



Déploiement d'une Ferme d'Analyse au CC-IN2P3

Réunion semestrielle LCG-France

Yvan Calas
Yvan.Calas@cc.in2p3.fr

Annecy, 19 mai 2009



Pourquoi une ferme d'analyse au CC ?



- Besoins nouveaux exprimés par les expériences pour les activités d'analyse:
 - Workshop « Analyse des données au CC-IN2P3 », Avril 2008.
 - Comité de Direction LCG-France, Mai 2008.
 - Comité de Direction du CC, Septembre 2008.
 - Forum Trimestriel LCG France, « Synthèse de l'étude d'une ferme d'analyse au CC-IN2P3 », Mars 2009.

▶ L'approche adoptée



- Deux axes pour l'étude de la ferme:
 - Bénéficiaire des travaux du groupe de travail créé au CERN (UAG):
 - *Création du groupe en Novembre 2008: ~20 membres avec Markus Schulz (CERN) comme coordinateur.*
 - *Objectifs:*
 - Synthétiser les modèles d'analyse des expériences LHC.
 - Comprendre l'impact de ces modèles sur les ressources fournies.
 - Proposer un modèle de ferme d'analyse.
 - Rapprochement avec les expériences LHC pour identifier les besoins et les solutions possibles.

▶ Conclusions de l'étude préliminaire



- Suggestion par les VOs elles-mêmes d'une ferme Xrootd/PROOF.
- Pas de demande d'utiliser un système de batch pour réguler les jobs d'analyse.
- Difficulté d'obtenir des données chiffrées provenant des expériences en terme de:
 - Nombre d'utilisateurs.
 - Taille des fichiers & accès aux données.
 - Bande passante.

Personnes impliquées dans le déploiement



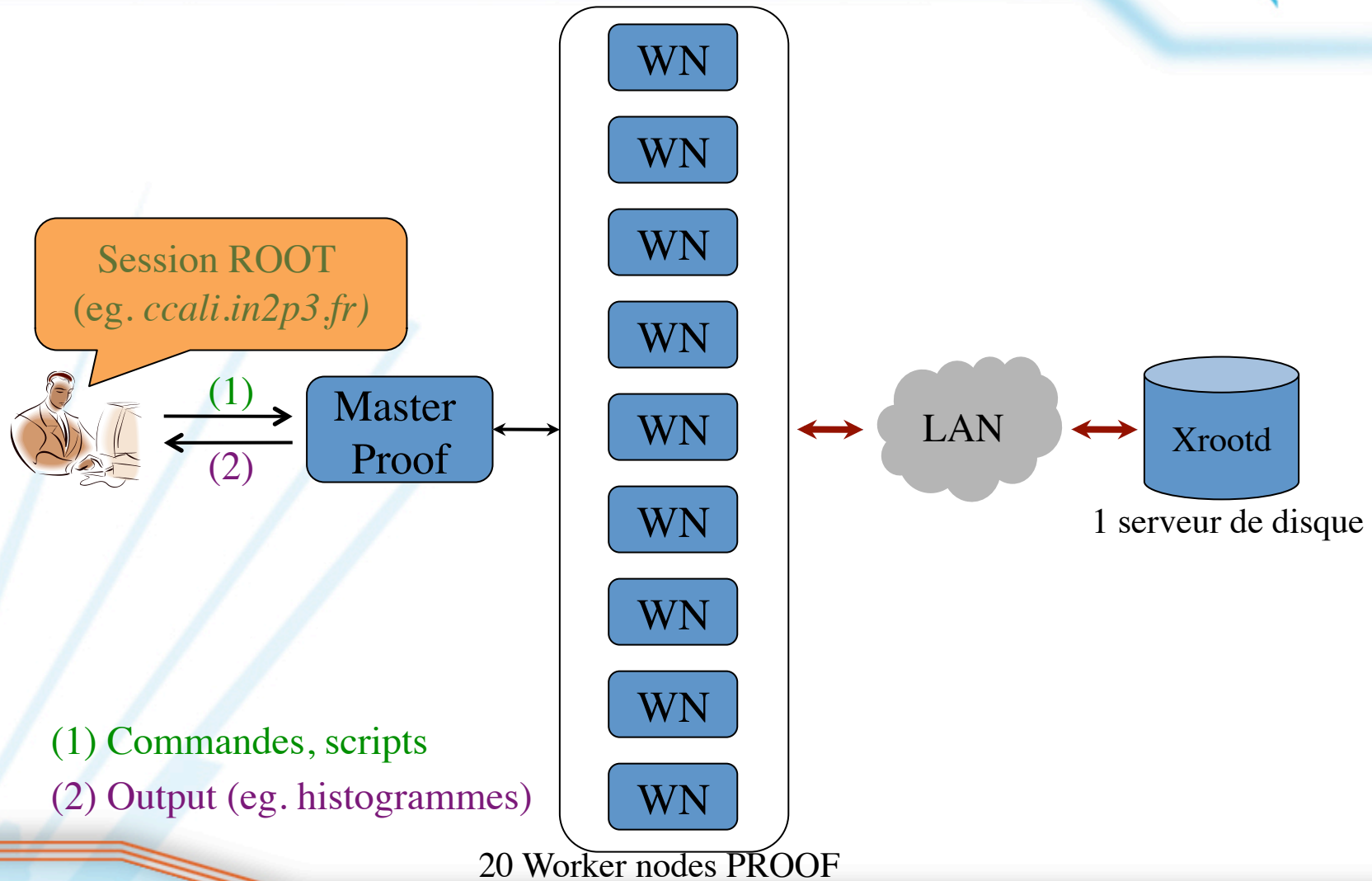
- Coordination générale:
 - Dominique Boutigny & Fabio Hernandez.
- Administration système:
 - Fabien Wernli.
- Stockage (Xrootd):
 - Jean-Yves Nief.
- Proof / déploiement / support:
 - Ghita Rahal (Support) et Yvan Calas.
- 1 contact pour chaque VO LHC.

