

Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules

# DOMA: Data Organization Management Access evolution

Objectifs 2021



- ▶ DOMA était jusqu'à présent organisé suivant trois thèmes principaux :
  - Activités liées aux services de stockage
    - QoS.
    - Aspects matériel (RAID vs EC,...).
    - Outils de management des données
  - Activités liées aux protocoles
    - Redirection (Third Party Copy).
    - HTTP vs xroot vs.....
  - Activités liées à l'accès à la donnée
    - Caching.
    - Backend de stockage (Ceph).
    - Performance de l'accès distant.

- ▶ DOMA 2021: Réorganisation/ changement d'objectifs. Compatibles avec les conclusions de la revue sur le calcul HL-LHC
  - Aller vers la mise ne place de prototypes de datalake pour valider /quantifier globalement les solutions et approches.
    - Probablement plusieurs prototypes au départ qui doivent converger.
    - Regroupe beaucoup des aspects liées au data.
      - Caching.
      - Performance.
      - Opérationnelle.
      - A ce jour des proposals de proto datalake ATLAS et CMS existent.
      - Le défi c'est la convergence.
    - L'activités liée à ALPAMED est en phases avec ca ( cf présentation stephane).
      - En faire plus ? Avec qui ?

- Réalisation de benchmarks de performances
  - Les benchmarks sur l'usage des tapes restent d'actualités
    - Idéalement ces benchmarks évoluent vers une approche plus multi application ( ATLAS + CMS+ ...)
    - Ils concernent uniquement le Tier1
    - A moyens terme ces benchmarks ont vocation à intégrer les prototypes datalake
  - Les benchmarks visant à mesurer les performances dans les transferts des données
    - Ils impliquent des mesures de performance simultanées des réseaux et des systèmes de stockage.
    - Ils visent à montrer que l'on peut atteindre des objectifs chiffrés ( XX de Gb/s entre sites de stockage à stockage)
    - Ils doivent se dérouler sur la période 2021-2023
    - Initialement ils visent à mesurer les capacités : Europe-US, Tier0-Tier1 et Tier1->Tier2

- L'objectifs de l'activité QoS glisse de la définition des différentes QoS vers la validation plus technique (hardware) de solutions qui permettraient d'atteindre ou proposer certaines QoS
  - Tests de solutions techniques.
  - Implique probablement des investissements.