

IN2P3-SUBATECH : Tier2 Alice

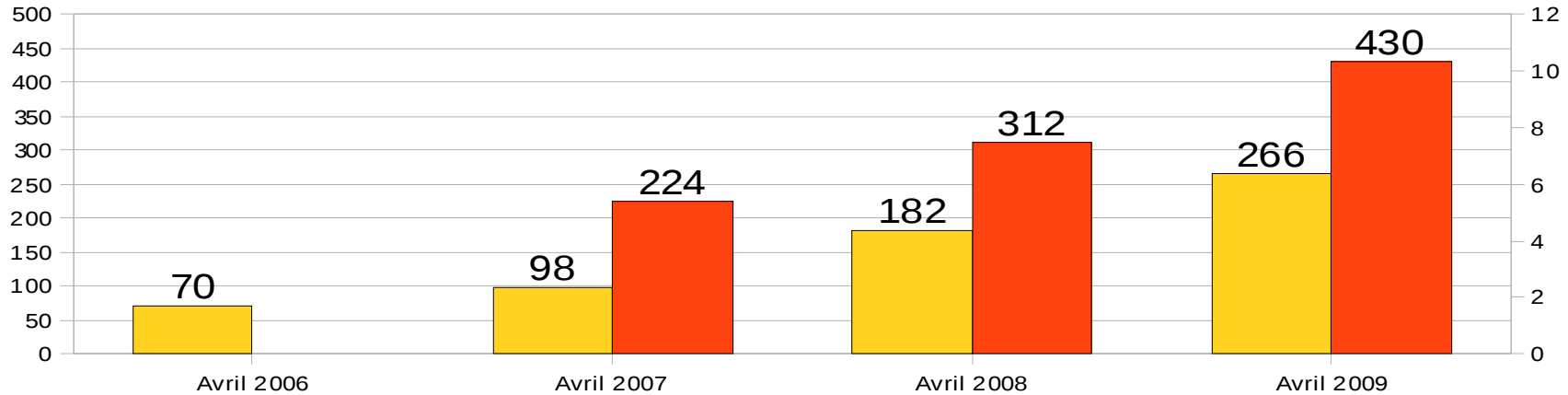
- Le Projet “Tier2” Subatech
- Situation actuelle
- Choix techniques
- Quelques résultats
- Experiences Xrootd natif et CREAM-CE
- Questions à l'étude
- Conclusion

Le Projet “Tier2” Subatech

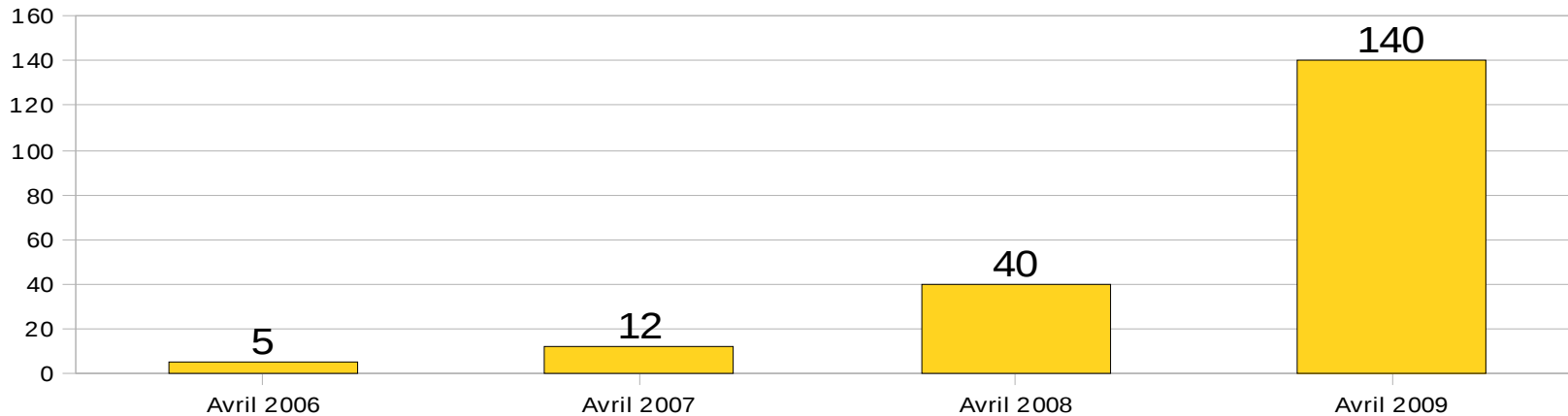
- Une seule VO LHC : Alice
- Dates clés
 - Avant LCG : Participation au PDC04 Alice avec AliEn
 - Site IN2P3 Subatech Certifié le 20 Janvier 2006
 - Premiers jobs Alice (sur gLite) : 23 Mai 2006
 - Stockage DPM-xrootd utilisé en production : 26 Mars 2008
- Financement : fonds propres et région Pays-Loire
- Utilisation locale à hauteur de 30% max du CPU
- Ressources Humaines : 1.4 ETP(*)
 - 1 CR CNRS : 0.1 ETP
 - 4 ITA CNRS : 1.3 ETP

(*) ETP = Equivalent Temps Plein

Ressources



■ Nombre de CPUs
■ Pledged KSI2K



■ Stockage en To

Notes : Nombre de CPU indiqué = 70% du total (30% usage local)
 KSI2K : Valeurs WLCG

