

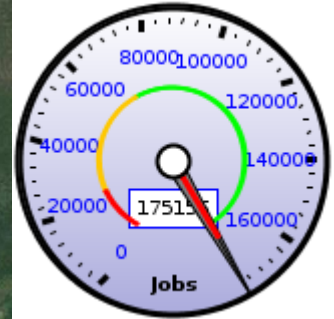
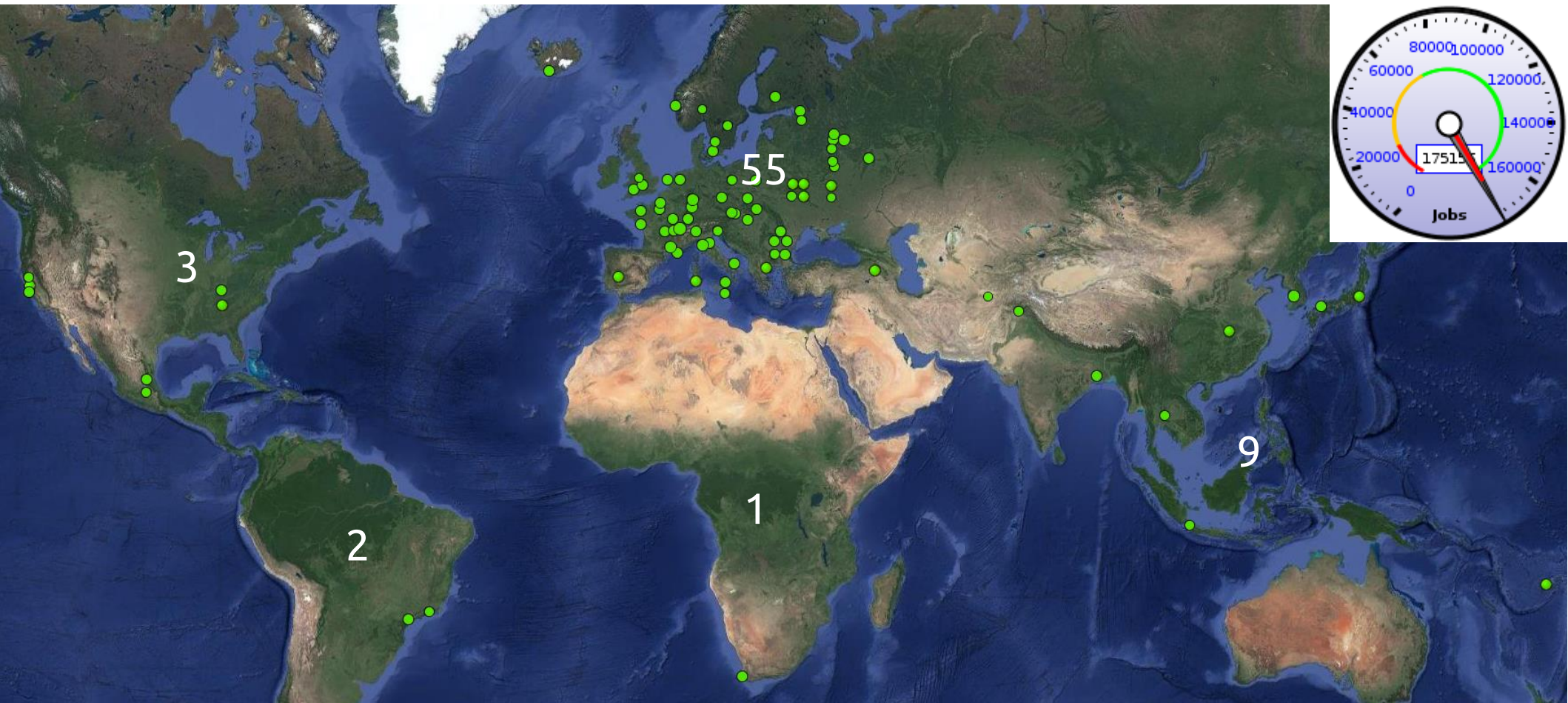
Résumé Workshop ALICE T1/T2

