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

Data Management for extreme scale computing

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€**

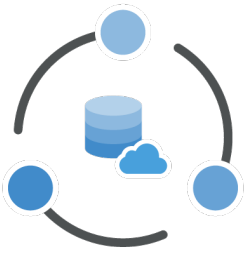
XDC Foundations

eXtreme DataCloud

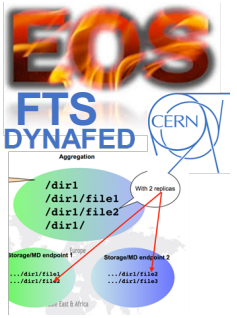
Stolen from Daniele Cesini

- **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

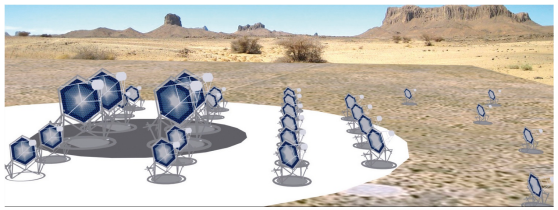
INDIGO PaaS Orchestrator

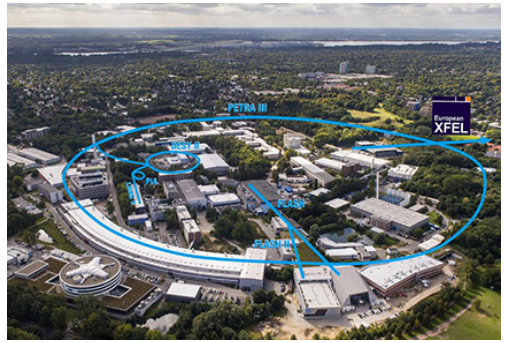


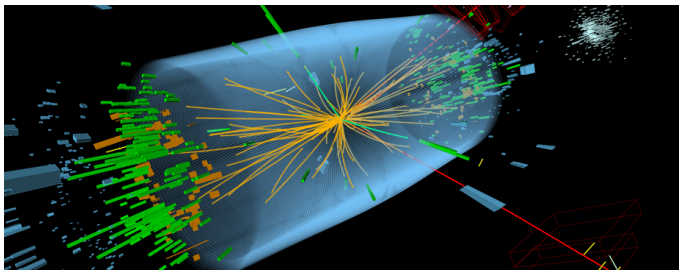
INDIGO CDMI Server

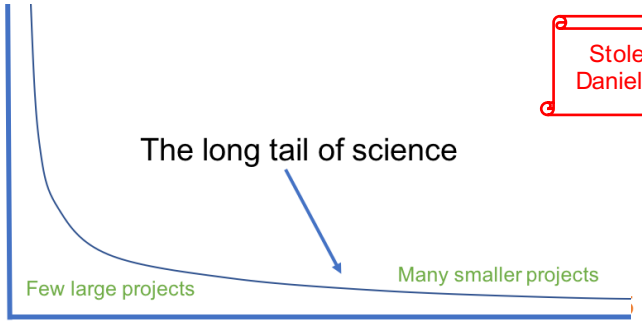


Represented research communities









XDC Consortium

eXtreme DataCloud

Stolen from Daniele Cesini

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

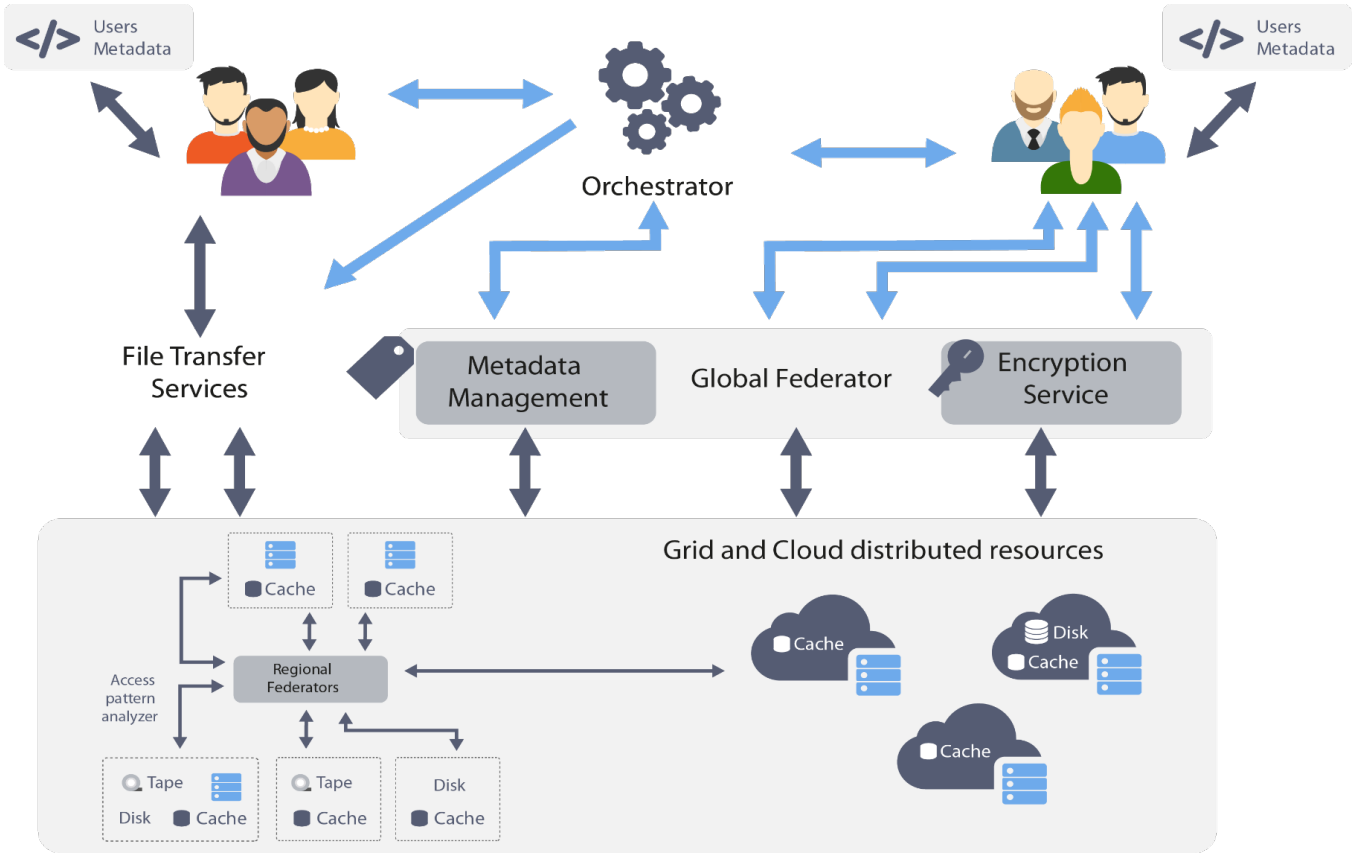
- 8 partners, 7 countries
- 7 research communities represented + EGI
- XDC Total Budget: 3.07Meuros
- XDC started on Nov 1st 2017 – will run for 27 months until Jan 31st 2020

Project Status

- Started on Nov 1st 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

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

Stolen from Daniele Cesini

What is this XDC WP 4 about ?

