

eXtreme-DataCloud project: Advanced data management services for distributed e-infrastructures

A. Costantini¹, D. Cesini¹, D.C. Duma¹, D. Michelotto¹, A. Falabella¹, L. Dell’Agnello¹, D.Salomoni¹, L. Morgantii¹, G. Grandi²

¹ INFN-CNAF, Bologna, Italy

² INFN Bologna, Bologna, Italy

E-mail: alessandro.costantini@cnaf.infn.it

Abstract. The development of new data management services able to cope with very large data resources is becoming a key challenge. Such capability, in fact, will allow the future e-infrastructures to address the needs of the next generation extreme scale scientific experiments. To face this challenge, in November 2017 the H2020 eXtreme DataCloud - XDC project has been launched. Lasting for 27 months and combining the expertise of 8 large European research organisations, the project aims at developing scalable technologies for federating storage resources and managing data in highly distributed computing environments. The project is use-case driven with a multidisciplinary approach, addressing requirements from research communities belonging to a wide range of scientific domains: High Energy Physics, Astronomy, Photon and Life Science, Medical research. XDC will implement data management scalable services, combining already established data management and orchestration tools, to address the following high level topics policy driven data management based on Quality-of-Service, Data Life-cycle management, smart placement of data with caching mechanisms to reduce access latency, meta-data with no predefined schema handling, execution of pre-processing applications during ingestion, data management and protection of sensitive data in distributed e-infrastructures, intelligent data placement based on access patterns. This contribution will introduce the project, presents the foreseen overall architecture and the contribution to the activities that are being carried on by INFN-CNAF personnel to achieve the project goals and objectives.

1. Introduction

Lead by INFN-CNAF, the eXtreme DataCloud (XDC) project [1] develops scalable technologies for federating storage resources and managing data in highly distributed computing environments. The provided services are capable of operating at the unprecedented scale required by the most demanding, data intensive, research experiments in Europe and Worldwide. The targeted platforms for the released products are the already existing and the next generation e-Infrastructures deployed in Europe, such as the European Open Science Cloud (EOSC) [2], the European Grid Infrastructure (EGI) [3], the Worldwide LHC Computing Grid (WLCG) [4] and the computing infrastructures that will be funded by the upcoming H2020 EINFRA-12 call. XDC is funded by the H2020 EINFRA-21-2017 Research and Innovation action under the topic Platform-driven e-Infrastructure innovation [5]. It is carried on by a Consortium that brings together technology providers with a proven long-standing experience in software development and large research communities belonging to diverse disciplines: Life Science,

Biodiversity, Clinical Research, Astrophysics, High Energy Physics and Photon Science. XDC started on 1st November 2017 and will run for 27 months until January 2020. The EU contribution for the project is 3.07 million euros. XDC is a use case driven development project and the Consortium has been built as a combination of technology providers, Research Communities and Infrastructure providers. New developments will be tested against real-life applications and use cases. Among the high level requirements collected from the Research Communities, the Consortium identified those considered more general (and hence exploitable by other communities), with the greatest impact on the user base and that can be implemented in a timespan compatible with the project duration and funding.

2. Project Objectives

The XDC project develops open, production quality, interoperable and manageable software that can be easily plugged into the target European e-Infrastructures and adopts state of the art standards in order to ensure interoperability. The building blocks of the high-level architecture foreseen by the project are organized in a manner to avoid duplication of development effort. All the interfaces and links to implement the XDC architecture are developed exploiting the most advanced techniques for authorization and authentication. Services are scalable to cope with most demanding, extreme scale scientific experiments like those run at the Large Hadron Collider at CERN and the Cherenkov Telescope Array (CTA), both of them represented in the project consortium. The project will enrich already existing data management services by adding missing functionalities as requested by the user communities. The project will continue the effort invested by the now ended INDIGO-DataCloud project [6] in the direction of providing friendly, web-based user interfaces and mobile access to the infrastructure data management services. The project will build on the INDIGO-DataCloud achievements in the field of Quality of Services and data lifecycle management developing smart orchestration tools to realize easily an effective policy driven data management. One of the main objectives of the project is to provide data management solutions for the following use cases:

- Dynamic extension of a computing center to a remote site providing transparent bidirectional access to the data stored in both locations.
- Dynamic inclusion of sites with limited storage capacity in a distributed infrastructure, proving transparent access to the data stored remotely.
- Federation of distributed storage endpoints, i.e. a so-called WLCG Data Lake, enabling fast and transparent access to their data without a-priori copy.

