



Gestion des GPUs dans le cluster

Philippe Séraphin

Réunion utilisateurs MUST, 10 mars 2022

Gpus disponibles

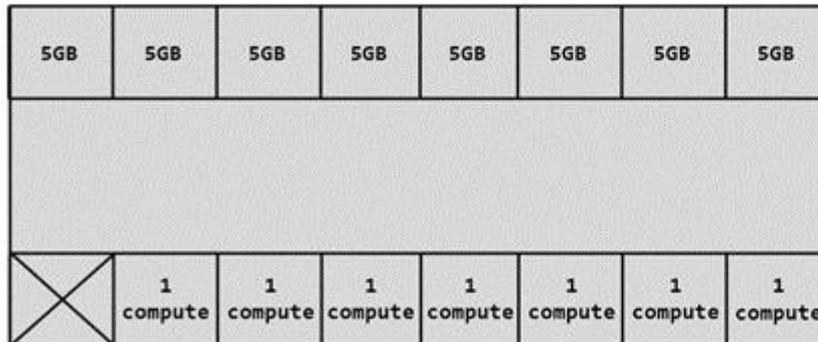


- 3 serveurs équipés chacun d'une carte K80 (vue comme deux cartes par le système) : wngpu001 à 003
- 1 serveur équipé d'une carte Tesla V100 : wngpu004
- 1 serveur équipé d'une carte Quadro P6000 : wngpu005
- 1 serveur équipé de 4 cartes T4 : wngpu006
- 2 serveurs équipés de 3 cartes A100 :
 - wngpu007, le MIG est activé dessus et chaque carte est vue comme 2 cartes distinctes par le système, soit 6 cartes en tout
 - wngpu008, le MIG n'est pas activé
- 1 serveur équipé d'une carte A100 , celle-ci étant réservée pour le Listic, acheteur de la machine : wngpu009

Nouveauté A100 : le MIG

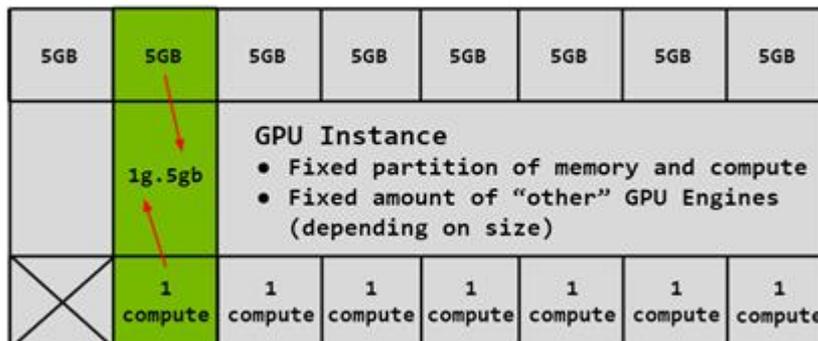
La A100 est composée de 7 unités de calcul (g) et de 40 Go (gb) de Mémoire (80 Go pour les prochaines)

Le MIG consiste à découper cette carte en plusieurs cartes avec un panachage possible des caractéristiques.



NVIDIA A100 (40GB)

- 8 x 5GB Memory Slices
- 7 Compute Slices



Nouveauté A100 : le MIG



Avec HTCondor

- Toutes les cartes d'un serveur doivent être identiques
- Donc pas de panachage possible avec le MIG
- Les possibilités de « MIG » sur MUST sont donc :
 - 1 carte 7g.40gb (équivalent A100 sans MIG)
 - 2 cartes 3g.20gb
 - 3 cartes 2g.10gb
 - 6 cartes 1g.5gb
- Le wngpu007 est actuellement configuré en 2 cartes 3g.20gb
- Pour reconfigurer le MIG, il faut stopper le service condor, et donc drainer la machine

Soumission simple

Request_gpus = 1

- Pourrait tourner sur n'importe quel GPU
 - ➔ Limité par nos soins aux cartes K80 et P6000 ou 1g.5gb
- Dans le cas de serveurs avec plusieurs cartes, HTCondor va réserver une ressource GPU
 - ✗ Ressource non bloquée sur le serveur
 - ✗ 2 jobs pourrait utiliser la même carte
 - ✓ L'utilisateur doit forcer son jobs à utiliser la ressource affectée par HTCondor (voir slide spécifique)

Demander un usage

A green callout bubble with a white border, containing the text "Méthode à privilégier !". The bubble is connected to the top by three smaller green circles of decreasing size, suggesting a path or flow.

