Évolutions à venir du modèle de calcul de CMS (2013)

Réunion T1 & AF, 14 mars 2013

PLAN

- Séparation disque et « tape » (T1)
- Worker Nodes sur LHCOPN (T1 & T2)
- Fédération (T2)

Séparation disque et « tape »

- Meilleure utilisation du stockage de masse
  - Plus de sauvegarde automatique des données à leur arrivée : seul celles validées sont mises en « tape »
  - Minimisation des besoins de « staging » : les données restent sur disque tant qu'elles sont utilisées (et jamais en cours de workflow)
  - Par conséquence, moins d'effacement (donc moins de compactification) nécessaire

- Accès plus souple aux données
  - Les données utiles restent sur disque tant qu'elles sont utilisées
  - Mise en place de fall-backs vers les autres sites (T1) en cas de données manquantes (plus rapide et efficace que le re-transfert)

Fournir à CMS deux points d'entrées vers le stockage :

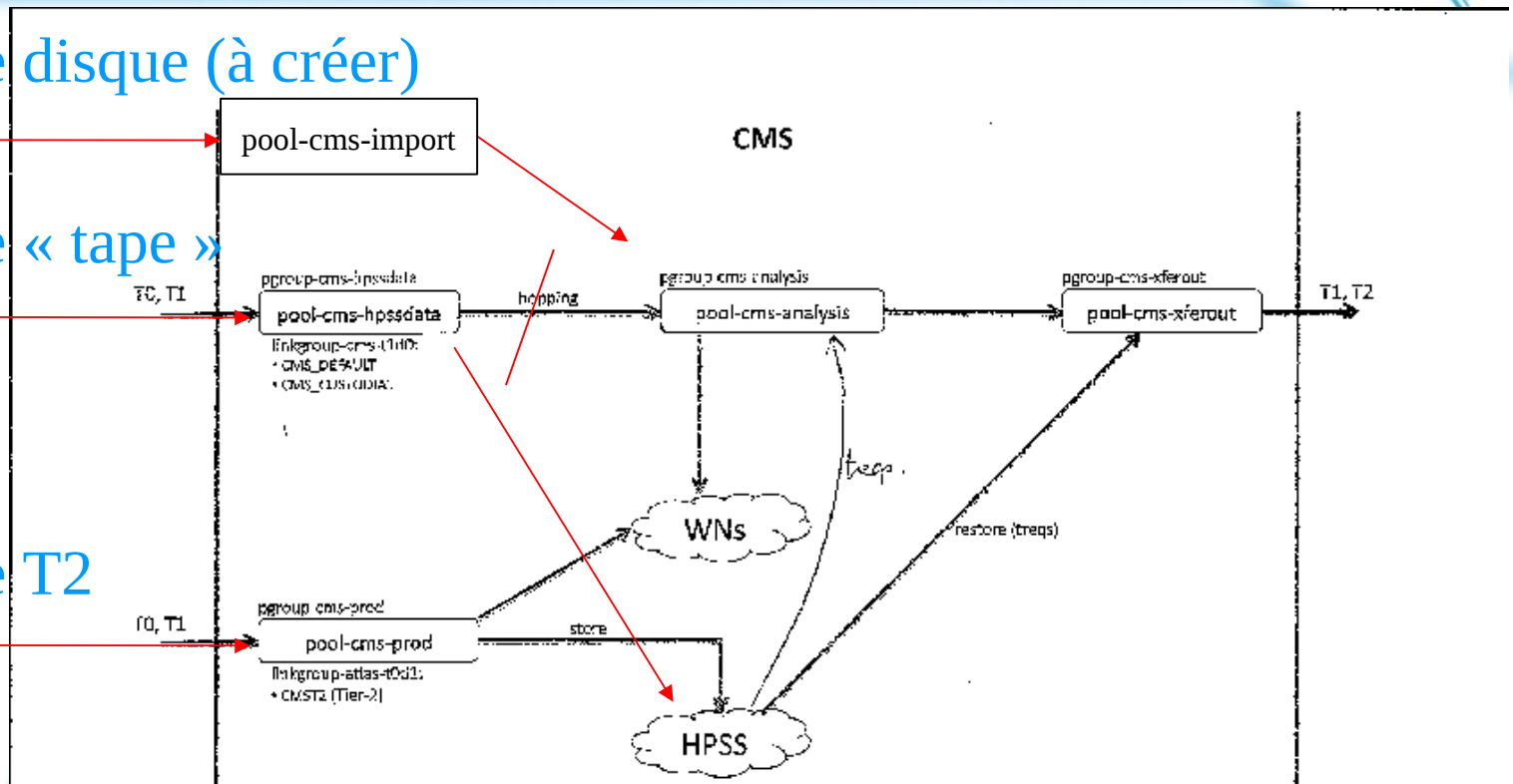
- Disque (t0d1)
- Tape (t1d0)

En pratique, au niveau dCache

Entrée disque (à créer)

Entrée « tape »

