

# DOMA: Service de stockage Retour sur la première année

JCAD 2019

- ▶ Le projet DOMA (Rappels)
- ▶ DOMA : Année 1
- ▶ Les activités en France
- ▶ Conclusion

- ▶ Voir la présentation faite aux dernières [JCAD](#)
- ▶ Projet international intégrant les différents aspects que pourraient prendre l'accès aux données dans les années à venir. Avec en filigrane la volonté d'en maîtriser le coût.
- ▶ Les objectifs sont :
  - Suivre les avancées et développements dans les différentes activités identifiées.
  - Etre un forum de partage d'informations.
  - Veiller à l'interopérabilité des différentes solutions de stockage.
- ▶ DOMA n'est pas un projet de DATALAKE ni un produit mais une organisation qui participera à définir ce que sera un DATALAKE ou autre.
- ▶ Bien qu'initié par les besoins futurs des expériences de physiques des hautes énergies, DOMA ambitionne de couvrir globalement la question du service de stockage scientifique, quelle que soit la thématique.
- ▶ Les participants sont et surtout doivent être divers:
  - Les expériences/utilisateurs ,les développeurs de piles logicielles, les fournisseurs de services de stockage, les diverses infrastructures impliquées.

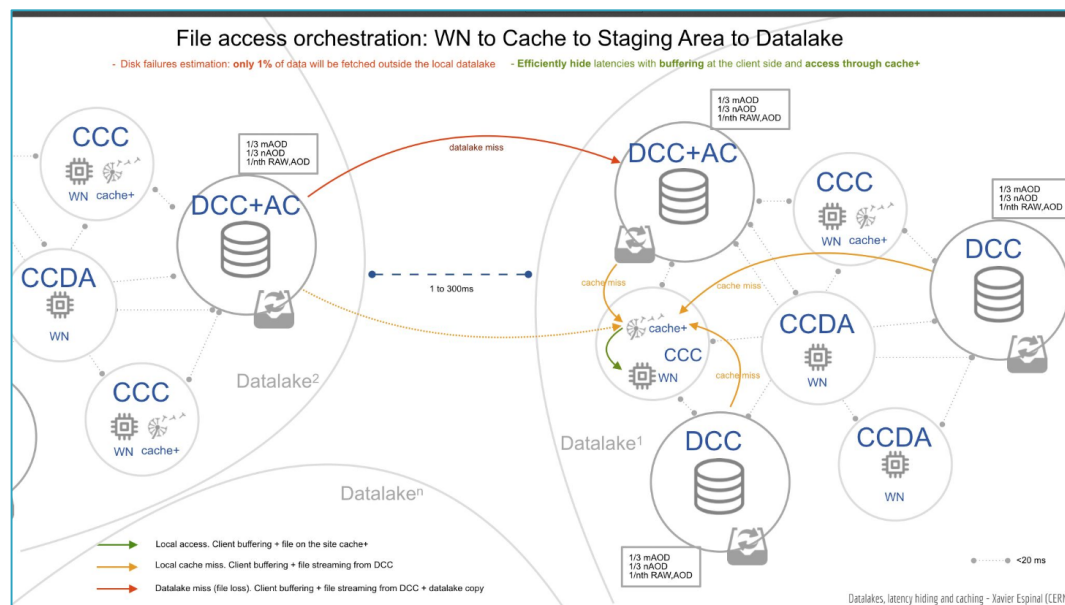
## ► DOMA-FR

- Organise les activités DOMA en France.
  - Une à deux rencontres par an.
- Met en avant et identifie toutes les activités qui sont en relation avec ces études.
- Sont impliqués des ITA, des chercheurs.
- DOMA-FR a vocation à nourrir le projet DOMA global mais aussi à intégrer et évoluer en fonction des pistes/réflexions qui sortiront de DOMA.
- Laboratoires qui sont partie prenante.
  - IPHC, CC-IN2P3, LAL, LAPP, LLR, LPSC, CPPM, LPC

## ► Plusieurs groupes de travail sont en place (WP)

### ◦ ACCESS

- Regarde les aspects datalake, cache, accès à la donnée.
- Sont concernés : les utilisateurs/expériences, les fournisseurs de ressources (sites), les piles logicielles.
- Ont notamment été étudiés les performances et stratégies de caches, les patterns d'accès aux données et fichiers et différents modèles de datalake.

