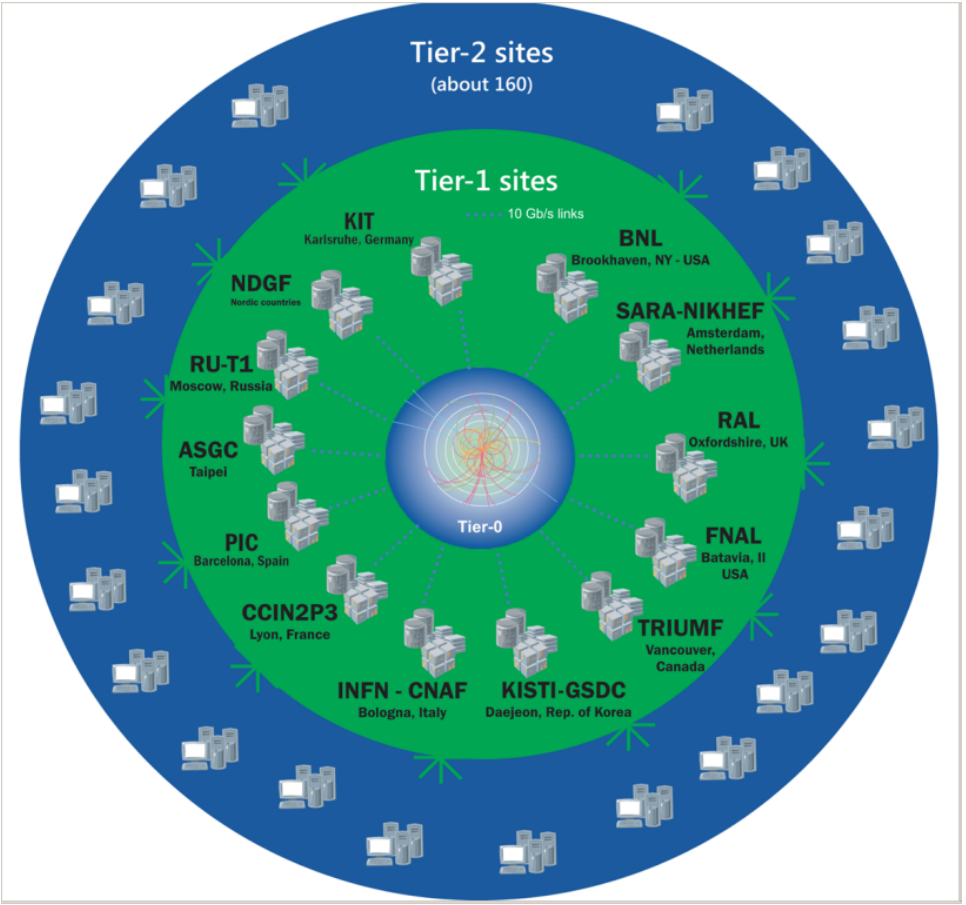
Introduction : avant le Run 1

- Modèle du tout début du Run1 très hiérarchisé : les Tier2 passent par les Tier1 pour échanger des données avec les autres sites. Pré-placement des données dans les sites. Reconstruction des données réelles uniquement au T1