These use cases will be addressed implementing intelligent, automatic and hierarchical caching mechanisms.

3. Overall project structure

The project is structured in five Work Packages (WPs). In details, there are two WPs devoted to networking activities (NA), one for service activities (SA) and two for development activities (JRA). The relationships among the WPs are represented in Figure 1: NA1 will supervise the activities of all the Work Packages and will deal with the management aspects by ensuring a smoothly progress of the Project activities. NA2, representing communities, will provide requirements that will guide the development activities carried out by JRA1 and JRA2. JRA1 and JRA2 will be responsible to provide new and integrated solutions to address the user requirements provided by NA2. SA1 will provide Quality Assurance policies and procedures and the Maintenance and Support coordination, including Release Management. NA2 will close the cycle validating the released components on the pilot testbeds made available for exploitation by SA1.

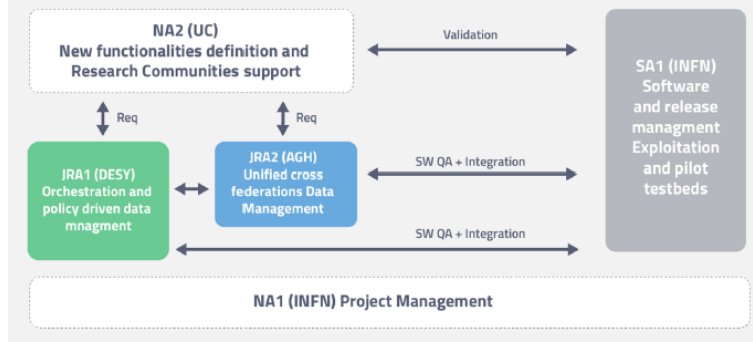


Figure 1. Relationships between the WPs in the XDC project.

Work Package 1 (NA1) bring the project towards the successful achievement of its objectives, efficiently coordinating the Consortium and ensuring that the activities progress smoothly. Coordinated by INFN-CNAF, WP1 is responsible for the financial administration of the project; it controls the effort reporting and the progress of the work ensuring that they adhere to the work plan and to the Grant Agreement. WP1 defines the Quality Assurance plan and reports periodically to the EC about the overall project progress. It is responsible for the resolution of internal conflicts and for the creation of a fruitful spirit of collaboration, ensuring that all the partners are engaged to fulfil the project objectives. WP1 communicates the project vision and mission to the relevant international events and interested institutions. In particular, INFN-CNAF has been in charge of the organization of the joint eXtreme-DataCloud (XDC) and the DEEP-HybridDataCloud kickoff meeting [7], hosted in Bologna (Italy) on 23-25 January, 2018.

Work Package 2 (NA2) identify new functionalities for the management of huge data volumes from the different Research Communities Use Cases providing requirements to the existing tools developers to enhance the user experience. WP2 is also responsible for testing the new developments and for providing adequate feedback about the user experience of the different services. WP2 analyzes the scalability requirements for the infrastructure taking into account the challenges expected in next years and the new frameworks of scientific data. INFN-CNAF mainly participates to WP2 activities by supporting the WLCG community and related LHC experiments through the figure of Champion who is a member of Research Communities/Infrastructures that understands very well the needs of the use case and has a general understanding of the available solutions and features. In particular, WLCG Champion at INFN-CNAF has the role to harmonize the XDC development activities with those carried on within the WLCG community having the objective to improve the software solutions in terms of scalability, usability, maintainability, interoperability and cost effectiveness, looking forward the HL-LHC high demanding data taking conditions.

Work Package 3 (SA1) provides software lifecycle management services together with pilot e-Infrastructures

- For the project developers in order to ensure the maintenance of the high quality of the released software components and services while adding new features
- Through Continuous Integration and Delivery, followed by deployment and operation/monitoring in production environment
- For the User communities, in order to ensure the delivered software consistently pass the customer acceptance criteria and continually improve its quality
- For the e-Infrastructure providers, in order to ensure an easy and smooth delivery of improved software services and components while guaranteeing the stability of their

production infrastructures.