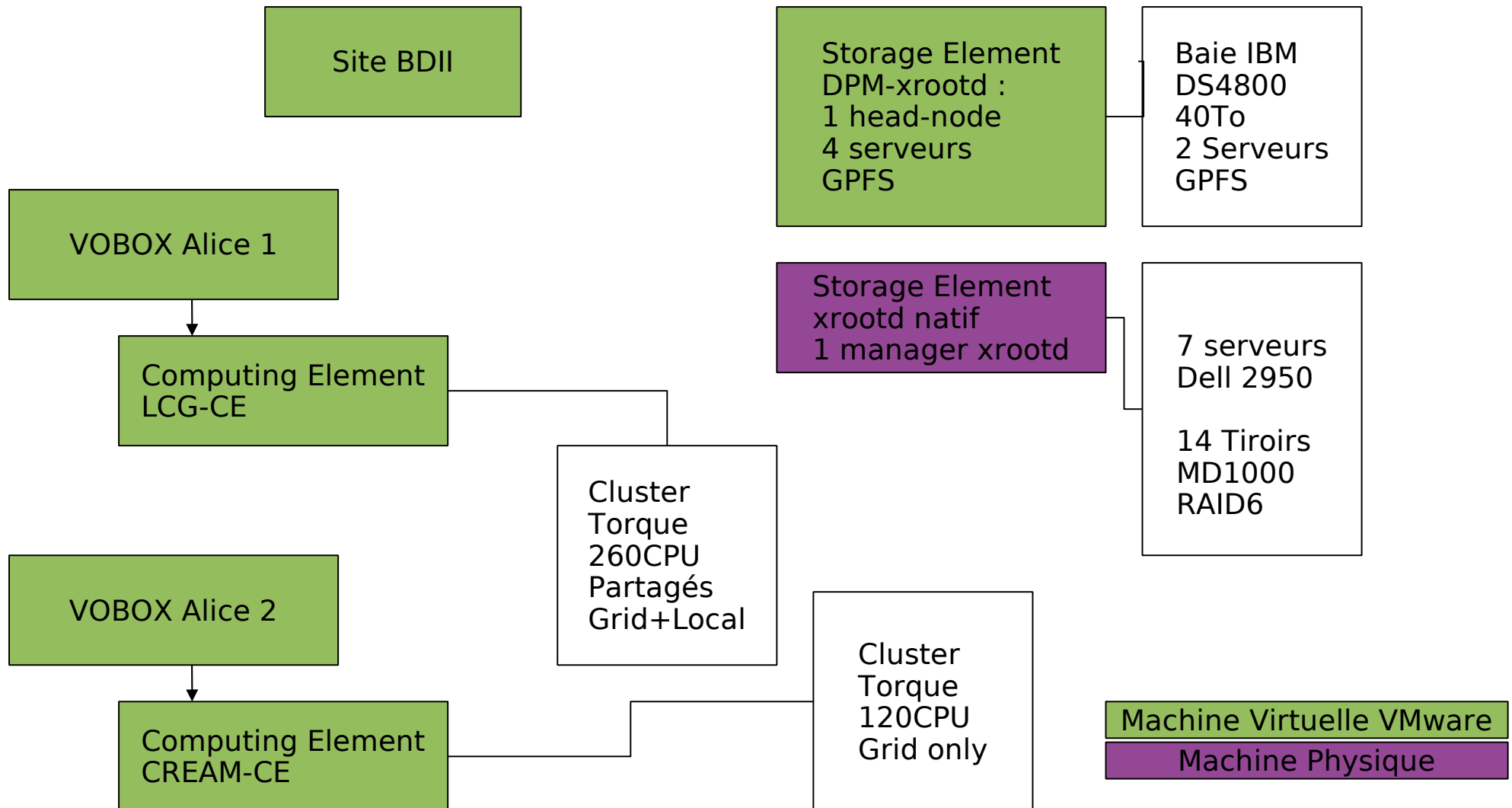
Choix techniques

- Intégration du site Grille avec les autres services opérés par le Services Informatique
 - Même Infrastructure, sécurité (firewall)
 - Intégration au monitoring Nagios Local et développement de tests spécifiques
- Utilisation de machines virtuelles VMWare
 - Snapshots, clones, déplacement à chaud
- Distribution “relogeable” pour le middleware
- Quattor
 - Le système d'exploitation des workers, pas le middleware
 - Intégration progressive des autres noeuds : Site BDII, xrootd

Infrastructures

- Salle 60-80m² avec faux-plancher
- 2 unités de climatisation
- 2 lignes électriques avec onduleur
 - Double alimentation sur la majorité des équipements
- Contrôle environnement
 - Boitier APC : détection température (3 sondes), fuite liquide,...
 - Transmetteur téléphonique

Situation Actuelle : Services





Quelques Résultats

- Reliability/Availability EGEE :
 - 2008 : Reliability : 94.7 Availability : 94.7
 - 2009 : Reliability : 98.0 Availability : 98.0

FI-HIP-T2 (Finland. NDGF/HIP Tier2)



Tier-2 Availability and Reliability Report

March 2009

Critical SAM Tests - <http://sam-docs.web.cern.ch/sam-docs/docs/html/docs/MANUserManual/node22.html>

Availability = Uptime / (Total_time - Time_status_was_unknown)
 Reliability = Uptime / (Total_time - Scheduled_downtime - Time_status_was_unknown)
 Reliability and Availability for Federation - Weighted average of all sites in the Federation (based on CPU count)