<https://indico.cern.ch/event/690294/>

CCIN2P3, Journées LCG France

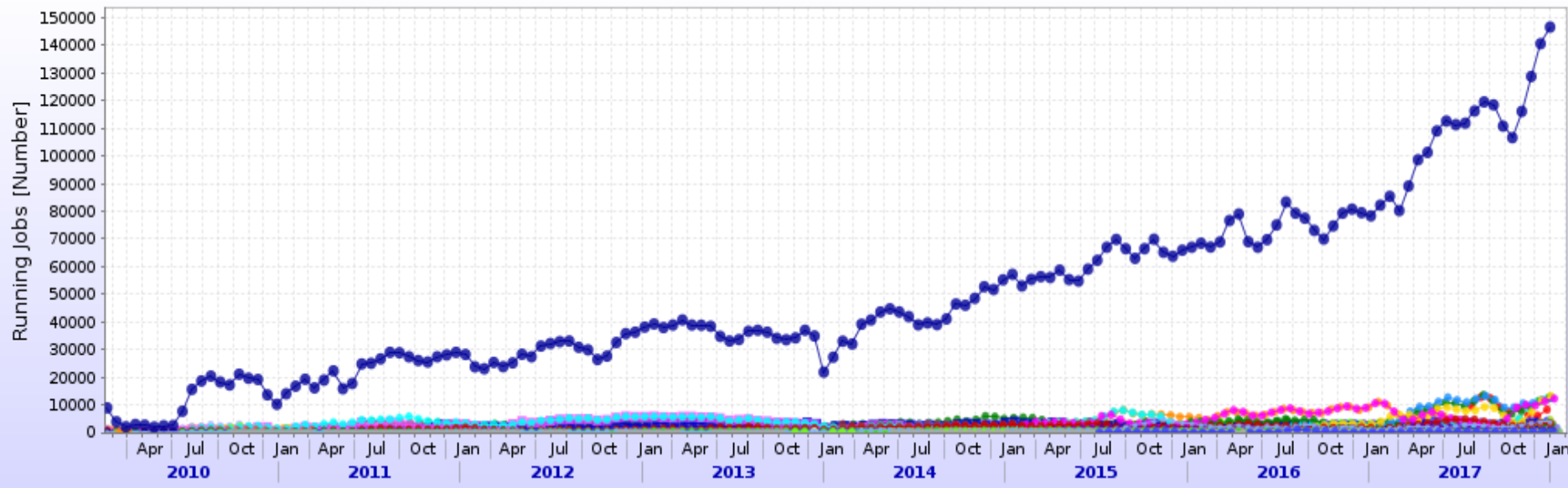
J-M. Barbet, C. Diarra, R. Vernet

La Grille ALICE aujourd'hui



Historiques ressources CPU

Running Jobs



Year on year
increase



+97%



+20%



+26%



+17%



+43%

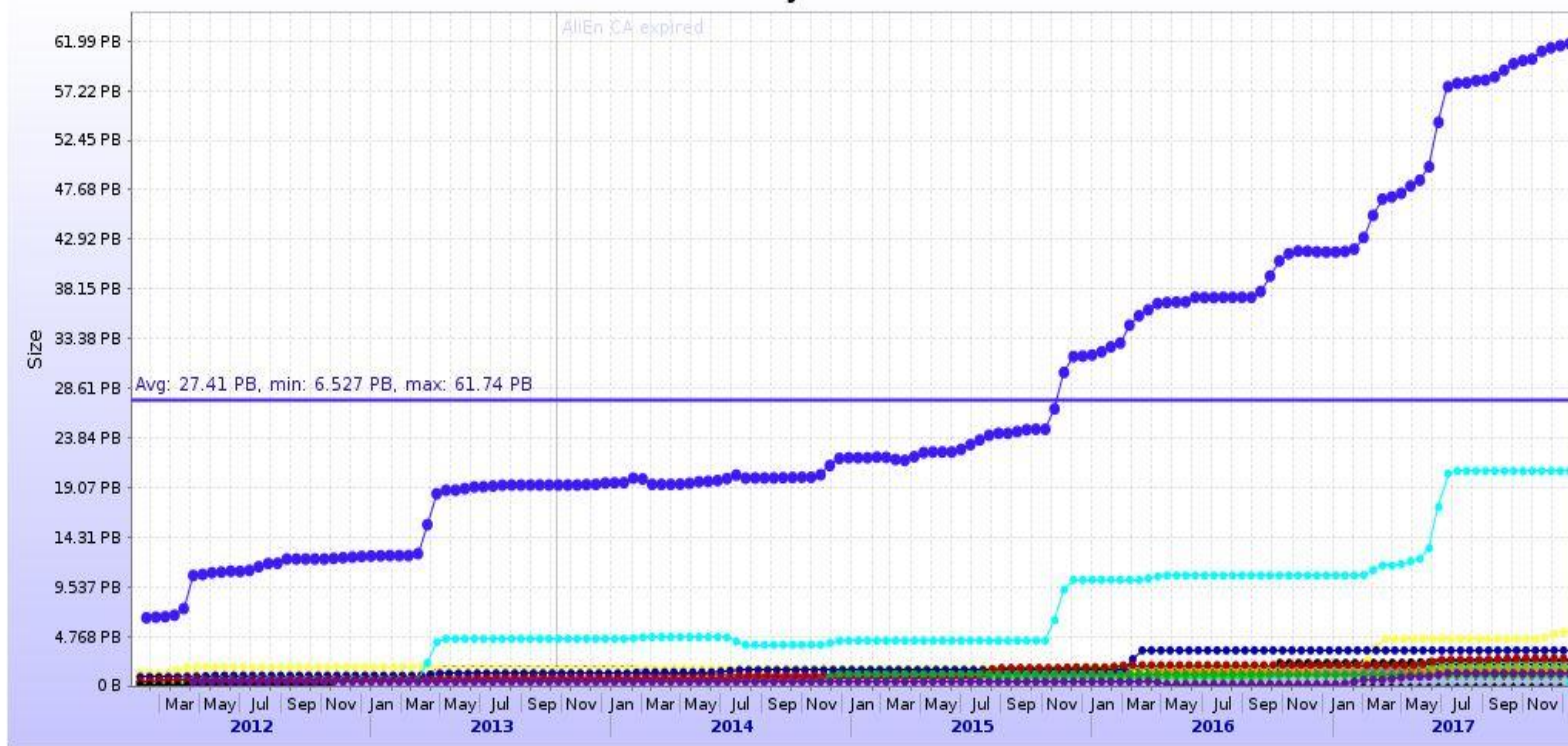


+21%



+46%

Historique ressources stockage



**Year on year
increase**

↑
+36%

↑
+3%

↑
+35%

↑
+25%

↑
+32%

Croissance et utilisation des ressources

