
numactl

Contrôle la stratégie NUMA pour les processus ou le partage mémoire

numactl lance les processus avec une planification NUMA spécifique ou stratégie de placement mémoire. La stratégie est définie pour la commande héritée par tous ses enfants. De plus il peut définir une stratégie persistante pour les segments de mémoire partagée ou fichiers.

Les nœuds peuvent être spécifiés comme N,N,N ou N-N ou N,N-N ou N-N,N-N. Les nœuds relatifs peuvent être spécifiés comme +N,N,N ou +N-N ou +N,N-N. + indique que les numéros de nœuds sont relatifs au jeu de processus de nœud permis dans son cpuset actuel. !N-N indique l'inverse de N-N. Au lieu d'un numéro de nœud, un nœud peut également être :

netdev :DEV Le nœud connecté au périphérique réseau dev

file :PATH Le périphérique block PATH

ip :HOST le nœud du périphérique réseau de HOST

block :PATH le nœud du périphérique block PATH

pci : [seg :] bus :dev [:func] le nœud du périphérique PCI

Noter que block résout les noms de périphériques block du kernel seulement pour les noms udev dans le fichier dans /dev.

Les paramètres de stratégie sont :</br>

- all, -a** reset le cpuset par défaut, afin qu'il soit possible d'utiliser tous les cpu/nœuds pour les paramètres de stratégie suivante.
- interleave=nodes, -i nodes** Définit une stratégie d'entrelacement mémoire. La mémoire sera allouée en utilisant un round robin dans les nœuds. Plusieurs nœuds peuvent être spécifiés dans `-interleave`, `-mbind` et `-cpunodebind`
- mbind=nodes, -m nodes** Alloue la mémoire uniquement dans les nœuds. L'allocation échoue s'il n'y a pas suffisamment de mémoire disponible dans ces nœuds.
- cpunodebind=nodes, -N nodes** Exécute le processus seulement sur les cpus spécifiés. Accepte les numéros de cpu comme affichés dans le champ processor de /proc/cpuinfo ou les cpus relatifs dans le cpuset courant. "all" est accepté, qui signifie tous les cpu du cpuset.
- localalloc, -l** Alloue toujours dans le nœud courant
- preferred=node** Alloue la mémoire de préférence dans le nœud spécifié, mais si la mémoire ne peut être allouée, utilise d'autres nœuds.
- show, -s** Affiche les paramètres de stratégie numa du processus courant
- hardware, -H** Affiche l'inventaire des nœuds disponibles dans le système.

numactl peut définir une stratégie pour un segment de mémoire partagée SYSV ou un fichier dans shmfs/hugetlbfs. Cette stratégie est persistante et sera toujours utilisée par tous les mappages de cette mémoire partagée. L'ordre des options a une signification ici. La spécification doit au moins inclure `-shm`, `-shmid`, `-file` pour spécifier le segment de mémoire partagée ou le fichier et une stratégie mémoire comme décrite plus haut (`-interleave`, `-localalloc`, `-preferred`, `-mbind`).

- huge** En créant un segment de mémoire partagée SYSV, utilise des grandes pages. Seulement valide avant `-shmid` ou `-shm`
- offset** Spécifie l'offset dans le segment de mémoire partagé. défaut : 0. Les unités valides sont m (Mo), g (Go), k (Ko), sinon est spécifié en octets.
- strict** Donne une erreur quand une page dans l'aire de stratégie dans le segment de mémoire partagée a déjà généré une faute avec un conflit de stratégie. par défaut, ignore silencieusement.
- shmmode shmmode** seulement valide avant `-shmid` ou `shm` en créant un segment de mémoire partagé avec un mode numérique shmmode.
- length length** Applique la stratégie à la longueur de plage du segment de mémoire partagée ou crée la longueur du segment. par défaut, utilise la longueur restante requise quand un segment de mémoire partagée est créé et spécifie la longueur du nouveau segment. Les unités de valeurs sont m, g, k ou en octet par défaut.

- shmid id** Créé ou utilise un segment de mémoire partagé avec l'id numérique spécifié
- shm shmkeyfile** Crée ou utilise un segment de mémoire partagé, avec l'id généré en utilisant `fofk(3)` depuis `shmkeyfile`.
- file tmpfsfile** Définis la stratégie pour un fichier dans `tmpfs` ou `hugetlbfs`
- touch** Touch les pages pour forcer la stratégie. Ne les touch pas par défaut, la stratégie est appliquée quand une application mappe et accède à une page.
- dump** Dump la stratégie dans la plage spécifiée
- dump-nodes** Dump tous les nœuds de la plage spécifique (très verbeux !)

Les spécifieurs de nœud valides sont :

- all** tous les nœuds
- number** Numéro de nœud
- number1{,number2}** nœud `number1` et nœud `number2`
- number1-number2** plage de nœuds
- ! nodes** Inverse la sélection

Exemples

Lance `myapplic` sur les cpu 0-4 et 8-12 du `cpuset` courant :

```
numactl -physcpubind=+0-4,8-12 myapplic arguments
```

Lance une grosse base avec sa mémoire entrelacée sur tous les CPU :

```
numactl -interleave=all bigdatabase arguments
```

Lance le process dans le nœud 0 avec la mémoire allouée dans les nœuds 0 et 1 :

```
numactl -cpunodebind=0 -membind=0,1 process
```

Idem, mais avec une option (-l) qui peut être confondue avec une option `numactl` :

```
numactl -cpunodebind=0 -membind=0,1 - process -l
```

Lance `network-server` dans le nœud de périphérique réseau `eth0` avec sa mémoire également dans le même nœud :

```
numactl -cpunodebind=netdev :eth0 -membind=netdev :eth0 network-server
```

Définis le nœud 1 comme préféré et affiche l'état résultant :

```
numactl -preferred=1 numactl -show
```

Entrelace toute la région mémoire partagée `sysv` spécifiée par `/tmp/shmkey` sur tous les nœuds :

```
numactl -interleave=all -shm /tmp/shmkey
```

Place un fichier `tmpfs` dans 2 nœuds :

```
numactl -membind=2 dd if=/dev/zero of=/dev/shm/A bs=1M count=1024
```

```
numactl -membind=3 dd if=/dev/zero of=/dev/shm/A seek=1024 bs=1M count=1024
```

Réinitialise la stratégie pour le fichier de mémoire partagée `file` à la stratégie par défaut `localalloc` :

```
numactl -localalloc /dev/shm/file
```

Notes

Les commande ne sont pas exécutée en utilisant un shell. Utiliser `sh -c` pour cela.

Définir la stratégie pour un fichier `hugetlbfs` ne fonctionnent pas actuellement.

Les segment de mémoire partagée plus grand que l'espace d'adressage ne peut pas être complètement sous stratégie.