INFN-CNAF is coordinating WP3 and the main activities are related to:

- **Software Lifecycle Management:** expressed in terms of Software Quality Assurance and Software Release and Maintenance, CNAF is coordinating the management of those software products that became officially part of the first XDC releases, codenamed Pulsar [8], foreseen for late 2018 and effectively released in January 2019. CNAF is also coordinating the implementation of the continuous software improvement process, following a DevOps approach, through the definition and realization of an innovative Continuous Integration (CI) and Delivery (CD) system.
- **Pilot Infrastructure Services:** CNAF, with the support of the project partners, is providing and maintaining the testbeds dedicated to developers, software integration and software preview. In particular, the activities were focused in implementing the services needed to support the software development and release management and included among others the source code repository, and continuous integration system.
- **Exploitation activities:** focused in bridging with infrastructure providers which are the targets for the XDC software together with the user communities (WP2). Among the exploitation activities, INFN-CNAF actively participated on a task devoted to creating a Service Providers Board and establish communication channels with providers.

Moreover, INFN-CANF is hosting and maintaining the XDC collaborative tools put in place for an effective project communication among partners (web site, INDICO agent, mailing lists, document repository, video conference, issue tracking system, project management and content collaboration).

Work Package 4 (JRA1) provides the semi or fully automated placement of scientific data in the Exabyte region on the site (IaaS) as well as on the federated storage level. In the context of this Work Package, placement may either refer to the media the data is stored on, to guarantee a requested Quality of Service, or the geographical location, to move data as close to the compute facility as possible to overcome latency issues in geographically distributed infrastructures. In the latter case, data might either be permanently moved, or temporarily cached. In WP4, INFN-CNAF contributes in the development of an HTTP caching system based on NGINX [9] web server. Serving as a content cache, the XDC-HTTP caching service, can be deployed in several WLCG data management workflows, given that many of the software solutions support the HTTP protocol for data operations. In particular, this activity carried on in WP4 aims to add support for VOMS proxy certificates exploiting the modularity of NGINX to develop an additional module that inspects the VOMS proxy certificate attributes. INFN-CNAF contributed also to the geographically scalability tests of the Xcache system, developed within WP4 activities, in particular in the deployment and related support of the software in the national testbed.

Work Package 5 (JRA2) provides enriched high-level data management services unifying access to heterogeneous storage resources and services, enable extreme scale data processing on both private as well as public Cloud computing resources using established interfaces and allowing usage of legacy applications without the need for rewriting them from scratch. These functionalities will be provided mainly by Onedata distributed virtual file-system platform. In particular, INFN-CANF is in charge to deploy and test the new features released by WP5 by adopting the services and solution provided by WP3.

4. General architecture

The XDC project aims at providing advanced data management capabilities that require the execution of several tasks and the interaction among several components and services. Those

capabilities should include but are not limited to QoS management, preprocessing at ingestion and automated data transfers. Therefore global orchestration layer is needed to take care of the execution of those complex workflows. Figure 2 highlights the main components and their role among the three different levels: Storage, Federation, and Orchestration.

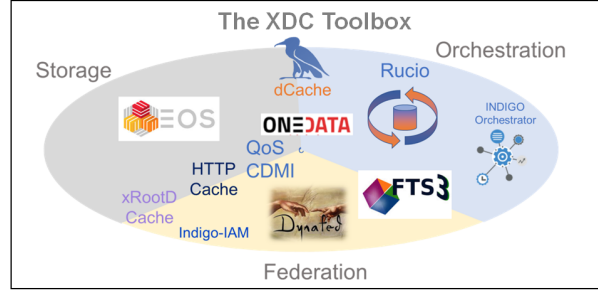


Figure 2. XDC main components and related roles.

Figure 3 highlights the high level architecture of the XDC project by describing the components and the related connections.