Méthode à privilégier !

`+wantGpuUsage = "Default"`

3 types d'usages :

- Default : K80 ou Quadro P6000 ou MIG 1g.5gb
- Inférence : T4, MIG 2g.10gb ou 1g.5gb
- Training : V100, A100 ou MIG 7g.40gb ou 3g.20gb

Demander un type de GPU



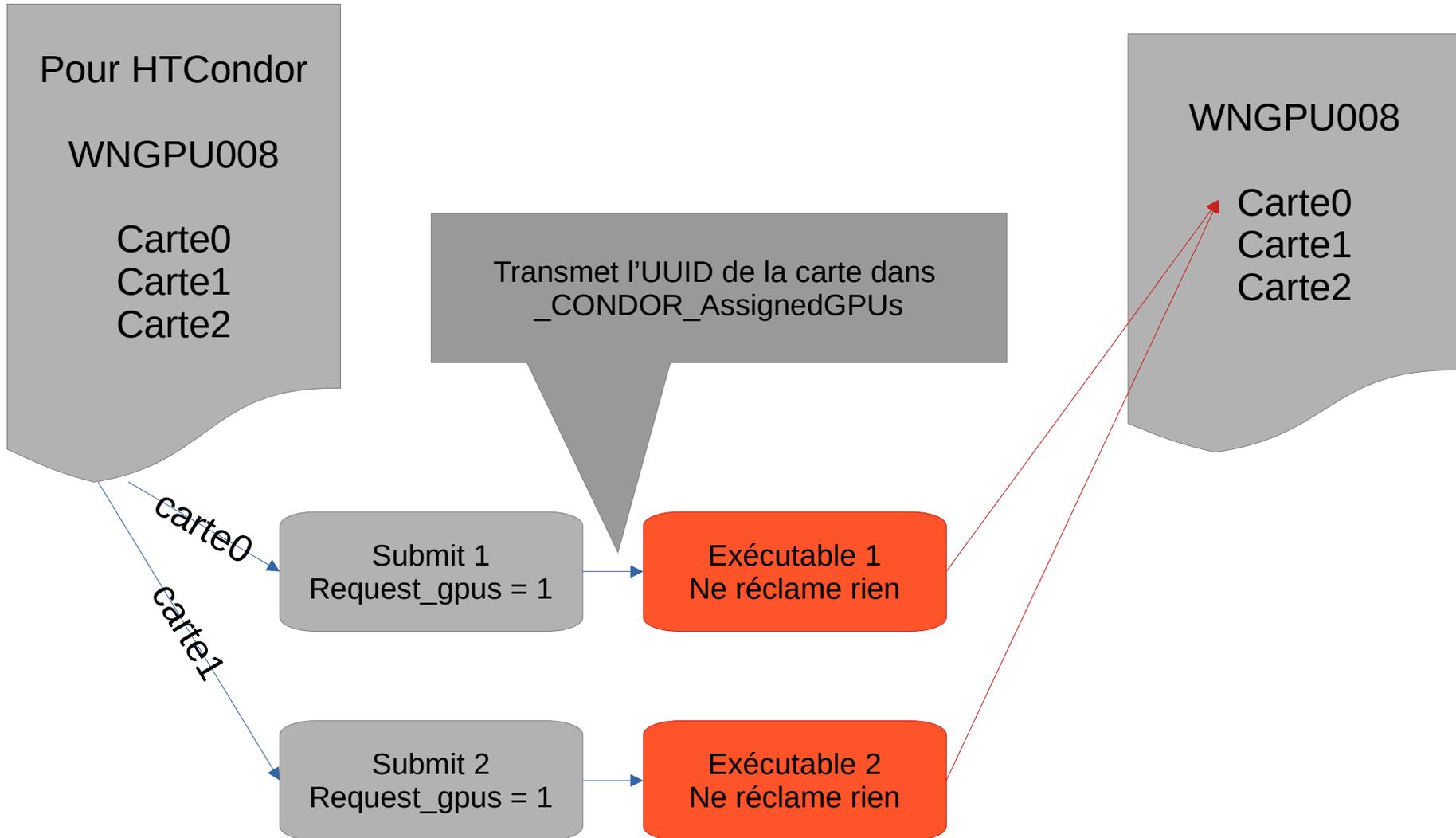
Si `+wantGpuType` est défini, il est prioritaire sur `+wantGpuUsage`