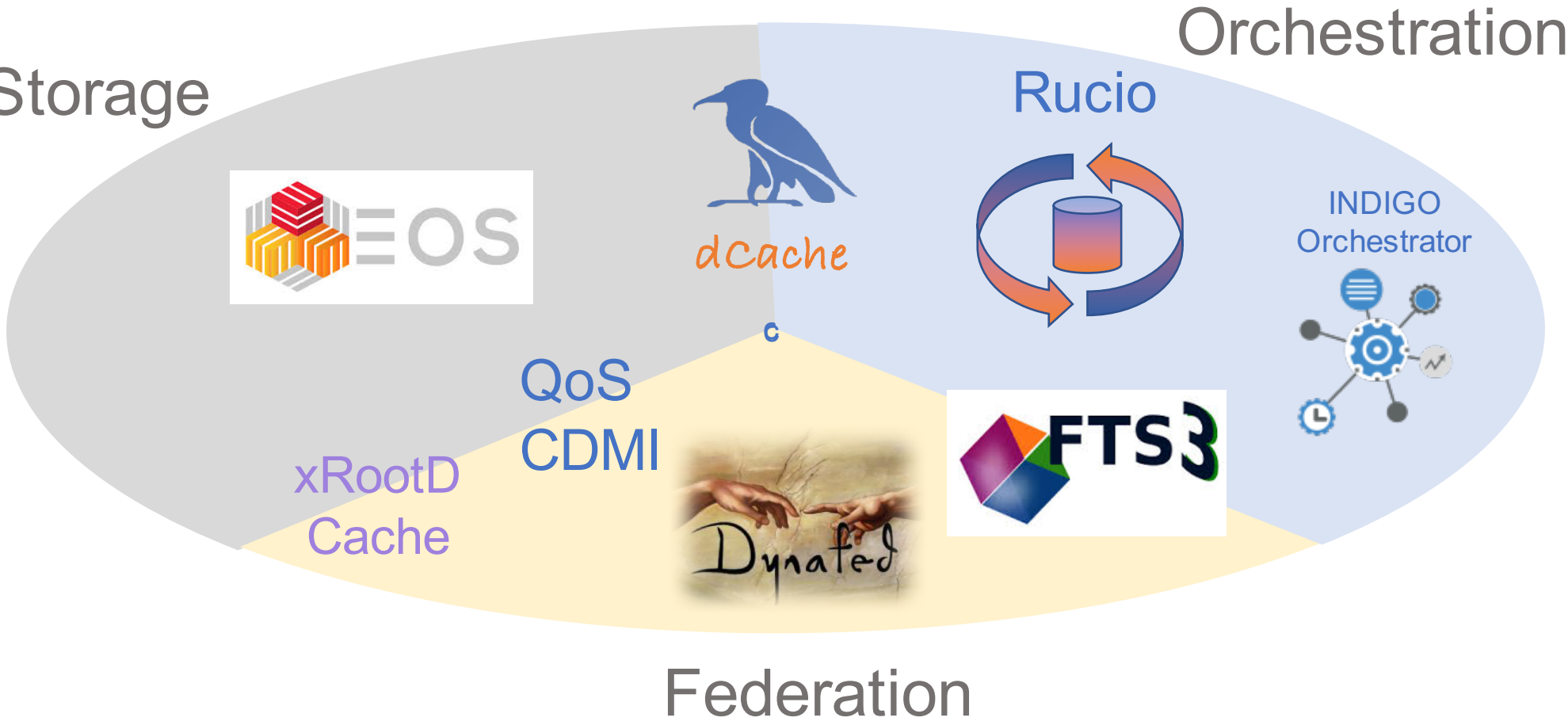
- Implementing a configurable data workflow orchestration, in terms of data location and storage quality (QoS).
- Providing managed and unmanaged data caching services at all levels.
- Providing event based interfaces to external systems
  - Generating events to the XDC orchestration services when data 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.