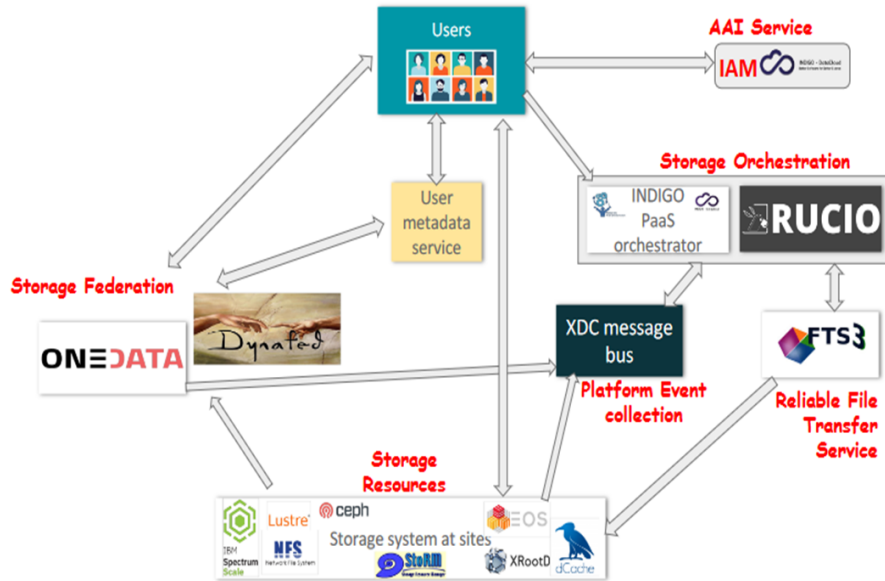


Figure 3. High level architecture of the XDC project.

4.1. XDC Orchestration system

In the XDC project the global orchestration layer is needed to take care of the execution of those complex workflows. The orchestration covers two essential aspects:

- The overall control, steering and bookkeeping including the connection to compute resources
- The orchestration of the data management activities like data transfers, and data federation.

Consequently we have decided to split the responsibilities between two different components: the INDIGO Orchestrator [10] and Rucio [11]. The INDIGO PaaS Orchestrator, the system wide

orchestration engine, is a component of the PaaS layer that allows to instantiate resources on Cloud Management Frameworks (like OpenStack and OpenNebula) and Mesos clusters. It takes the deployment requests, expressed through templates written in TOSCA YAML Profile 1.0 [12], and deploys them on the best cloud site available. The Rucio project, the data management orchestration subsystem, is the new version of ATLAS Distributed Data Management (DDM) system services for allowing to manage the large volumes of data, both taken by the detector as well as generated or derived, in the ATLAS distributed computing system. Rucio is used to manage accounts, files, datasets and distributed storage systems. Those two components, the PaaS Orchestrator and Rucio, provide different capabilities and can complement each other to offer a full set of features to meet the XDC requirements. Rucio implements the data management functionalities missing in the INDIGO Orchestrator: the Orchestrator will make use of those capabilities to orchestrate the data movement based on policies.

4.2. XDC Quality-of-Service implementation

The idea to provide scientific communities or individuals with the ability to specify a particular quality of service when storing data, e.g. the maximum access latency or minimum retention policy, was introduced within the INDIGO-DataCloud project. In XDC, the QoS concept is envisioned to consistently compliment all data related activities. In other words, whenever storage space is requested, either manually by a user or programmatically by a framework, the quality of that space can be negotiated between the requesting entity and the storage provider

4.3. Caching within XDC

In this section we consider how the XDC architecture treats the storage and access of data, building a hierarchy of components whose goal is to maximise the accessibility of data to clients while minimising global infrastructure costs. The architecture considers a set of multi-site storage systems, potentially accessed through caches, both of which are aggregated globally through a federation. To such purpose, various technologies are available to the project to serve as the basis of an implementation:

- The system runs native dCache [13] or EOS, but operates in a "caching mode" staging data in when a cache miss occurs.
- A service such as Dynafed [14] will be augmented to initiate data movement. While it would only hold only metadata, it would use a local storage system for holding the data
- A standalone HTTP cache could be built from existing web technology, such as NGINX, modified for horizontal scalability and relevant AAI support.

