

Le project XDC



Data Management for extreme scale computing



Présentation par Baptiste Grenier - baptiste.grenier@eqi.eu

Contenu de Daniele Cesini, Patrick Fuhrmann, Marica Antonacci, Oliver Keeble, Paul Millar and Giacinto Don Vito



eXtreme DataCloud is co-funded by the Horizon2020 Framework Program – Grant Agreement 777367

Copyright © Members of the XDC Collaboration, 2017-2020

XDC: Les objectifs

- ✘ eXtreme DataCloud: un projet développement et intégration
 - EINFRA-21-2017 (Public Procurement of Innovative systems, Research and Innovation Actions, starting from TRL6 products)

- ✘ Développer des technologies **évolutives et extensibles** pour fédérer des ressources de stockage et gérer les données dans un environnement de calcul hautement distribué
 - Data Management avec un focus sur l'efficacité, basé sur les politiques et la Qualité de Service (QoS)

- ✘ Les plateformes ciblées sont les e-infrastructures courantes et next-gen déployées en Europe
 - European Open Science Cloud (EOSC)
 - Les e-infrastructures utilisées par les communautés représentées dans le projet

XDC: Les fondations

X XDC se base sur

- Les activités de gestion de données d'INDIGO-DataCloud
- L'expérience des partenaires du projet en "data-management"

X Améliorer des services préexistant de gestion fédérée de données et de niveau production

- Par l'ajout de **fonctionnalités manquantes** requises par les communautés de recherche
- Harmonisé dans les e-infrastructures de recherche européennes



INDIGO PaaS
Orchestrator



INDIGO CDMI
Server



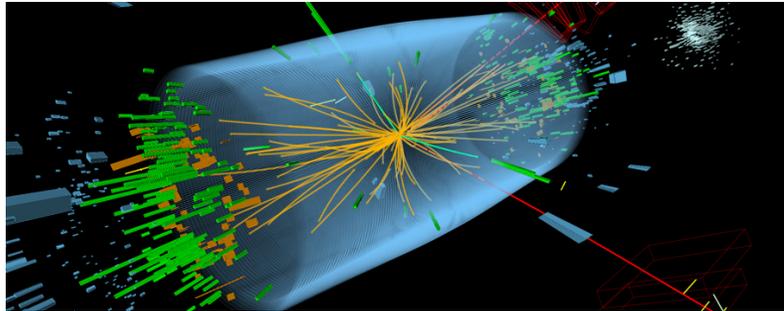
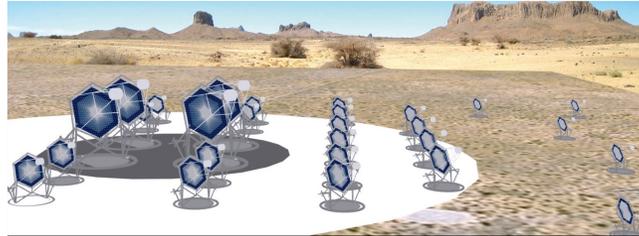
XDC: Le consortium

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	

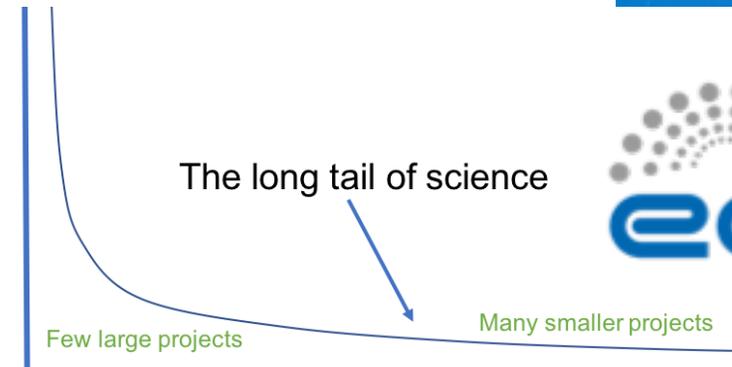


- ✘ 8 partenaires, 7 pays
- ✘ 7 communautés de recherche + EGI
- ✘ Budget Total: 3.07 M euros
- ✘ Démarre le 1er Novembre 2017 – pour 27 mois jusqu'au 31 Janvier 2020

Communautés de recherche représentées



ECRIN supports multinational clinical trials in Europe



XDC: Sujets Techniques

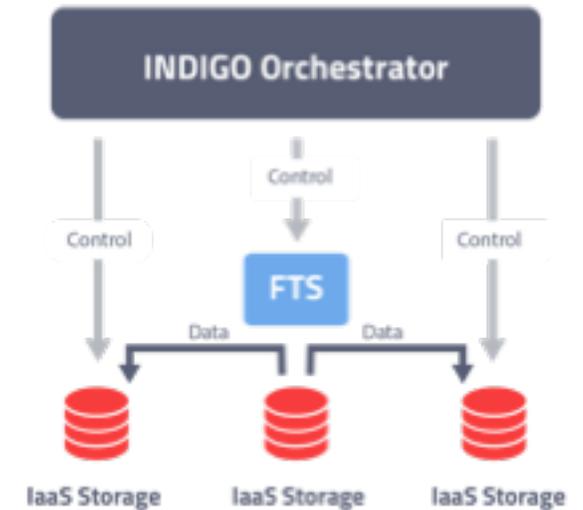
- ✘ Distribution intelligente et automatisée de Dataset
 - ☞→ Orchestration pour une gestion de données axée sur des politiques
 - ☞→ Politiques de distribution des données basée sur la Qualité de Service (i.e. disques dur vs bande vs SSD) **supportant des ressources distribuées géographiquement** (cross-sites)
 - ☞→ Gestion du cycle de vie des logiciels
- ✘ Traitement préalable de données durant l'ingestion
- ✘ Gestion des données basée sur des schémas d'accès
 - ☞→ Déplacer les données inutilisées sur un stockage de type "glacier", déplacer les données très demandées sur un stockage rapide
 - ☞→ Au niveau de l'infrastructure
- ✘ Mise en cache intelligente
 - ☞→ Accès aux données distantes de façon transparente sans besoin de copie a priori
- ✘ Gestion des Métadonnées
- ✘ Gestion des données sensibles
 - ☞→ Stockage sécurisé et chiffrement

