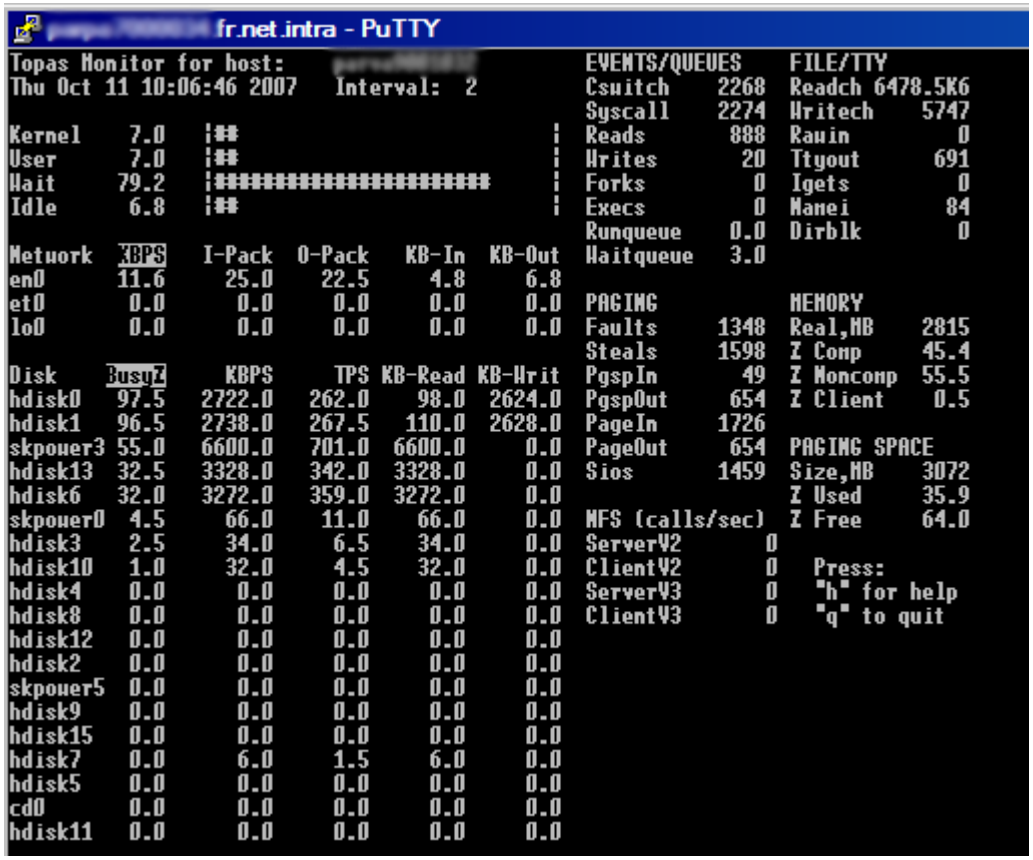
Table des matières

Généralités ... 3
Activité de swapping ... 3
Activité disques ... 4
Activité réseau ... 4
CPU ... 4
Liens ... 4
Mémoire ... 4
minperm, maxperm, lru_file_repage ... 4
Liens ... 5
Optimisation ramdisk ... 5
Détecter une fuite mémoire ... 5
Sous-système disques ... 5
Les I/O asynchrones ... 5
Les WAIT IO ... 5
Liens ... 5
Documentations ... 5

Infos plutôt valables pour des partoches physiques. Maintenant, avec PowerVM et AIX 7.x, les infos ci-dessous sont moins pertinentes.

Généralités

Lancer **topas** pour avoir un état global de la machine :



Activité de swapping

- *PAGING SPACE* : indique si la machine a swappé depuis le boot. Ici on voit que la machine a utilisé *35.9%* de swap soit 1 Go environ. Cette valeur n'est remise à zéro qu'au reboot.
- *PgspIn / PgspOut* : indique le nombre de pages swappées au moment du topas (pages de 4 ko). PgspIn → pages lues du swap, PgspOut → pages écrites dans le swap.

Dans ce cas on voit que la machine swappe. Pour expliquer les baisses de performances on peut voir que les devices de swap sont situés (généralement) sur les disques internes (hdisk0 & hdisk1 → *lscfg -vl hdisk0* pour avoir des infos sur les disques). Pour déterminer sur quel(s) disque(s) se trouve l'espace de pagination on peut utiliser les commandes suivantes :

```
mrcmid@SomeMachine:/$ lsvg -l rootvg |grep paging
hd6                 paging     48    96    2    open/syncd    N/A
```

```
mrcmid@SomeMachine:/$ lslv -l hd6
hd6:N/A
PV               COPIES        IN BAND       DISTRIBUTION
hdisk0           048:000:000   100%          000:042:006:000:000
```

```
hdisk1            048:000:000   100%          000:048:000:000:000
```

L'activité de swapping sollicite beaucoup ces disques (d'où le *Wait* à 79.2%) ce qui pénalise la machine car le CPU passe son temps à déterminer quoi lire et écrire dans le swap au détriment du reste de l'activité système.

