

Stockage campus

1 - Description :

Serveur SUN X4240, 8 cœurs AMD64 2 GHz, 16 Go de mémoire

3 contrôleurs RAID Intel StorageTek 2 ports 4 canaux SAS 3 Gbps

4 baies de 24 disques SATA multipath de 1 To. Chacune est raccordée par un lien 4 canaux SAS au serveur.

Volumes RAID6, avec un disque « spare » par baie.

8 interfaces ethernet 1 Gbps, une par utilisateur.

CentOS 5.7

2 - Principe :

Mettre à disposition de chaque unité un volume disque iSCSI de 11 To par un réseau de niveau 2 (vlan) dédié, et hermétique. Il sera acheminé jusqu'à votre bâtiment, si possible en utilisant un lien physique dédié afin de cibler une bande passante maximale. Eviter tout routage de niveau 3 qui diminuerait les performances.

Distribuer un volume plutôt qu'un système de fichier pour vous laisser la liberté de sa gestion, de son organisation. Le système de fichier est de votre responsabilité, nous n'y aurons pas accès. Il vous appartient d'en assurer la pérennité.

Terminologie iSCSI : « target » désigne le serveur de volume – c'est-à-dire le serveur campus – et « initiator » désigne le client – c'est-à-dire votre serveur.

3 - Mise en œuvre par les unités :

Le tableau ci-dessous résume toutes les informations nécessaires à la mise en œuvre de votre initiator iSCSI, en prenant soin de mettre une de ses interfaces dans le bon VLAN.

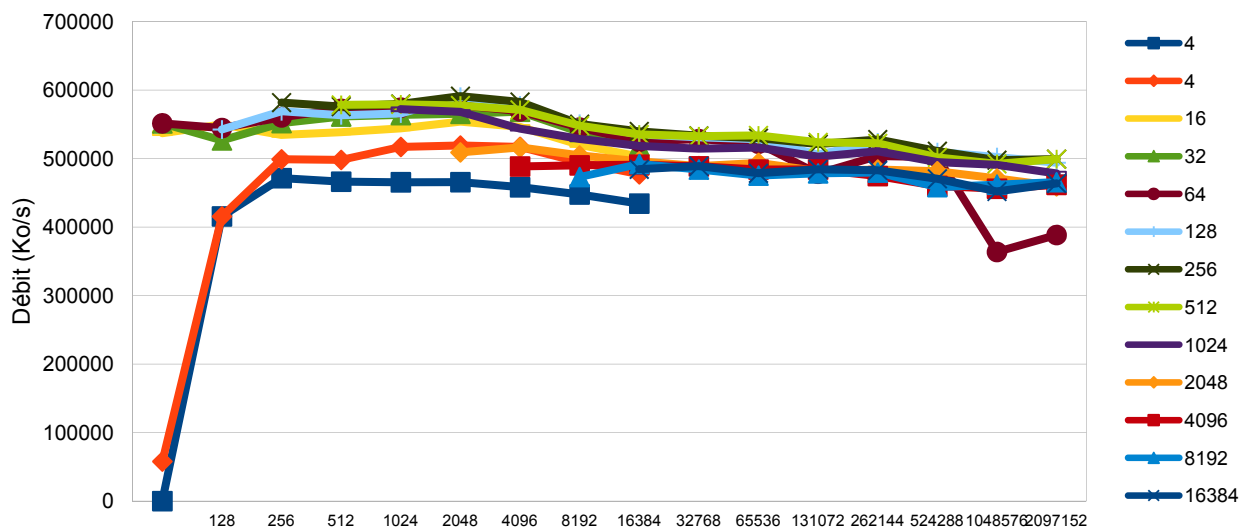
Unité	Nom du volume	Numéro du vlan	Adresse du target iSCSI	Adresse de votre initiator iSCSI	Taille brute
DR13	iqn.2011-01.fr.cnrs.montp:deleg	241	10.25.1.2	10.25.1.3	11 To
SIC	iqn.2011-01.fr.cnrs.montp:sic	242	10.25.2.2	10.25.2.3	11 To
IGF	iqn.2011-01.fr.cnrs.montp:igf	243	10.25.3.2	10.25.3.3	11 To
IGH	iqn.2011-01.fr.cnrs.montp:igh	244	10.25.4.2	10.25.4.3	11 To
CEFE	iqn.2011-01.fr.cnrs.montp:cefe	245	10.25.5.2	10.25.5.3	11 To
MRI	iqn.2011-01.fr.cnrs.montp:mri	246	10.25.6.2	10.25.6.3	11 To
SRC	iqn.2011-01.fr.cnrs.montp:src	247	10.25.7.2	10.25.7.3	6 To