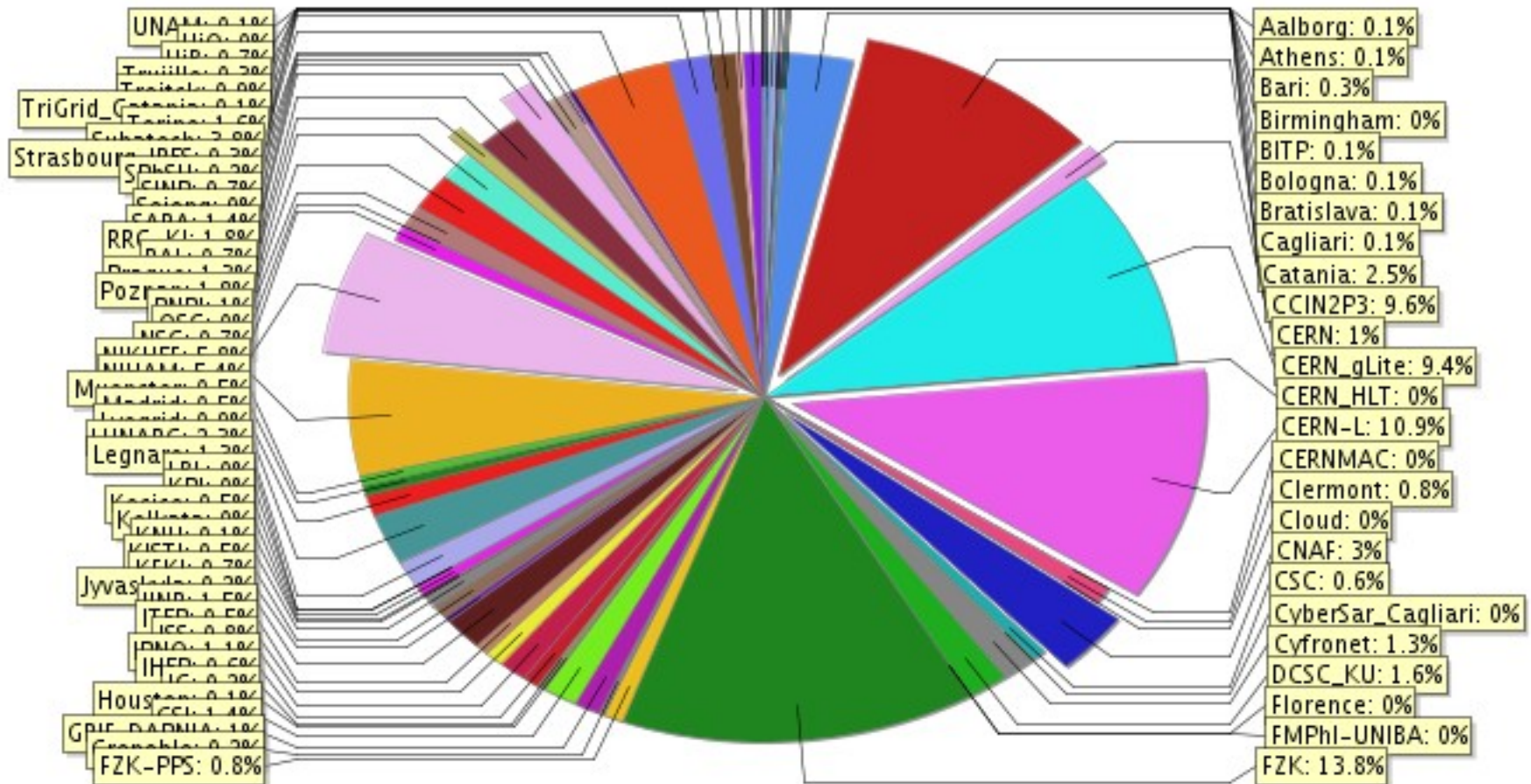
Colour coding : N/A < 30% < 60% < 90% >= 90%

Federation	Site	CPU Count	Reli-ability	Avail-ability	Reliability History		
					Dec-08	Jan-09	Feb-09
FR-IN2P3-SUBATECH	IN2P3-SUBATECH	276	99 %	99 %	94 %	99 %	96 %