▶ Topologie actuelle (1/2)



- 1 master proof: `ccapl0001.in2p3.fr`
- 20 worker nodes: `ccapl00[02-21].in2p3.fr`
 - 8 cores Intel Xeon, 2.66GHz, SL4 64 bits.
 - 16GB RAM, 160GB disque.
 - 1GB/s Ethernet.
- 1 serveur Xrootd dédié: `ccxrasn001.in2p3.fr`
 - 17TB disque.
 - 2 x 1GB/s Ethernet.

▶ Topologie actuelle (2/2)



▶ Accès à la ferme d'analyse



```
[xxxxx-16:23:55] /Users/ycalas/ROOT/work (512) > root  
[.....]  
ROOT 5.22/00 (trunk@26997, De 18 2008, 10:17:00 on macosx)
```

```
CINT/ROOT C/C++ Interpreter version 5.16.29, Jan 08, 2008  
Type ? for help. Commands must be C++ statements.
```

```
Enclose multiple statements between { }.
```

```
root [0] TProof * p = TProof::Open("ycalas@ccapl0001.in2p3.fr")
```

```
Starting master: opening connection ...
```

```
Starting master: OK
```

```
Opening connections to workers: OK (152 workers)
```

```
Setting up worker servers: OK (152 workers)
```

```
PROOF set to parallel mode (152 workers)
```

```
root [1] p.Process("/Users/ycalas/ROOT/root/tutorials/proof/ProofSimple.C", 10000000)
```

```
Mst-0: grand total: sent 102 objects, size: 94413 bytes
```

```
(Long64_t)0
```

```
root [2] TFile* myfile =TFile::Open("root://ccxrasn001:1094//xrootd/essai1.root")
```

```
root [3]
```

Connexion au « master proof »
(ici « ccapl0001 »)