Le VLAN pour votre unité sera acheminé jusqu'au commutateur réseau campus de votre bâtiment, votre serveur y sera également raccordé pour garantir la bande passante.

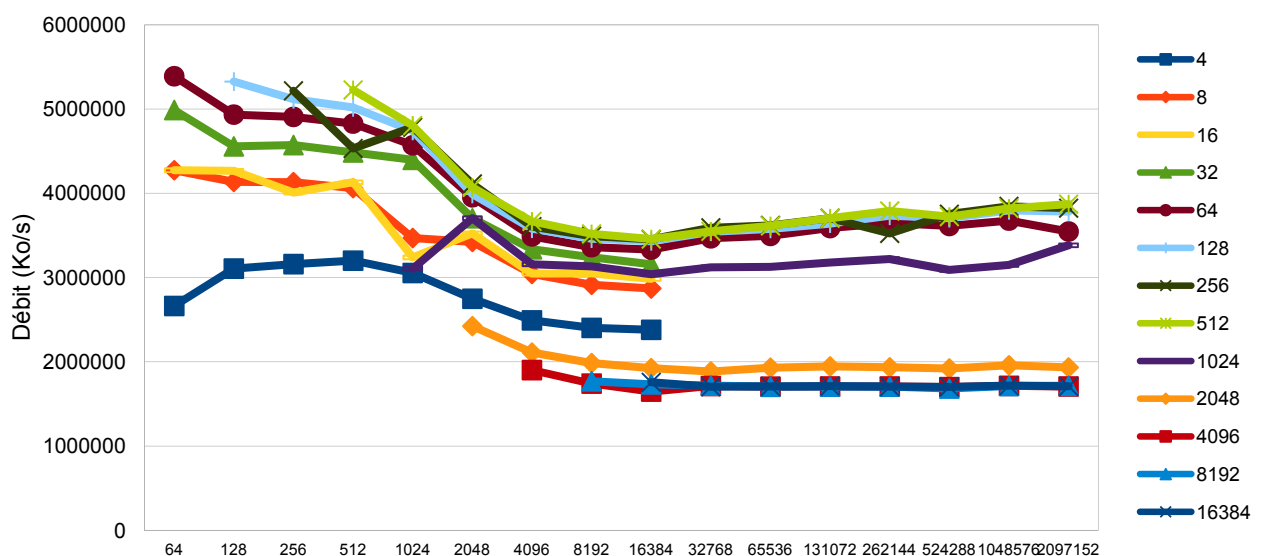
4 - Performances :

La première mesure concerne les performances entre le serveur iSCSI campus et les baies. Tests effectués avec l'outil iotop, référence dans le domaine, qui procède à des milliers d'opérations de lecture/écriture avec des tailles de fichiers différentes et des tailles de tampon différentes. Voir <http://www.iozone.org/>. Les données brutes sont disponibles.

