
Équilibrage des proxies selon la charge

Le pool virtuel envoie les machines à GPU
là où vos employés en ont besoin

Dossier d'explication pour la DSI

Cellule-PRO, version 0.2.158

À retenir

Vos serveurs Cellule-PRO (les *pools*) sont des serveurs simples. La puissance de calcul, ce sont les machines à carte graphique (les *proxies*). Plutôt que d'attacher chaque machine à un site pour toujours, vous pouvez la laisser se présenter au *pool virtuel* du cluster, qui l'envoie là où l'on attend des réponses : d'abord vers un pool qui n'a plus aucun moteur, ensuite vers le pool le plus chargé. Quand le calme revient, elle est rendue. **Rien ne bouge sans votre accord** : tout est désactivé par défaut.

Table des matières

1	En deux minutes	2
2	Comment ça marche	2
2.1	Les trois règles du pool virtuel	2
2.2	Le trajet d'un proxy	3
2.3	Et si un pool tombe ?	3
3	Les réglages, pas à pas	4
3.1	Étape 1 — choisir les proxies qui peuvent bouger	4
3.2	Étape 2 — activer le déplacement selon la charge	4
3.3	Étape 3 — le délai de bascule en cas de panne	5
3.4	Étape 4 — vérifier dans la Supervision	5
4	Avantages et inconvénients	5
4.1	Avantages	5
4.2	Inconvénients et limites	5
5	Faut-il l'activer chez vous ?	6
6	Questions fréquentes	6

1 En deux minutes

Imaginez une compagnie de taxis avec trois agences. Chaque agence a son guichet (le *pool*), mais les voitures (les *proxies*, vos machines à carte graphique) ne sont pas obligées de rester garées devant le même guichet toute la journée. Le matin, une agence est débordée pendant qu'une autre attend ; le central radio (le *pool virtuel*) envoie alors une voiture libre là où les clients font la queue. Le soir, la voiture revient au central.

Cellule-PRO fait exactement cela avec la puissance de calcul :

- **Les pools** reçoivent les questions des employés de leur site. Ce sont des serveurs Linux ordinaires, sans carte graphique obligatoire.
- **Les proxies** sont les machines puissantes qui produisent les réponses (une ou plusieurs cartes graphiques, un ou plusieurs modèles).
- **Le pool virtuel** est l'adresse unique du cluster (l'adresse IP virtuelle, ou VIP, portée à tour de rôle par l'un des pools). C'est à cette adresse que les proxies se présentent.

À retenir

- **Seuls les proxies bougent.** Jamais les postes des employés, jamais les pools, jamais les données.
- **Aucun pool n'est laissé de côté.** Un pool qui n'a plus aucun moteur d'inférence est servi en premier.
- **Chaque pool garde au moins un proxy par modèle :** on ne déshabille pas un site pour en habiller un autre.
- **Deux interrupteurs, désactivés par défaut :** un pour le cluster, un par proxy.

2 Comment ça marche

2.1 Les trois règles du pool virtuel

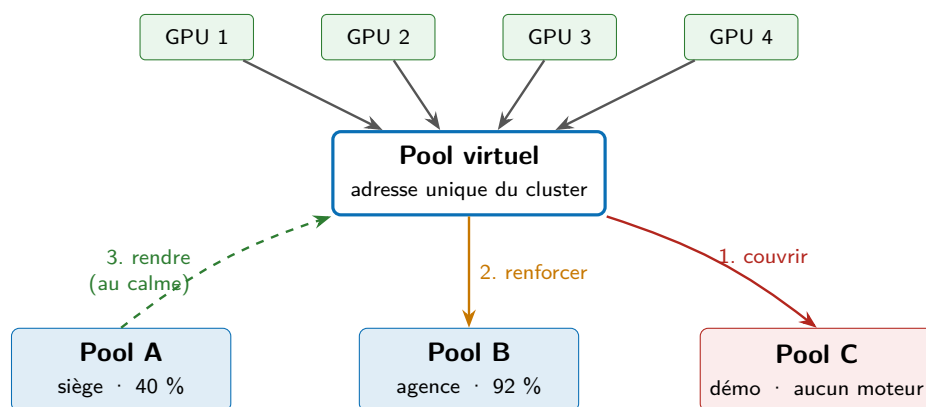
Toutes les minutes, le pool qui porte l'adresse virtuelle regarde les proxies *arrivés par cette adresse* et la charge de chaque pool du cluster. Il applique trois règles, dans cet ordre :

Règle	Ce qui se passe
1. Couvrir	Un pool du cluster n'a aucun moteur d'inférence (pool de démonstration sans carte graphique, pool tout juste installé, pool dont les machines sont parties) : il reçoit un proxy mobile en priorité , même si le pool virtuel est lui-même chargé.
2. Renforcer	Un pool dépasse le seuil haut (80 % par défaut) : il reçoit un proxy mobile, à condition que le pool virtuel reste lui-même sous le seuil haut sans ce proxy.