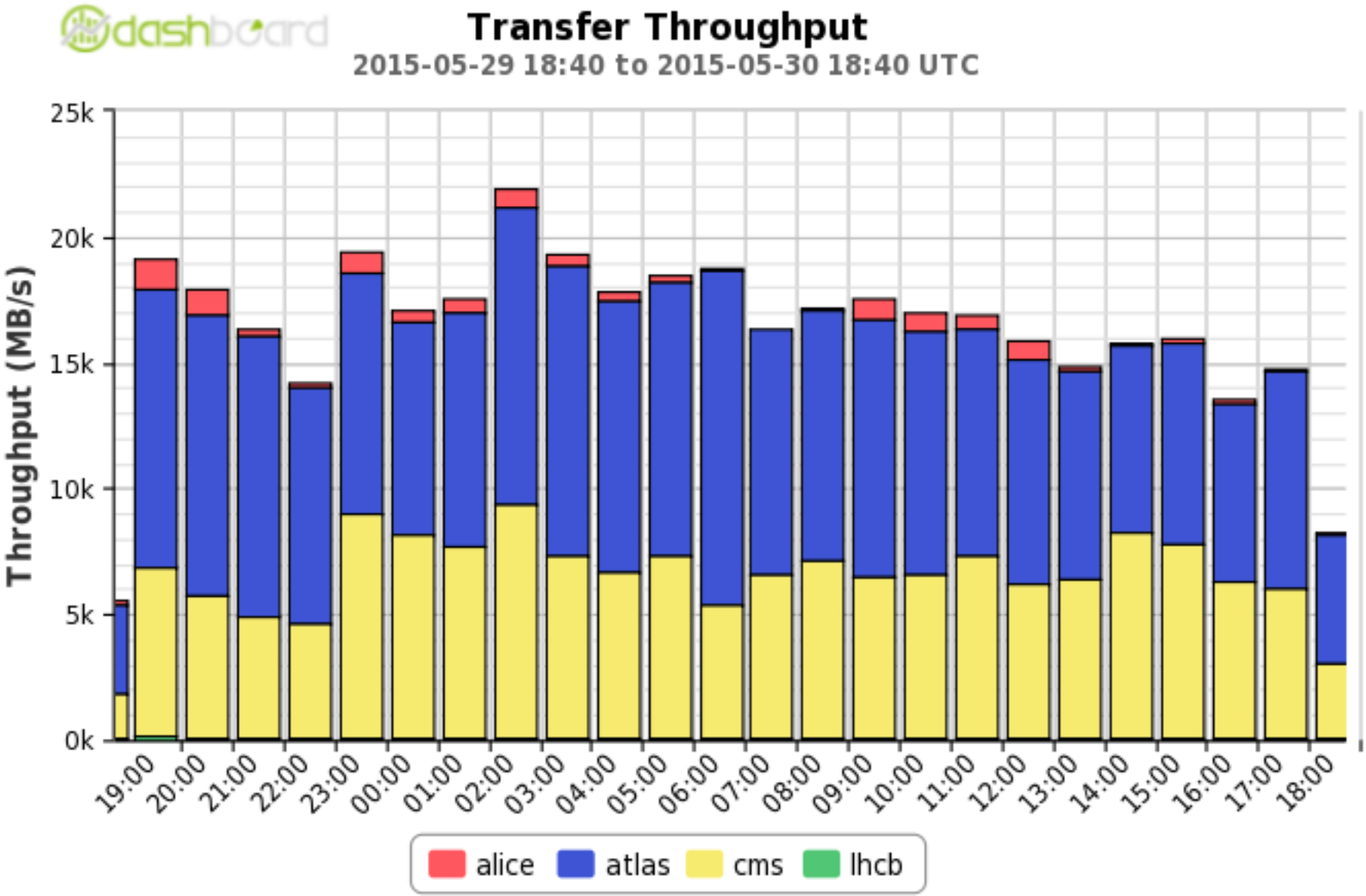


NB: avec des variations selon les expériences

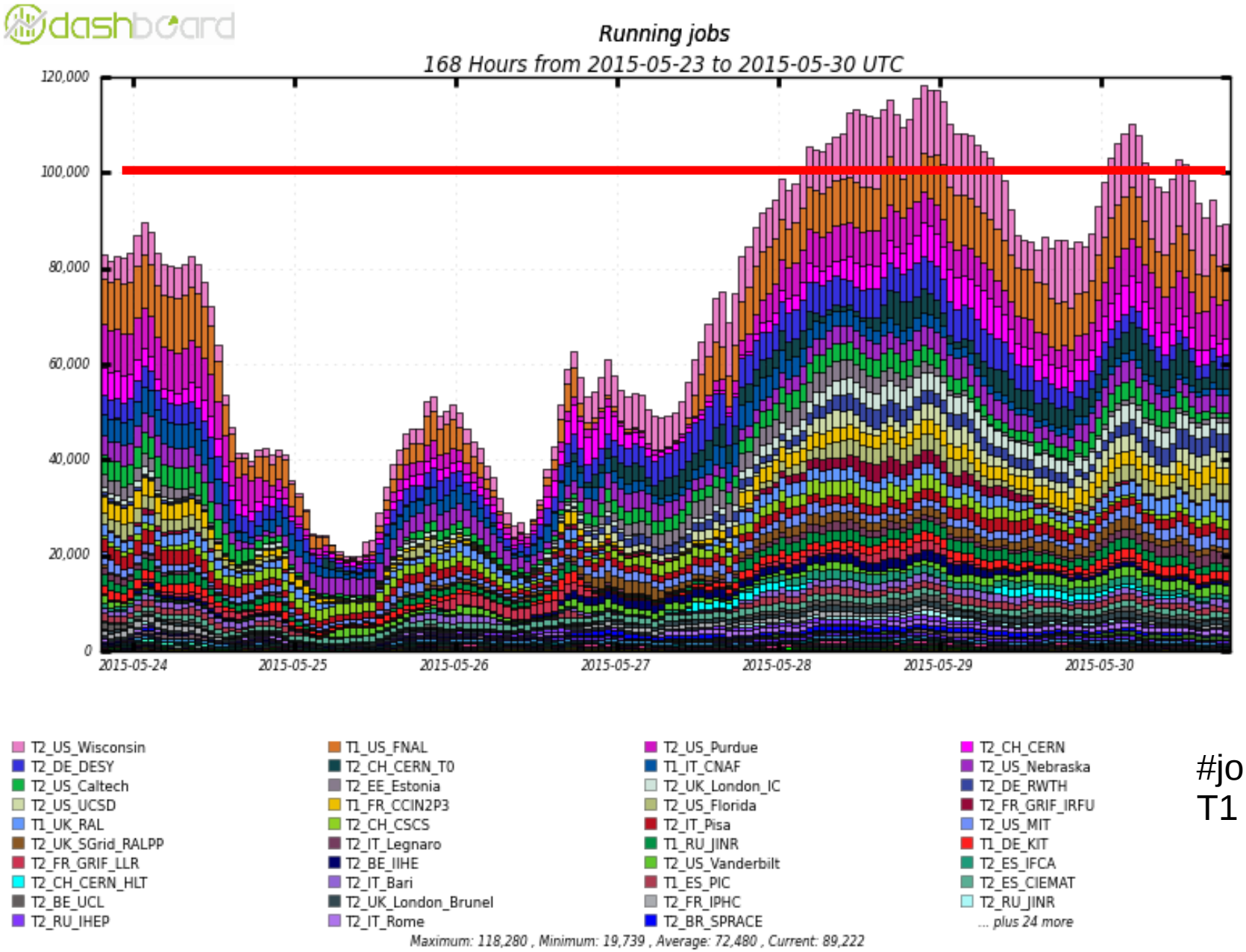
Introduction : évolutions pendant le Run 1

- Placement des données plus dynamique, en fonction de leur « popularité »
- Introduction des « bons T2 » (fiables et bien connectés) => mouvements de données sans passer par les T1. Reconstruction de données réelles possible pour certaines expériences.
- Introduction de l'accès de données à distance [les jobs ne vont plus aux données] principalement pour palier aux problèmes locaux des sites => mode « failover »
- Investigation de l'utilisation opportuniste des sites non grille : @HOME, HPC, cloud
  - Pas encore à très grande échelle mais par ex dans ATLAS @HOME est ~ Tier2