Quelques Résultats

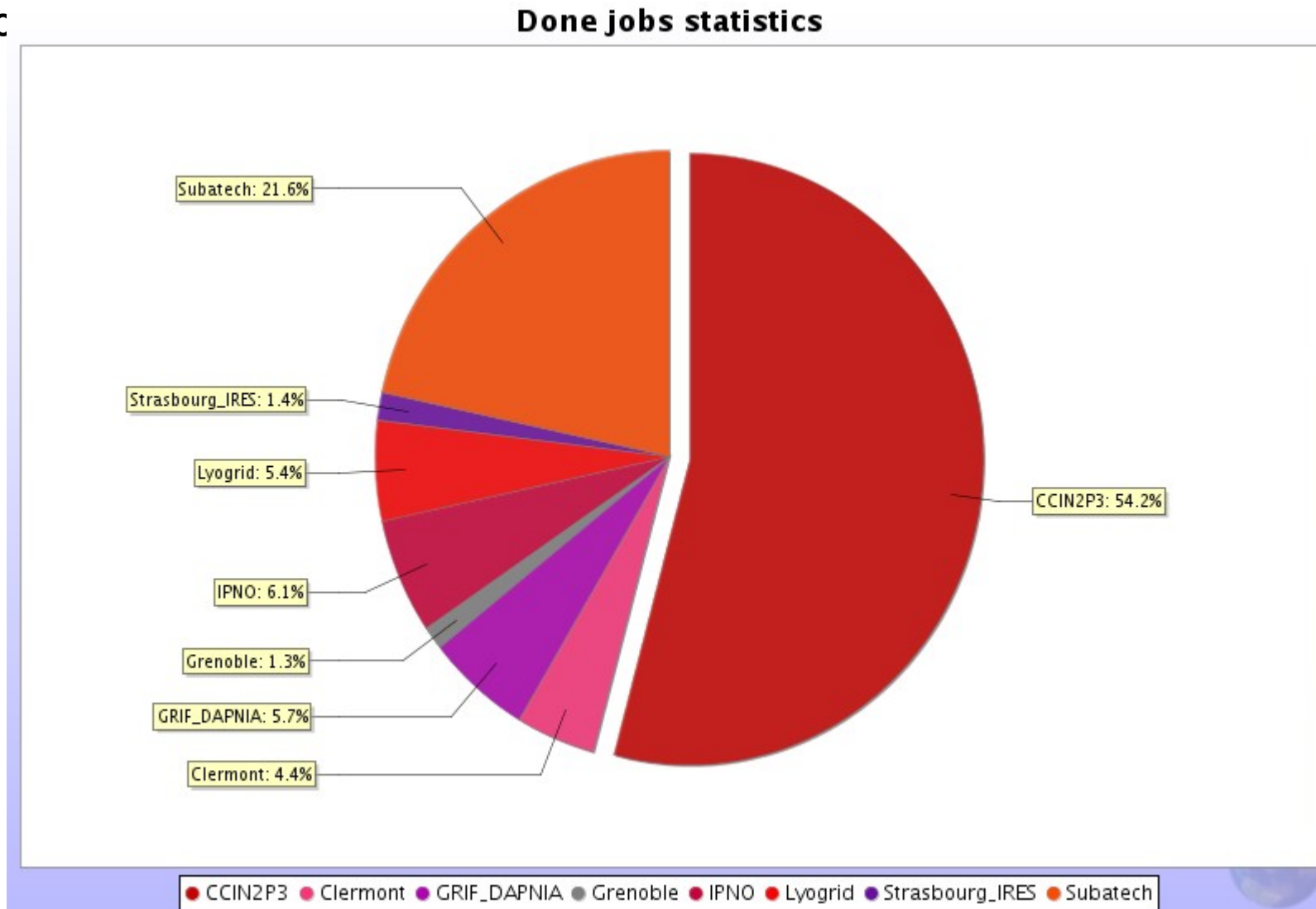
- Part “Done Jobs” Subatech dans Alice en 2008 : 4%

