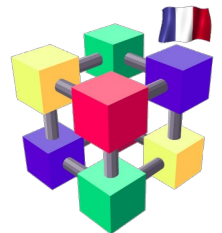




# HSF-WLCG workshop : retour

---

Journées LCG-France – 18-20 juin 2024 – LPCA  
David Bouvet, Laurent Duflot



## HSF-WLCG workshop – 13 au 17 mai 2024 – DESY, Hambourg

- <https://indico.cern.ch/event/1369601/timetable>



## Processeurs

- 3 industriels seulement pour la fabrication des processeurs (Samsung, TSMC, Intel)
- ARM
  - depuis 10 ans en augmentation mais seulement 8 % du marché mondial
  - évolution à suivre
- Coûts CPU
  - ralentissement sur les 2 dernières années dans le facteur d'amélioration prix/performance (moy. 20 %/an sur les 5 dernières années)
    - instabilités et prédiction coût difficile pour les années à venir ?
  - passage à carte 25 Gb pour le HL-LHC à intégrer
  - ralentissement gain efficacité énergétique CPU : gain moy. 15 %/an sur 10 ans
  - stratégie environnementale ⚡ gain financier
- GPU
  - marché très volatile et très haut
  - usage *online* ALICE, CMS et LHCb partiellement passé sur offres commerciales spéciales → TCO difficile à évaluer
    - clarification d'ici 2027 sur TCO ainsi que sur mix CPU/GPU  
(opération, effet secondaire/complication sur puissance et urbanisation, livraison, état du marché)



## Stockage

- Disque
  - 3 industriels dominant le marché (Toshiba, Seagate, WesternDigital)
  - peu d'améliorations en performance à venir, tout sur la capacité (disque 30 TB arrive sur le marché)
    - besoin en performance 2029 > x10  
⇒ pb IO attendu
    - compensation par
      - plus de serveurs ⇒ plus cher, et trop d'espace
      - SSD capacitifs : capacité nécessaire disponible en 2028 ? (60 TB utile pour le moment)
      - changement d'architecture : CPU + SSD pour mix calcul et stockage avec carte réseau 400 Gb ?
- Coûts
  - même tendance que pour le CPU avec plutôt augmentation de 10 % cette année
  - augmentation plus forte possible en 2028 du fait pb IO ?



## Abandon GridFTP dans FTS

- Utilisé de façon occasionnelle (backup)
- ALICE, ATLAS, CMS: plus d'utilisation
- LHCb et BELLE II : encore des endpoints GridFTP en secours → peut être supprimé
- DUNE : uniquement entre FNAL et NERSC  
→ arrêt du support de GridFTP par les sites après le 30/06/2024

## IPv6 sur WN

- Déploiement à finaliser d'ici le 30/06/2024 → sites FR ✓

## Benchmark

- Besoin de participants au niveau développement et opérationnel

## Middleware et EL9

- LCMAPS disponible dans EPEL
- Reste pas prêt lors du workshop principalement dû à UMD-5 non disponible  
→ paquets disponibles dans le [repo WLCG](#) depuis



## Sécurité

- Introduction de nouvelles collaboration sur LHCONE ⇒ + de sites connectés ⇒ réduction confiance potentielle
  - ⇒ **Projet MultiONE**
  - préfixe d'identification ajouté et annoncé au réseau par des *tags* (*BGP communities*) pour identifier les communautés que le site sert.
  - pas de VPN supplémentaire, pas de changement au niveau des sites quand un nouveau site est ajouté à LHCONE.
- Calendrier
  - mise en place des *tags* pour printemps 2025
  - début filtrage 2026 en préparation du DC26

## Jumbo frames

- Avantages pour les transferts longue distance mais problèmes opérationnels
  - tests approfondis à venir à plus grande échelle
    - gains de performance
    - problèmes opérationnels
  - déploiement sur 50 % des sites d'ici DC26