Quelques chiffres



Quelques chiffres



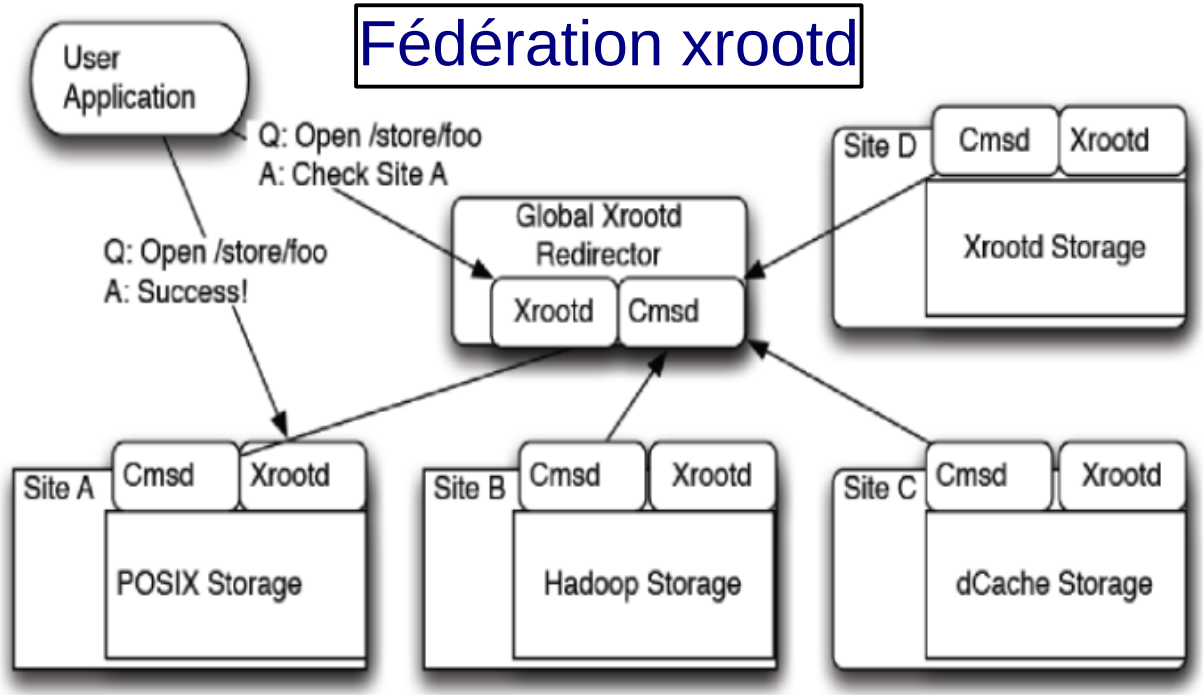
#jobs CMS
T1 + T2

Fédération de données

Pseudo-definition pour HEP

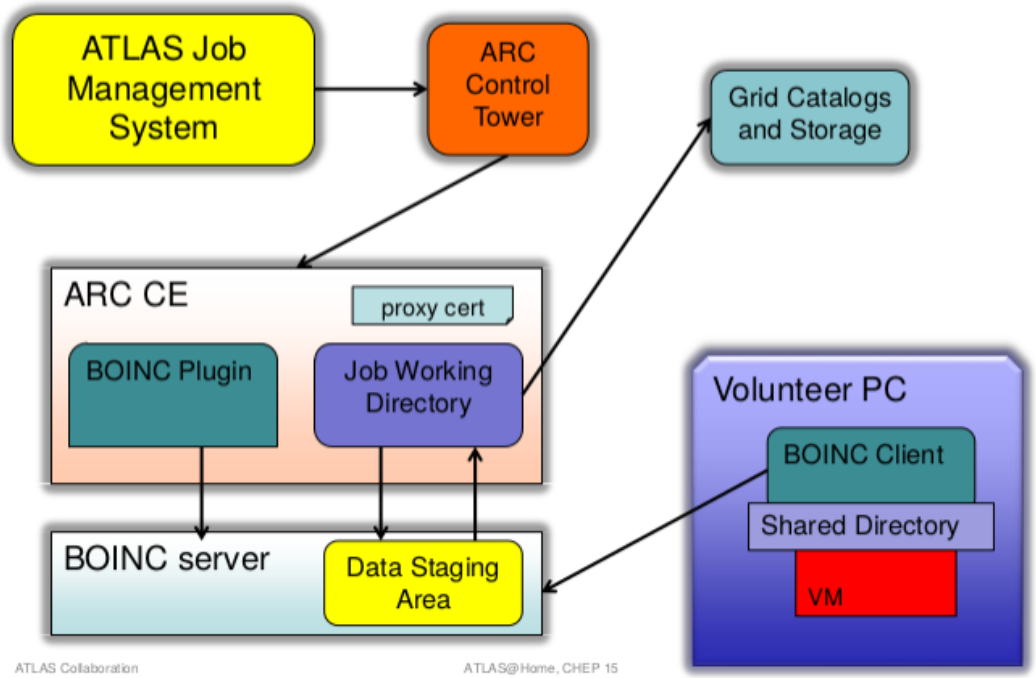
- Des sites indépendants de stockage de données qui autorisent un accès distant.
- Fédérés à travers un catalogue de données qui permet de diriger une demande d'accès vers un site ayant les données.