`+wantGpuType = "3g.20gb"`

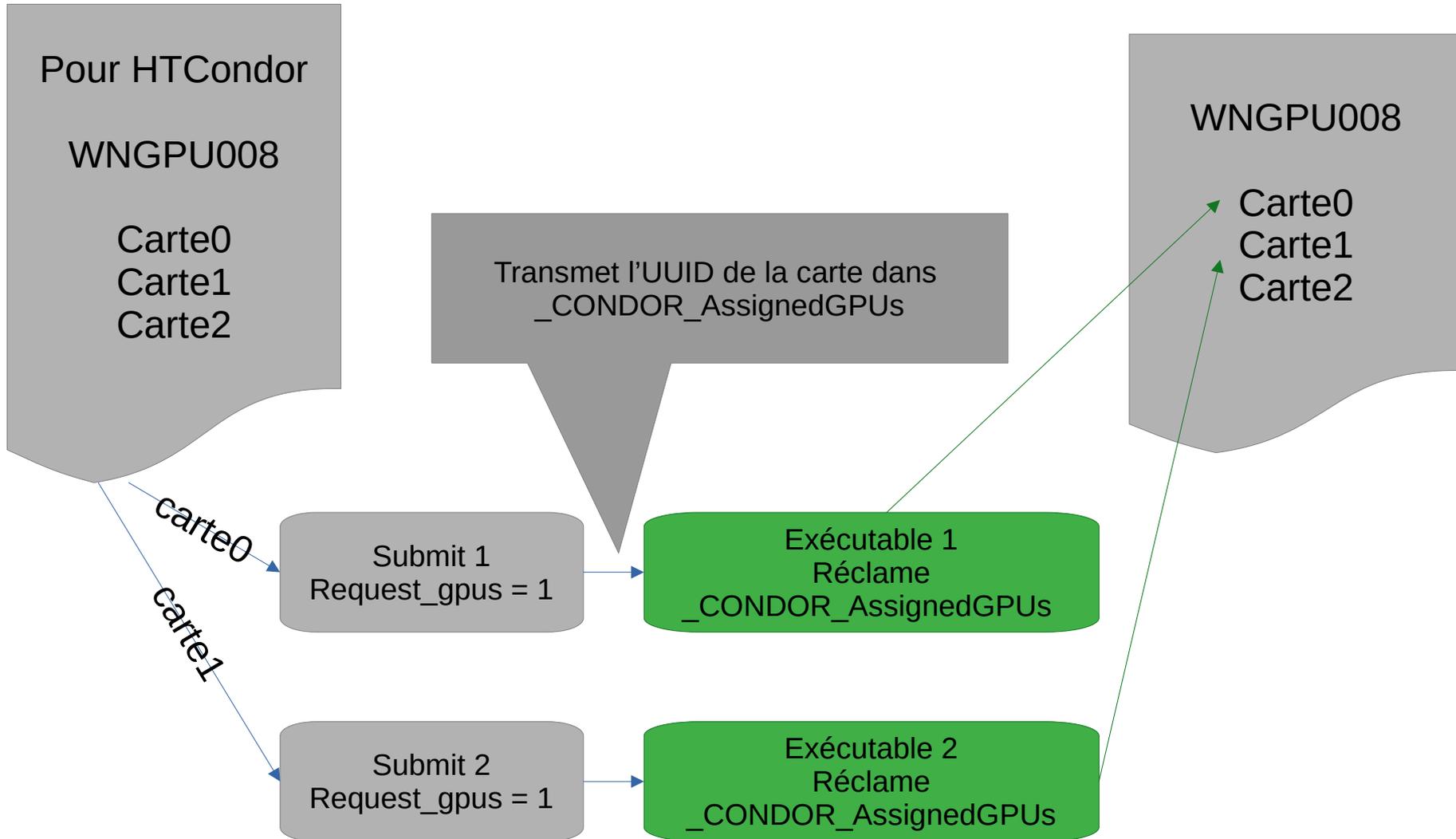
Valeurs possibles :

- K80
- P6000
- V100
- T4
- A100
- Pour le MIG : 1g.5gb, 2g.10gb, 3g.20gb ou 7g.40gb

Soumission job GPU : mauvaise méthode



Soumission job GPU : bonne méthode



Merci de votre attention.