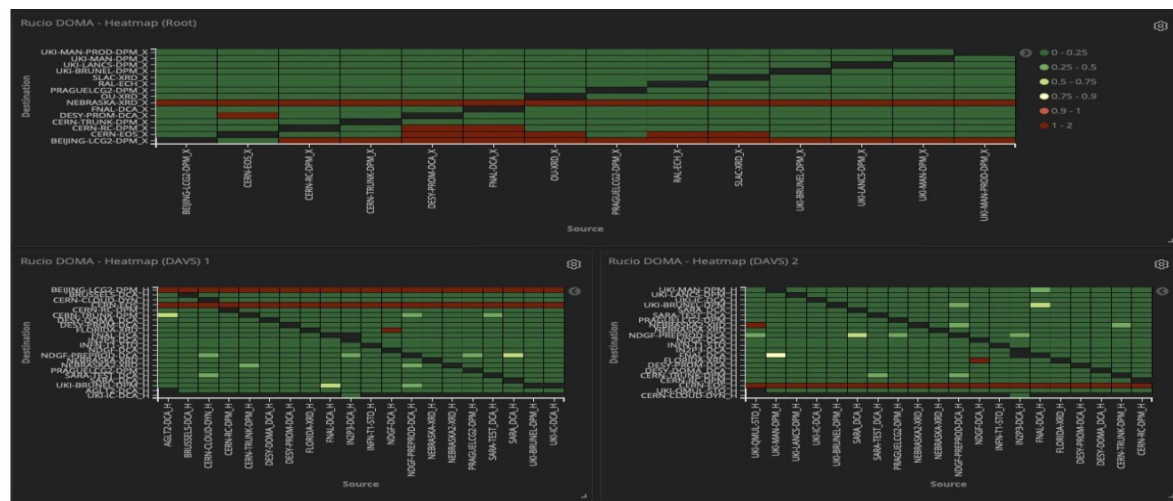


## Datalake Straw Model v1

Graph from Xavier Espinal (DOMA)

## ◦ Third Party Copy (TPC)

- Regarde les aspects protocoles et copie tierce.
- La copie tierce est un mécanisme permettant la copie de données de site à site directement sans passer par le site ordonnanceur de la requête.
  - C'est une fonctionnalité liée au protocole (http, xrootd,...) qui met en oeuvre des mécanismes de délégation de droits.
- Les phases de validation technique et de passage à l'échelle ont été validées. Début 2020, les sites de grilles (WLCG) devront offrir en production un service de stockage compatible avec le TPC.



