

Machine VMWare avec un fort taux de CPU Stolen (%ready + %costop)

Signification du CPU Stolen

Le CPU Stolen est remonté par les VM Tools installés sur une VM VMWare. Il indique le temps qui est perdu par la VM à cause de l'ESXi, et réduit d'autant les performances de la machine.

- Il représente le temps perdu par la VM par rapport au cas où la machine serait une machine physique. C'est donc le temps perdu par toute la couche de virtualisation pour diverses raisons.
- Un taux élevé de CPU Stolen (>10%) est un problème qui doit être adressé au plus vite, car les démons Shinken sont fortement ralentis.

La documentation VMWare concernant cet indicateur est disponible en téléchargement sur le site <https://code.vmware.com/web/sdk/7.0/vsphere-guest> (lien : [Guest and HA Application Monitoring SDK Programming Guide](#)).

- L'indicateur est présenté page 30.

Une autre explication qui très claire de ces concepts est disponible sur le site <https://www.virtualease.fr/vmware-mecanismes-de-gestion-cpu/>.

- Vous pouvez également l'avoir sous forme de PDF : [VMware – Les mécanismes de gestion du CPU Virtualease.pdf](#)

Composantes du CPU Stolen

CPU Stolen: le temps perdu par la VM par rapport au cas où la machine aurait été physique sans ESXi

Le CPU Stolen remonté par les VMWare tools est décomposé en deux indicateurs sur VSphere :

- **%ready** (*prêt*) : temps où la VM n'a aucun CPU de disponible pour commencer à travailler
- **%costop** (*Arrêt simultané*) : temps cumulé de :
 - temps de recherches de lecture dans les snapshots disques
 - et ceux où les VCpus ont commencé à travailler, mais qu'il leur manque encore des CPUs physiques pour finir leur tâche
 - en résumé il y trop de VCpus utilisés par rapport au CPUs physiques disponibles.

Voir les indicateurs de perte de temps dans VSphere

Ces indicateurs sont récupérables dans l'interface VSphere pour la VM. Dans l'affichage des performances de la VM, il faut sélectionner les indicateurs :

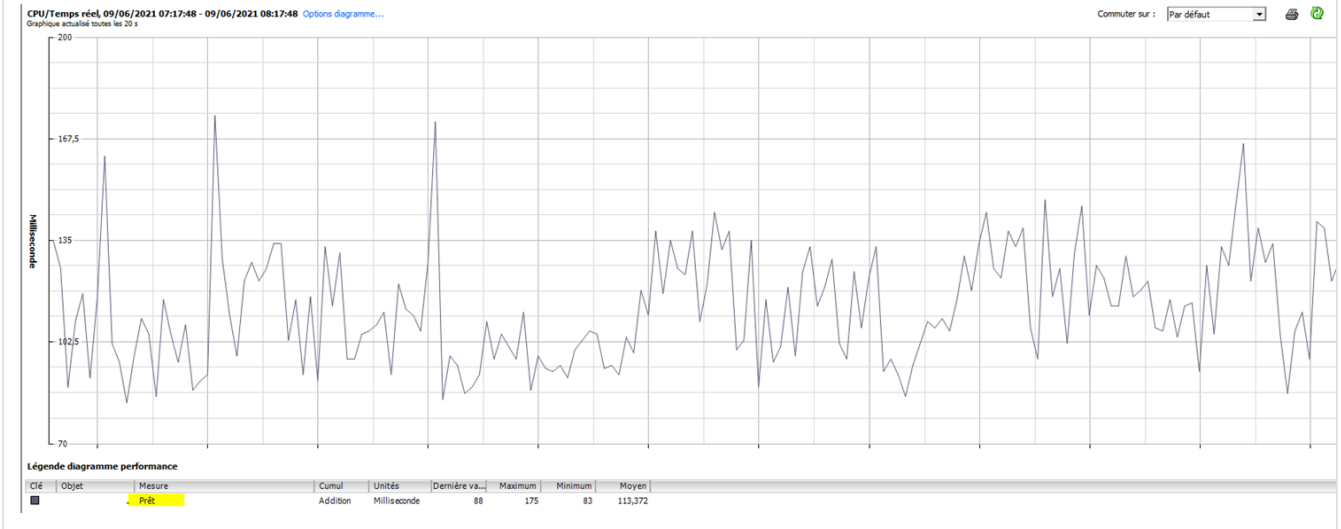
- **Prêt** (*ready*)
- **Arrêt simultané** (*costop*)

Description	Cumul	Unités	Nom int
<input type="checkbox"/> Attente	Addition	Milliseconde	wait
<input type="checkbox"/> Prêt	Addition	Milliseconde	ready
<input type="checkbox"/> Utilisé	Addition	Milliseconde	used
<input type="checkbox"/> Arrêt simultané	Addition	Milliseconde	costop
<input type="checkbox"/> Système	Addition	Milliseconde	system