Done jobs statistics



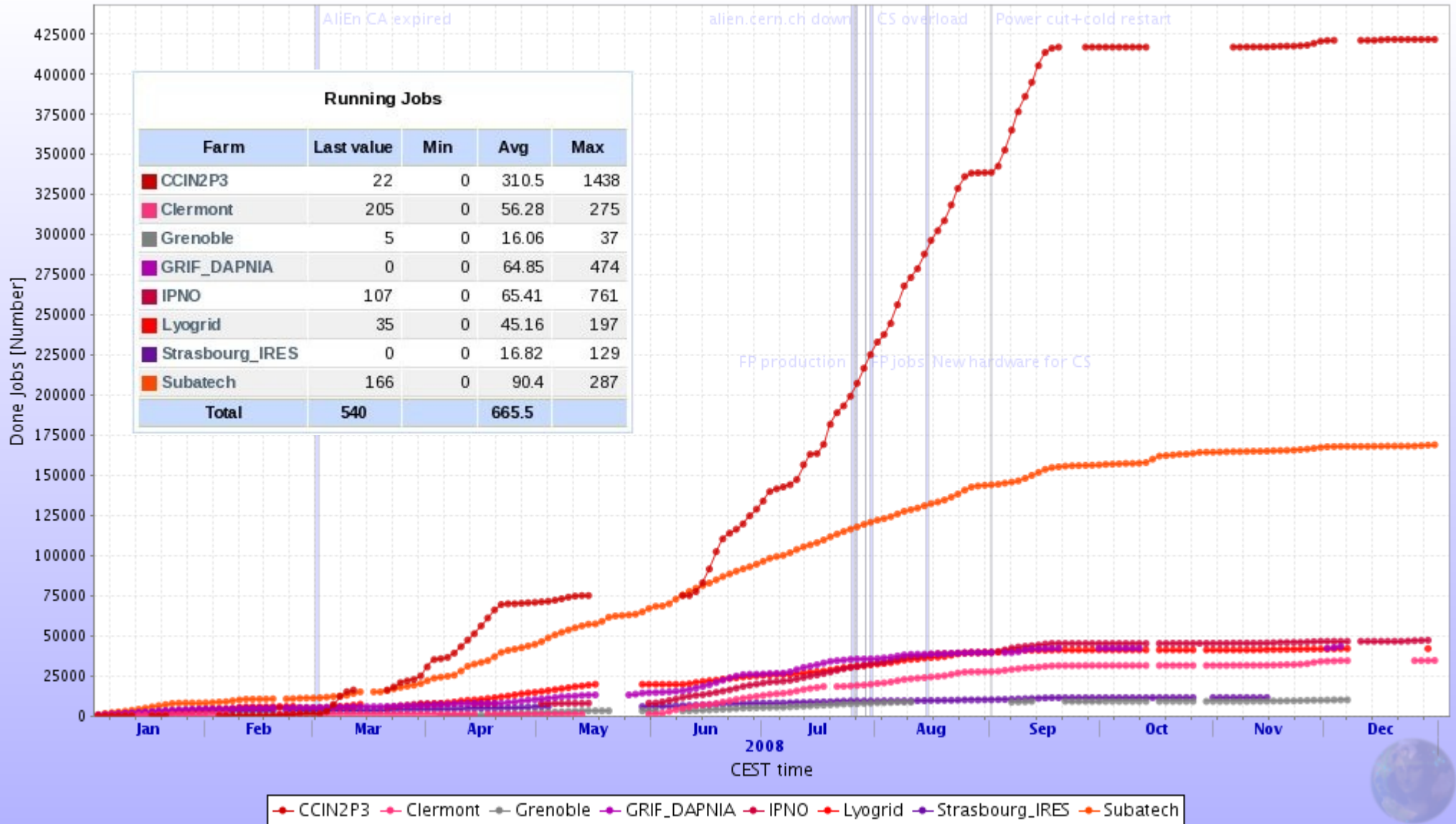
Quelques Résultats

- “Done Jobs” Subatech en 2008 dans LCG-FR : 21.6 %



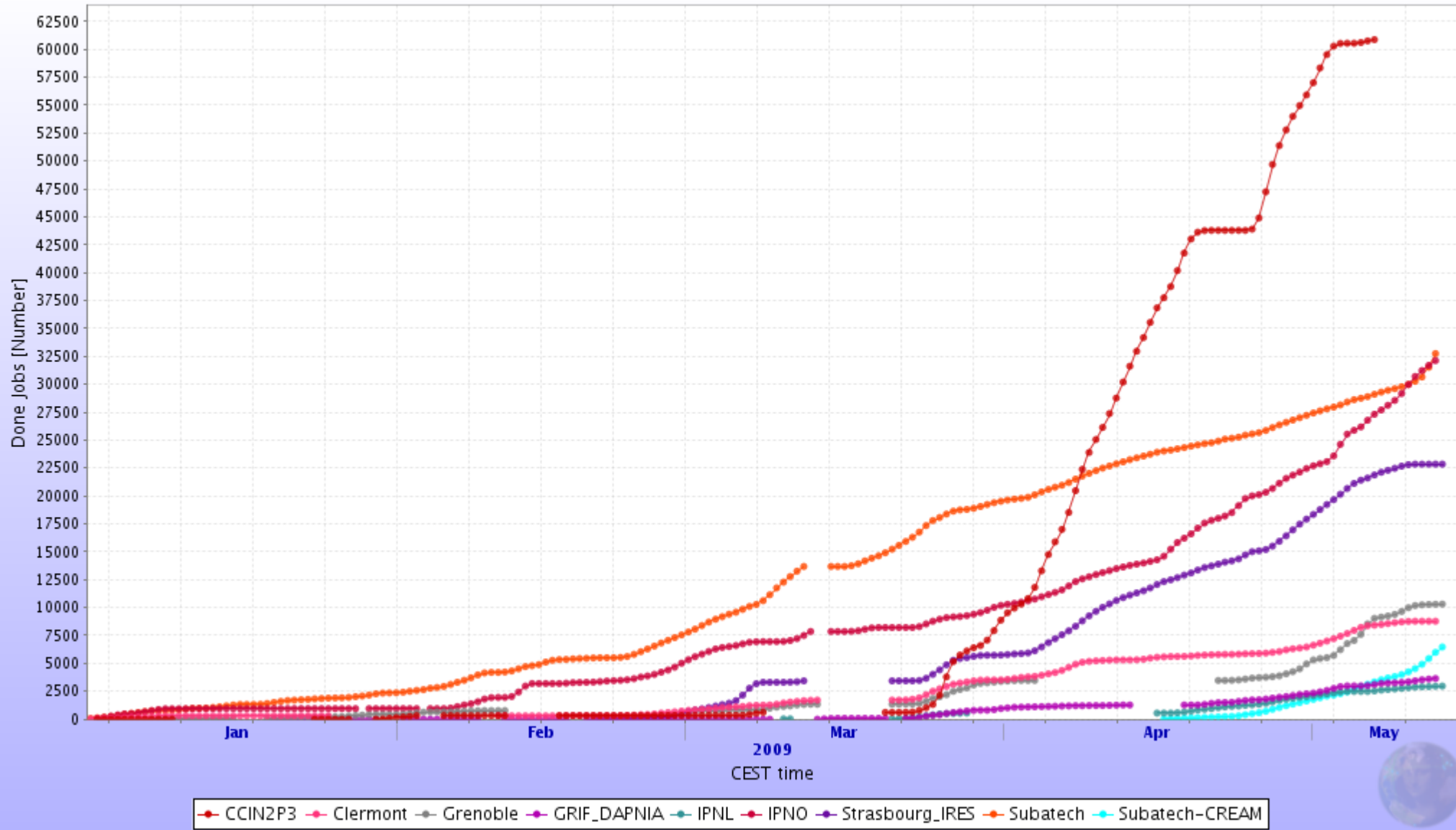
Sites Français 2008

Done Jobs



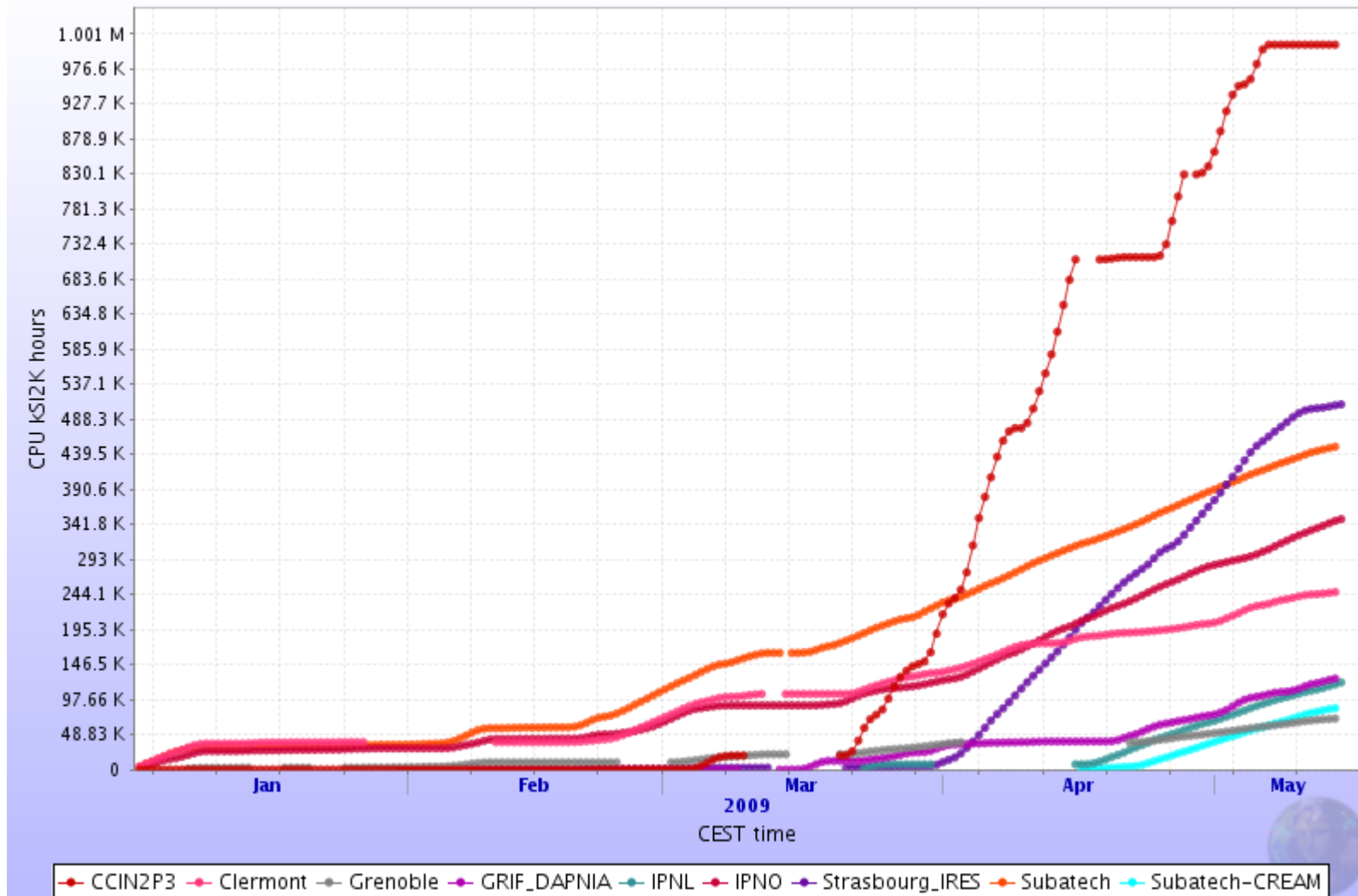
Sites Français 2009

Done Jobs



Sites Français 2009 en KSI2K

Total CPU kSI2K hours for ALICE jobs

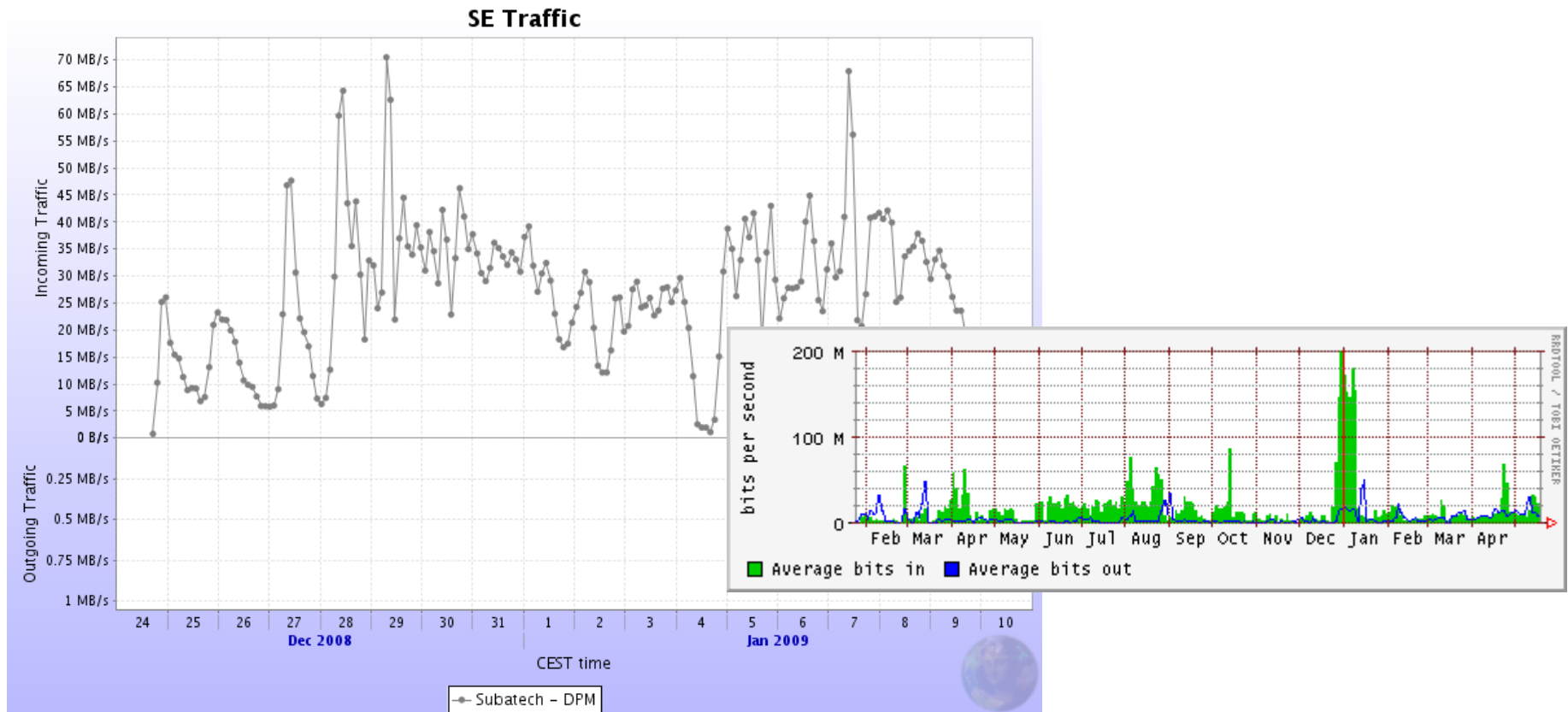


DPM-xrootd

- Disponible Novembre 2006
