Smart Policy Driven Data Management and Data Federations

enabled by the H2020 eXtended Data Cloud project

Presenter: Patrick Fuhrmann
With contributions by Daniele, Marica, Oliver, Paul and Giacinto

eXtreme DataCloud is co-funded by the Horizon2020 Framework Program – Grant Agreement 777367
Copyright © Members of the XDC Collaboration, 2017-2020
XDC Objectives

The eXtreme DataCloud is a software development and integration project

Develops **scalable** technologies for federating storage resources and managing data in highly distributed computing environments

- Focus efficient, policy driven and Quality of Service based DM

The targeted platforms are the current and next generation e-Infrastructures deployed in Europe

- European Open Science Cloud (EOSC)
- The e-infrastructures used by the represented communities
The Einfra-21-2017 Call

(a) Support to Public Procurement of innovative HPC systems, PPI
(b) Research and Innovation Actions for e-Infrastructure prototypes

1 - Universal discoverability of data objects and provenance
2 – Computing e-infrastructure with extreme large datasets

Service prototypes should follow common interfaces to access and analyse underlying data collected/stored in different platforms, formats, locations and e-infrastructures [...] tested against requirements of very large or highly heterogeneous research data sets.

Funds development of service prototypes at TRL6+

Bring to TRL8 and include in a unified service catalogue in 2018+

Budget per proposal: 2.5-3M€
XDC Foundations

XDC take the move from

- the INDIGO Data management activity
- the experience of the project partners on data-management

Improve already existing, production quality, Federated Data Management services

- By adding missing functionalities requested by research communities
- Must be coherently harmonized in the European e-Infrastructures
Represented research communities

[Images of research communities: CTA, LSST, WLCG, ECRIN, European XFEL, LifeWatch ERIC]

Stolen from Daniele Cesini

The long tail of science

Few large projects

Many smaller projects
## XDC Consortium

<table>
<thead>
<tr>
<th>ID</th>
<th>Partner</th>
<th>Country</th>
<th>Represented Community</th>
<th>Tools and system</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INFN (Lead)</td>
<td>IT</td>
<td>HEP/WLCG</td>
<td>INDIGO-Orchestrator, INDIGO-CDMI(*)</td>
</tr>
<tr>
<td>2</td>
<td>DESY</td>
<td>DE</td>
<td>Research with Photons (XFEL)</td>
<td>dCache</td>
</tr>
<tr>
<td>3</td>
<td>CERN</td>
<td>CH</td>
<td>HEP/WLCG</td>
<td>EOS, DYNAFED, FTS</td>
</tr>
<tr>
<td>4</td>
<td>AGH</td>
<td>PL</td>
<td></td>
<td>ONEDATA</td>
</tr>
<tr>
<td>5</td>
<td>ECRIN</td>
<td>[ERIC]</td>
<td>Medical data</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>UC</td>
<td>ES</td>
<td>Lifewatch</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CNRS</td>
<td>FR</td>
<td>Astro [CTA and LSST]</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>EGI.eu</td>
<td>NL</td>
<td>EGI communities</td>
<td></td>
</tr>
</tbody>
</table>

- 8 partners, 7 countries
- 7 research communities represented + EGI
- XDC Total Budget: 3.07Meuros
- XDC started on Nov 1\(^{st}\) 2017 – will run for 27 months until Jan 31\(^{st}\) 2020

Stolen from Daniele Cesini
Project Status

- Started on Nov 1\textsuperscript{st} 2017
- Detailed requirements collection from user communities completed
- Definition of the detailed architecture ready in May
- Creation of the Pilot Testbed started
  - Currently reserved for internal communities
  - Under discussion the possibility to open to external users
- Liaisons initiated with other DM development projects and EOSC-related initiatives
  - EOSC-Hub
  - EUDAT
  - DEEP-HybridDataCloud
  - RUCIO development team
  - All EINFRA-21 projects
  - StoRM

Stolen from Daniele Cesini
XDC high level architecture

Stolen from Daniele Cesini
The Release Plan

Event with User Communities –  Jun 18-22 2018, Santander – joint with DEEP

XDC reference releases – 1  -  Oct-Nov 2018

XDC reference releases – 2  -  Oct-Nov 2019

Functionalities and scalability demonstrated  -  Jan 2020
What is this XDC WP 4 about?

- Implementing a **configurable data workflow orchestration**, in terms of data location and storage quality (QoS).
- Providing managed and unmanaged **data caching services at all levels**.
- Providing **event based interfaces to external systems**
  - Generating events to the XDC orchestration services when data is entering the XDC system.
  - Generating events to external compute clusters when data is ready to be processed.
- Federating heterogeneous **data sources**, building a virtual horizontal infrastructure-specific data space.
Architecture Discussions are still ongoing
Some words on the toolbox
WP4 Production Level Components

Storage
- EOS
- xRootD
- CDMI

Federation

Orchestration
- Rucio
- INDIGO Orchestrator

QoS
Storage, Cache ...

Don’t need further introduction
Federation, Orchestration

Dynafed

- Federates endpoints to a virtual global namespace.
- Supports http, WebDAV, S3
- Can make use of Name Mapping Databases

FTS

- Orchestrates reliable wide area data transfers
- Supports multiple protocols, GridFTP, http, xrootd
- Can select most appropriate source
- Has knowledge on network topology and health
Rucio in a shoebox

The Daemons

- **Conveyor**: Transfer Daemon – in charge of file transfers
- **Reaper**: File Deletion Service
- **Undertaker**: Data Expiration manager
- **Transmogrifier**: Data Placement Policies / Subscriptions
- **Judge**: Replication Rule Engine
- **Hermes**: Messaging
- **Auditor**: Consistency Manager
- **Data Rebalancing**

Stolen from the Rucio Tutorial
The INDIGO PaaS Orchestrator

- The INDIGO PaaS Orchestrator 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 and deploys them on the best cloud site available. In order to do that:
  - it gathers SLAs, monitoring info and other data from other platform services,
  - it asks to the cloud provider ranker for a list of the best cloud sites.
- The exposed REST APIs are consumed by the Future Gateway portal.