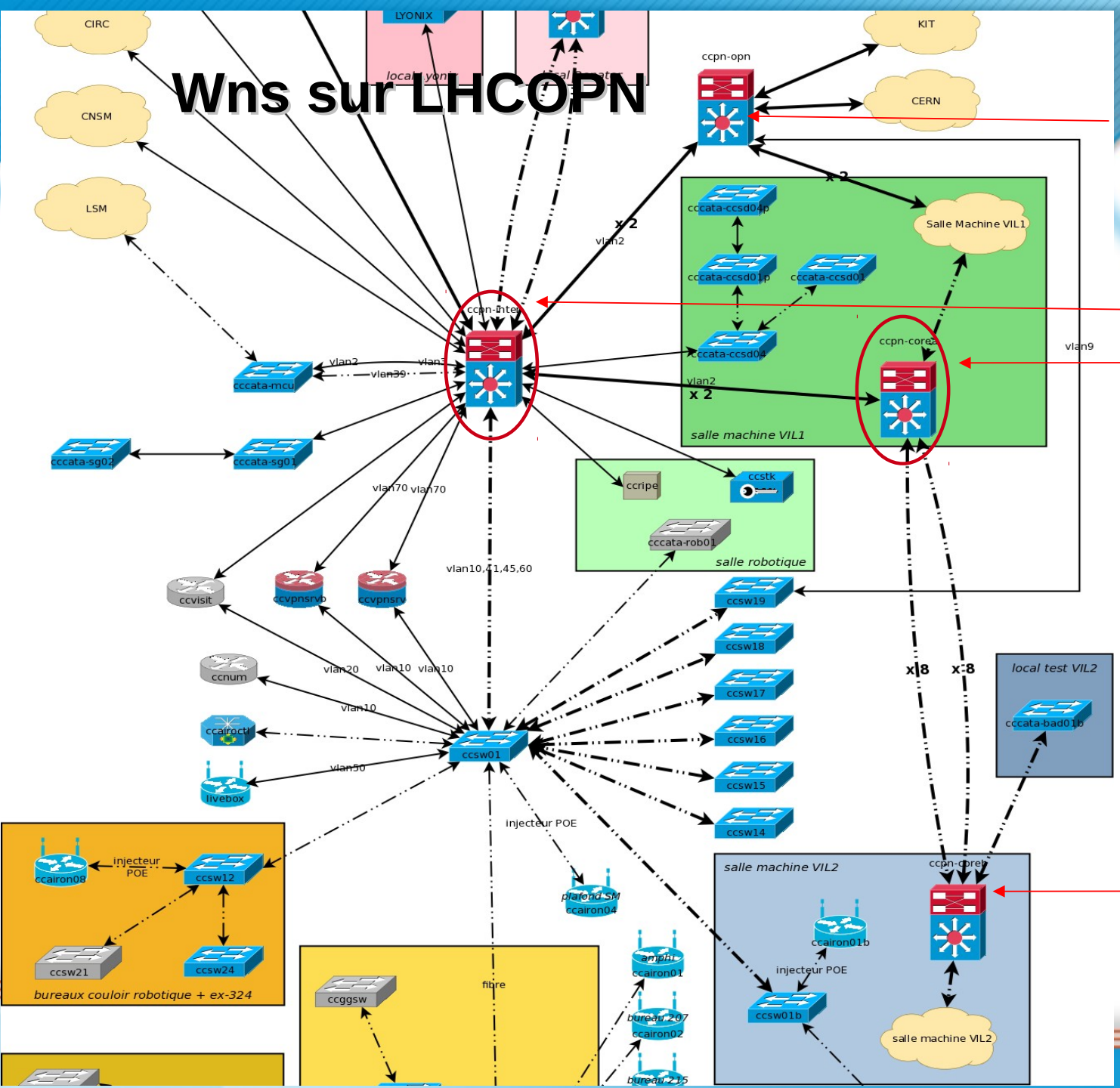
Entrée T2



- Il faudra aussi vérifier les règles en fonction des LFNs des données [moi].
- Étudier la taille des buffers d'import (disque et tape) [moi]

- Discussion avec les dCache-masters : à mettre en place de préférence avant l'arrêt de juin (migration vers la dernière « golden release »).

Wns sur LHCOPN



LHCOPN

ccpn-inter

ccpn-corea

ccpn-corel
(Salle machine Vil2)