Règle	Ce qui se passe
3. Rendre	Un pool reste sous le seuil bas (60 % par défaut) pendant le temps minimum (5 min) : il rend au pool virtuel un proxy qu'on lui avait confié. Jamais le dernier proxy d'un modèle, jamais l'unique proxy d'un pool.

La **charge d'un pool** est la part de ses moteurs d'inférence occupés à répondre : 3 moteurs occupés sur 4 = 75 %. Un pool qui n'a aucun moteur compte pour 100 % : il ne peut répondre à personne.

2.2 Le trajet d'un proxy



Les employés de chaque site continuent d'écrire à leur pool ; seules les machines à GPU changent de place.

Le pool s'en charge seul

Un proxy qui reçoit l'ordre de partir **termine d'abord les réponses en cours** (jusqu'à une minute), puis se reconnecte au pool d'arrivée en quelques secondes, par une liaison chiffrée. Les employés ne voient rien d'autre qu'une réponse qui arrive plus vite.

2.3 Et si un pool tombe ?

- **Le pool qui porte l'adresse virtuelle tombe** : l'adresse passe en quelques secondes sur un autre pool (keepalived). Les proxies, toujours en marche, se reconnectent à la même adresse ; le nouveau pool virtuel les remplace. Aucune base centrale n'est perdue : il n'y en a pas.
- **Un autre pool tombe** : les proxies qu'il hébergeait basculent sur un autre pool du cluster après le délai de bascule (5 min par défaut, réglable) ; ils continuent donc de répondre, et le pool virtuel remplace ceux qui reviennent par son adresse.
- **Un pool revient sans aucun moteur** : la règle « couvrir » lui renvoie un proxy dès la minute suivante.

3 Les réglages, pas à pas

Quatre réglages, dans cet ordre. Les trois premiers se font **une seule fois, depuis n'importe quel pool** : ils sont recopiés sur tous les pools du cluster.

Avant tout : installer les proxies sur l'adresse du pool virtuel

À l'installation d'un proxy, indiquez comme adresse du pool **l'adresse virtuelle du cluster** (ou son nom de domaine), pas l'adresse d'un pool en particulier. Un proxy installé sur l'adresse d'un pool précis reste attaché à ce pool : il fonctionne, mais le pool virtuel ne le placera pas.

3.1 Étape 1 — choisir les proxies qui peuvent bouger

Onglet « Workers connectés », carte « Proxies déplaçables selon la charge ». Cochez la colonne « Mobile » pour chaque machine que vous acceptez de déplacer.

Workers connectés › Proxies déplaçables selon la charge			
Proxy	Modèle	Connecté	Mobile
gpu-salle-serveur-1	35B	à ce pool	☒ déplaçable
gpu-salle-serveur-2	35B	à ce pool	☒ déplaçable
gpu-vision	vision	à ce pool	☐ reste sur son pool

Conseil : laissez *sur son pool* une machine qu'un site doit garder quoi qu'il arrive (par exemple la seule machine de lecture de documents d'un site), et rendez mobiles les machines « de renfort ».

3.2 Étape 2 — activer le déplacement selon la charge

Onglet « Répartition workers », carte « Charge du cluster et placement des proxies ». Cochez « Déplacer les proxies selon la charge (cluster) », vérifiez les seuils, puis enregistrez.

Répartition workers › Charge du cluster et placement des proxies

☒ **Déplacer les proxies selon la charge (cluster)** Activé (n'agit que sur les proxies « mobiles »)

Seuil haut % Seuil bas % Temps minimum sur un pool s

Réglage	Défaut	Effet
Seuil haut	80 %	Au-dessus, un pool reçoit du renfort. Plus bas = renfort plus tôt, plus de déplacements.
Seuil bas	60 %	En dessous, un pool rend un renfort. Gardez au moins 15 à 20 points d'écart avec le seuil haut pour éviter les allers-retours.
Temps minimum sur un pool	300 s	Durée pendant laquelle un proxy reste là où on l'a envoyé avant de pouvoir repartir. Plus long = plus stable.

Le tableau au-dessus de ces réglages montre, pour chaque pool, sa charge, ses questions en attente et ses moteurs (occupés / total) : regardez-le quelques jours avant d'activer, pour choisir des seuils adaptés à votre activité.

3.3 Étape 3 — le délai de bascule en cas de panne

Onglet « Paramètres avancés », catégorie « Assistant de la DSI (alertes) », réglage « Délai avant qu'un proxy bascule sur un autre pool (secondes) ».

Paramètres avancés › Assistant de la DSI (alertes)

Délai avant qu'un proxy bascule sur un autre pool (secondes) de 30 à 3600

Court (30 à 60 s) : bascule rapide en cas de vraie panne, mais une coupure réseau de quelques secondes peut déplacer des proxies. Long (5 min et plus) : plus de stabilité. Ce délai ne concerne pas les proxies installés sur l'adresse du pool virtuel, qui suivent l'adresse en quelques secondes.