Disclaimer

Architecture Discussions are still ongoing

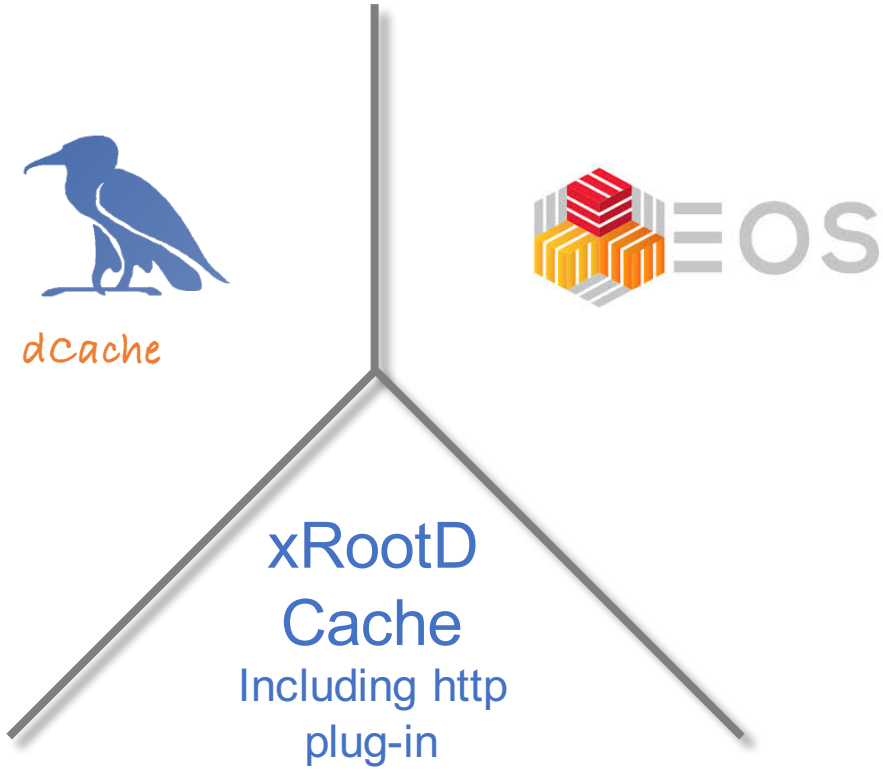
Some words on the toolbox

WP4 Production Level Components



Storage, Cache …

Don’t need further introduction



Federation, Orchestration

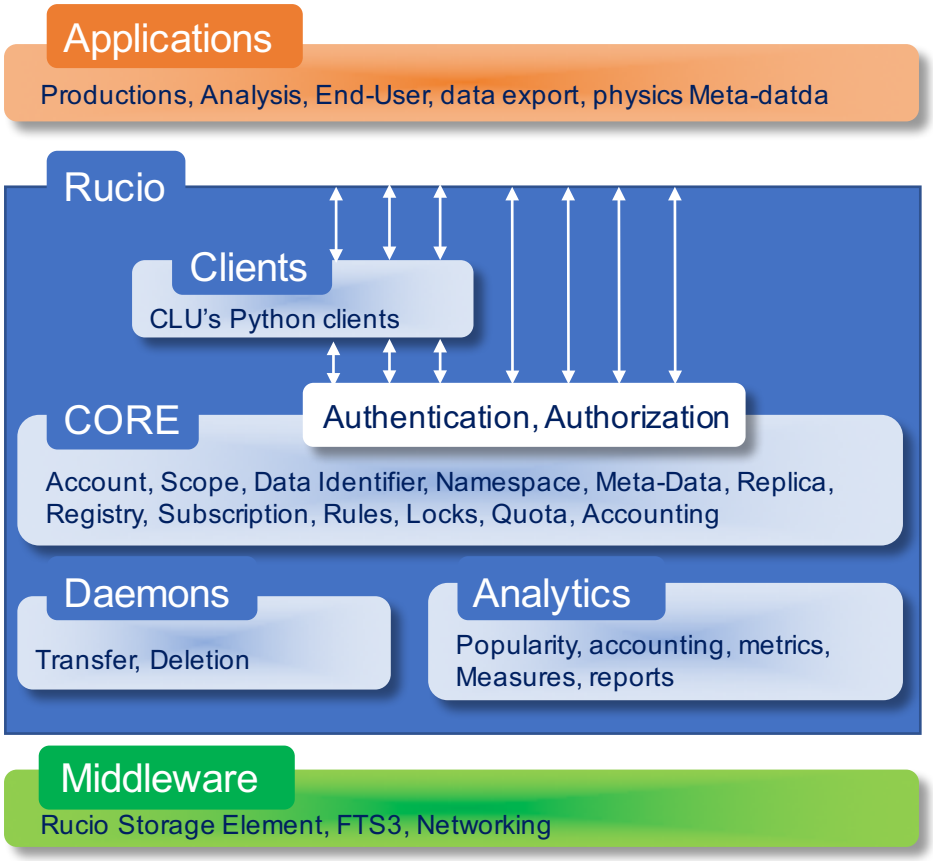
Dynafed

- 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



Stolen from the Rucio Tutorial

The Daemons

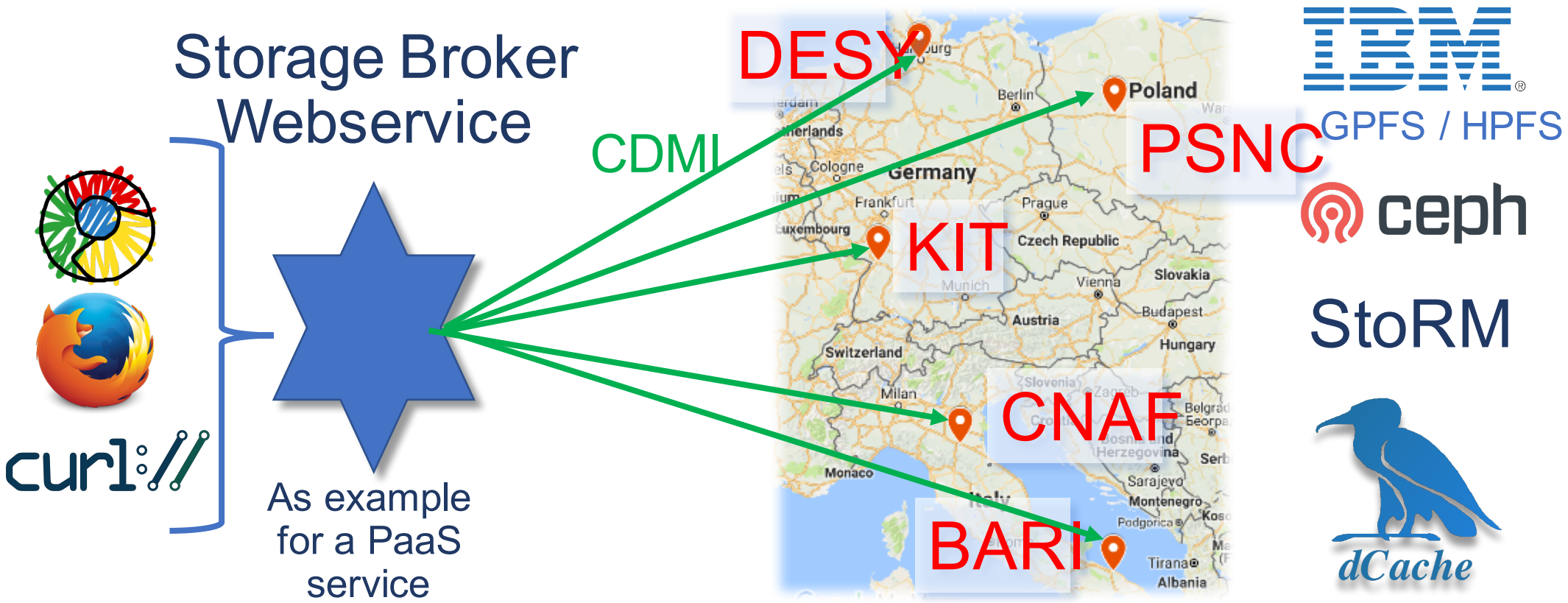
- Conveyor: Transfer Daemon – in charge of file transfers
- Reaper: File Deletion Service
- Undertaker: Data Expiration manager
- Transmogrifier: Data Placement Policies / Subscriptions
- Judge: Replication Rule Engine
- Hermes: Messaging
- Auditor: Consistency Manager
- Data Rebalancing

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

INDIGO-DataCloud CDMI QoS



INDIGO-DataCloud CDMI QoS

Quality of Service in storage (Broker Page)

Logout

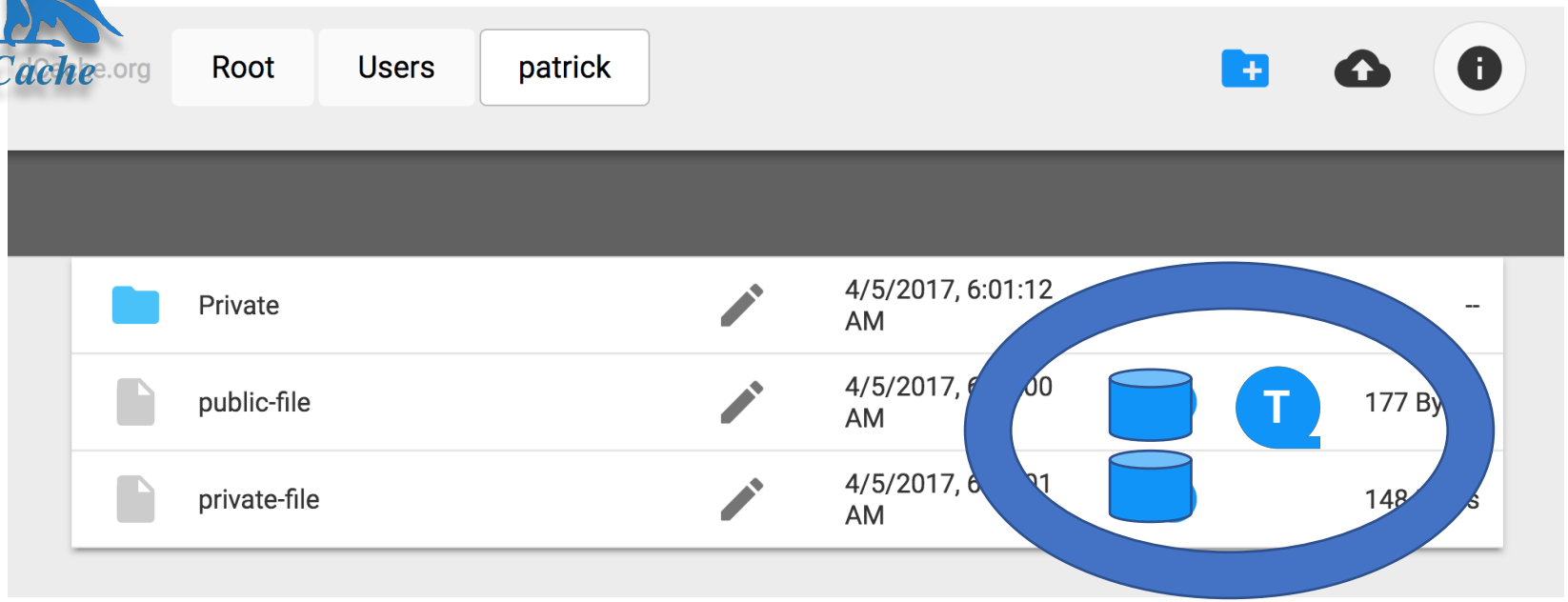
Available Qualities of

- Access Latency [ms]
- Number of Copies
- Storage Lifetime
- Location
- Available Transitions