Fédération xrootd



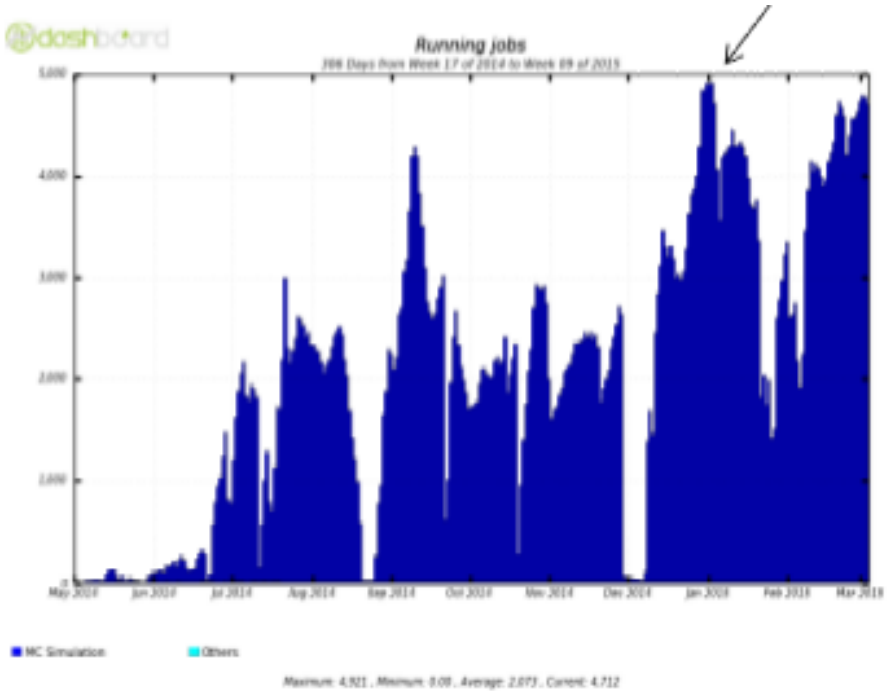
- Xrootd est un protocole purement HEP -> tout réimplémenter (catalogue, redirecteur, répartition de charge, mesures du réseau)
- Certains poussent pour des solutions de l'industrie (http)

Utilisation de @HOME dans ATLAS



D. Cameron CHEP 2015

Infrastructure dédiée pour interfacer le monde ATLAS et BOINC



Dans les derniers mois, ~5000 jobs en permance, ce qui rend @HOME équivalent à un T2

- Continuous 4000-5000 running jobs
- almost 900k completed jobs
- 1M CPU hours, 20M events
- 50% CPU efficiency