3.4 Étape 4 — vérifier dans la Supervision

Ouvrez la « Supervision » : une machine déplacée apparaît sous son nouveau pool à l'actualisation suivante. L'historique des ordres (proxy, départ, arrivée, état) reste consultable dans l'onglet « Répartition workers », carte « Historique des migrations », et chaque décision du pool virtuel est inscrite dans son journal avec sa raison (*couvrir*, *renforcer* ou *rendre*).

Attention

Le profil d'installation s'applique aussi au placement : un proxy n'est jamais envoyé vers un pool d'un autre profil (dossier « Profils d'installation »). Un cluster en air-gap total ne prête donc jamais de machine à un pool ouvert sur l'extérieur.

4 Avantages et inconvénients

4.1 Avantages

Moins de cartes graphiques à acheter	Une machine sert le site qui en a besoin à l'instant, au lieu de dormir dans un site calme. Vous dimensionnez le parc sur la pointe du cluster, pas sur la somme des pointes de chaque site.
Aucun site sans réponse	Un pool qui perd ses machines, ou un nouveau site sans carte graphique, est servi en priorité.
Résistance aux pannes	Un pool tombe : ses machines continuent de tourner et sont redistribuées. Aucun serveur central, aucune base unique.
Une seule adresse à donner	Chaque proxy est installé avec l'adresse du cluster ; ajouter un site ne demande de réinstaller aucune machine.
Vous gardez la main	Deux interrupteurs désactivés par défaut, un choix machine par machine, des seuils réglables, au plus 6 déplacements par heure, un historique de chaque ordre.
Rien d'autre ne bouge	Ni postes, ni données, ni comptes : les employés écrivent toujours à leur pool.

4.2 Inconvénients et limites

Le réseau entre sites	Une machine envoyée vers un site distant répond à travers votre liaison inter-sites (chiffrée). Sur une liaison lente, chaque réponse arrive un peu plus tard qu'en local ; le texte des questions et des réponses circule sur cette liaison.
Quelques secondes de transition	Pendant un déplacement, la machine termine ses réponses puis se reconnecte : elle ne prend pas de nouvelle question pendant ce court moment.
Une mesure encore globale	La charge est mesurée pour tout le pool, pas modèle par modèle : un pool saturé sur la lecture de documents mais calme sur le texte peut ne pas recevoir le bon renfort. Mesure par modèle prévue dans une prochaine version.
Il faut une adresse virtuelle	Sans adresse virtuelle (un seul pool, ou pas de keepalived), il n'y a pas de pool virtuel, donc pas de placement.
Rien ne se fait sans activation	Interrupteur coupé : un pool sans moteur reste sans moteur. Une alerte « ce pool n'a aucun moteur d'inférence » est prévue dans la Supervision.
Des machines qui changent de site	Pour un audit, la machine qui a répondu à une question n'est plus forcément celle du site. L'historique des ordres permet de le retrouver.

5 Faut-il l'activer chez vous ?

Votre situation	Conseil
Un seul pool	Inutile : il n'y a nulle part où envoyer une machine.
Plusieurs pools sur le même site (même bâtiment, réseau rapide)	Oui. Le gain est maximal et le coût réseau nul.
Plusieurs sites reliés par une bonne liaison	Oui , en gardant sur chaque site une machine « sur son pool », et en rendant mobiles les machines de renfort.
Sites reliés par une liaison lente	Avec prudence : relevez le seuil haut (85 à 90 %) et allongez le temps minimum, pour ne déplacer qu'en cas de vraie saturation ou de pool sans moteur.
Pool de démonstration ou d'évaluation	Oui : il n'a souvent pas de carte graphique, la règle « couvrir » lui prête une machine.

6 Questions fréquentes

Une machine peut-elle partir alors que ses utilisateurs sont en pleine conversation ?

Elle termine les réponses en cours avant de partir (jusqu'à une minute). La conversation elle-même est conservée par le pool de l'employé, qui ne bouge pas.

Peut-on interdire à une machine précise de bouger ?

Oui : laissez sa case « Mobile » décochée. C'est l'état par défaut de toute nouvelle machine.

Le pool virtuel est-il un « maître » ?

Non. C'est un rôle temporaire lié à l'adresse virtuelle : si le pool qui la porte s'arrête, un autre la reprend et continue, sans rien perdre, car il n'y a aucun état central.

Que se passe-t-il si deux pools sont très chargés en même temps ?

Le pool virtuel sert d'abord un pool sans aucun moteur, puis le plus chargé, un ordre à la fois, sans jamais se mettre lui-même au-dessus du seuil haut et sans retirer le dernier proxy d'un modèle.

Mes postes de travail (CPU des employés) sont-ils déplacés ?

Jamais. Seuls les proxies, ces machines dédiées à carte graphique, peuvent être déplacés.