13

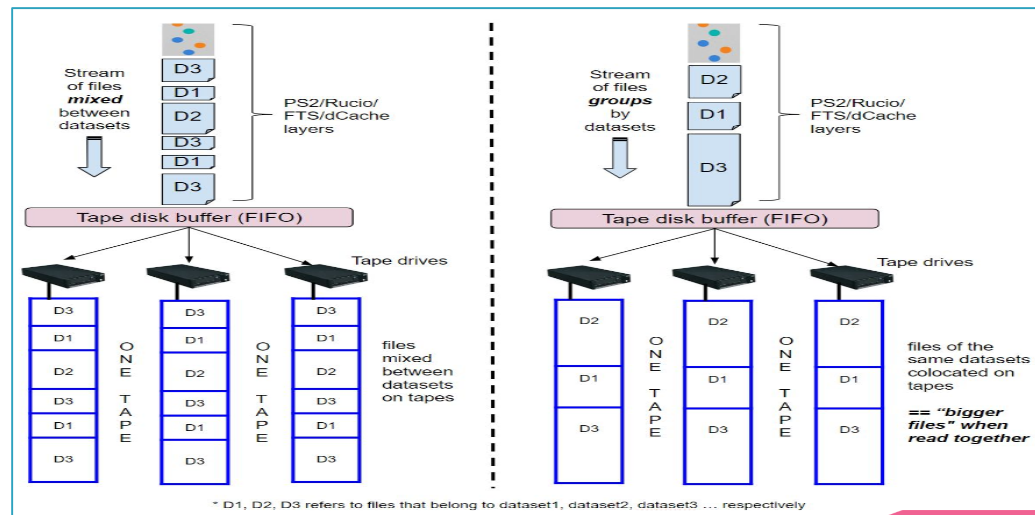
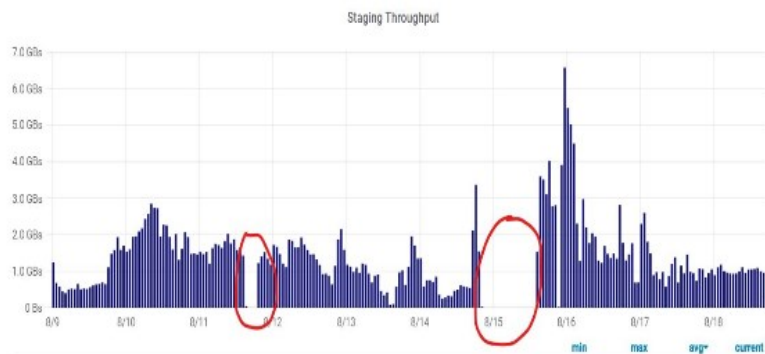
- Quality of Service ( QoS)
  - Se focalise sur les aspects définition et implémentation de notion de qualité de service (QoS) au niveau des services de stockage.
  - L'objectif est de passer d'une notion de « disk and tape » à des classes de QoS plus riches, intégrant des notions de performances, coûts, résilience, sécurité,....
  - Actuellement le but recherché est de :
    - Formaliser un vocabulaire commun afin que fournisseurs de services et utilisateurs mettent derrière certains mots la même définition.
    - Définir une liste d'objectifs prioritaires.
- Networking
  - Activité principalement axée sur la mise en place de broker de réseau.
  - Data Transfert Node, SDN, réseau dédié aux services de transferts (MultiONE)
  -
- AAI : Authorisation et Identification

- Fédération de stockage
  - Mise en place d'une fédération de stockage à l'échelle nationale
    - Annecy/Clermont-Ferrand/Grenoble/Marseille.
  - Cet unique point de stockage est utilisé pour des tests mais aussi en production (par l'expérience du CERN ATLAS).
  - Permet d'aborder les problématiques liées notamment à la latence réseau et les approches de caching envisageables.
- File système distribué à large échelle
  - Création d'une instance CEPH de stockage à l'échelle du « campus » sur le plateau de Saclay et extension jusqu'à Strasbourg.
- Third party Copy
  - Déploiement dans le cadre d'un testbed international de endpoints compatibles avec des protocoles sachant faire du TPC.
    - Phase de validation technique et passage à échelle validé lors des 12 derniers mois.
    - Plusieurs Po de donnée échangés chaque jours à travers ce testbed.
  - Déploiement d'ici 2020 sur les services de stockage grille en production