DESY
KIT
PSNC
INFN

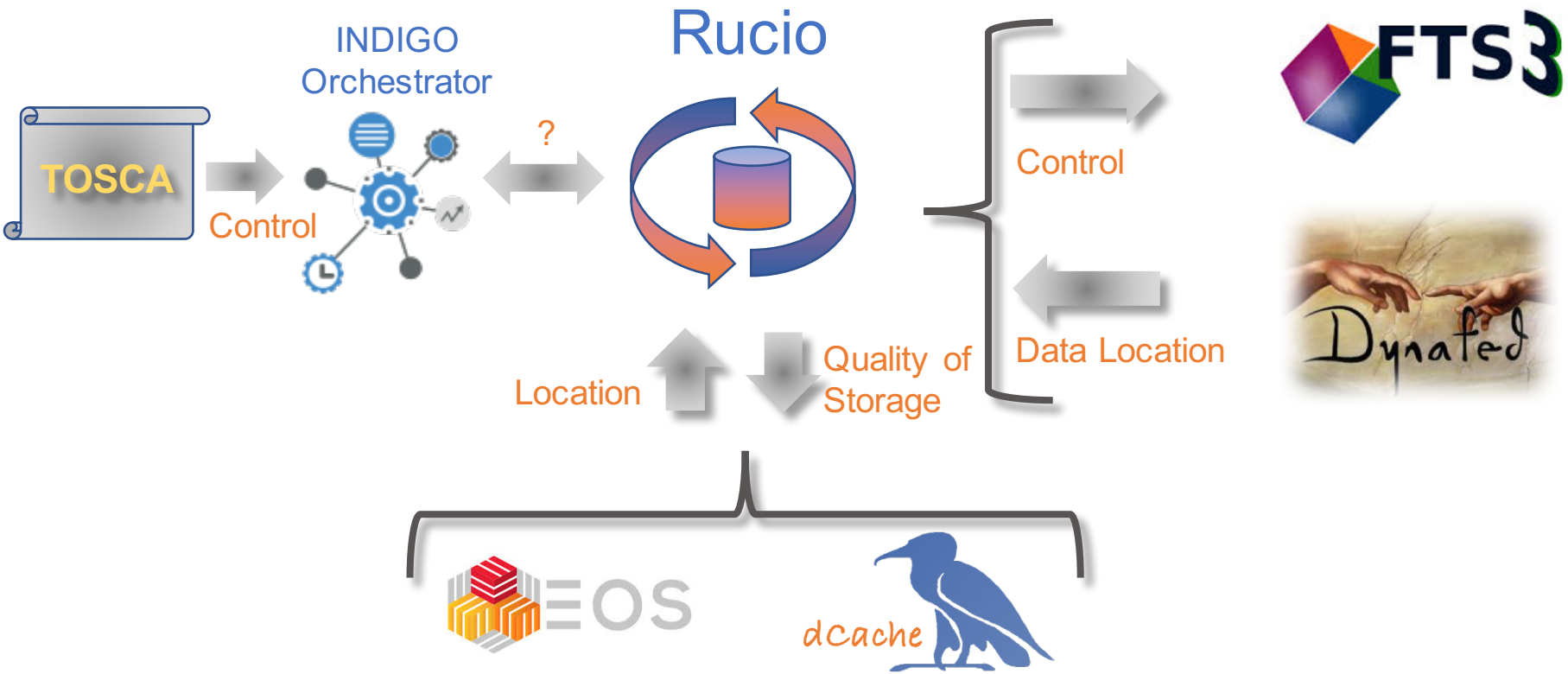
Name	Access Latency [ms]	[illegible]	[illegible]tion	Storage type	Available Transitions
disk	100	[illegible]		Processing	tape, disk+tape
disk+tape	100	2		Processing	tape
DiskAndTape	50	3		Processing	TapeOnly
DiskAndTape	50	2	IT	Processing	
DiskOnly	50	3	20 years DE	Processing	
DiskOnly	50	1	IT	Processing	
profile1	10	3	20 years DE	Processing	profile2
profile2	10000	2	DE	Archival	profile1
SSDDisk	10	1	IT	Processing	StandardDisk, Tape
StandardDisk	1000	3	20 years IT	Archival	SSDDisk, Tape