Gestion des données axée sur des Politiques

✘ Distribution Intelligente et Automatisée de Dataset

☛ Un Workflow typique

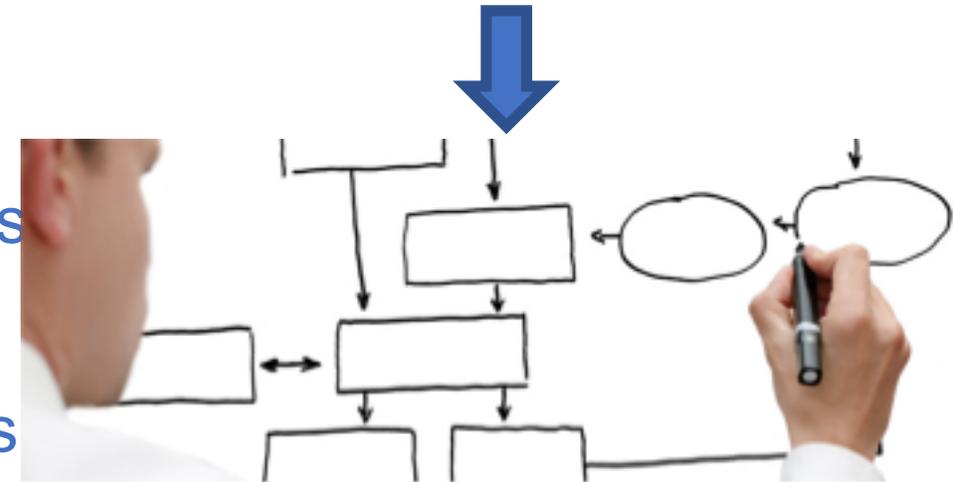
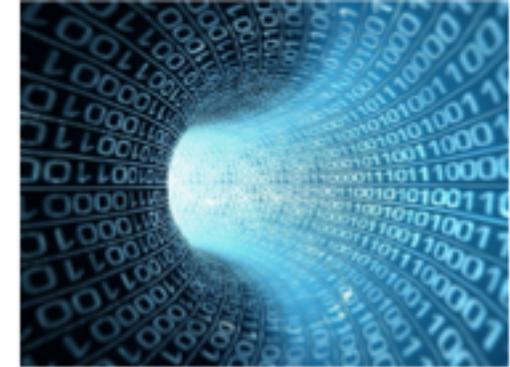
- ☛ **Stockage** des données sur des périphériques à **faible latence** pour des **accès rapides**
 - ☛ Pour la **sécurité** des données, **réplication** sur un second stockage et migration sur du stockage **long terme** type **bande** ou **S3**
 - ☛ Les utilisateurs éligibles ont la **permission** de **restaurer** les données archives, **si nécessaire**
 - ☛ **Après** une période d'**embargo**, le contrôle d'**accès** sera changé de "**privé**" à "**libre**"
- ### ☛ Gestion basée sur les schémas d'accès



Prétraitement des données

✘ Prétraitement Durant l'ingestion

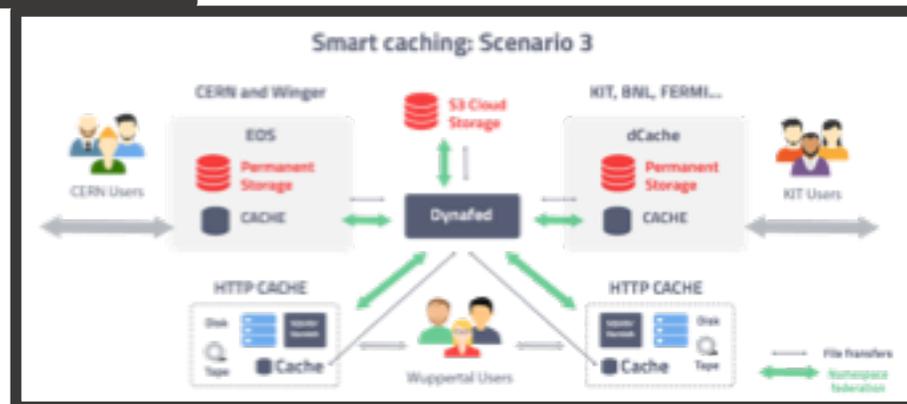
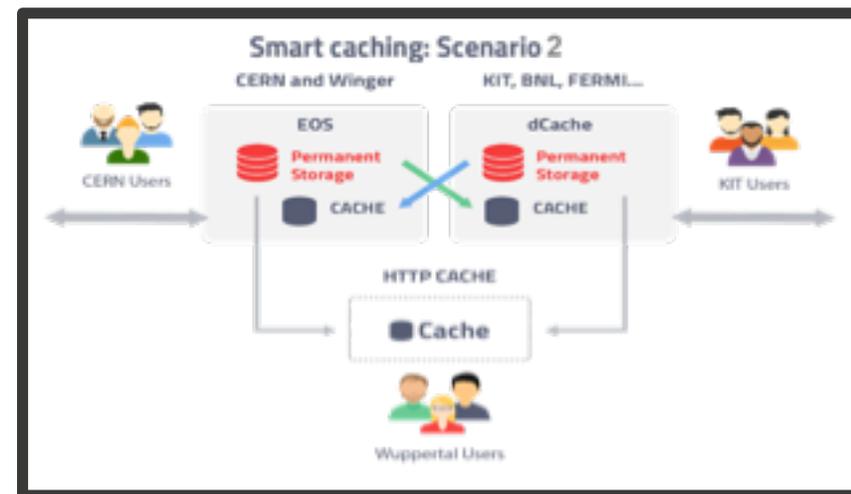
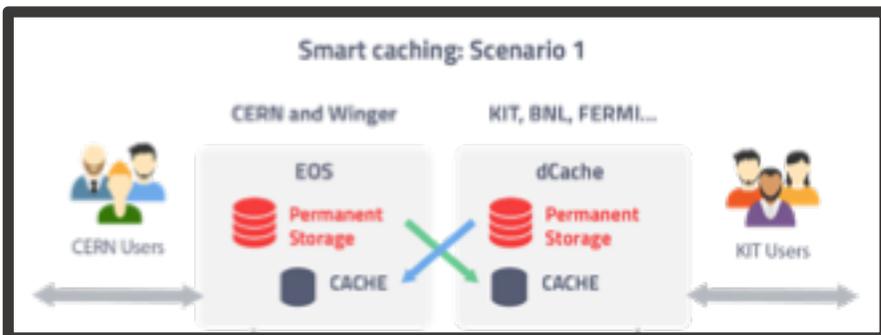
- Exécuter des applications et workflows définis par l'utilisateur lors de l'upload
 - i.e. pour filtrer, indexer, extraire des métadonnées, tester la cohérence
- Implémenter une solution pour la découverte de nouvelles données à des emplacements spécifiques
- Créer les fonctions permettant à l'INDIGO PaaS Orchestrator d'exécuter des applications spécifiques sur l'infrastructure
- Implémenter un moteur de workflow de haut niveau, pour exécuter les applications définies par les utilisateurs
- Implémenter le transfert des données transformées vers leur destination finale



Mise en cache intelligente

✗ Mise en cache intelligente

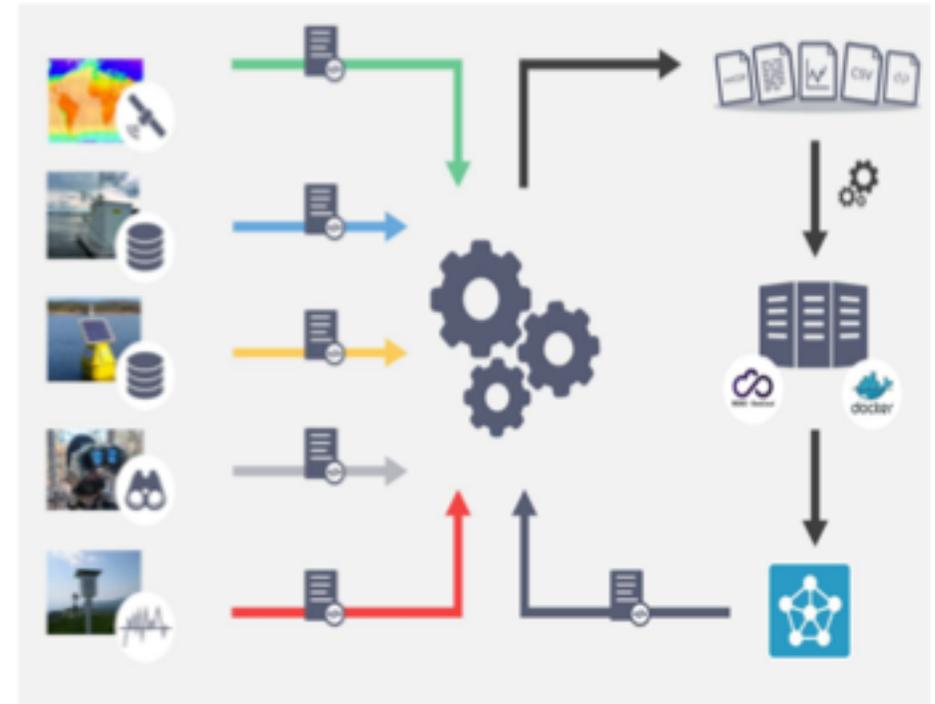
- ...→ Développer une infrastructure globale de cache avec les bases suivantes:
 - ...→ Intégrer dynamiquement des sites satellites dans les data centres existants
 - ...→ Création d'un cache indépendant façonné sur les solutions web
 - ...→ Fédération des services en une infrastructure de cache à grande échelle