- Bonne croissance d'année en année
- Globalement T2 moindre qu'attendue

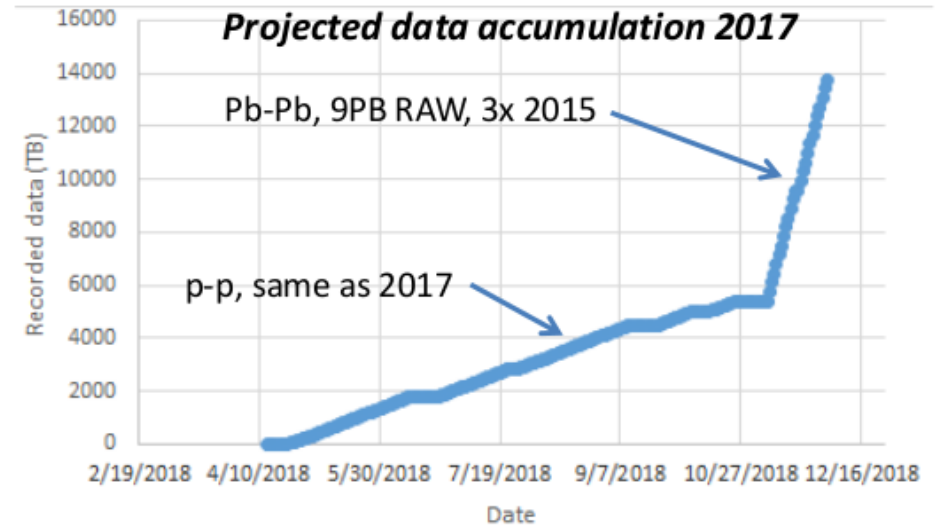
- Read/Write = 10:1
- Accès WAN ~ 3 % (lecture)
- Lecture totale de la capacité disque en 1.5 mois
 - Constant
- ~ 1 milliard de fichiers repartis sur la grille

Années à venir

- Dernier Run en 2018 avant upgrade
- Déficit de ressources déjà présent
 - Bataille avec CRSG
- Redémarrage en 2021 (2022 ?)
- Shutdown 2019-2020
 - Croissances nécessaires