Stolen from the INDIGO PaaS Orchestrator GitHub

2018-05-28 Patrick Fuhrmann eXtreme DataCloud @ dCache Workshop Hamburg
INDIGO-DataCloud CDMI QoS

Storage Broker Webservice

As example for a PaaS service

CDMI

DESY

PSNC

KIT

CNAF

BARI

StoRM

curl://

2018-05-28 Patrick Fuhrmann eXtreme DataCloud @ dCache Workshop Hamburg
### Quality of Service in storage (Broker Page)

#### Available Qualities of Service

<table>
<thead>
<tr>
<th>Name</th>
<th>Access Latency [ms]</th>
<th>Number of Copies</th>
<th>Storage Lifetime</th>
<th>Location</th>
<th>Available Transitions</th>
<th>Storage type</th>
</tr>
</thead>
<tbody>
<tr>
<td>disk</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td>Processing</td>
<td>tape, disk+tape</td>
</tr>
<tr>
<td>disk+tape</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td>Processing</td>
<td>tape</td>
</tr>
<tr>
<td>DiskAndTape</td>
<td>50</td>
<td>2</td>
<td>20 years</td>
<td>DE</td>
<td>Processing</td>
<td>TapeOnly</td>
</tr>
<tr>
<td>DiskAndTape</td>
<td>50</td>
<td>2</td>
<td></td>
<td>IT</td>
<td>Processing</td>
<td></td>
</tr>
<tr>
<td>DiskOnly</td>
<td>50</td>
<td>3</td>
<td>20 years</td>
<td>DE</td>
<td>Processing</td>
<td>profile2</td>
</tr>
<tr>
<td>DiskOnly</td>
<td>50</td>
<td>1</td>
<td></td>
<td>IT</td>
<td>Processing</td>
<td>profile1</td>
</tr>
<tr>
<td>profile1</td>
<td>10</td>
<td>3</td>
<td>20 years</td>
<td>DE</td>
<td>Processing</td>
<td>StandardDisk, Tape</td>
</tr>
<tr>
<td>profile2</td>
<td>10000</td>
<td>2</td>
<td></td>
<td>DE</td>
<td>Processing</td>
<td>SSDDisk, Tape</td>
</tr>
<tr>
<td>SSDDisk</td>
<td>10</td>
<td>1</td>
<td></td>
<td>IT</td>
<td>Processing</td>
<td>StandardDisk, Tape</td>
</tr>
<tr>
<td>StandardDisk</td>
<td>1000</td>
<td>3</td>
<td>20 years</td>
<td>IT</td>
<td>Archival</td>
<td>SSDDisk, Tape</td>
</tr>
</tbody>
</table>

- Access Latency [ms]
- Number of Copies
- Storage Lifetime
- Location
- Available Transitions
QoS in dCache
The orchestration
Expected Control Flow

TOSCA → INDIGO Orchestrator → Rucio

Control → ? → Location

Quality of Storage → Data Location

Rucio → Control

FTS3

EOS → dCache
The simple X-FEL Use Case

2018-05-28    Patrick Fuhrmann

The XDC-Orchestration Team

Rucio

“new file” Event

Online System

Burst Handling

Transfer file command.

Start ‘job’ commands

QoS command

XDC-Orchestration Team

INDIGO Orchestrator

Transfer file command.

QoS command

“new file” Event

Online System

Burst Handling

Transfer file command.

QoS command

“new file” Event

Online System

Burst Handling

Transfer file command.

QoS command
Now on the caching part
Prerequisite considerations

EOS, dCache

Headnode + Namespace + ResourceMGR + n * Data Pools
Access: control and data flow
Cached Access: control and data flow

In production for years e.g. NDGF, Michigan

Somewhere else

Center
Examples

- Denmark
- Sweden
- Norway
- Finland
- CERN
- Budapest (Wigner)
Current situation (naturally)

Proprietary
However, it would be nice to have ...
Better would be ....

CERN

Sync Namespace

Request

DESY
Advantages

- Same software stack as we currently have at the sites.
- After the data has been transferred to the local storage system, a name space entry has been created locally and the data is available at the local site independently of the remote network link and the availability of the central service.
Putting it together

xRootD Cache
Including http plug-in
Not discussed

✗ The interactions with Dynafed:

- at all levels, Dynafed can federate primary and cached storage spaces.
- It can be accessed directly by clients or it can be used to feed caches with a federation of primary storage spaces.

✗ Beside the caching mechanisms discussed above, WP4 will provide a squid like cache based on the xrootd framework including its http plug-in. This will provide a zero-maintenance storage space.

✗ The QoS in Storage definition within RDA and its implementation.

✗ The interactions of OneData (XDC WP5) with the components discussed above.
Summary

- XDC WP4 will provide the basic software stack to support a European wide, heterogeneous data management infrastructure, including data flow orchestration and smart caching. WP4 is particularly targeting experiments with extremely high volume and data throughput requirements.

- WP4 will be build upon a toolbox of well established and production grade software components.

- All involved contributors have a long tradition in providing professional big data software and to run high performance, data centric computer centers.

- The output of XDC WP4 is already considers in a set of proposals currently being submitted to the H2020 program.