Les cas d'utilisation XDC

Cas d'utilisation LifeWatch

- ✗ **Problématique:** Gestion du cycle de vie des données se rapportant à la **Qualité de l'eau**, impliquant des **sources de données hétérogènes**
 - Satellite, surveillance temps-réel, stations météo.
- ✗ **But:** Intégrer les sources de données et différents outils de modelage pour simuler des masses d'eau douce dans un environnement **FAIR**
 - Utilisation de standards tel que EML (Ecological Metadata Language)
- ✗ **Solution:**
 - Onedata
 - Gestion et découverte des métadonnées, émission d'identifiant digital, stockage
 - Orchestrateur INDIGO PaaS
 - Prétraitement automatique pour l'harmonisation des données et le déploiement de modèles



Cas d'utilisation CTA

✘ **Problématique:** Gestion Big Data complexe dans un environnement distribué. Assurance Qualité des données.

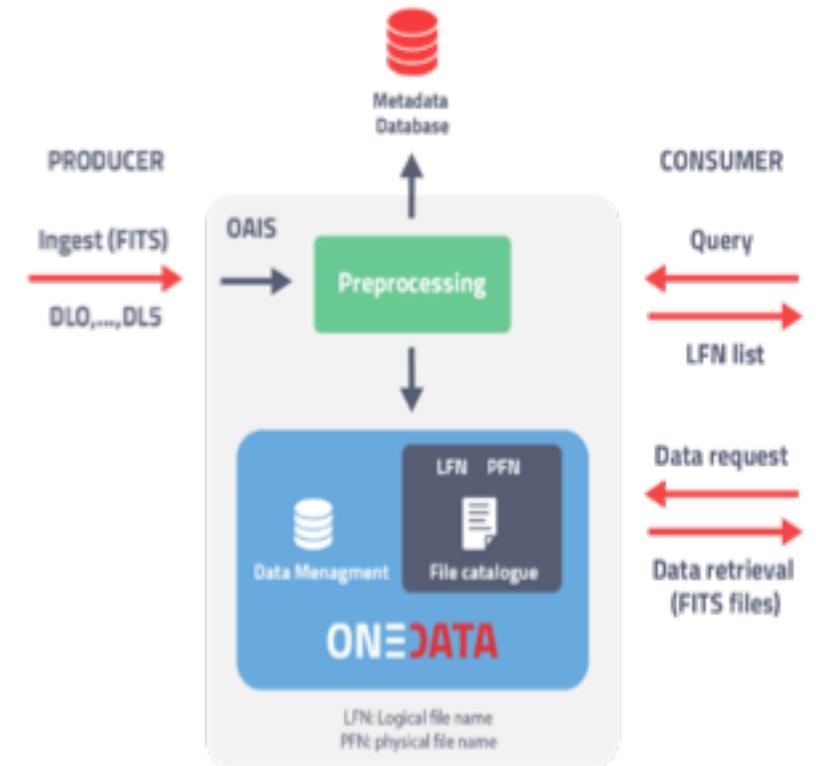
- L'archive distribuée CTA repose sur le standard ISO OAIS: "Open Archival Information System".
- Données enregistrées dans des fichiers (FITS format) contenant toutes les métadonnées.

✘ **But:** Extraire les métadonnées des fichiers ingérés pour automatiquement remplir la base des métadonnées.

- Les métadonnées seront utilisées pour interroger l'archive.
- Le système doit pouvoir **gérer les réplicas**, les bandes, les disques, etc. avec des données de bas ou haut niveau.

✘ **Solution:**

- Onedata
 - Gestion et découvertes des métadonnées
- INDIGO PaaS Orchestrator + QoS



Cas d'utilisation ECRIN

- ✘ **Problématique:** Fichiers et objets de données distribués à travers différents dépôts. Métadonnées hétérogènes. Données sensibles.
- ✘ **But:** Un environnement unique afin de pouvoir partager des données d'essais cliniques. Les sources sont réparties à travers:
 - Différents mécanismes d'accès
 - Plusieurs locations
 - Augmentation du nombre de dépôts généraux et spécialisés
 - Publications
 - L'institution originale des chercheurs
- ✘ **Solution:**
 - Onedata
 - Gestion et découvertes des métadonnées
 - Stockage sécurisé

Cas d'utilisation WLCG

- ✘ **Problématique:** Des besoins en stockage croissant
 - ☛→ Jusqu'à 900 PB en 2027
 - ☛→ Données prêtes a être exploitées dans un environnement hautement distribué
- ✘ **But:** Réduire les couts, réduire les besoins en disques, allocation intelligente des données.
- ✘ **Solution:**
 - ☛→ Systèmes de Cache intelligent
 - ☛→ Fédération des données
 - ☛→ Stockage multi-site – concept de “Data Lake”.
 - ☛→ Extension dynamique des sites a des locations distantes
 - ☛→ EOS, XrootD, FTS, Rucio...

Cas d'utilisation XFEL

✘ **Problématique:** Gestion du cycle de vie des données pour les installations XFEL, ACL et contrôle de la période d'embargo

⋯→ **En ligne:**

⋯→ Access exclusive au cluster accès au cluster durant le temps de faisceau uniquement depuis les salles d'expérimentation. Analyse préliminaire à la volée