4.4. XDC data management and new developments

Data management functionality for end users will be also available via the Onezone component of the Onedata platform [15]. Onezone will provide single-sign on authentication and authorization for users, which will be able to create access tokens to perform data access activities via the Web browser, REST API or using Onedata POSIX virtual filesystem. Onezone will enable federating multiple storage sites by deploying Oneprovider services on top of the actual storage resources provisioned by the sites. For the purpose of job scheduling and orchestration, Onedata will communicate with Indigo Orchestrator component by means of a message bus, allowing the orchestrator to subscribe for events related to data transfers and data access. This will allow Orchestrator to react to changes in the overall system state (e.g. a new file in a specific directory or space, data distribution changes initiated by manual transfers, cache invalidation or on-the-fly block transfers). Onedata will be also responsible for definition of federation level authentication and authorization aspects of data access, based on OpenID Connect [16]. On the data access layer, Onedata will provide a WebDAV [17] storage interface, to enable integration of other

HTTP transfer based components such as FTS [18] or EOS [19] to make the data managed by these components to be accessible in a unified manner via the POSIX virtual filesystem provided by Onedata. Furthermore, Onezone, the entry point to the data management aspects of the platform, will allow for semi-automated creation of data discovery portals, based on metadata stored in the federated Oneprovider instances and on a centralized ElasticSearch engine indexing the metadata. This solution will allow the communities to create custom indexes on the data and metadata, provide customizable styles and icons for their users and defining custom authorization rights based on user classes (public access, access on login, group access, etc.).

5. Conclusions

In the present contribution the XDC objectives, starting from the technology gaps that currently prevent effective exploitation of distributed computing and storage resources by many scientific communities, have been discussed and presented together with the activities (and related contributions) carried on by INFN-CNAF for each WP of the project. Those objectives are the real driver of the project and derive directly from use cases, and the related needs, presented by the scientific communities involved in the project itself, covering areas such as Physics, Astrophysics, Bioinformatics, and others. Starting from the above assumptions, the overall structure of the project have been presented by emphasizing its components, typically based upon or extend established open source solutions, and the relations among them.

For the second part of project, the activities carried on at INFN-CNAF will continue to ensure the fulfilment of the project objectives. In particular, the already available software solutions will be enriched by advanced functionalities (provided by JRAs) aimed at addressing the use case requirements provided by NA2. The implementation and related testing of those new solutions will be performed in the testbeds maintained by SA1. SA1 will also continue its activities aimed at further validate the software, its robustness and scalability and will follow the preparation of the second project release, codenamed Quasar, foreseen for the second half of 2019.

Moreover, XDC project can complement and integrate with other running projects and communities and with existing multi-national, multi-community infrastructures. As an example, XDC is collaborating with the Designing and Enabling E-Infrastructures for intensive Processing in Hybrid Data Clouds (DEEP-Hybrid-DataCloud) [20] project aimed at promoting the integration of specialized, and expensive, hardware under a Hybrid Cloud platform, and targeting the evolution of the corresponding Cloud services supporting these intensive computing techniques to production level.

6. References

- [1] Web site: www.extreme-datacloud.eu
- [2] Web site: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>
- [3] Web site: <https://www.egi.eu/>
- [4] Web site: wlcg.web.cern.ch/
- [5] Web site: <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/einfra-21-2017.html>
- [6] Web site: <https://www.indigo-datacloud.eu>
- [7] Web site: <http://www.extreme-datacloud.eu/kickoff/>
- [8] Web site: <http://www.extreme-datacloud.eu/pulsar-out/>
- [9] Web site: <https://www.nginx.com/>
- [10] Web site: www.indigo-datacloud.eu/paas-orchestrator
- [11] Web site: <https://rucio.cern.ch/>
- [12] TOSCA Simple Profile in YAML Version 1.0. Edited by Derek Palma, Matt Rutkowski, and Thomas Spatzier. 27 August 2015. OASIS Committee Specification Draft 04 / Public Review Draft 01
- [13] Web site: www.dcache.org
- [14] Web site: lcgdm.web.cern.ch/dynafed-dynamic-federation-project
- [15] Web site: onedata.org
- [16] Web site: <https://openid.net/connect/>

- [17] Web site: www.webdav.org/
- [18] Web site: information-technology.web.cern.ch/services/file-transfer
- [19] Web site: eos.web.cern.ch
- [20] Web site: <https://deep-hybrid-datacloud.eu/>