## STFC chargé de

- Développement
  - APEL client and server
  - APEL messaging client
- Opération
  - APEL server

## Réorganisation de l'équipe à STFC

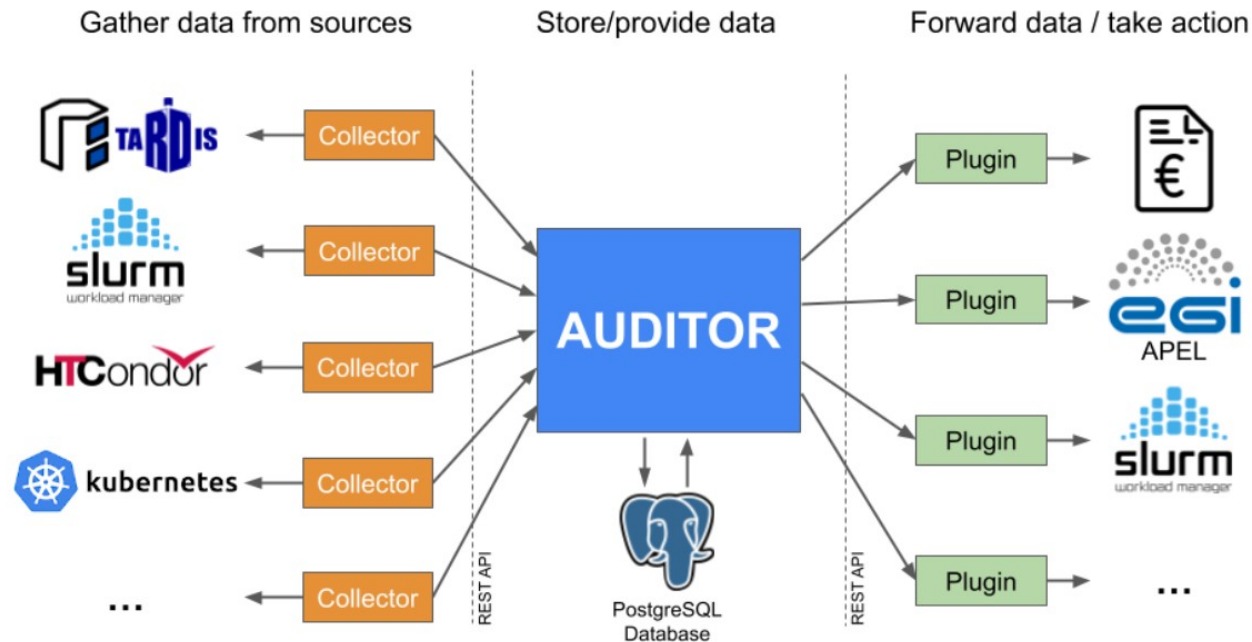
## Calendrier de développement

- mai-juillet
  - passage à python3
  - finalisation et passage en production de la comptabilité HEPscore23
  - nouveau système de publication/synchronisation
  - migration de l'infra du serveur APEL vers Rocky9
- juillet-août
  - authentification par jeton
  - comptabilité GPU et stockage



AUDITOR : AccoUning Data handling Toolbox for Opportunistic Resources  
(<https://github.com/ALU-Schumacher/AUDITOR>)

- Permet comptabilité ressources différentes partagées par un OBS (Overlay Batch System)





# Sondage sur outils et procédures *accounting* et *pledges*

Sondage réalisé en juin 2023 auprès des sites et VO

(APEL, portail EGI, WAU, procédure de validation de *l'accounting*, rapports mensuels, procédure pour les *pledges*, CRIC UI pour les *pledges*)