Besoins, en gros

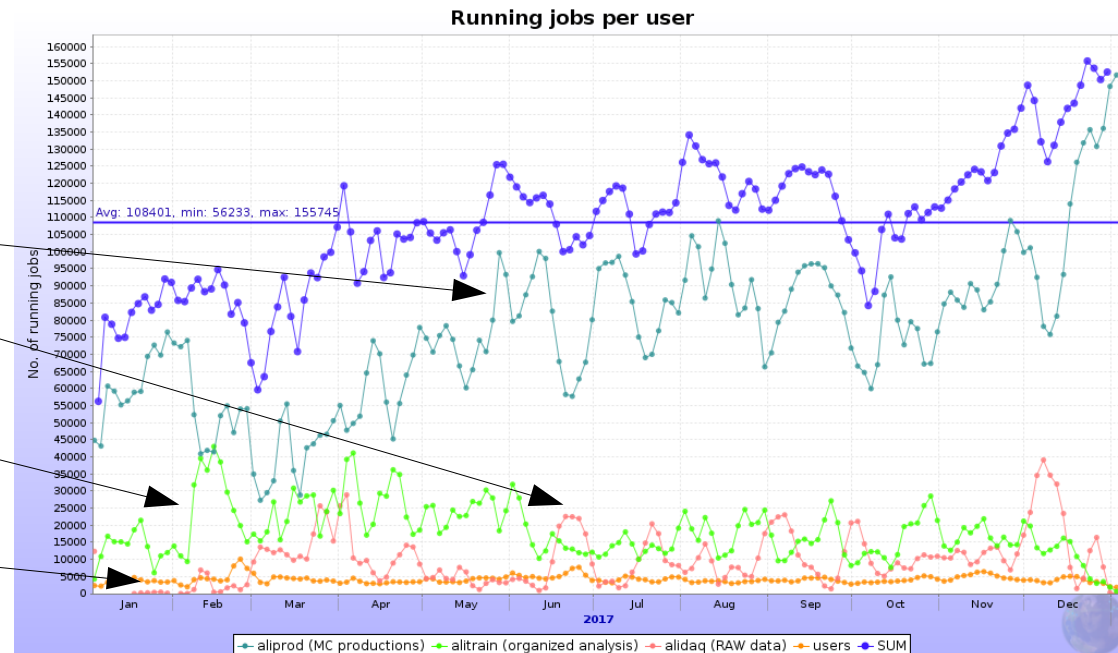
	CPU	Disk	Tape
2019	+20 %	+20 %	-
2020	-	+20 %	-



Opérations 2017

- Backlog données 2015+2016
 - Absorbé → plus de backlog !
- 2017 :
 - 4 PB raw + 4 PB répliqués aux T1
- De plus en plus d'analyse organisée

Monte Carlo : 70 %
Raw processing : 10 %
Analyse organisée : 17 %
Analyse individuelle : 3 %

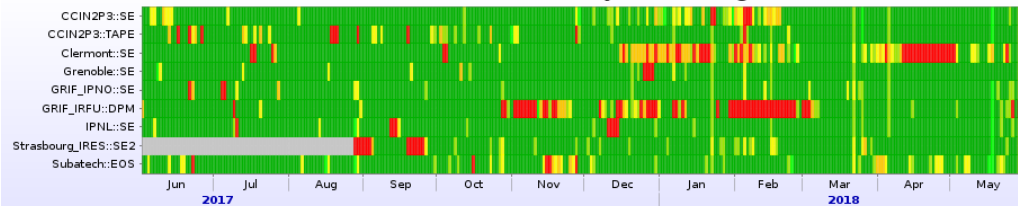


Opérations 2017 (stockage)

- Disponibilité demandée > 95 %
 - Beaucoup de sites en-deçà
- France : pas si mal

