

Gestion de la Production :

Partage avec les utilisateurs locaux

- Cas du Laboratoire SUBATECH
- Généralisation
- Problèmes
- Solutions possibles

Cas du Laboratoire SUBATECH

- Historique
- Situation actuelle
- Problèmes rencontrés

Ferme SUBATECH : Historique

- Version 1 : 1999
 - 9 Workers sur un subnet spécifique
 - Machine «maî tre» : accès interactif, serveur NFS
 - Gestionnaire de batch : DQS
 - Comptes utilisateurs spécifiques
 - Queues suivant temps CPU
- Version 2 : 2003
 - Workers sur le réseau principal
 - OpenPBS
 - Comptes via NIS, accès aux disques utilisateurs via NFS
 - Queue unique : default

Principaux choix

- Simplicité d'accès
 - accessible immédiatement à tout utilisateur
- Simplicité d'utilisation
 - écrire un shell-script monscript.sh
 - qsub monscript.sh
- Usage très libre
 - pas de comptabilité formelle
 - on compte sur le civisme des utilisateurs

Situation actuelle

- Version 3 : torque+MAUI
- Queues : default,alice,dteam,ops
- Fair-Share MAUI :
 - 70% alice, 30% locaux
- Limites :
 - FS+10% en occupation des CPU
 - Max 10 travaux/par utilisateur local

Configuration torque

```
set server max_running = 140  
set server default_queue = default
```

```
set queue ops Priority = 80  
set queue ops resources_max.cput = 02:00:00  
set queue ops resources_max.walltime = 04:00:00
```

```
set queue alice Priority = 60  
set queue alice max_queuable = 140  
set queue alice resources_max.cput = 48:00:00  
set queue alice resources_max.walltime = 72:00:00  
set queue alice resources_default.cput = 48:00:00
```

```
set queue dteam Priority = 80  
set queue dteam max_queuable = 10  
set queue dteam resources_max.cput = 48:00:00  
set queue dteam resources_max.walltime = 72:00:00
```

```
set queue default Priority = 60  
set queue default max_queuable = 150  
set queue default resources_max.cput = 240:00:00  
set queue default resources_max.walltime = 999:00:00
```

Configuration MAUI

```
RMPOLLINTERVAL    00:00:10
NODEALLOCATIONPOLICY FIRSTAVAILABLE
```

```
QUEUETIMEWEIGHT 1
FSWEIGHT         10
CREDWEIGHT       10
CLASSWEIGHT      1
```

```
FSPOLICY          DEDICATEDPES
FSINTERVAL        24:00:00
FSDEPTH           14
FSDECAY           0.99
```

```
SYSCFG PRIORITY=100 MAXPROC=140 MAXJOB[USER]=10
```

```
CLASSCFG[default] PRIORITY=100 FSTARGET=30+ MAXPROC=56 MAXJOB[USER]=10
CLASSCFG[alice]   PRIORITY=100 FSTARGET=70+ MAXPROC=112
CLASSCFG[ops]     PRIORITY=1000 FSTARGET=1+  MAXPROC=2
CLASSCFG[dteam]   PRIORITY=1000 FSTARGET=1+  MAXPROC=2
```

Problèmes rencontrés

- Remplissage / Disponibilité
 - Si on remplit, la disponibilité n'est pas immédiate
- Problème des travaux long
 - certains travaux locaux > 1 semaine
- Partage entre utilisateurs locaux
 - trouver une règle équitable

Généralisation

- Partition / Partage
 - partition = séparation complète
 - partage = compter uniquement sur le FS ?
 - préférence des VO : Alice souhaite que l'on garantisse une puissance de calcul et qu'il n'y ait pas d'oscillations.
- Conflits d'environnement ?
 - pas rencontrés jusqu'à présent

Problèmes à résoudre

- Garantir les ressources promises à chaque VO
- Remplir la ferme tout en préservant une certaine disponibilité ?
- Faire passer rapidement ops (dteam?)
- Comptabiliser (accounting) l'utilisation
 - pour valider les réglages, le FS en particulier
 - pour confronter avec les mesures LCG (APEL) et des Vos (Monalisa pour Alice)

Solutions possibles

- Intérêt d'une VO locale
- Connaître le temps CPU prévu pour un job permet l'optimisation du remplissage de la ferme tout en préservant la disponibilité
 - une queue pour les travaux courts ?
- Possibilités MAUI :
 - réservations
 - travaux en sur-couche (job overlay)

Discussion

- Idées / expériences
- Extraire et mettre en forme les données accounting PBS
- Comment publier correctement dans APEL
 - Conversion temps CPU en KSI2K
- Performances (efficiency = ?)

Références

- Documentation torque/MAUI
<http://www.clusterresources.com/products/torque/docs20/torqueadmin.shtml>
- Présentation Steve Traylen,
Tutorial on how to solve most common problems at sites
WLCG 22-26 Janvier 2007
<http://indico.cern.ch/conferenceDisplay.py?confId=3738>
-