Utilisation de la ferme HLT* dans CMS

Using the HLT for Processing & Production

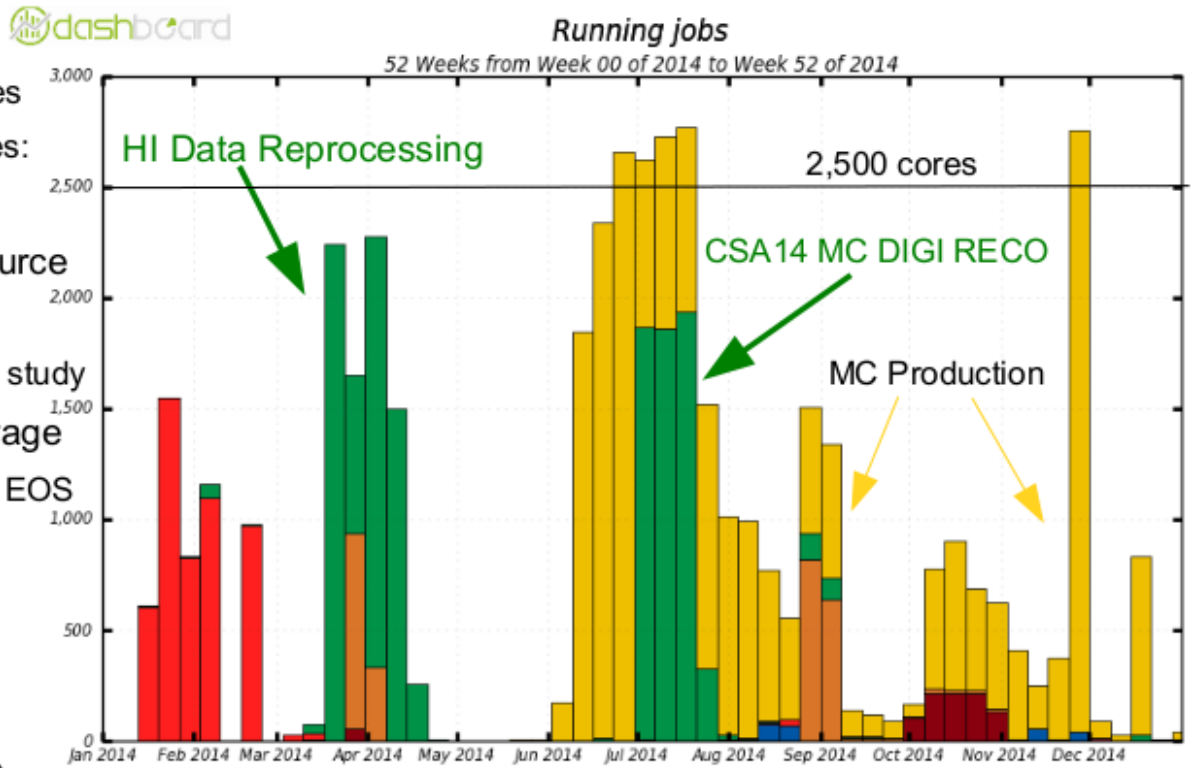
Sizeable resource:
- >200kHS06 / ~15k cores
- All Tier-1 WLCG pledges: ~300kHS06 (2015)

Openstack Cloud resource
- During Technical Stops
- Interfill mode still under study

No attached mass storage
- Reading/Writing from/to EOS

WLCG 2015 Strong network link
- 60Gbit/s
- 120Gbit/s in progress

Workflows
- MC production (little I/O)
- Data RECO ~50kB/s in/out per core [Can scale to full HLT] ✓
- MC DIGI-RECO with high PU adds ~50kB/s per unit of PU ✗



CHEP talk: D.Colling - The diverse use of clouds by CMS

*HLT = High Level Trigger

NB : aussi ALICE, ATLAS, LHCb (les + « agressifs »)

Les modèles pour le début du Run 2

- Il s'agit surtout de pousser plus loin les solutions développées à la fin du Run 1 et pendant le shutdown :
  - Généralisation du processing multicoeur pour contenir la mémoire/job
  - Généralisation de l'utilisation des fédérations de données, plus seulement comme solution de secours pour un job mais comme solution première pour 10-20 % des jobs (selon les expériences)
    - Monitoring des réseaux, transferts, sites encore plus important
  - Utilisation systématique des fermes HLT
  - Utiliser au maximum des ressources opportunistes : HPC, cloud, @HOME
    - Optimisation des frameworks de processing, de production, agilité des jobs face à des ressources volatiles.