- 1ère installation : Avril 2007, Prêt en Mars 2008
- Participation à la rédaction de la documentation
- Utilisé en production en Juin 2008
 - Mode Cible pour la production MC : OK
 - Mode Analyse : forte charge MySQL constatée lors de travaux de “merging” impliquant de nombreux accès.
 - De nombreux accès font gonfler les tables qui conservent l'historique des requêtes (nettoyage nécessaire)

• DPM-xrootd

- 90% plein suite production en décembre 2008

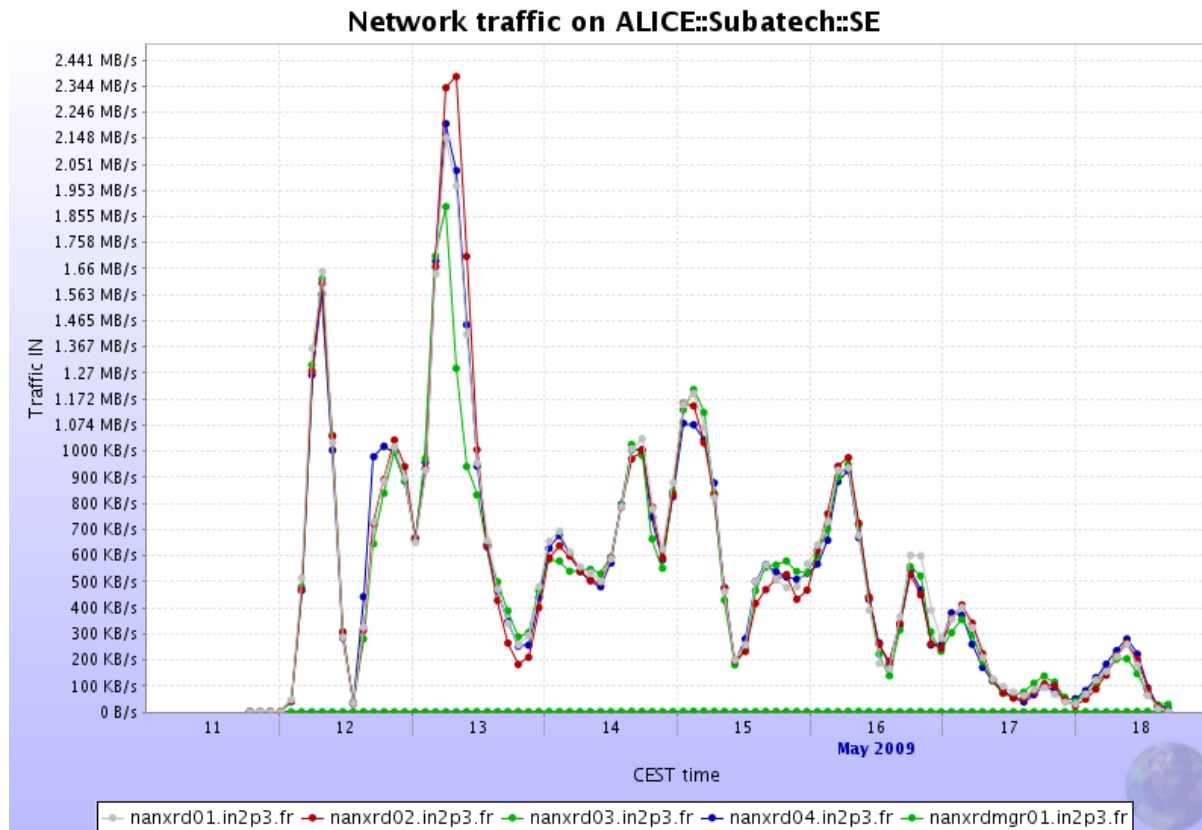


xrootd natif

- Premier cluster xrootd natif de test : Août 2008
- Test des versions successives et du packaging RPM : Septembre 2008
 - En collaboration avec F.Furano
 - Test débit : 400Mbits/s avec un seul server/manager
- Installation du stockage actuel (~100To) :
 - 1 manager et 4 serveurs installés via Quattor
 - Serveurs: Dell 2950 + 2 tiroirs MD1000 soit 24To RAID6