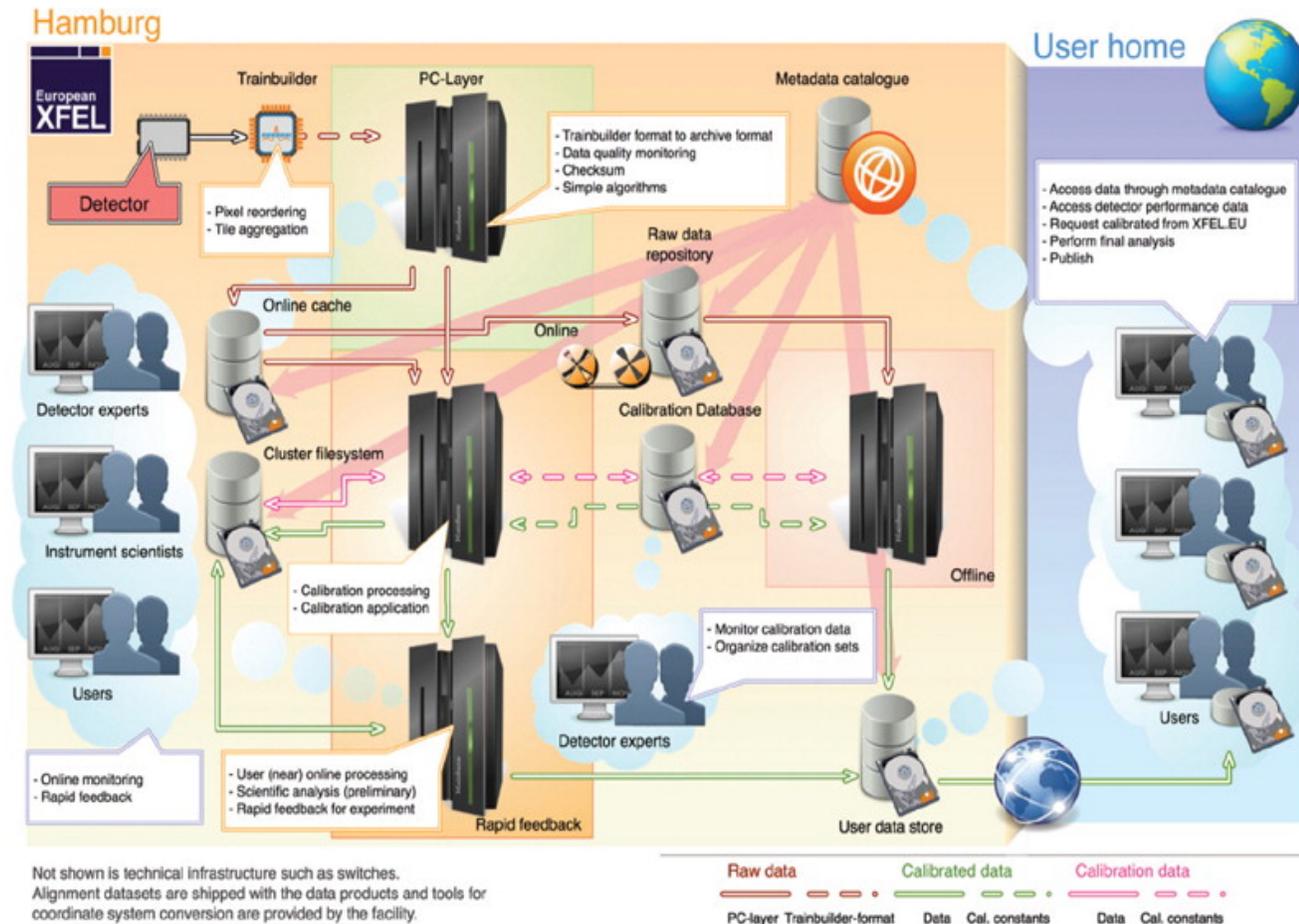
⋯→ **Hors ligne:**

⋯→ Fourniture de données calibrées pour analyse sur le cluster HPC et dans l'environnement cloud OpenStack.

✘ **Solution:**

⋯→ Plateforme de stockage distribué dCache

⋯→ INDIGO PaaS Orchestrator

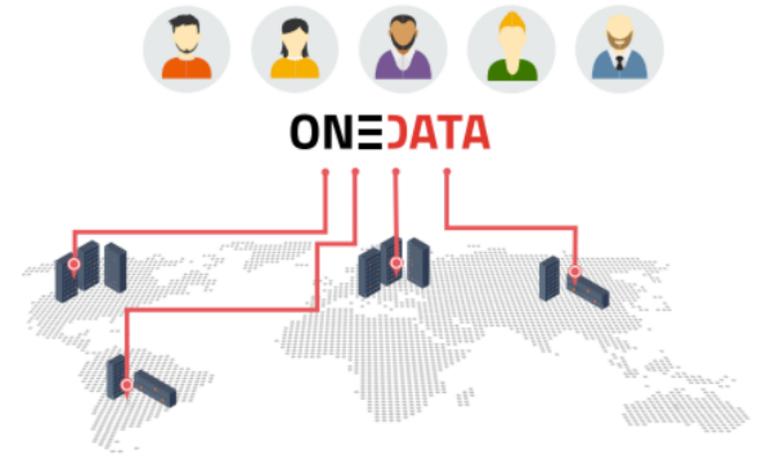


La toolbox XDC

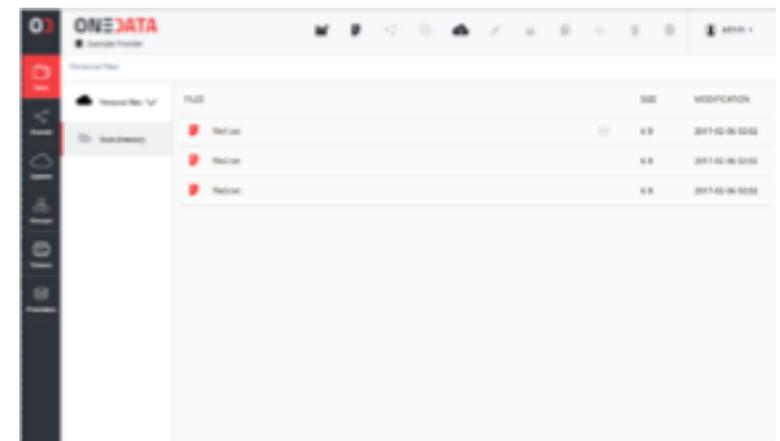
Onedata (WP5)

- ✘ Une fédération du stockage permettant aux utilisateurs de stocker, traiter et publier leurs données en utilisant un stockage global s'appuyant sur des fournisseurs de ressources dans le monde entier

- Les fournisseurs déploient les services *Oneprovider* à proximité du stockage physique
- Les utilisateurs utilisent les interface Web *Onezone*
 - APIs disponibles
 - Possibilité de monter les données localement sur les machines des utilisateurs
- Le stockage est organisé en **Zones**
 - Fédérations de fournisseurs
 - Création de communautés interconnectées ou isolées