Pour détecter les process qui swappent le plus :

```
root@SomeMachine:/apps/sys/log/stats/swap$ svmon -gP -t 10 |grep -p Pid|grep '^.*[0-9]'
     Pid Command          Inuse      Pin     Pgsp  Virtual 64-bit Mthrd LPage
 1359954 oracle         1029357   251501      365   259596      Y     N     N
     Pid Command          Inuse      Pin     Pgsp  Virtual 64-bit Mthrd LPage
 2375768 oracle         1029417   251501      349   259596      Y     N     N
     Pid Command          Inuse      Pin     Pgsp  Virtual 64-bit Mthrd LPage
 1740830 oracle         1034135   251501      349   259621      Y     N     N
```

Activité disques

- *%Busy* : taux d'occupation des disques, ici on voit les hdisk0/1 (swap) et le hdiskpower3 qui travaillent. *hdiskpower3* est un disque SAN composé de 2 *hdiskX* qu'on peut voir avec la commande *powermt display dev=hdiskpower3*. On retrouve les *hdisk13* et *hdisk6* comme le montre le topas.
- *KBPS* : taux de transfert (en ko/s)
- *TPS* : nombre de transferts par secondes
- *KB-Read / KB-Writ* : Ko lus et écrits

Avec la commande *lspv* on peut déterminer les LVs et FSs présents sur le disque pour, éventuellement, effectuer une meilleur répartition sur des disques moins sollicités.

Activité réseau

. . . à suivre . . .

CPU

Liens

http://www.ibm.com/developerworks/views/aix/libraryview.jsp?search_by=Optimizing+AIX+5L+performance
http://www.ibm.com/developerworks/aix/library/au-aix5_cpu/index.html
http://www.ibm.com/developerworks/aix/library/au-processinfinity.html

Mémoire

minperm, maxperm, lru_file_repage

- **minperm** : seuil minimal (en %) de la quantité de mémoire à utiliser pour le cache filesystem
- **maxperm** : seuil maximal (en %) de la quantité de mémoire à utiliser pour le cache filesystem
- **%comp** : mémoire réellement utilisée par le système (applis + OS)
- **%noncomp** : cache filesystem
- **%client** : mémoire utilisée par le JFS2 et NFS (le cas échéant)
- **numperm** : % de mémoire réellement utilisée par le cache filesystem

- Depuis AIX 5.2 ML4, il existe un paramètre (quasi) imparable, le *lru_file_repage*. Ce paramètre mis à 0, le serveur essaie de dégager uniquement les file pages. Du coup, si le minperm est à 5/10%, le maxperm à 80% et le système gere (en jfs2 ou pas) les pages mémoires. Une belle courbe vmtune se traduit par un *numperm* qui navigue entre *minperm* et *maxperm*.

Attention tout de même en cas de *jfs2* ou *NFS* fortement sollicité le maxclient peut littéralement exploser et atteindre les 80% du maxperm% et faire swapper le serveur, dans le cas maxperm%=maxclient%. Il est peut être appréciable de mettre un maxclient% à 30%.

Rapidement, l'utilisation du lru_file repage=0 permet de se libérer des contraintes min/maxfree% (intervalle qui encadre la garantie d'avoir de la place dans la free list, depuis 5.3, ces valeurs sont appliquées sur chaque memory pool).

- AIX utilise des pages de 4k par défaut. L'utilisation des pages de 16M (*lgpg_size*) ne sera fera uniquement par les applications sachant les utiliser (Sybase par ex.) et ne seront pas “libérable” pour les autres process. *lgpg_regions* permet de fixer le nombre de régions qui seront de cette taille.

Sinon, pour les larges pages, il faut aussi

1. Activer le *vpin_shr*
2. Le user doit être *large page* capable (c'est un attribut du user)

- Pour suivre la conso :

```
vmstat -l
```

- Voir le découpage par type de page :

```
vmstat -P all
```

Liens

http://www.ibm.com/developerworks/views/aix/libraryview.jsp?search_by=Tuning+your+memory+settings
http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.prftungd/doc/prftungd/aix_mem_aff_support.htm

Optimisation ramdisk

Il faut les monter avec l'option *rbrw* pour ne pas cacher 2 fois (ram & disques).

Détecter une fuite mémoire

Une doc tirée du site IBM.

Sous-système disques

Les I/O asynchrones

Par défaut les I/O asynchrones sont mal configurés. Le *min server* doit être fixé au nombre potentiellement accédé de disques avec les I/O asynchrones. Le *max server* doit etre égal au minimum multiplié par le nombre de CPU. Attention en AIX 5.3, le *max server* est d'office multiplié par l'AIX donc min=max. Le nombre de requêtes doit être à 16384 (4*4096). Il ne faut pas dépasser 1000 pour le max (ça ne sert à rien de monter au-delà) Ici, on parle bien de disques physique et non de LUN logique.

Petit détail piquant, le *min server* est le nombre de server d'I/O asynchrones lancés au démarrage du serveur avec les droits root. Toutes les applications peuvent s'en servir. Par contre, entre le min et le max, donc lancé au besoin, ils appartiennent à l'utilisateur qui en a besoin et ne sont jamais arrêtés. Je vous laisse imaginer lorsqu'il y à Sybase et Oracle sur le même serveur.

Les WAIT IO

Un PDF contenant des infos provenant d'IBM, disponible ici.

Liens

http://www.ibm.com/developerworks/aix/library/au-DirectIO.html
http://www-03.ibm.com/servers/aix/whitepapers/db_perf_aix.pdf

Documentations

Intitulé	Format
AIX 5L Performance Tools Handbook	
Tuning AIX pour Oracle	
Database Performance Tuning on AIX	
Certification Study Guide - pSeries AIX System Administration	
AIX 5L Practical Performance Tools and Tuning Guide	
AIX 5L Differences Guide	
Problem Solving and Troubleshooting in AIX 5L	
Aix Performance Tuning	

- Sites WEB :

Docs AX490AF

Outils pour les perfs AIX

Autres outils (non-AIX)

From:
https://unix-bck.ndlp.info/ - **Where there is a shell, there is a way**

Permanent link:
https://unix-bck.ndlp.info/doku.php/informatique:nix:ibm:ibm_aix_tuning

Last update: **2015/11/25 14:42**