are transferred to CNAF through GridFTP services. In order to know whether required files are already at CNAF and decide which SamWEB server interact with, a user can exploit the `samweb locate-file` command. If the answer for a single file contains the CNAF storage mount point, it means that the file is already at CNAF and there is no need to transfer it from FNAL.

3.1 Relevant results and future steps

A couple of CDF users have been selected in order to perform some tests of interactive job execution with positive results. Based on this experience and on CDF users' feedbacks, CNAF is planning to write a complete user guide explaining how to access CDF data through the CNAF data handling system.

Through the CDF Oracle database and the SAM stations installed at CNAF, all CNAF CDF data files were recently re-read from tape and the ADLER32 checksum of each file was calculated and compared with the original one, saved in the database. All non-consistent files have now been successfully retransferred from FNAL Enstore/dCache.

Next steps in the development of the CDF LTDP infrastructure at CNAF include the integration of the newly configured HTCondor [16] batch system for CNAF computing farm. CDF users at FNAL are provided with JobSub [17], which is a custom framework developed at Fermilab to submit Condor jobs to remote Condor pools. JobSub manages job submission through a Glidein Based WMS (Workload Management System) located at UCSD (University of California San Diego) that works on top of HTCondor, providing a simple way to access remote Grid resources. Some issues with pilot jobs authentication still happen: the Glidein Based WMS proxy-delegation mechanism does not follow VOMS standard [18], preventing CNAF HTCondor CEs from recognizing CDF Virtual Organization in the proxy attributes. However, these problems should be hopefully solved collaborating with OSG HTC group.

4 Conclusions

The CNAF CDF LTDP system is able to meet the requirements of the CDF scientific collaboration, not only for data preservation, but even for framework maintenance, meaning all the services and software needed to access and deal with data, in a long term perspective. In conclusion, this system is capable of providing different categories of users with the possibility of accessing and analyzing CDF data and can be a valid example of a data framework preservation model for other scientific experiments.

References

- [1] Sill A. et al., *Globally Distributed User Analysis Computing at CDF*, Proceedings of CHEP '04, 1139-1144 (2004)
- [2] Ricci P. P. et al., *Last developments of the INFN CNAF Long Term Data Preservation (LTDP) project: the CDF data recover and safekeeping*, IOP Conf. Series: Journal of Physics: Conf. Series **1085**, 032050 (2018)
- [3] Terekhov I. and White V., *Distributed Data Access in the Sequential Access Model in the D0 Experiment at Fermilab*, High-Performance Distributed Computing Proceedings, 310-311 (2000)
- [4] Brun R. and Rademakers F., *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Inst. & Meth. in Phys. Res. **A389**, 81-86 (1997)

- [5] Baisley W. et al., *Fermilab's Multi-Petabyte Scalable Mass Storage System*, Proceedings of the 22nd IEEE/13th NASA Goddard Conference on Mass Storage Systems and Technologies, 101-108 (2005)
- [6] The dCache project website <http://www.dcache.org/> accessed: 2020-03-13
- [7] Bakken J. et al., *The Fermilab Data Storage Infrastructure*, Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 101-104 (2003)
- [8] Votava M. et al., *UPS UNIX Product Support*, Seventh Conference Real Time '91 on Computer Applications in Nuclear, Particle and Plasma Physics Conference Record, 156-159 (1991)
- [9] Henning M. and Vinoski S., *Advanced CORBA Programming with C++* (Addison Wesley Longman, Boston, 1999)
- [10] Illingworth R., Mengel M. and Norman A., *Re-engineering SAM or Changing the Engine in the Train While it is Running*, Journal of Physics Conference Series **664**, 6 (2015)
- [11] Basney J., Fleury T. and Gaynor J., *CILogon: A federated X.509 certification authority for cyberinfrastructure logon*, Concurrency and Computation: Practice and Experience **26** (2014)
- [12] Aguado Sanchez C. et al., *CVMFS - a file system for the CernVM virtual appliance*, Proceedings of XII Advanced Computing and Analysis Techniques in Physics Research **1**, 52 (2008)
- [13] Kurtzer G. M., Sochat V. and Bauer M. W., *Singularity: Scientific containers for mobility of compute*, PLoS ONE **12**, 5 (2017)
- [14] Cavalli A., Cesini D., Fattibene E., Prosperini A. and Sapunen V., *HSM and backup services at INFN-CNAF*, EPJ Web of Conferences **214**, 04009 (2019)
- [15] Foster I., *The Globus Toolkit for Grid Computing*, Proceedings of the 1st International Symposium on Cluster Computing and the Grid, 1 (2001)
- [16] The HTCondor High Throughput Computing website <https://research.cs.wisc.edu/htcondor/> accessed: 2020-03-13
- [17] Box D., *FIFE-Jobsub: a grid submission system for intensity frontier experiments at Fermilab*, Journal of Physics: Conference Series **513**, 032010 (2014)
- [18] Alfieri R. et al., *VOMS, an Authorization System for Virtual Organizations*, Grid Computing, 33-40 (2004)