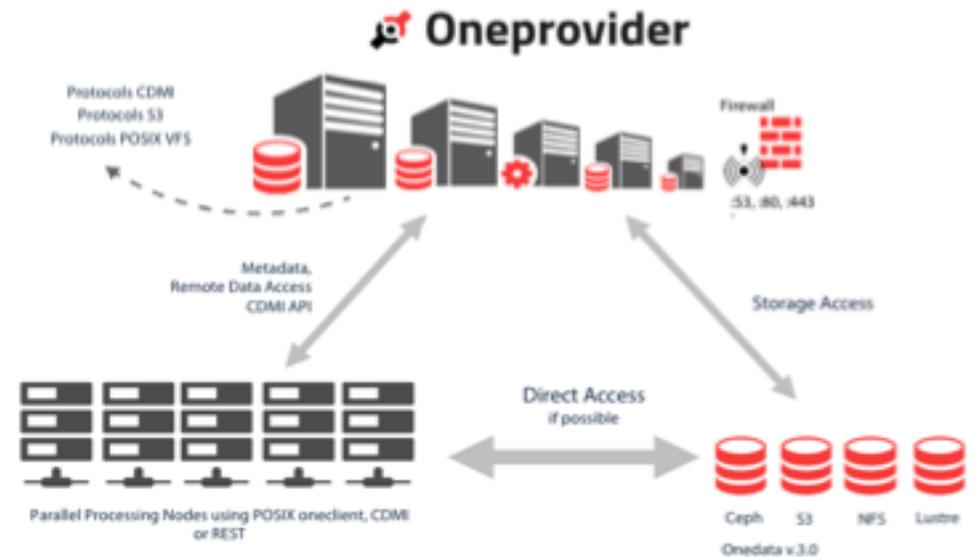
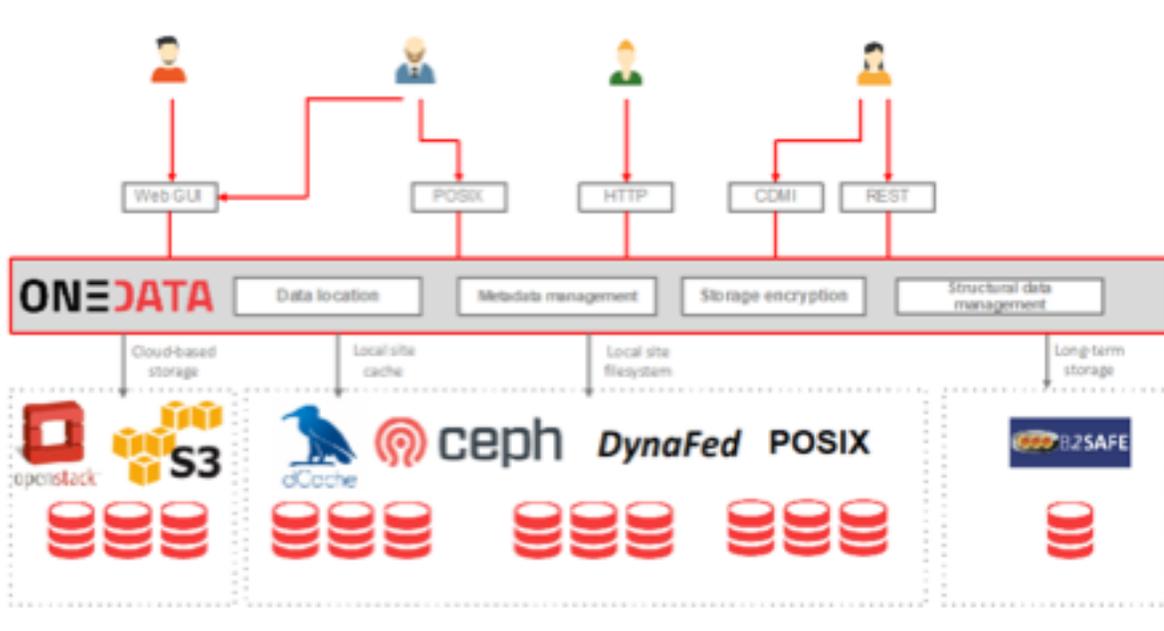


<https://onedata.org>



Onedata: développements

- ✗ Plateforme d'accès unifié aux données au niveau du PaSS et pour l'Exascale
- ✗ Gestion avancées des métadonnées sans schéma prédéfini
- ✗ Services de chiffrement et de stockage sécurisé
- ✗ Gestion des données sensible et stockage des clefs dans Onedata

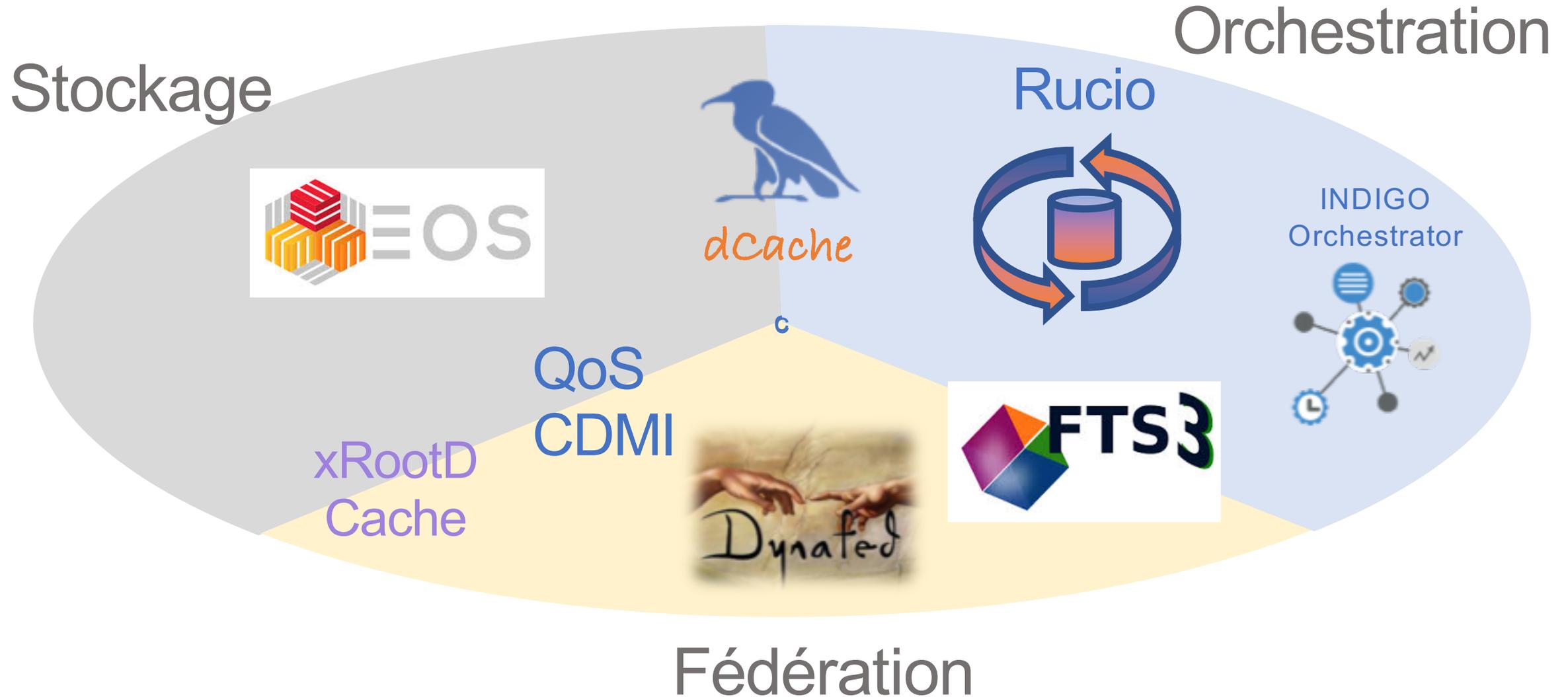


Orchestration et gestion des données axée sur les politiques (WP4)