Lancement de l'analyse

Lecture d'un fichier sur
le serveur xrootd

► Output d'un test simple



PROOF Query Progress: ycalas@ccyvca.in2p3.fr

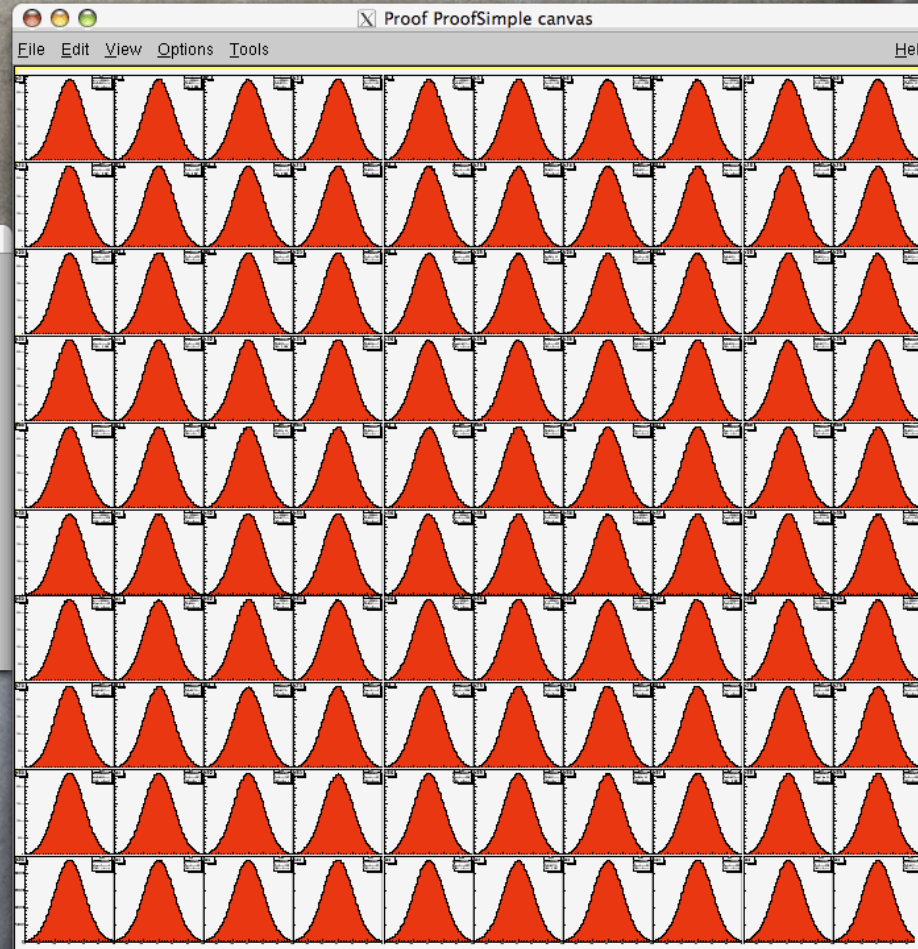
Executing on PROOF cluster "ccyvca.in2p3.fr" with 2 parallel workers:
Selector: /Users/ycalas/ROOT/root/tutorials/proof/ProofSimple.C
0 files, number of events 1000000, starting event 0

Initialization time: 0.5 secs
Processed: 1000000 events (0.00 MBs) in 1 min 15 sec
Processing rate: 13223.7 evts/sec (0.0 MBs/sec)