QoS in dCache

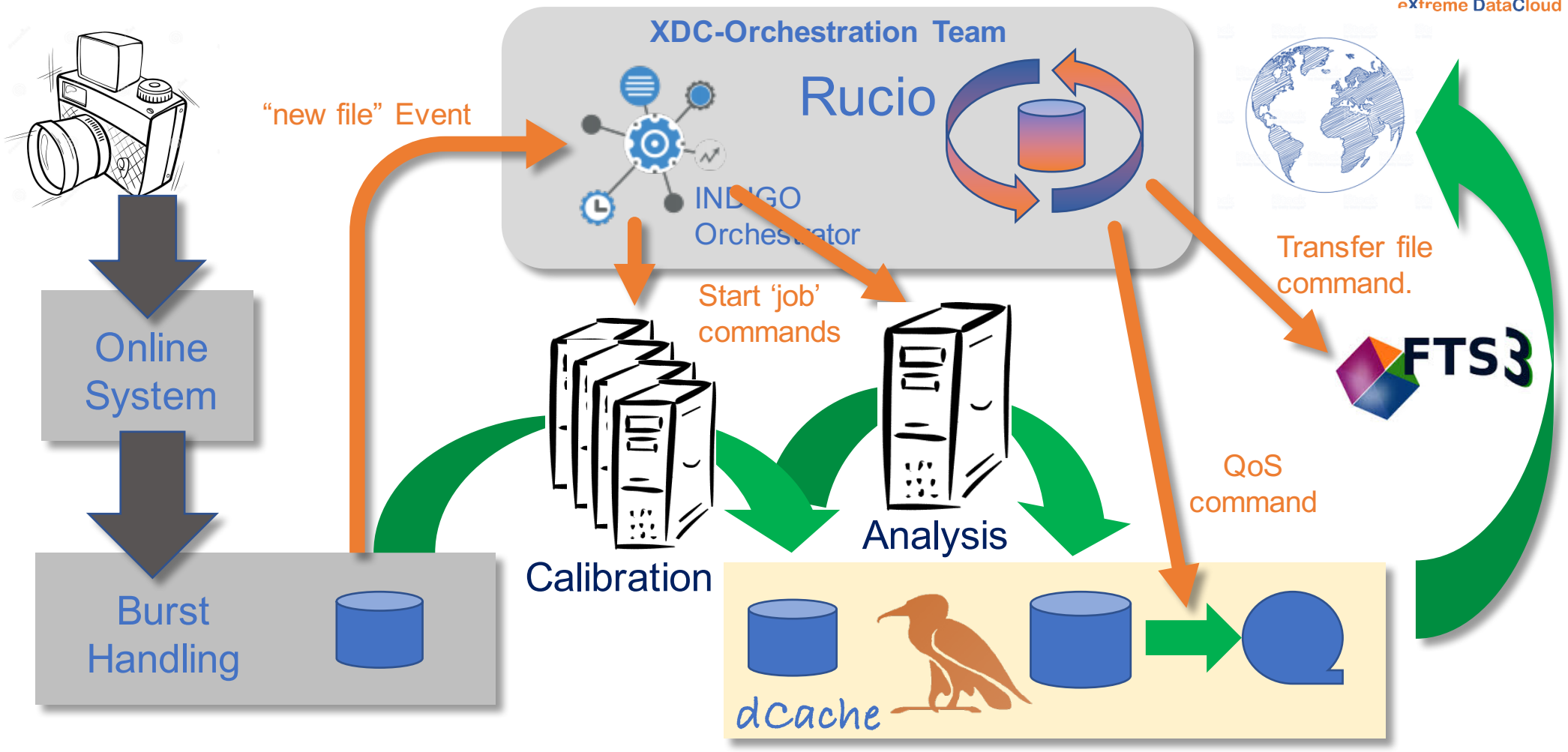


The orchestration

Expected Control Flow

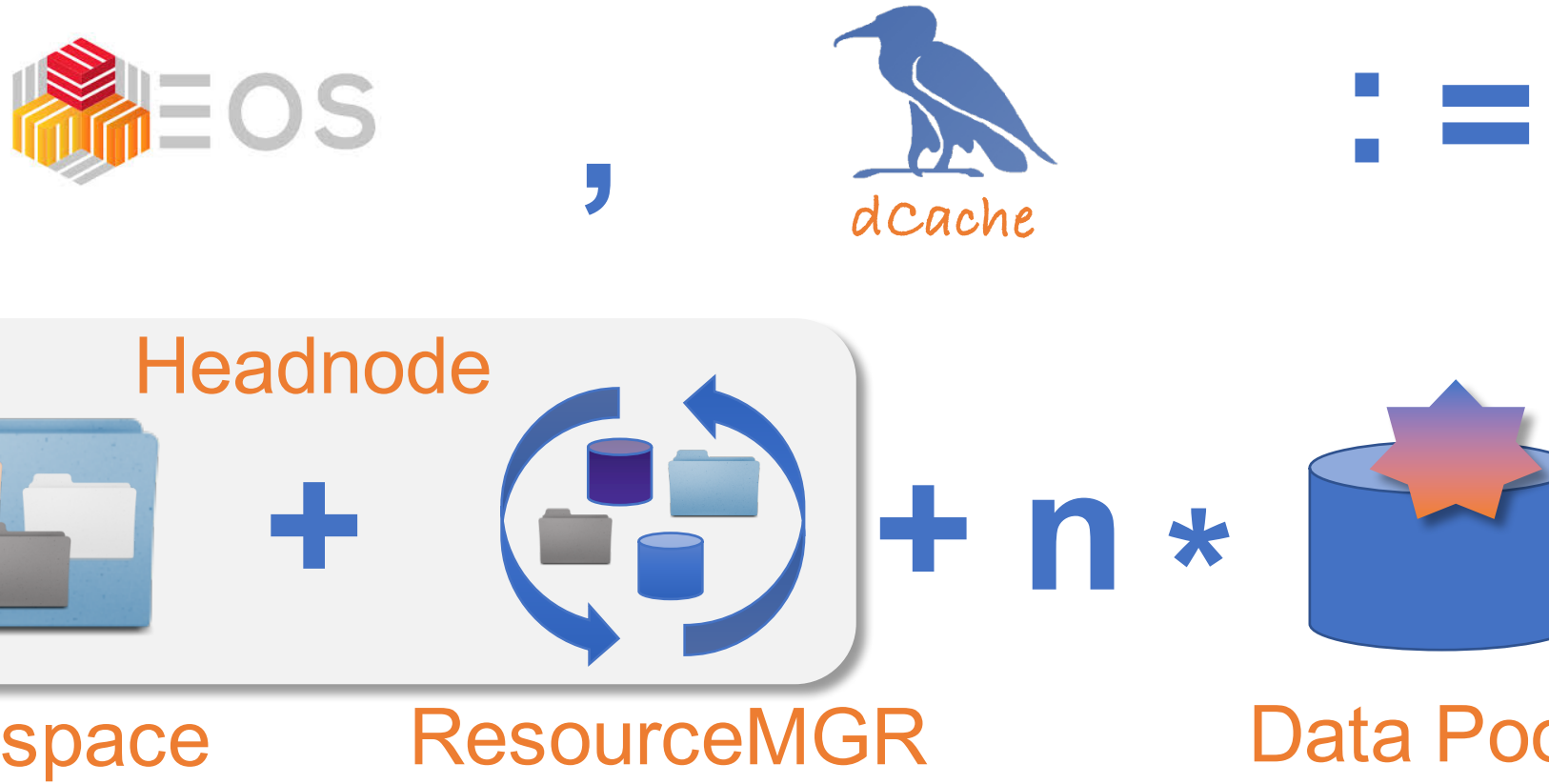


The simple X-FEL Use Case

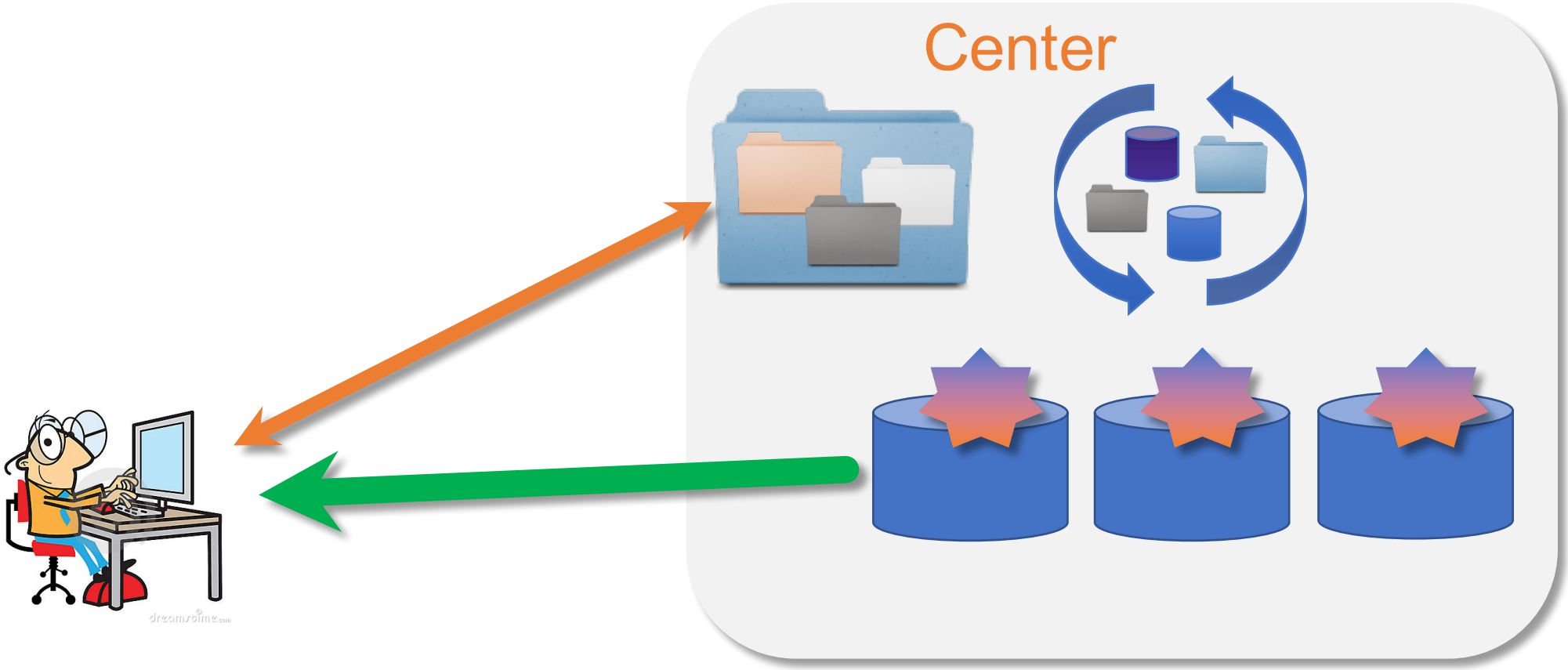


Now on the caching part