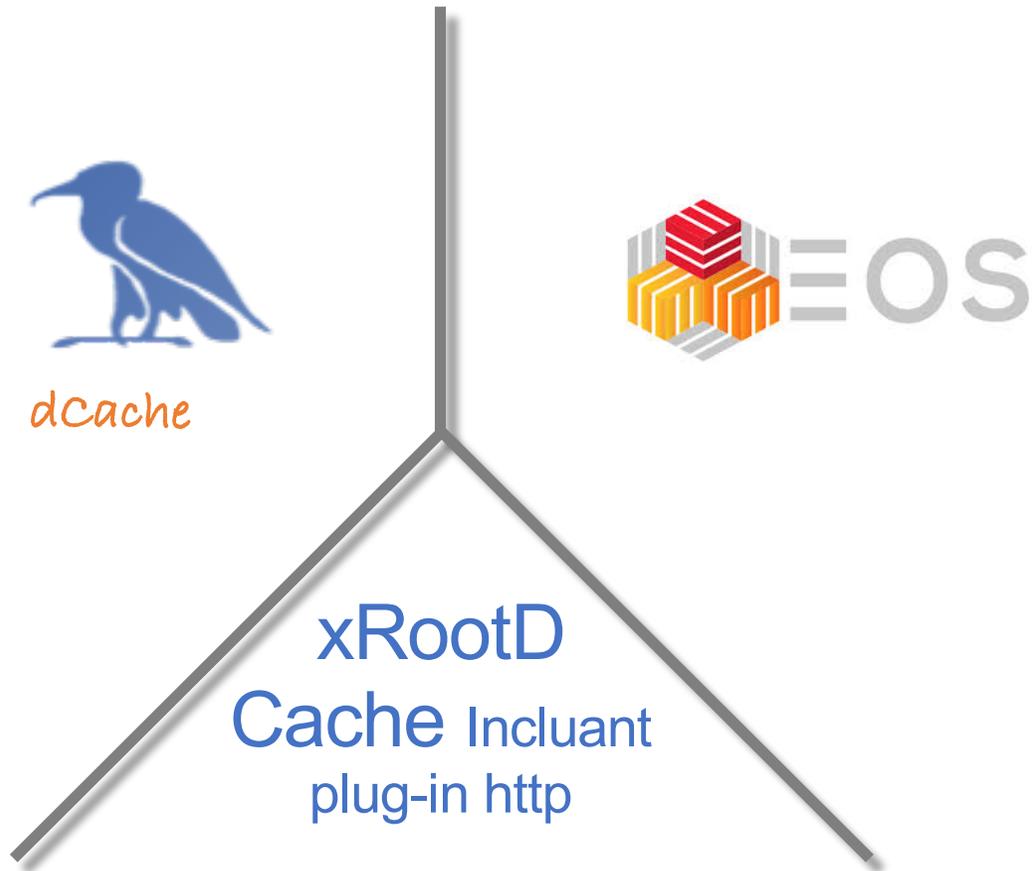


- ✗ Implémenter une **orchestration des «data workflow » configurable**, en termes de location des données et de qualité de stockage (QoS).
- ✗ Fournir des **services de cache de données** à tout les niveaux.
- ✗ Présenter **des interfaces basées sur les évènements** aux systèmes externes
 - ⋯→ Générer des événements pour les services d'orchestration XDC quand les données entrent le système XDC.
 - ⋯→ Générer des événements pour les clusters de calculs externes quand les données sont prêtes a être traitées.
- ✗ **Federer des sources de données hétérogènes**, formant un espace de données virtuel horizontal spécifique a l'infrastructure
- ✗ Architecture encore en cours de discussions

Les composants WP4



Stockage, Cache ...



Federation, Orchestration

Dynafed



- ✗ Federer des services dans un espace de nom virtuel global.
- ✗ Supporte http, WebDAV, S3
- ✗ Peut utiliser des bases pour mapper les noms

FTS

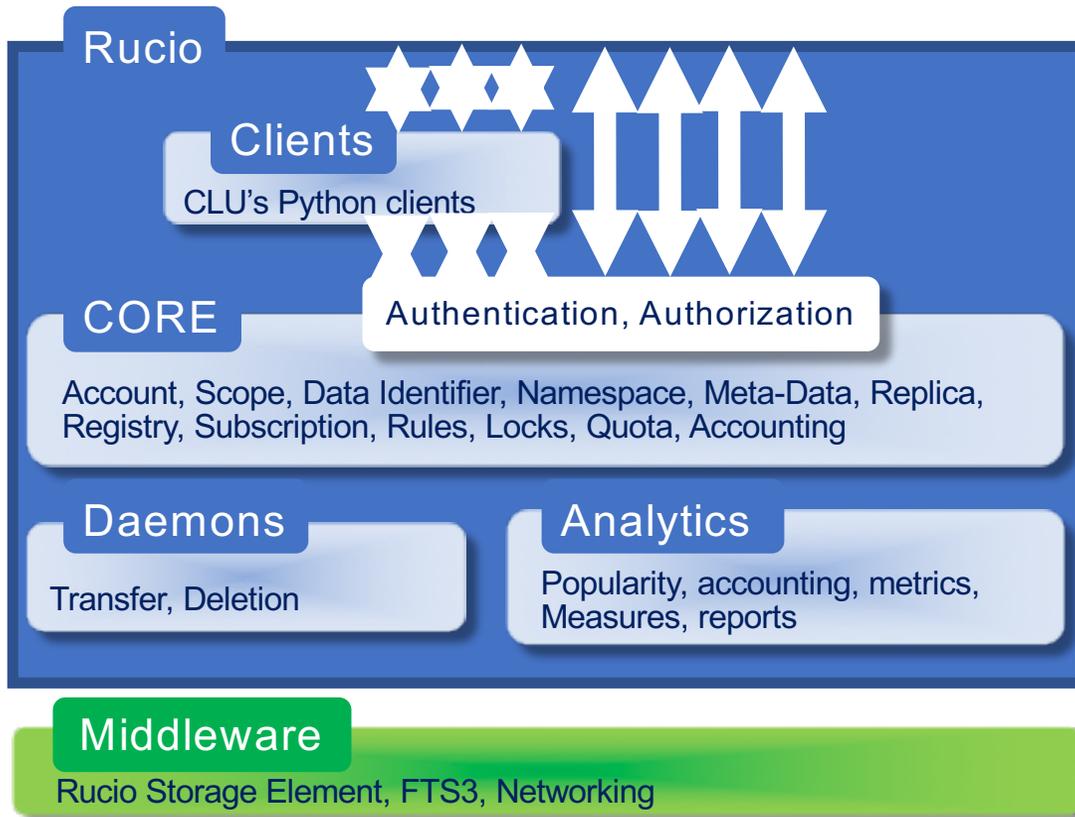


- ✗ Orchestration fiable de transfert sur des grandes distances
- ✗ Supporte plusieurs protocoles, GridFTP, http, xrootd
- ✗ Peut choisir la sources la plus appropriée
- ✗ Tiens compte des topologies réseau et du congestionnement

Rucio

Applications

Productions, Analysis, End-User, data export, physics Meta-datda



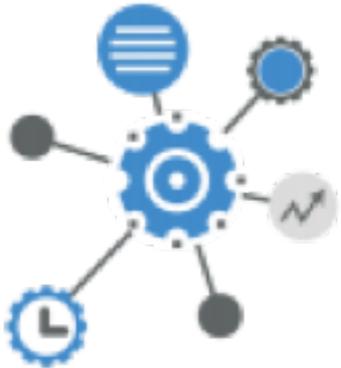
From the Rucio Tutorial

Les démons



- ✗ **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**

L'orchestrateur INDIGO PaaS Orchestrator



- ✗ L'orchestrateur INDIGO PaaS Orchestrator fait partie de la couche PaaS et sert à **instancier des ressources sur les solutions de Cloud** (comme **OpenStack** et **OpenNebula**) et les clusters **Mesos**.
- ✗ Les demandes de déploiement sont définies avec des **templates écrites en TOSCA** YAML Profile 1.0 et sont déployées sur le meilleur sites disponible. Afin de pouvoir effectuer cette tâche
 - Les SLAs sont collectées ainsi que les infos de supervision et des données d'autres services,
 - En fonction de quoi le "Cloud Provider Ranker" renverra la liste des meilleurs sites de clouds.
- ✗ Les **APIs REST exposées** sont consommés par le portail FutureGateway.