Close dialog when processing is complete

Show Logs Rate plot Memory Plot

Stop Cancel Close



▶ Tests effectués par les expériences



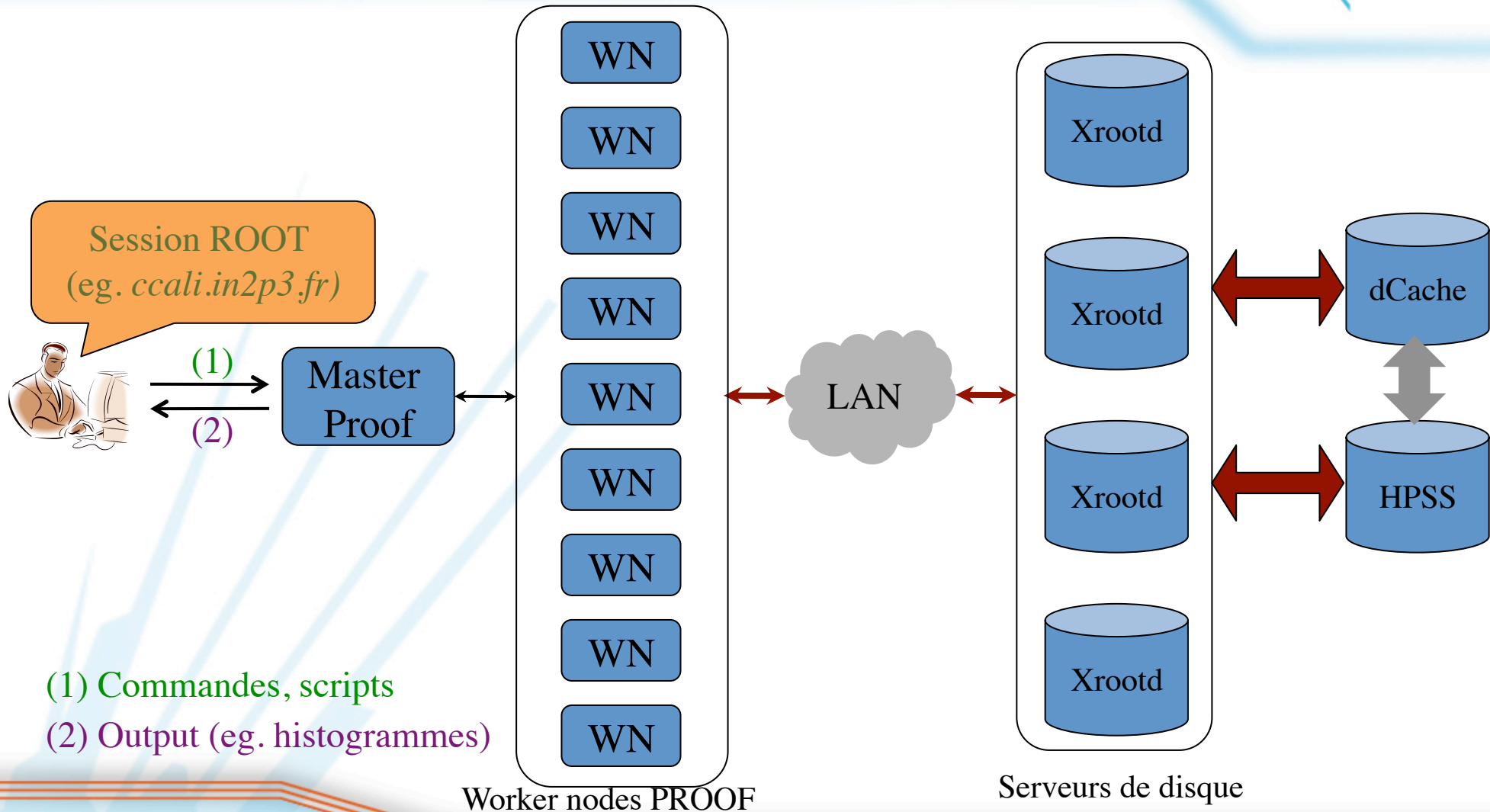
- CMS
 - Contact: Michal Bluj (CERN).
 - Test simple (*ProofSimple.C*) effectué avec succès.
- Alice
 - Contact: Laurent Aphecetche (Subatech).
 - Installation dédiée de ROOT, AliRoot et Alien.
 - Premiers tests « à la Alice » effectués avec succès la semaine dernière:
 - *Voir présentation de Laurent pour les détails.*

Evaluation des performances de la ferme



- Approche incrémentale en impliquant fortement les représentants des VOs.
 - **Cycle:**
 - *Déploiement.*
 - *Evaluation des performances + analyse des bottlenecks.*
 - *Proposition d'amélioration.*
 - **Améliorations possibles:**
 - *D'ordre matérielle:*
 - Infiniband: connexion haut débit entre WNs et serveurs Xrootd.
 - Disque SSD: cache disque de 1^{er} niveau.
 - *D'ordre topologique.*

Topologie future de la ferme Variantes (1/3)