Demandes des exp. à court terme

	2016	2017
ATLAS disk	+30 %	+30 %
ATLAS tape	+80 %	+80 %
ATLAS CPU	+20 %	+20 %
CMS disk	+30-40 %	+30-40 %
CMS tape		
CMS CPU	+30-40 %	+30-40 %
LHCb disk	+3- %	+30 %
LHCb tape	+70 %	+70 %
LHCb CPU	+30 %	+30 %
ALICE	???	???

⊕ 60 % unpledged

Semblent très difficiles à satisfaire dans le cadre d'un budget constant !

HNSciCloud

- Proposition issue de UE/CERN : externalisation de calcul vers des clouds commerciaux (majorité LHC mais pas uniquement)
- HNSciCloud est la partie « pilote » qui doit établir le cahier des charges, les aspects techniques, l'appel d'offre et l'évaluation du produit fini. À ce titre, il peut y a voir une grosse contribution financière de l'UE.
- L'IN2P3 participe à HNSciCloud :
  - Qu'attend-on de nous dans la période « production » ?
  - Veut-on utiliser cette opportunité pour développer en France les compétence sur l'interface WLCG – Cloud ?

Le côté américain : OSG

- La philosophie de OSG semble nettement plus agressive en terme d'accès distribué
  - Mise sur les réseau pour des accès performants à distance
  - Des projets pilotes pour des caches de données distribués
  - A dans son « portefeuille » pas mal de sites « de campus » d'universités qui ne sont pas dans les expériences et donc uniquement exploitables comme « batch system sans réel stockage dédié »
- OSG mise aussi beaucoup sur l'utilisation de ressources opportunistes :
  - HPC dans les labo nationaux (fortes incitations DOE/NSF)
  - Cloud, en particulier cloud commerciaux : accords avec Amazon, points d'accès dédié ESNet-Amazon

Évolutions vers la fin du Run 2 ou pour le Run 3

Attention : très spéculatif !

- Les expériences auraient plutôt tendances à encore plus gommer les différences a priori entre sites (T1 vs T2 par ex, sauf pour ce qui est du mass storage).
- Les expériences préfèrent moins de sites mais plus gros ou plus spécialisés
  - Très gros data centers, ressources de « bulk » processing locales ou commerciales.
  - Autres sites pour la simulation ou l'analyse des données réduites (à distance)
    - À la limite : « diskless » ou cloud
  - Pas toujours conciliable avec la diversité des sources de financement
- Ouverture à des modèles cloud, standardisation de machines virtuelles communes aux expériences (e.g. uCernvm+CVMFS)

Évolutions vers la fin du Run 2 ou pour le Run 3

Attention : très spéculatif !

- Exploration des solutions techniques de l'industrie des Content Delivery Networks (e.g. Netflix) :
  - caches de donné
- Réseaux dynamiques s'adaptant à la charge
  - Selon les transferts planifiés (volume et points de départ/arrivée, le réseau reconfigure sa bande passante et son routage
- Named Data Network : accès aux données par leur nom, le réseau s'adapte pour l'accès (caches)
  - Se développe dans l'industrie.

Évolutions WLCG / LCG-France : discussions

- Accord WLCG - Belle-II : ouvertures vers d'autres exp HEP ? Neutrinos ? Nucléaire ? Astroparticules ?
- Utilisation des ressources opportunistes : pour l'instant chaque expérience fait sa propre interface, y a-t-il moyen ou envie de rendre les choses plus communes ? Promouvoir en France les actions vers HPC ou cloud ?
- Externalisation des ressources, une réelle piste ?
- Évolution des sites LCG-France : Pourquoi casser ce qui marche ? moins ? Plus spécialisés ? Diskless ? Cloud ?
- Participation aux R&D computing ?