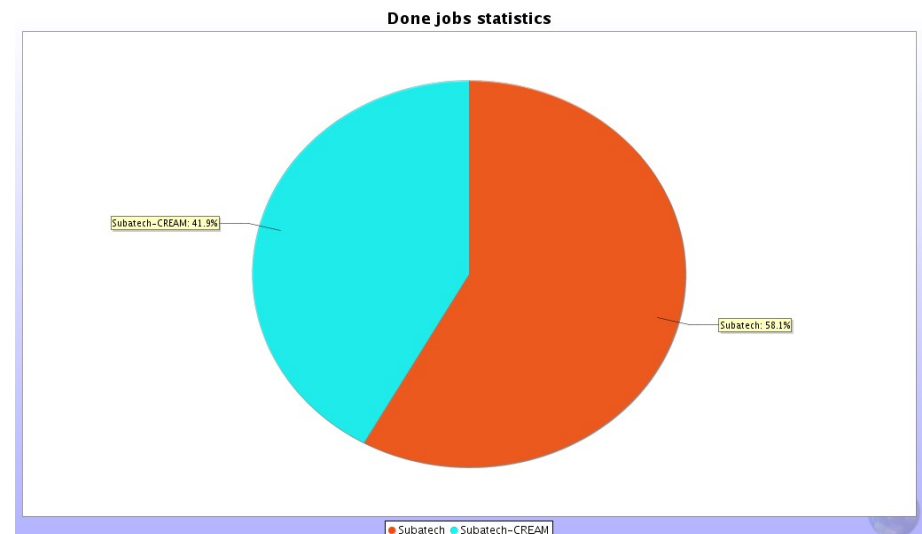
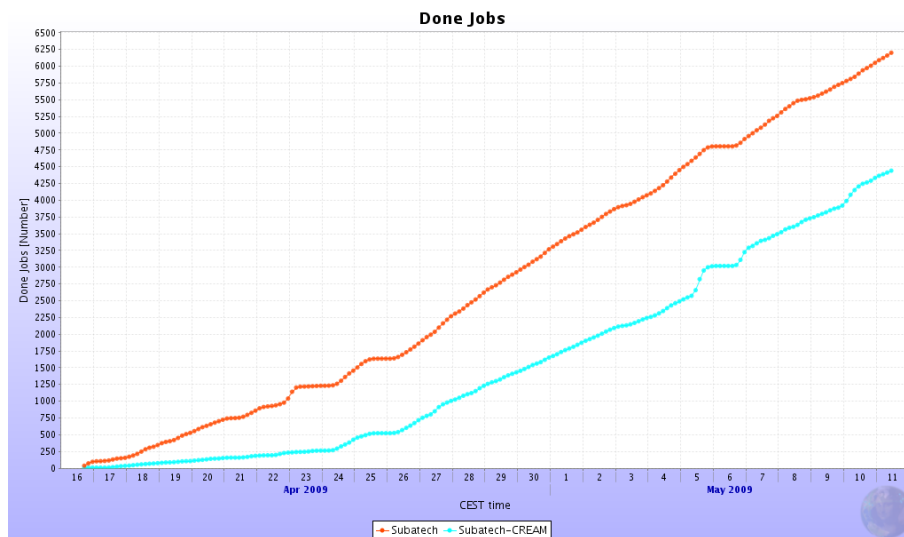
Xrootd natif

- En production depuis le 11 Mai 2009
 - Jusqu'à présent (14/05) le trafic réseau maximum a atteint 12Mo/s pour l'ensemble et 2.5Mo/s par serveur.



Experience CREAM-CE

- Premiers tests Nov 2008
- Nécessite une 2nde VOBOX
 - crée par clonage le 1er Avril 2009
- Cluster distinct constitué de 120 cores
- En production depuis le 16/04



Questions à l'étude

- Le partage d'un cluster avec les locaux
 - Permet aux utilisateurs locaux de profiter des périodes de creux de la production. Fair-share Maui réglé à 70%/30%
 - Mais il faut s'accorder sur l'OS et l'architecture
 - Des problèmes récents ont permis de prendre conscience des risques pour la sécurité de ce choix
- L'interprétation du “Job efficiency”
 - Enquête en Juin 2008 pour tenter d'expliquer certaines différences systématiques entre les sites au niveau de la métrique “Job Efficiency”.
 - Ces différences se sont estompées avec la production actuelle
 - Néanmoins, la surveillance et la compréhension de cette métrique sont indispensables

Questions à l'étude

- Comment comprendre ce qui se passe ?
 - Le “site manager” qui suit ce qui se passe sur son site doit pouvoir interpréter les incidents de production afin de savoir si c'est le site qui est en cause
 - Il nous faut répertorier les différents tableaux de bord et sites de monitoring permettant d'analyser la situation

Rappel :

T1/T2 ALICE Tutorial: Services Management
before the real data taking
CERN 26 et 27 Mai 2009

Boite à Outils pour un site Alice

- Documentation :
<http://alien.cern.ch/twiki/bin/view/AliEn/Home>
- Monitoring MonaLisa :
<http://pcalimonitor.cern.ch>
- Souscription aux alertes MonaLisa
- Alice LCG Task Force Meeting
<http://indico.cern.ch/categoryDisplay.py?categId=31151>
- WLCG Daily Operations Meeting
<https://twiki.cern.ch/twiki/bin/view/LCG/WLCGOperationsMeetings>
- Alice Dashboard :
<http://dashboard.cern.ch/alice/index.html>

Conclusion

- Se concentrer sur une seule VO LHC a permis de s'investir et de chercher à offrir le meilleur service possible,
- Les autres groupes de recherche du laboratoire ont bénéficié de l'augmentation des ressources de calcul
- La participation au projet LCG a demandé un effort supplémentaire très significatif au Service Informatique mais a permis aux ASRs de découvrir de nouvelles technologies et de progresser.