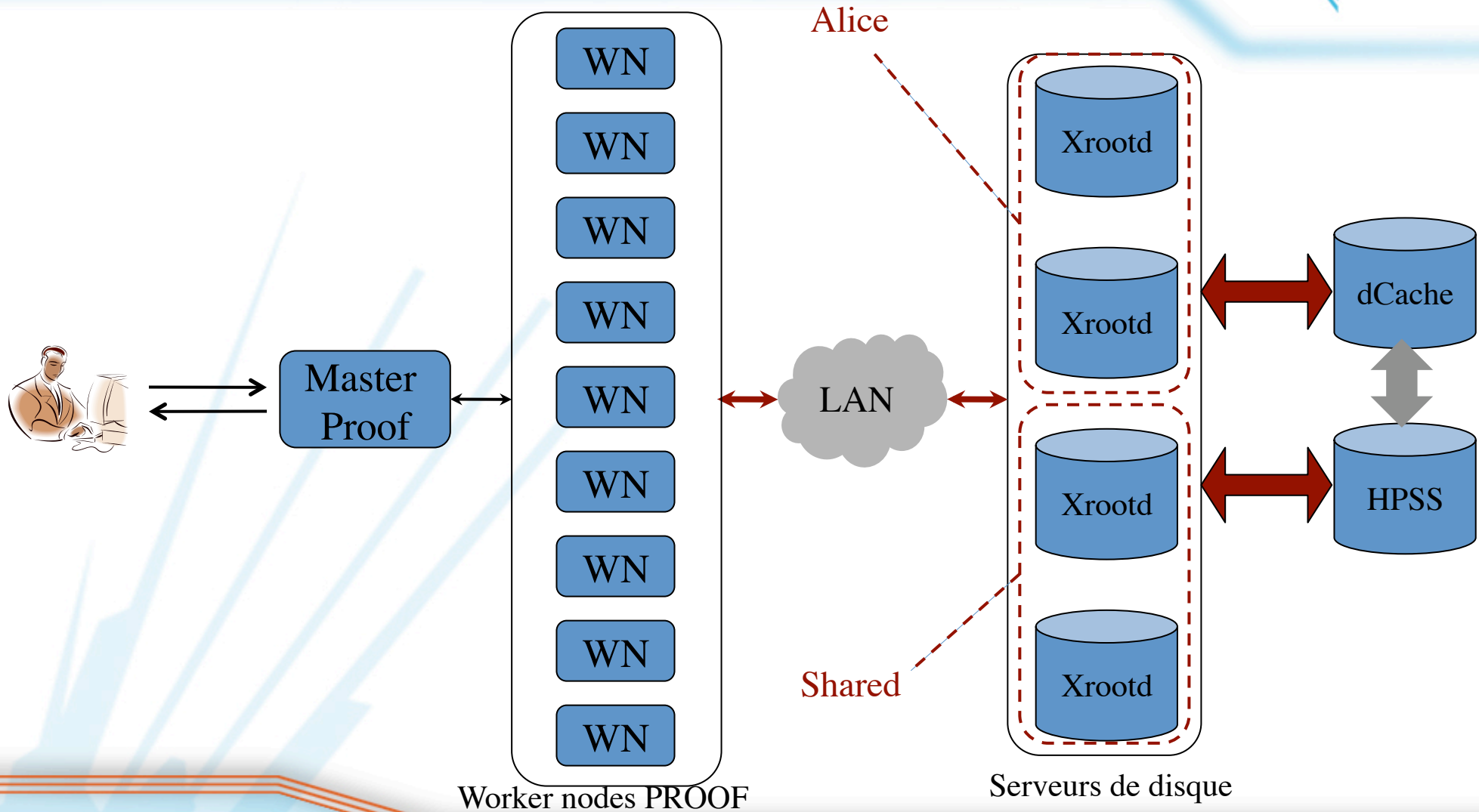
(1) Commandes, scripts

(2) Output (eg. histogrammes)