## ◦ Data Carrousel

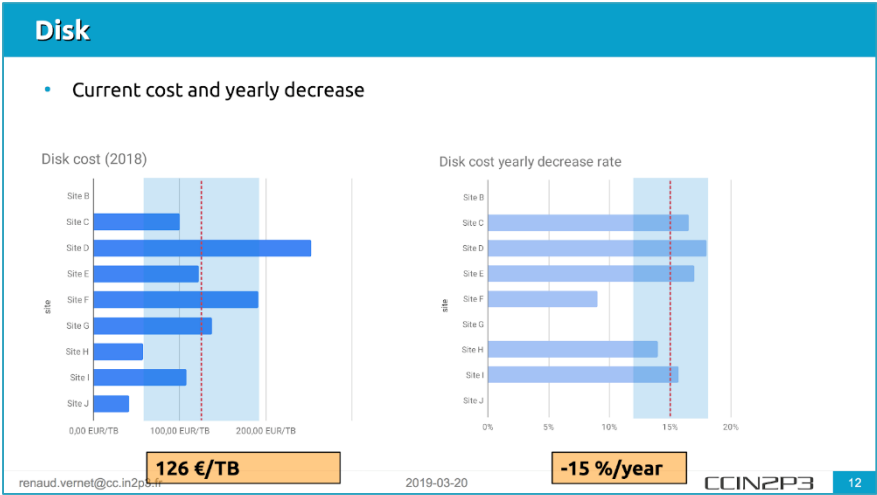
- Tests et validation de nouvelles approches dans l'usage des solutions de stockage de masse (les bandes magnétiques).
- Validation des performances intrinsèques des solutions hardwares et des sites.
- Validation et optimisations de différents pattern de gestion des données.
- Couvrir un large spectre de compétences (hardware jusqu'à la pile applicative).



## Tape organization writing

Graph from Xavier Espinal (DOMA)

- Modélisation des coûts
  - L'objectif de cette activité, plus large que la simple problématique du stockage, est de créer un modèle de coûts réaliste et global qui permette de faire des projections notamment pour anticiper le coût de tel ou tel modèle de stockage.



Coûts et décroissance des coûts disque  
sur quelques sites important de grille

| Change                                         | Effort Sites                            | Effort Users     | Gain                                          |
|------------------------------------------------|-----------------------------------------|------------------|-----------------------------------------------|
| managed storage on 15 sites + caches elsewhere | Some on large sites/gain on small sites | little           | 40% decrease in operations effort for storage |
| Caches at most sites (dataLake strawman)       | Some everywhere                         | Frameworks some  | 15% of storage                                |
| Reduced data redundancy                        | Some large sites                        | Frameworks some  | 30-50% disk costs                             |
| Reduced data replication and cold data         | little                                  | Frameworks some  | 30% disk costs                                |
| Compact data formats for analysis              | none                                    | Some             | >15% disk costs                               |
| Scheduling and site inefficiencies             | Some                                    | Some             | 10-20% gain CPU                               |
| Reduced job failure rates                      | Little                                  | Some-Massive     | 5-10% CPU                                     |
| Compiler and build improvements                | None                                    | Little-Some      | 15-20% CPU                                    |
| Improved memory access/management              | None                                    | Realistic        | 10%-15% CPU                                   |
| Exploiting modern CPU architectures            | None                                    | Massive          | 100% CPU                                      |
| Paradigm shift algorithms                      | Some                                    | Massive-Infinite | Factor 2-100 CPU                              |
| Paradigm shift online/offline data             | Little                                  | Massive-Infinite | 2-10 CPU 10-20 Storage                        |

Potentiels gains

Graph from « systems performance and cost modeling working group

- DOMA est une collaboration axée sur le partage d'expertises, d'expériences et de tests.
- De nombreux aspects ( opérationnels, techniques, organisationnels) sont abordés.
- Elle s'inscrit dans un contexte international et multi communautés scientifiques.
- Elle se décline en France aujourd'hui autour essentiellement d'acteurs issues de la physique des hautes énergies. Mais il y a une vraie volonté/besoin de l'ouvrir à d'autres acteurs.
  - Venez nous voir
- Vous êtes intéressé par certains aspects techniques mentionnés dans la présentation
  - Venez nous voir