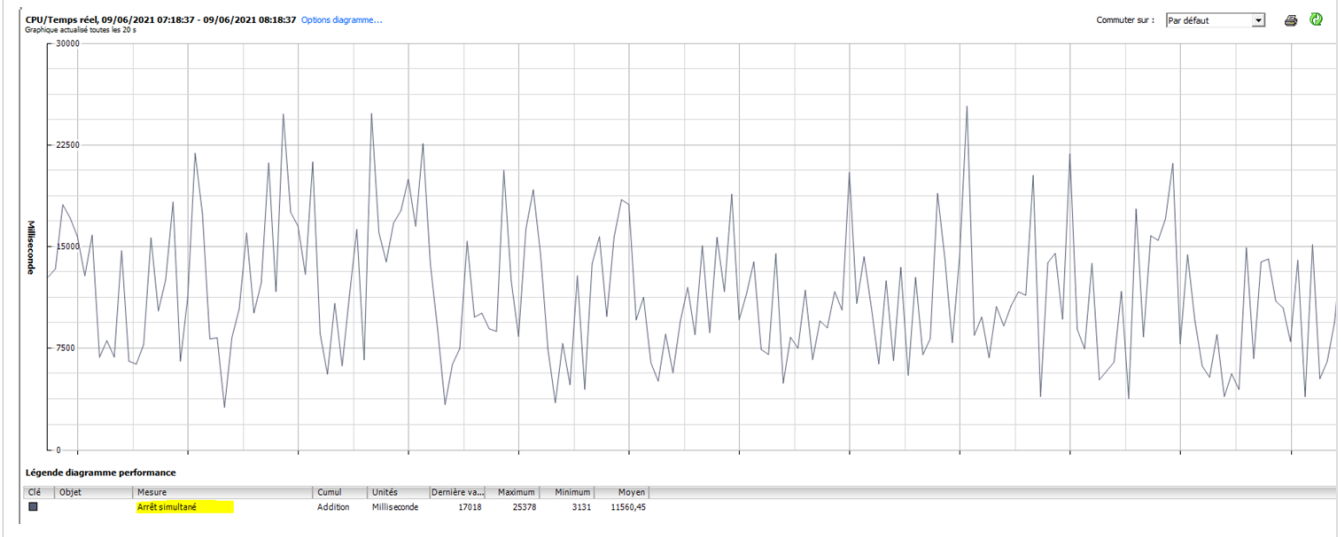
Ceci va donner deux courbes qui sont :

- en ms cumulées sur le dernier intervalle de mesure de VSphere, à savoir **20s**,
 - et donc pas en pourcentage facile à lire directement, il faut faire la conversion soit même
- et ce pour l'ensemble des VCpus (*ici une machine à 8VCpus*)

Indicateur ready dans VSphere, en ms écoulée sur 20s



Indicateur costop dans VSphere, en ms écoulée sur 20s



Calcul des ready et costop en pourcentage à partir des données dans VSphere

Il faut transformer les valeurs cumulées de VSphere en pourcentage, pour cela, il faut faire l'opération suivante : **pourcentage = moyenne / (nb_cpus * 20s * 1000ms)**

Ainsi dans notre cas, nous avons donc en moyenne :

- $113\text{ms} / (8\text{VCpus} * 20000\text{ms}) = 0.07\%$
- $11560\text{ms} / (8\text{VCpus} * 20000\text{ms}) = 7.2\%$

Dans notre cas la machine perd donc un peu plus de **7%** de son temps à cause de la couche de virtualisation.

Valeurs d'avertissement et critiques concernant le CPU Stolen

Le **CPU Stolen** étant du temps perdu à cause de la couche de virtualisation, il ne faut pas que ce temps soit trop important, sous peine d'avoir des démons Shinken qui tourneront au ralenti, et ne pourront pas effectuer les tâches qu'on leur a demandées.

Ainsi, les seuils d'alerte du CPU Stolen sont :

- **WARNING** si on est supérieur à **5%**
- **CRITICAL** si on est supérieur à **10%**

Pistes de résolution d'un taux de CPU Stolen élevé



Si une de vos VM a un fort taux de CPU Stolen, c'est un problème qui doit être réglé sans attendre.

Les valeurs de **%ready** et **%costop** dans l'interface de VSphere vont aider à trouver les bonnes pistes de résolution :

- si le taux de **%ready** est élevé, ceci signifie que le nombre de VCpus de la VM et des autres VMs de l'ESXi est trop élevée par rapport au CPU physiques de la machine
 - une VM ne peut fonctionner que si elle dédie pendant son temps d'exécution les CPU physiques de la machine
 - si par exemple elle a 8VCpus pendant qu'elle fonctionne (*par pas de 50ms par défaut dans ESXi*), elle va demander et bloquer 8 Cpus physiques
 - pendant ce temps les autres VMs ne peuvent pas les consommer si d'autres cœurs ne sont pas disponibles
 - **il faut donc diminuer le nombre de VCpus sur la VM et les autres de l'ESXi, ou rajouter des CPUs physiques sur l'ESXi**
- si le taux de **%costop** est élevé, il y a deux pistes possibles :
 - s'il y a des snapshots sur la VM chaque lecture doit être faite dans chaque snapshot : ce temps est trop important
 - il faut alors **supprimer les snapshots inutiles**
 - on peut avoir un cas similaire au %ready, car le temps costop peut montrer le temps qu'un VCpus a fini, mais qu'au moins un des autres VCpus de la VM n'a pas encore fini et bloque donc l'exécution des autres VMs
 - dans ce cas, comme dans le %ready, **il faut diminuer le nombre de VCpus de la VM et des autres VMs, ou rajouter des CPU physiques sur l'ESXi**

Cas important du sizing de l'ESXi avec de l'hyperthreading: gain de 20% de performance seulement (pas 100%)

L'hyperthreading permet à un même cœur physique de gérer deux threads en même temps.

- Il apparaît donc comme un cœur dans l'interface de VSphere.

Cependant, il ne faut pas considérer que les performances de la machine sont multipliées par deux. VMware, dans son document concernant les VMware tools vu plus haut (*page 31*) considère que l'activation de l'hyperthreading permet d'augmenter les performances de la machine de **20% (et non pas 100% donc)**.

Nous vous conseillons donc :

- De bien faire attention à ce point.
- Ne surtout pas faire le sizing de l'ESXi et des VCpus en vous basant uniquement sur le nombre affiché de CPU physiques.
- Ne prendre en compte que la moitié + 20% si le serveur a de l'hyperthreading d'activé.