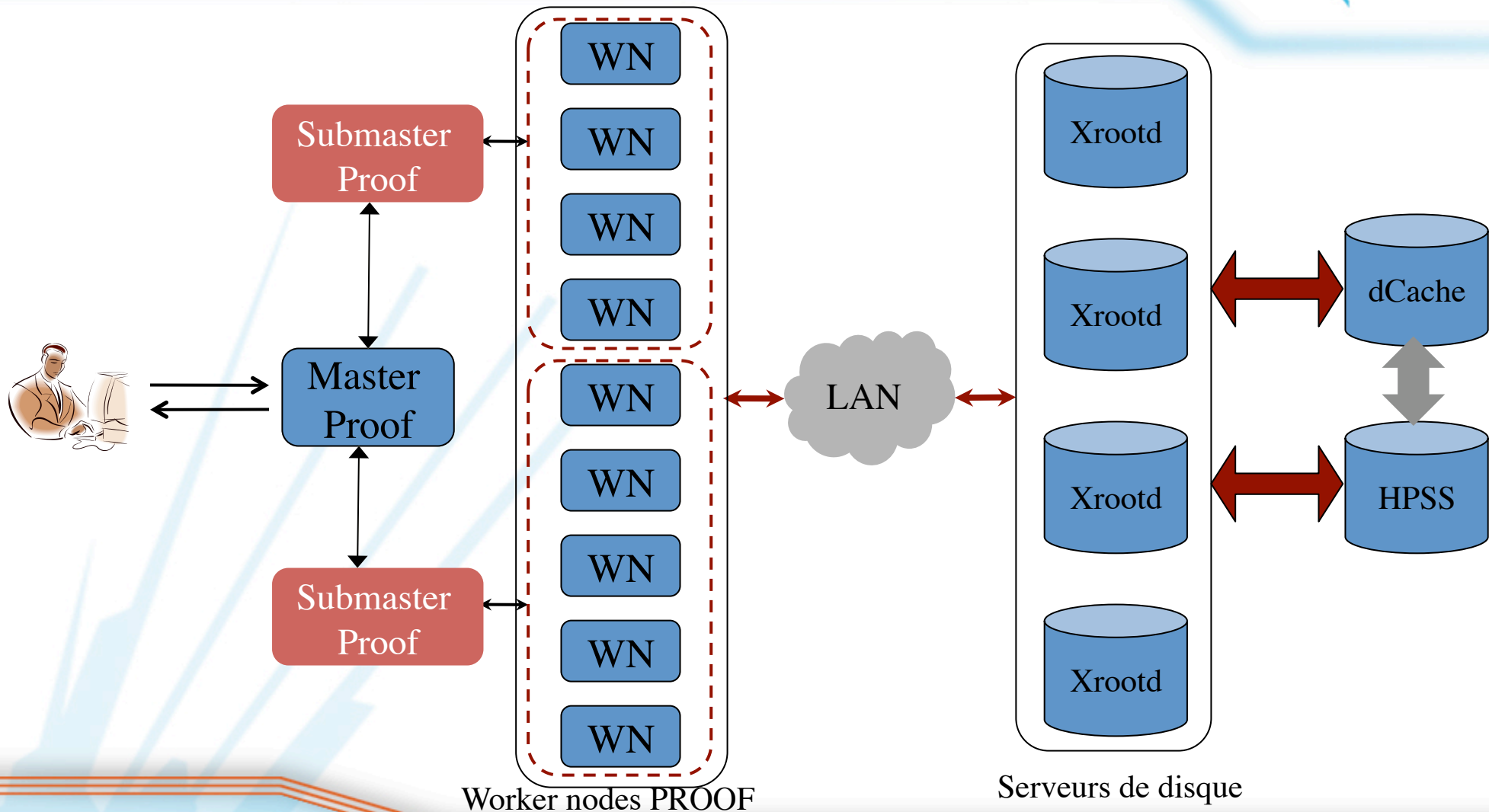
Worker nodes PROOF

Serveurs de disque

Topologie future de la ferme Variantes (2/3)



Topologie future de la ferme – Variantes (3/3)



▶ Autres points à améliorer



- Limitations actuelles:
 - Accès à la ferme (pas d'authentification stricte):
 - *Identification par password AFS et/ou certificat Grille.*
 - Automatiser le plus possible l'installation des WN.
 - Accès aux données si non disponible sur Xrootd.
 - Beaucoup de paramètres de configuration dans PROOF qui nécessitent d'être étudiés.
- Bons contacts avec les développeurs (notamment Gerardo Ganis), permettant de résoudre rapidement les pbs rencontrés.

▶ Suivi opérations et documentation



- Suivi des opérations & documentation:
 - Elog (bugs, incidents, modifications config, etc.).
 - Site Web à créer contenant:
 - *Les différentes présentations effectuées (LCG France, Workshops, etc.).*
 - *Documentation concernant:*
 - Les topologies étudiées, leurs performances et conclusions associées.
 - Description de l'accès à la ferme d'analyse (création de compte utilisateurs, tests à effectuer, etc.).
 - (Nécessité d'un twiki???)