INDIGO-DataCloud CDMI QoS

Qualité de Service pour le stockage (Page du Broker)

Logout

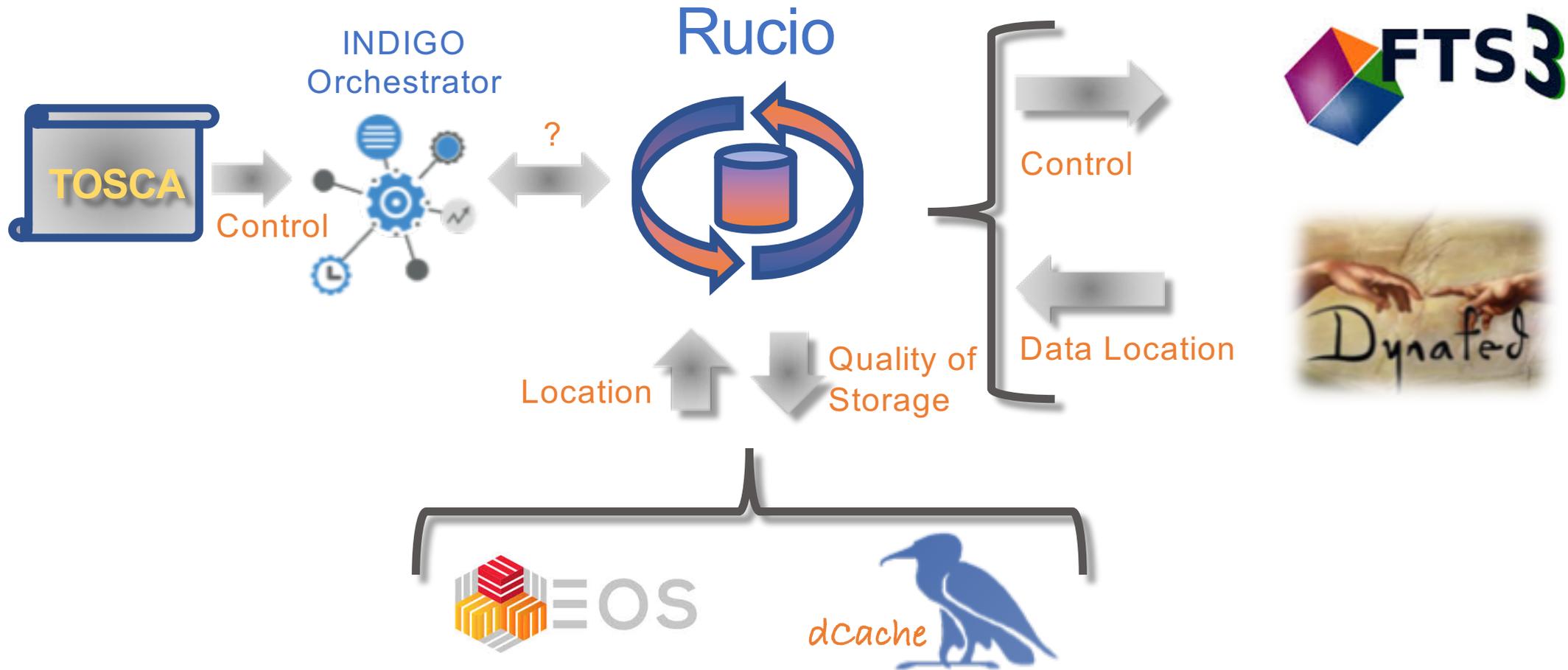
Available Qualities of Service

- Latence [ms]
- Nombre de Copies
- Durée de stockage
- Location
- Transitions disponibles

Name	Access Latency [ms]	Number of Copies	Retention	Storage type	Available Transitions
disk	100	3	IT	Processing	tape, disk+tape
disk+tape	100	2	IT	Processing	tape
DiskAndTape	50	3	IT	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	10000	2	DE	Archival	profile1
SSDDisk	10	1	IT	Processing	StandardDisk, Tape
StandardDisk	1000	3	20 years	IT	Archival