- 23 sites (5 T1, 16 T2, 1 T3)
- APEL
  - satisfaction sauf pour le support.
    - nouvelle équipe ⇒ amélioration ?
- Portail EGI
  - satisfaction. Suggestions d'amélioration.
- WAU (WLCG Accounting Utility)
  - peu de personnes connaissent cette plateforme.
  - VO souhaitent comparer leur données avec portail EGI
    - données expériences dans WAU + **comparaison disponible** avec portail EGI et données validées CRIC
- CRIC UI
  - suggestions d'amélioration + demandes du WLCG Project Office
    - nouvelle interface + historique des *pledges/VO requirements* dans **Grafana**
- Validation *accounting*
  - retour positif



## GGUS remplacé par instance Zammad

- Période de **transition** (coexistence GGUS et Zammad) : **01/10/2024** → **déc. 2024**
- Retrait GGUS : fin 2024 passage en lecture seule

## Zammad

- Beaucoup de travail pour intégrer spécificités système ticket de WLCG
  - hiérarchie *support units*
  - plusieurs type de ticket
  - schéma de notification complexe
  - nombreux niveaux différents : groupes, institutions, VO, *support units*, rôles...
  - intégration AAI
  - administration utilisateurs privilégiés par les VO
- Révision et adaptation de l'interface et des *workflows*
- Instance de test déployée : <https://helpdesk-dev.ggus.eu/>
- Instance de production définie : <https://helpdesk.ggus.eu/>
- Contact période transition : [help@ggus.eu](mailto:help@ggus.eu)

## Migration

- Documentation et statut : <https://confluence.egi.eu/display/EGIHLPDSK/EGIHLPDSK+Home>
- A la mise en production, chaque utilisateur :
  - se connecte
  - vérifie ses rôles et droits, et demande droits complémentaires si besoin
  - teste



## Questionnaire d'avril 2024

- 57 sites (T0, 9 T1, 46 T2, 1 T3)

## High memory

- Beaucoup de sites prêts et capables de supporter de jobs de 4 à 8 GB/cœur.
  - à condition que le besoin mémoire soit explicitement spécifié.
- Gestion allocation
  - préférence pour une utilisation pleine des cœurs
  - machines dédiées
  - efficacité dans la gestion des slots pour une utilisation optimale des ressources
- Préoccupation récurrente
  - comptabilisation de ressources (cœur inactif, CPU pour jobs *high memory*)
    - proposition d'adapter la comptabilité : tenir compte de l'utilisation CPU et mémoire
- Discussion/collaboration avec VO
  - clarification nécessaire sur le besoin en ressources
  - impact sur les *pledges* CPU si beaucoup de jobs *high memory*



## Whole-node

- Efficacité
  - doute exprimé sur l'efficacité des jobs *whole-node* / aux jobs plus petits.
- Optimisation ressources
  - gestion jobs *whole-node* complexe : draining, gestion des queues, partage des ressources en particulier pour les sites multi-VO.
  - perte de flexibilité pour les jobs plus petits.
- Besoin de tests pour comprendre les implications d'une telle utilisation.

## 16-core

- Préoccupations sur la compatibilité à l'architecture CPU avec NUMA (Non-Uniform Memory Access)
- Sites demandent que l'efficacité soit équivalente aux jobs 8 cœurs
- Compatibilité avec infrastructure actuelle (nœud non divisible par 16)
  - perte d'efficacité de *scheduling*, et d'optimisation dans l'utilisation des ressources
- Besoin de tests supplémentaire pour déterminer le nombre de cœurs le plus efficient



White Paper publié en avril (<https://arxiv.org/abs/2404.02100>)

- Issu du HSF AF forum

## Questionnaire LHCC

- LHCC a chargé WLCG d'établir un questionnaire sur l'AF à destination des VO
  - session pour compléter le document

## 200 Gb/s analysis demonstrator

- Challenge durant le DC24
  - à 25 % du HL-LHC → lecture à 200 Gbps
    - tenu avec données ATLAS et CMS

## Discussions autour

- des besoins/attentes utilisateur
- l'accès aux données
- monitoring pour les utilisateurs