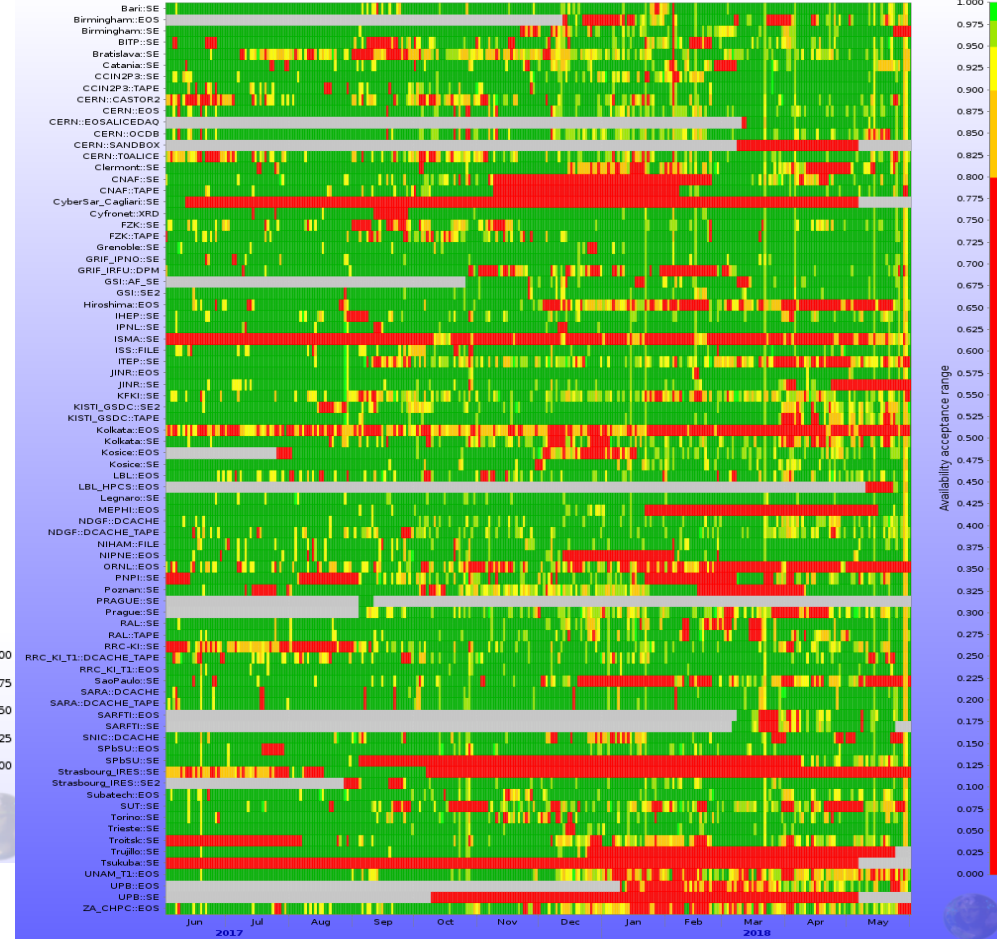
En France

AliEn SEs availability for reading



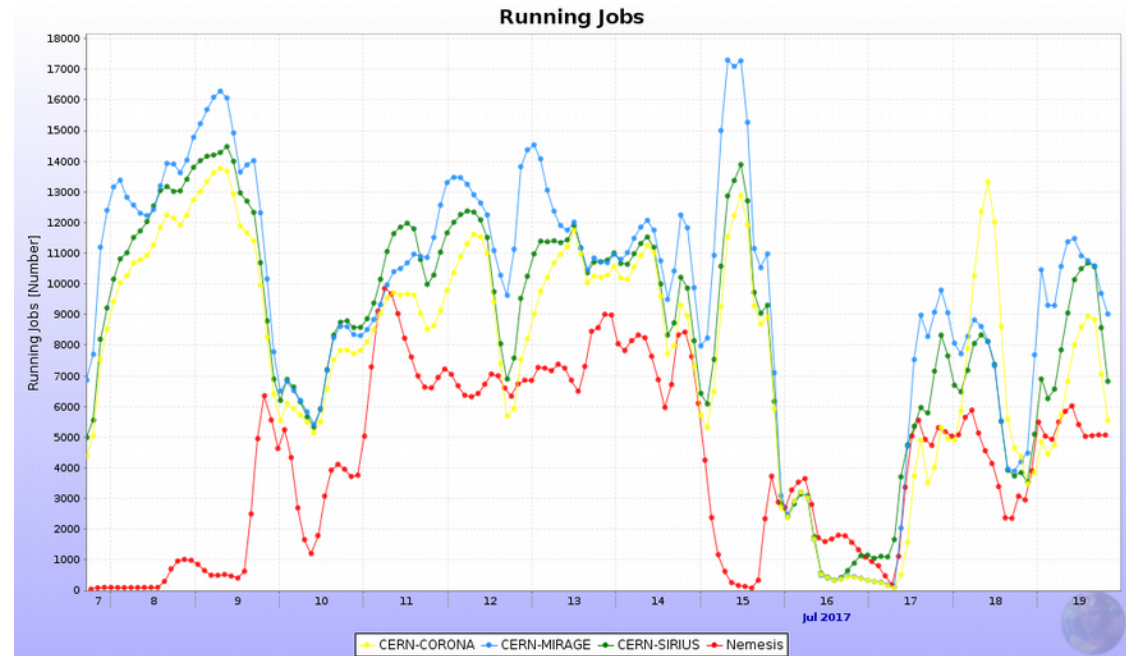
<http://alimonitor.cern.ch/status/index.jsp>