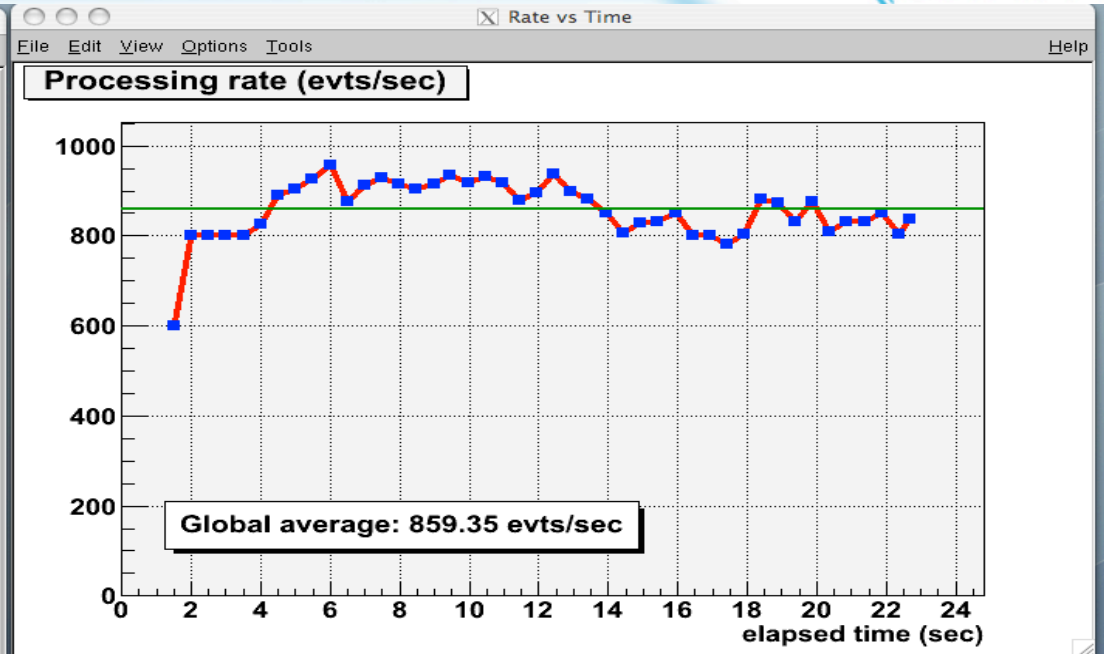
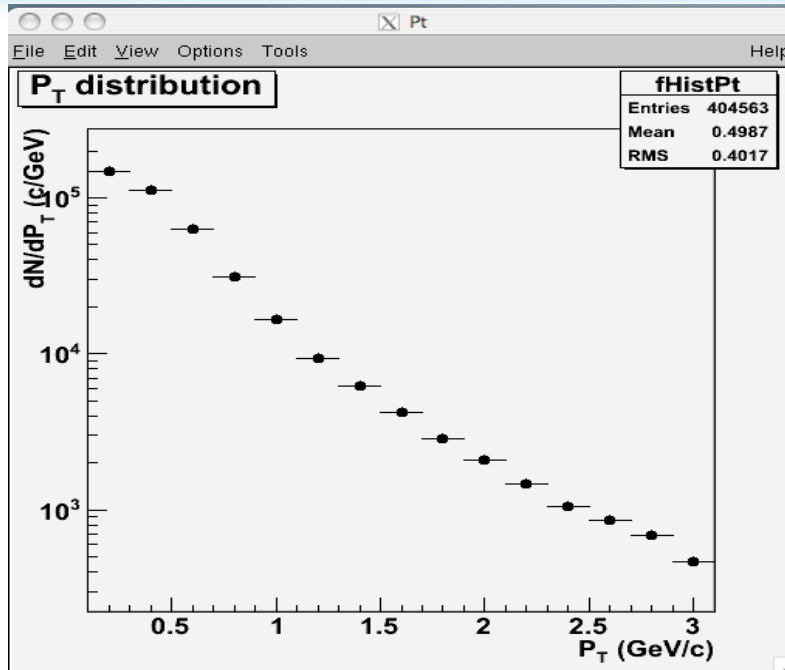
▶ Conclusions



- Premier prototype de ferme en place.
- Actuellement bonne interaction avec les expériences (notamment Alice), ce qui permettra à court terme:
 - D'obtenir rapidement des résultats concernant les différentes topologies.
 - De profiter de leur expérience déjà acquise sur d'autres fermes d'analyse (e.g. CAF) et d'en comparer les résultats.

QUESTIONS ?

► Premiers tests effectués par Alice



PROOF Query Progress: aphecetc@ccapl0001.in2p3.fr

Executing on PROOF cluster "ccapl0001.in2p3.fr" with 16 parallel workers:
Selector: AliAnalysisSelector
195 files, number of events 19500, starting event 0

Initialization time: 5.3 secs
Processed: 19500 events (1351.10 MBs) in 22 sec
Processing rate: 859.4 evts/sec (59.5 MBs/sec)

Close dialog when processing is complete

Show Logs Rate plot Memory Plot

Stop Cancel Close