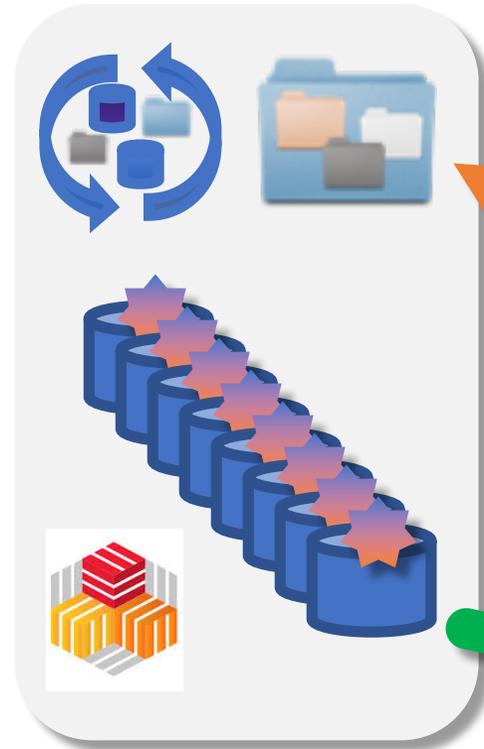
DESY
KIT
PSNC
INFN

Orchestration, flux de contrôle



Implémentation potentielle d'un cache avancé

CERN



Namespace synchro

Requête

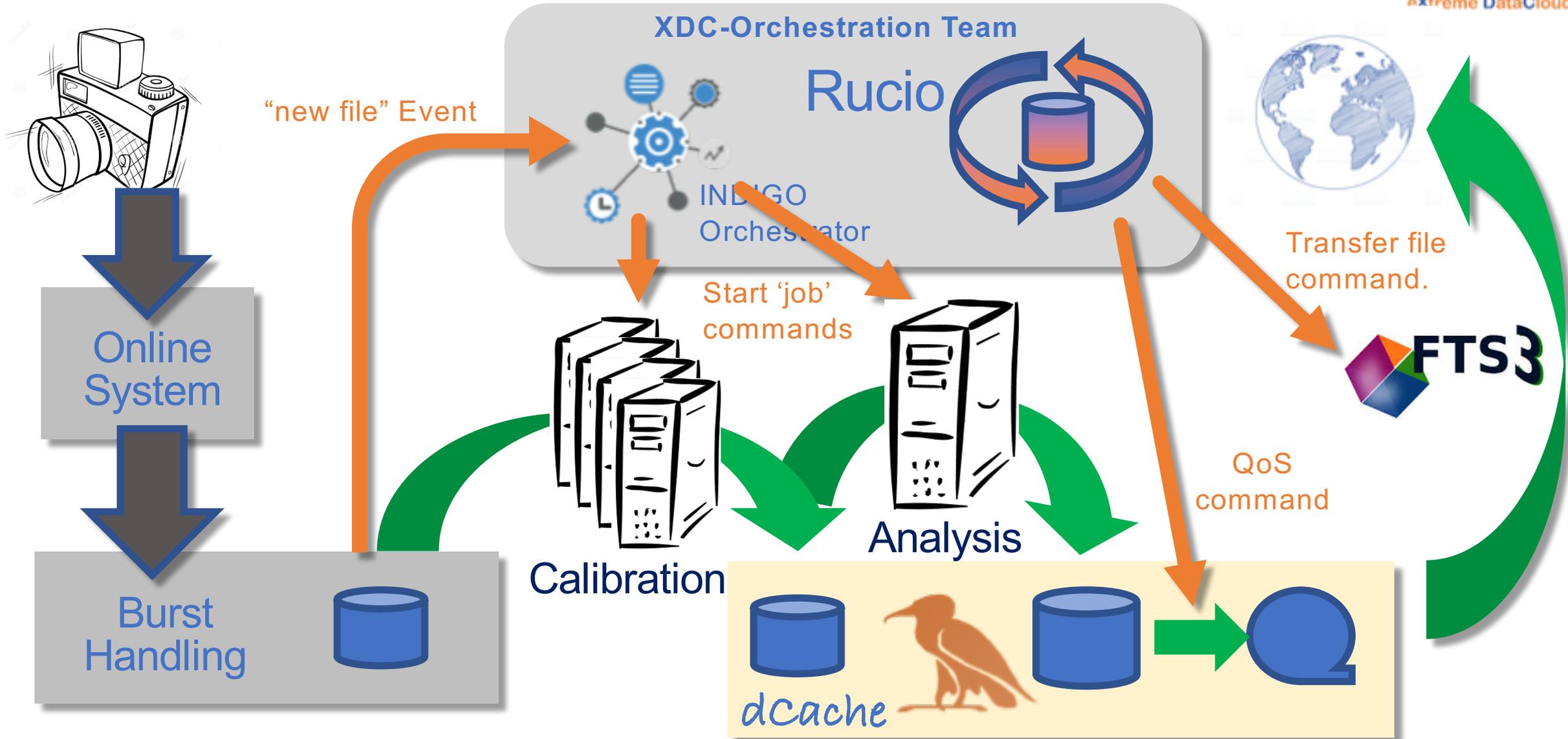
DESY



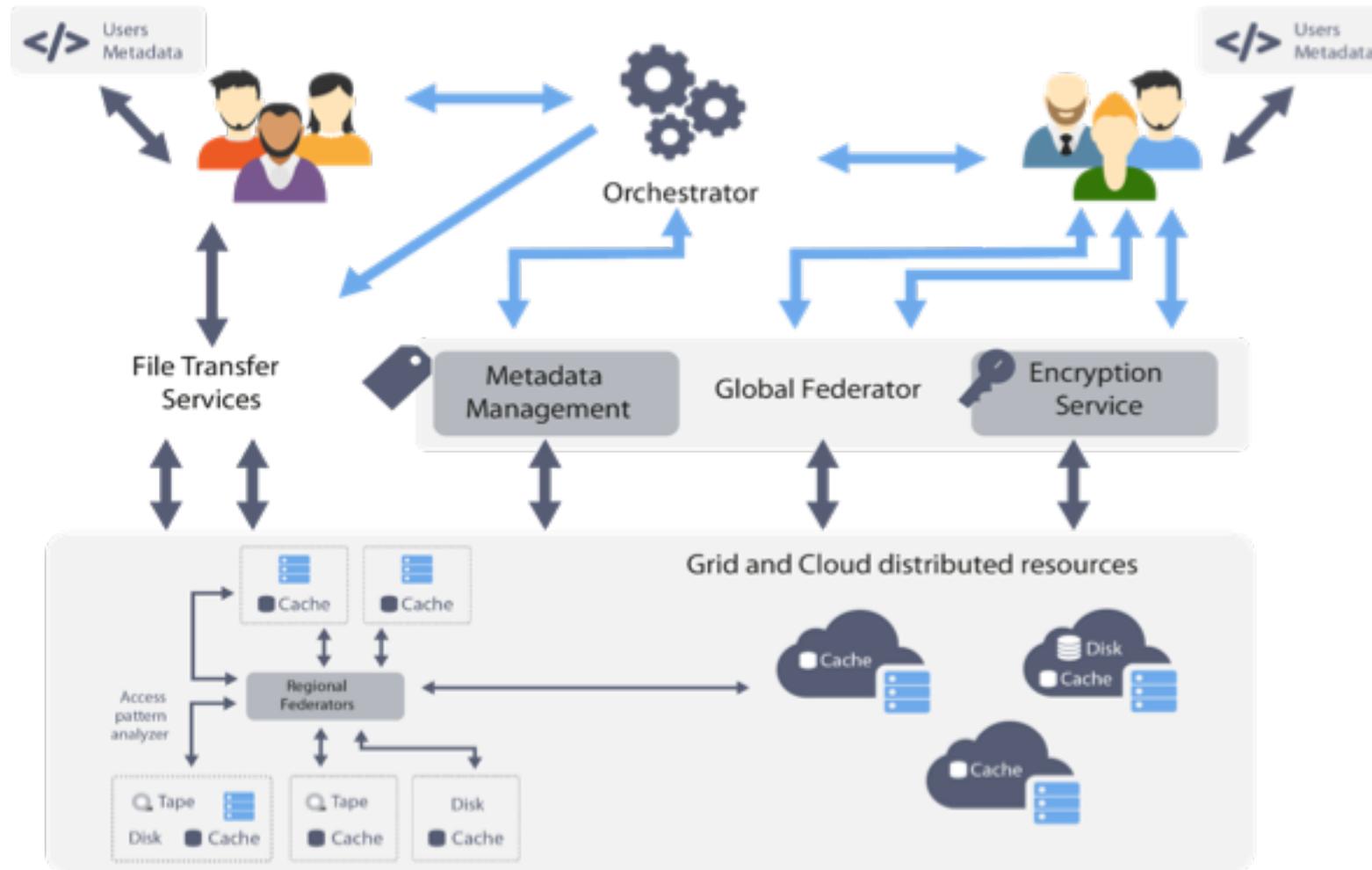
Avantages

- ✗ La même pile logicielle déjà utilisée sur les sites.
- ✗ Une fois que les données sont transférées localement, un namespace a été créé localement et les données sont disponibles localement indépendamment du lien avec le service central

Le cas d'utilisation simple X-FEL



XDC: architecture haut niveau



Statuts du Projet

- ✘ Commencé le 1er Novembre 2017
- ✘ Collecte des besoins détaillés des communautés de recherche effectuée
- ✘ Définition de l'architecture détaillée en cours
- ✘ Création d'un testbed pilote
 -→ Pour l'instant réservé aux communautés internes
 -→ Ouverture à des utilisateurs externes en discussion
- ✘ Liaisons initiées avec d'autres initiatives de gestion des données et liées a EOSC
 -→ Data Organisation Management Access (DOMA) WG a propos des "data lakes »
 -→ EOSC-Hub
 -→ EUDAT
 -→ DEEP-HybridDataCloud
 -→ RUCIO
 -→ Tout les projets EINFRA-21

Conclusion

- ✘ XDC a in programme de développement ambitieux pour des services de gestion des données
 - ☛→ Nous voulons supporter des besoins et cas d'utilisation très variés
- ✘ Nous utiliserons autant que possible des protocoles standard pour rendre les solutions les plus générales possible
- ✘ Première version prévue en Octobre-Novembre 2018
- ✘ Durabilité a long terme des produits
 - ☛→ Pousser en amont vers les projets originaux toutes les fonctionnalités développées par XDC
 - ☛→ Impliquer les communautés d'utilisateurs pour exploiter les réalisations XDC dans leurs environnements de production
 - ☛→ Publier les développements XDC dans l'annuaire de services EOSC
 - ☛→ Mettre en place des liaisons avec les fournisseurs d'e-infrastructures et les autres projets de développement de la gestion des données

XDC Contacts



✂ Website: www.extreme-datacloud.eu

✂ [@XtremeDataCloud](https://twitter.com/XtremeDataCloud) on Twitter

✂ Mailing list: [info<at>extreme-datacloud.eu](mailto:info@extreme-datacloud.eu)