Prerequisite considerations



Access: control and data flow

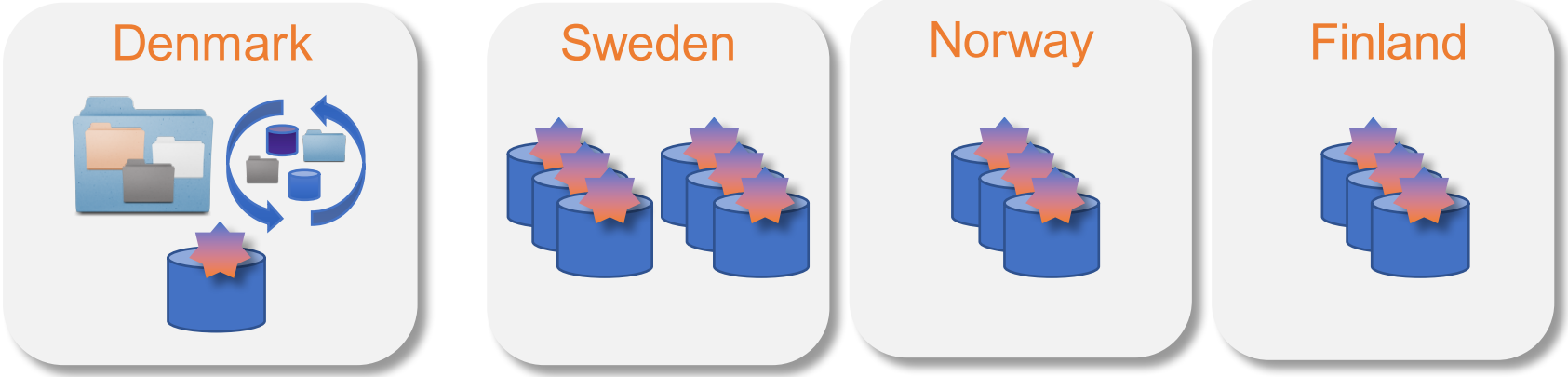


Cached Access: control and data flow

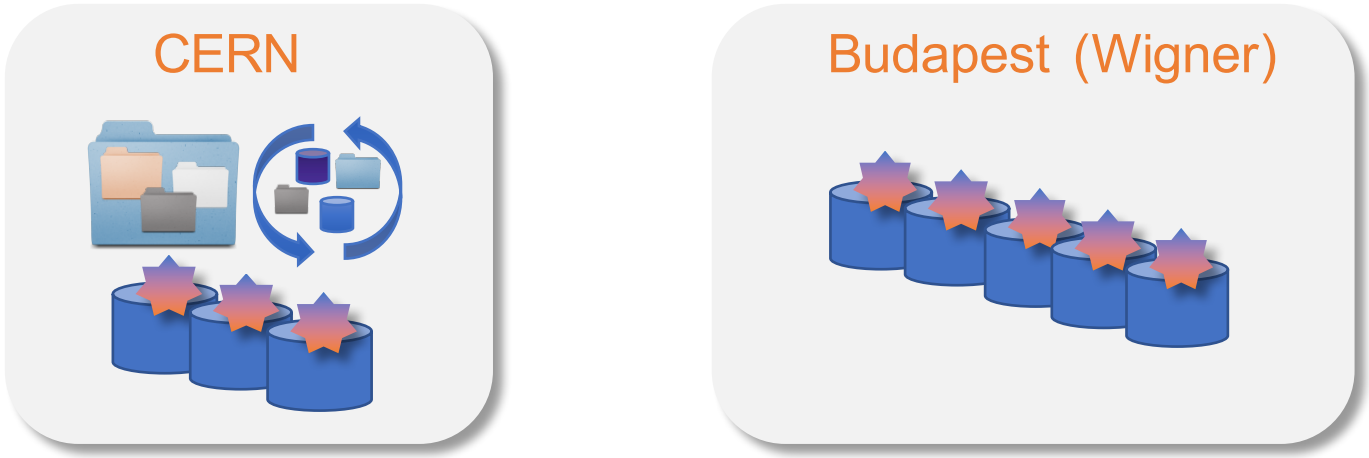
In production for years
e.g. NDGF, Michigan