AliEn SEs availability for reading



Virtual sites

- Containered vobox (production)
- HNSciCloud
 - Sites virtuels existent, peu utilisés pour le moment
- Oracle Bare Metal Cloud
 - 9k jobs run



http://alimonitor.cern.ch/display?page=jobs_per_site

Plan IPv6 stockage

- T1 (doivent être) en IPv6 depuis avril
 - CCIN2P3 OK
- T2 devront l'être début 2019
- ~ Tous les SE devront être en dual stack avant l'utilisation de IPv6-only

- EL7 : OK
 - On peut mettre à jour les WN
- SAM tests sur SE impactent la disponibilité/fiabilité
- Upgrade vers xrootd \geq 4.1 souhaité
 - Plugins ALICE ici : <http://linuxsoft.cern.ch/wlcg/>
- Problèmes des sites
 - Suivi ici : <http://alimonitor.cern.ch/siteinfo/issues.jsp>
- Notifications
 - Inscription ici : <http://alimonitor.cern.ch/xml.jsp>

- CERN : stockage des 4 expériences + CERNbox
 - 16 instances, 1400 nœuds, 50k disques, 250 PB
- IPv6 avec Citrine, basé sur xrootd v4
- Amélioration du support
 - Communautaire : <https://eos-community.web.cern.ch/>
 - eos-support@cern.ch
- Scalabilité Citrine
 - >2 managers, DB répartie (QuarkDB)
- Projets :
 - Unboxed setup = EOS + Cernbox + CVMFS + Swan → conteneurs Docker
 - Stockage repartit distant : Wigner, Eulake, AARNet
 - Geotagging ... datalake
 - Cache dynamique (→ petits sites)
- Dans ALICE : 30 % des sites, 50 % des données (en augmentation)

- Réécriture de AliEn en Java
- Sera déployé un 1 fichier .jar
 - JRE > v8
- Concerne :
 - JobAgent
 - Services AliEn des sites (AliEn CE, ...) et services centraux
 - Catalogue de fichiers
- Plan de déploiement
 - AliEn et jAliEn en parallèle aux services centraux (mêmes machines)
 - Remplacement des API servers (centraux) sans modifier les clients
 - Modification du JobAgent → connexion jAliEn
 - MachineJobFatures → exploitation tout ou partie du nœud par la payload
 - Possible utilisation de conteneurs
 - VOBOX sous jAliEn

Reseau :
8097/tcp et 8098/tcp

Data transfer & integrity

- Procédure pour décommissionner un SE
 - ou en cas de perte de données
- Monitoring via <http://alimonitor.cern.ch/transfers/>
- Vue des SE : <http://alimonitor.cern.ch/stats?page=SE/table>
 - Catalogue vs reporting xrootd
 - → fichiers orphelins
 - → dark data
- Outils en place centralement
- Prendre contact si besoin

Retours sites (highlights)

- Birmingham T2
 - 1000 cores
 - 1.4 PB
- Migration WN vers VAC
 - Peu de services à gérer par le site (CE, batch, ~~vobox~~)
- Jobs tournent sur site virtuel Altaria (CERN)
- Close SE = CERN - > problèmes de perf I/O

Retours sites (highlights)

- GSI Analysis Facility
 - Prototype pour Run3
 - 1000 cores, 600 TB
- Plugin Xrootd pour accès Lustre
 - Mapping PFN-LFN

Retours sites (highlights)

- NERSC (T2)
- Démonstrateur pour jobs ALICE sur HPC (Cori)
- Analisa : outil python pour lancer plusieurs jobs séquentiels comme jobs MPI
- I/O par xrootd (proxy cache)
- Pas de SE local → pre-staging pourrait aider (?)
- ATLAS & CMS utilisent déjà NERSC pour simulation

Conclusions

- Expérience au modèle simple et qui fonctionne efficacement
- Disque = partie critique
 - Manque général de capacité
 - Besoin d'une meilleure disponibilité
- Plusieurs évolutions techniques récentes et prévues
 - Sites virtuels, GSI AF, HPC/Cori, jAliEn
- Liens forts avec EOS
 - Sites ALICE invités à tester EOS, bénéfices attendus des deux côtés
- La France est bien intégrée et visible dans la grille ALICE