WNs sur LHCOPN

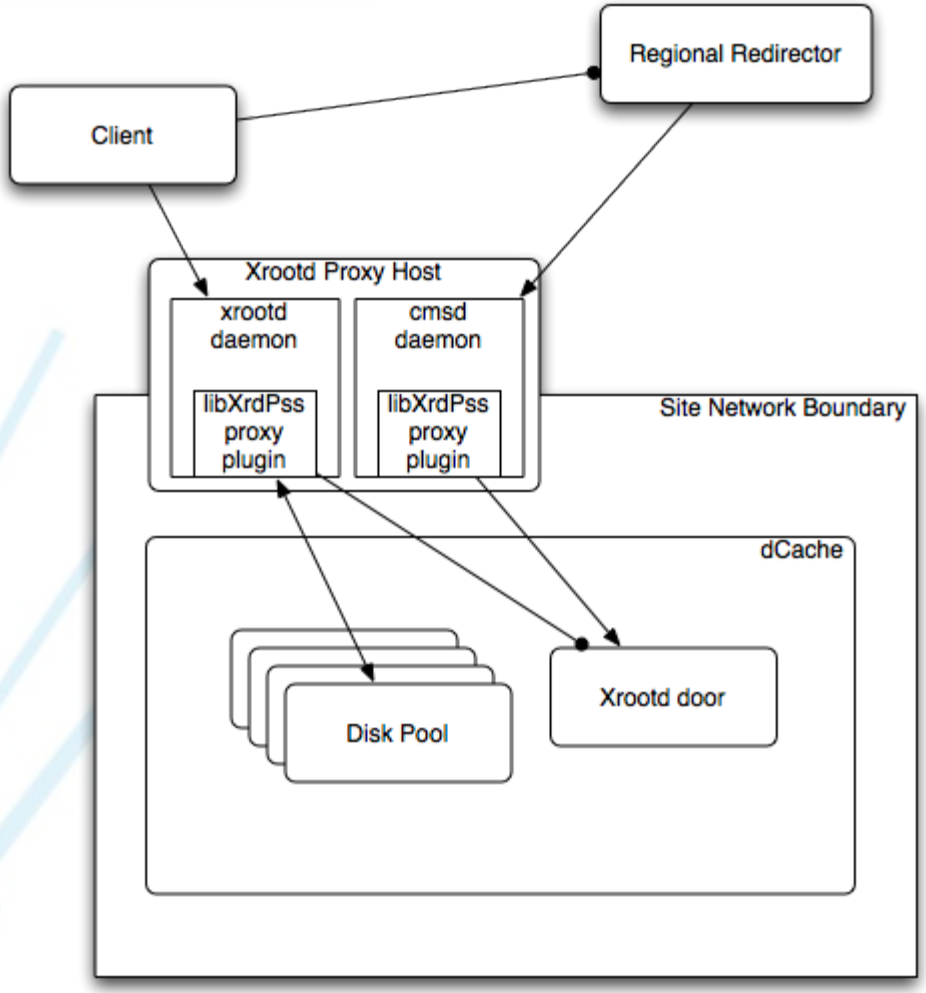
- Techniquement réalisable (vu avec Jérôme et Laurent)
  - Reconfiguration des routeurs ccpn-corea et ccpn-inter (routage vers LHCOPN)
  - Potentiellement renforcer la liaison entre ces deux routeurs
  - Reconfiguration du routeur ccpn-opn
- Difficulté identifiée
  - Nombre d'IPs (publiques) limité (suffisamment d'IPs tant qu'on ne virtualise pas une partie de la ferme).
  - Résolution (potentielle) avec le passage en IPv6 (WNs en dernier, migration complète estimée d'ici 2 ans)
  - Certains services ne supportent pas l'IPv6 (AFS), donc l'IPv4 sera toujours nécessaire (période transitoire en dual stack IPv4-IPv6)...

Passage à la fédération

- Le choix d'implémentation passe par la mise en place d'un « Proxy Host »
  - Permet de protéger dCache derrière un Firewall (et également de maîtriser la charge réseau par le bridage de la bande passante à 1 GB/s)
  - Une machine va être récupérée pour cette usage (d'ici 2 à 3 mois, Jean-Yves)
  - Possibilité de contrôler la charge réseau via un logiciel CMS qu'on installera également sur cette machine

- CMS recommande 3 machines (2 pour le « load-balancing », et 1 en « fail-over »)
  - Période de test dans la configuration initiale (6 mois)
  - Augmentation de la bande passante et ajout d'une machine pour le « load-balancing » (configuration minimale préconisée par CMS)
  - Le contrôle de la charge réseau se fera toujours par le biais du logiciel CMS dédié (si test concluant pendant la période d'essai)
  - Question : est-il possible de faire acheter les machines par le biais de LCG ?

Passage à la fédération



Conclusion

- Séparation disque et « tape »
  - Techniquement faisable en modifiant la structure actuelle de dCache
  - Travail technique relativement important (changement important de la configuration côté dCache)
  - Échéance estimée : avant l'arrêt de juin (mise-à-jour vers la dernière « golden release »)

- WNs sur LHCOPN
  - Relativement simple (reconfiguration de 3 routeurs de notre côté)
  - Problème en fonction du nombre d'IPs nécessaire (en cas de virtualisation partielle de la ferme)
  - Le passage à l'IPv6 ne résout pas tout (services non compatibles IPv6)

- Fédération
  - Choix de l'implémentation via un « Proxy Host »
  - Test dans la configuration de départ pendant 6 mois
  - Après la période de test, ajout d'une (au moins) machine pour le « load-balancing » et débridage de la bande passante (si tests concluants avec le logiciel de contrôle de la bande passante)