Examples

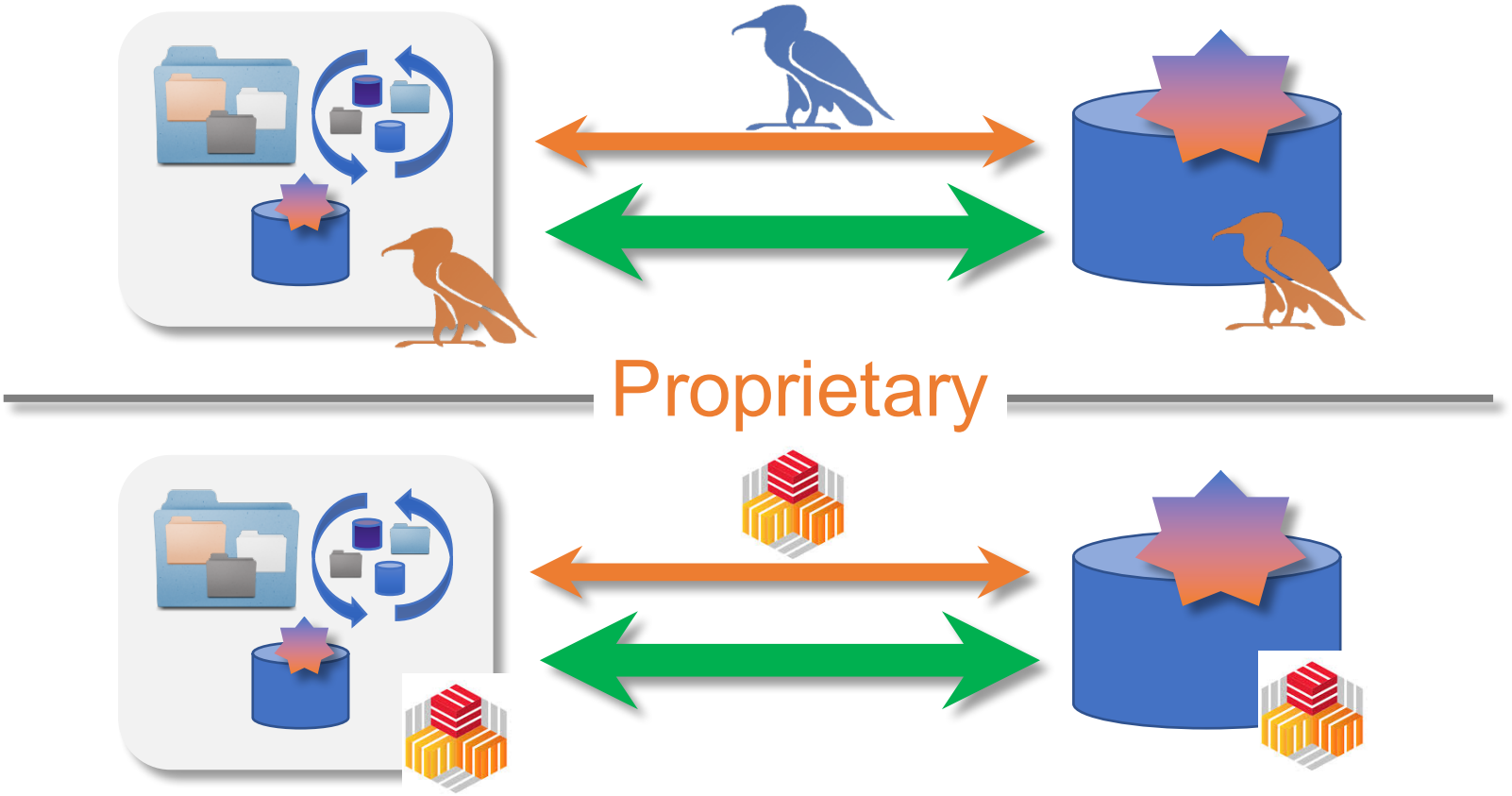

dCache


Denmark
Sweden
Norway
Finland

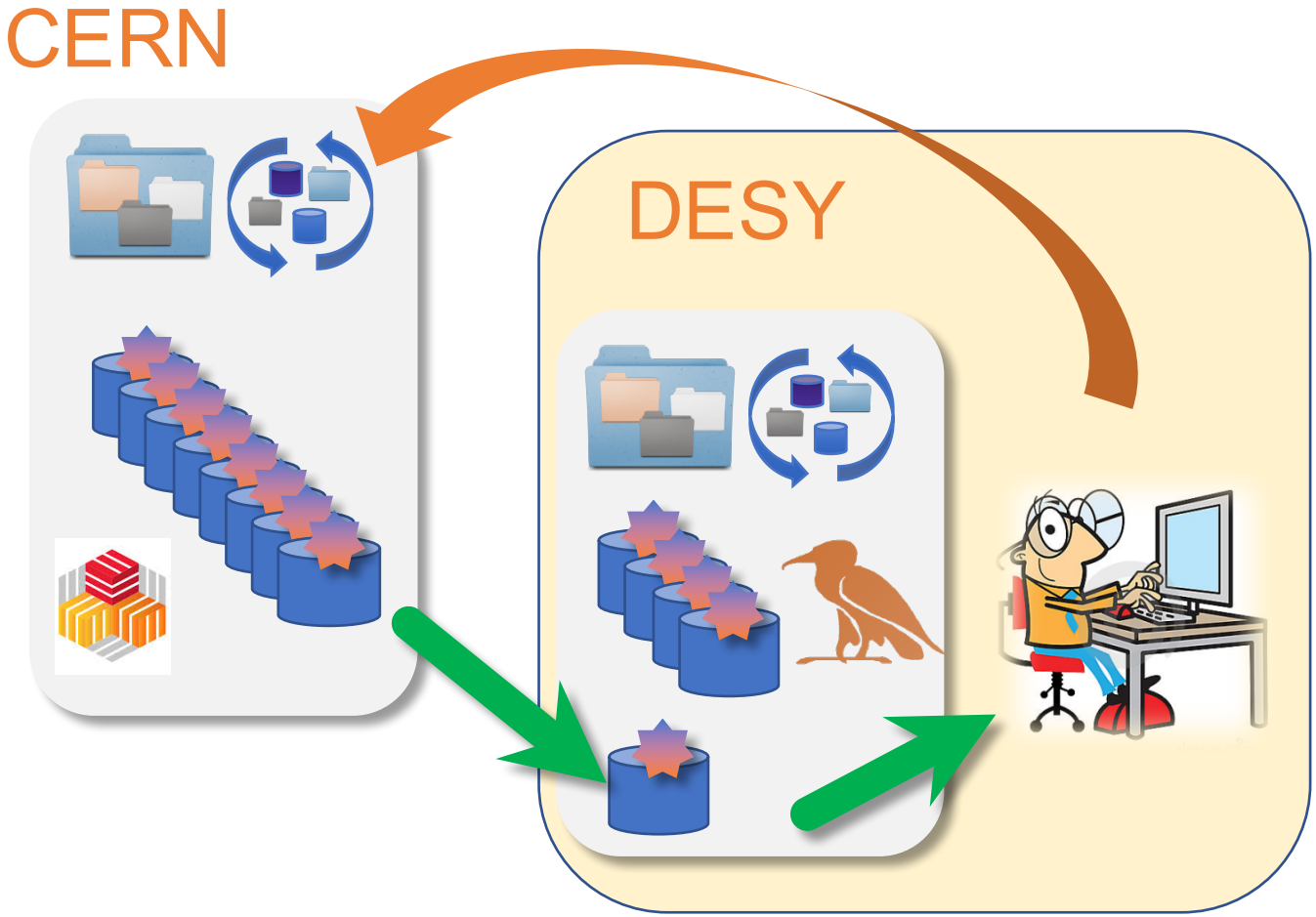

EOS


CERN
Budapest (Wigner)

Current situation (naturally)



However, it would be nice to have …



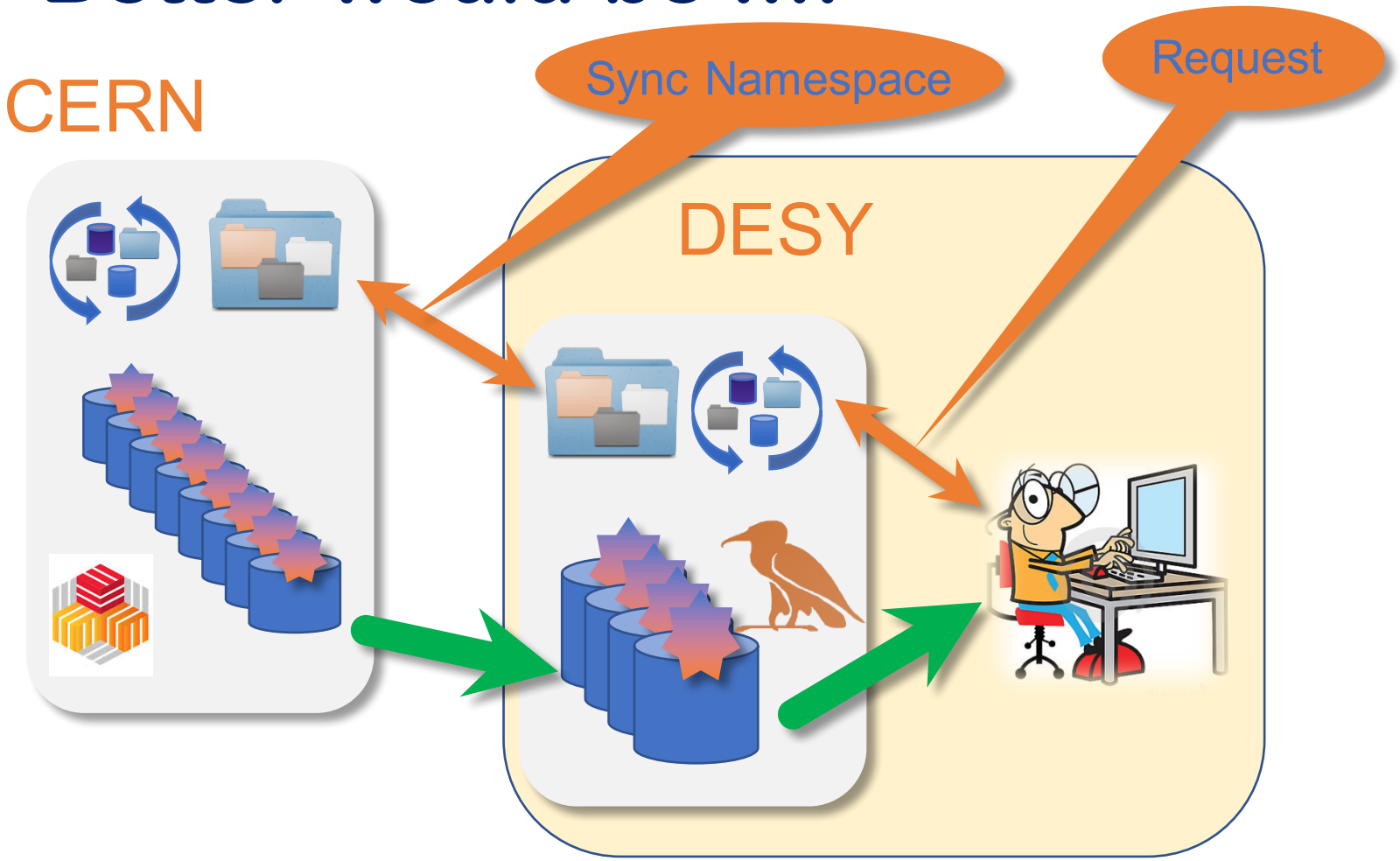
Advantages

- ✗ no additional software stack needed at sites.

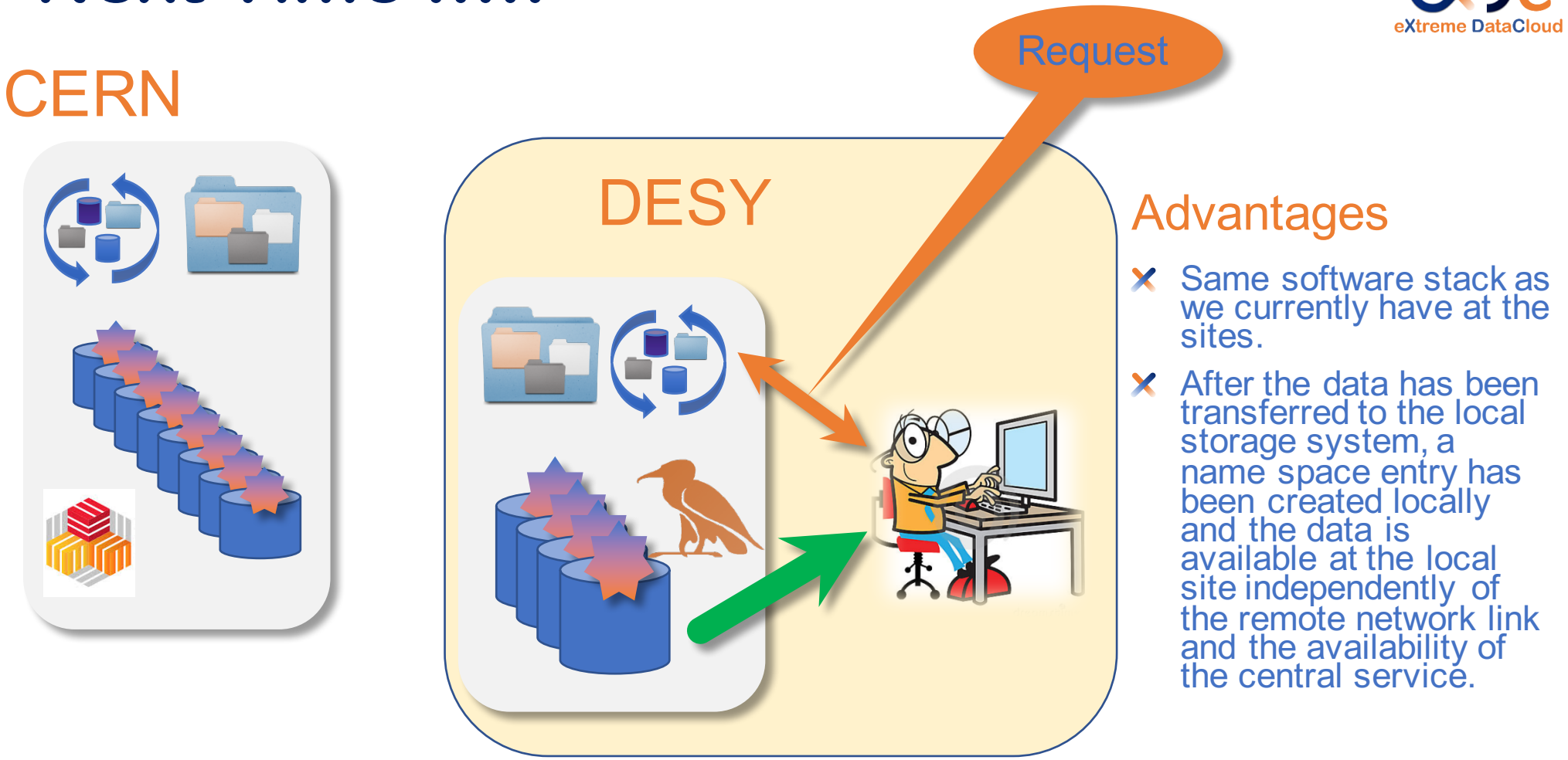
Still disadvantage

- ✗ Local data not accessible in case data link is down or central service not available

Better would be ....



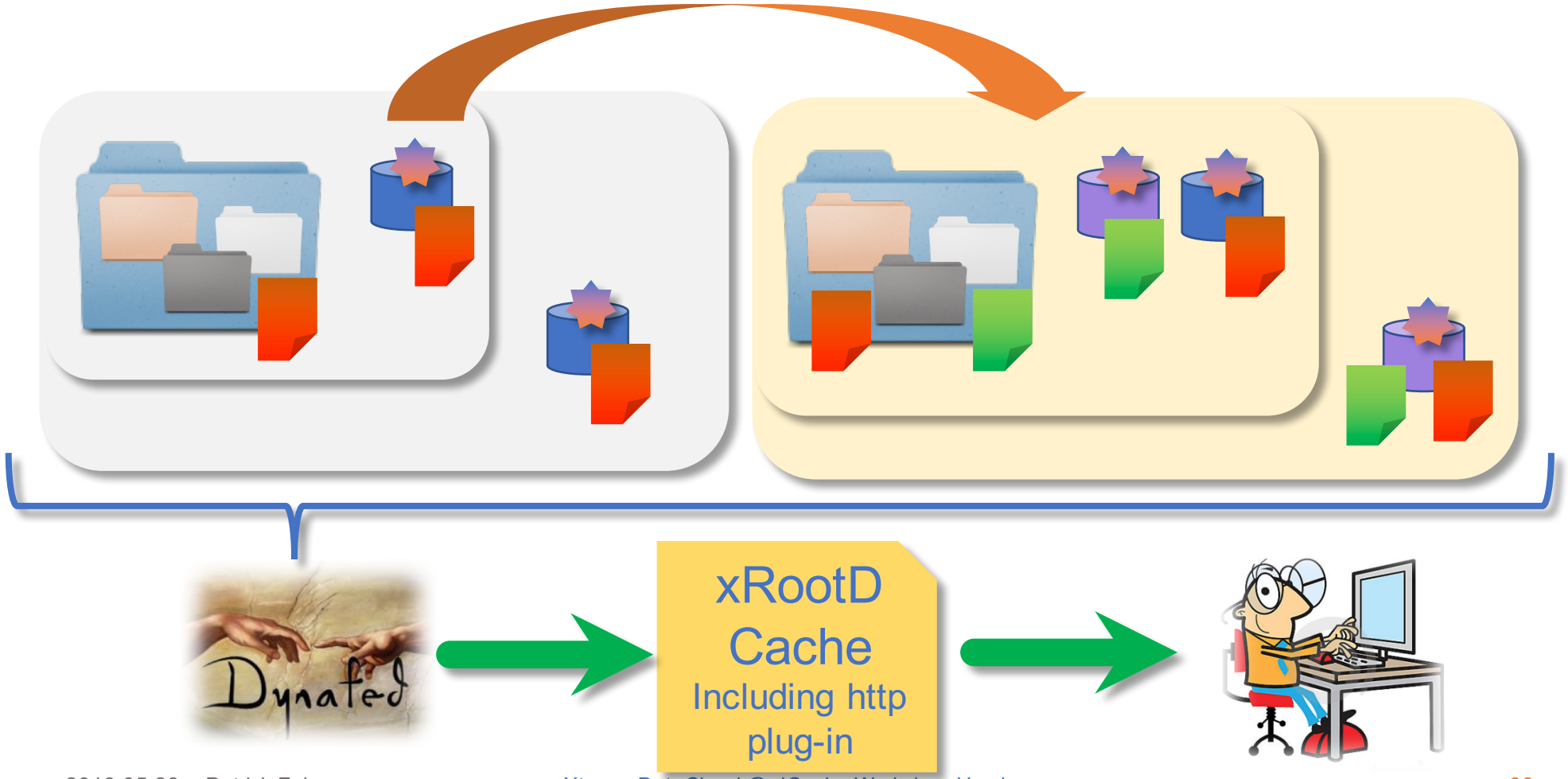
Next Time .....



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



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.