2 ports SAS - Ecriture - Raid 6

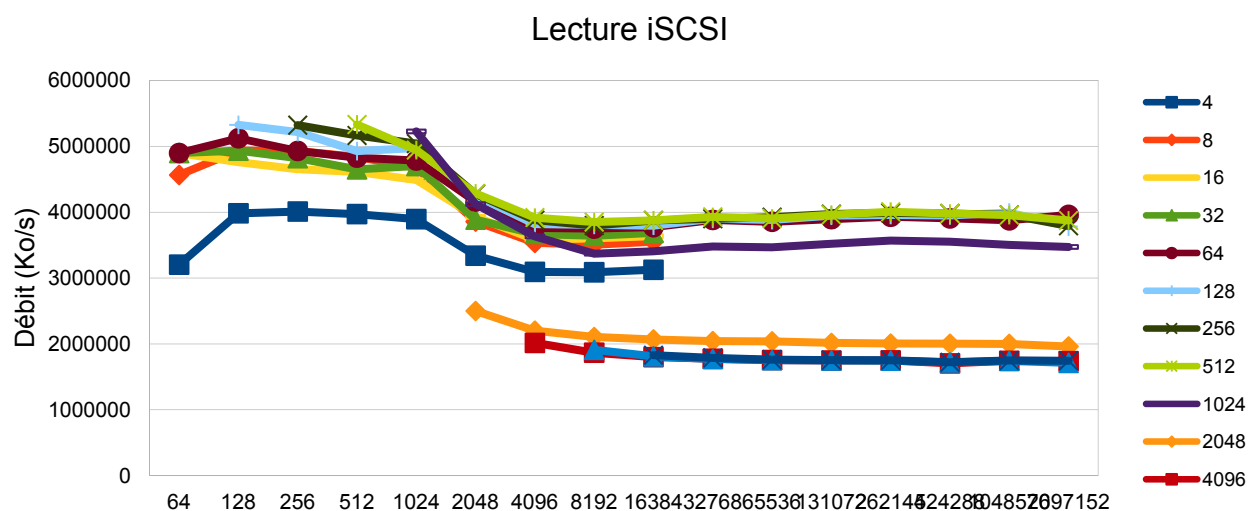
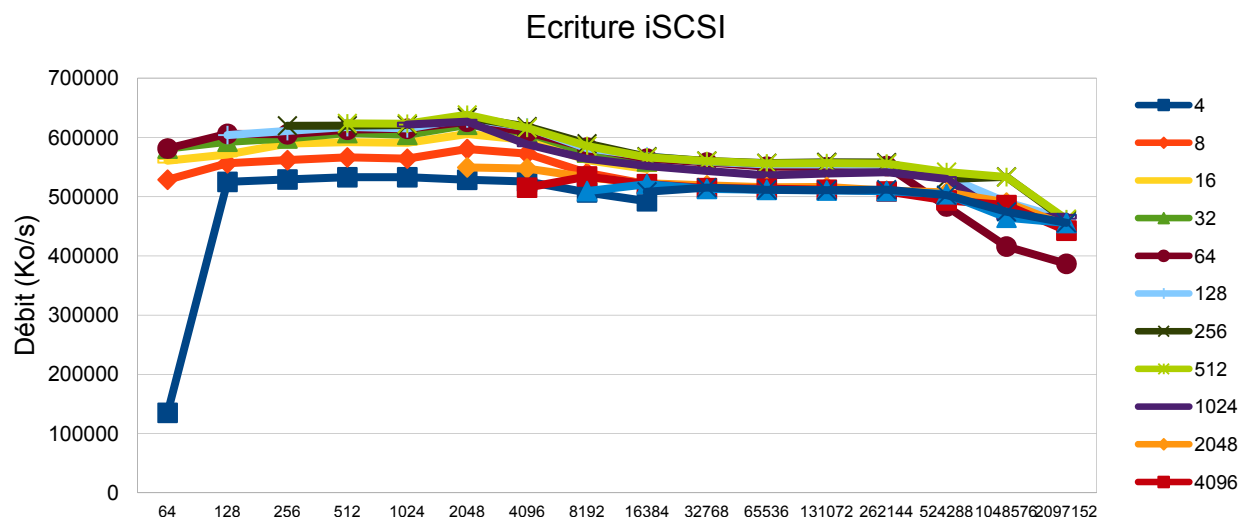


2 ports SAS - Lecture - Raid 6

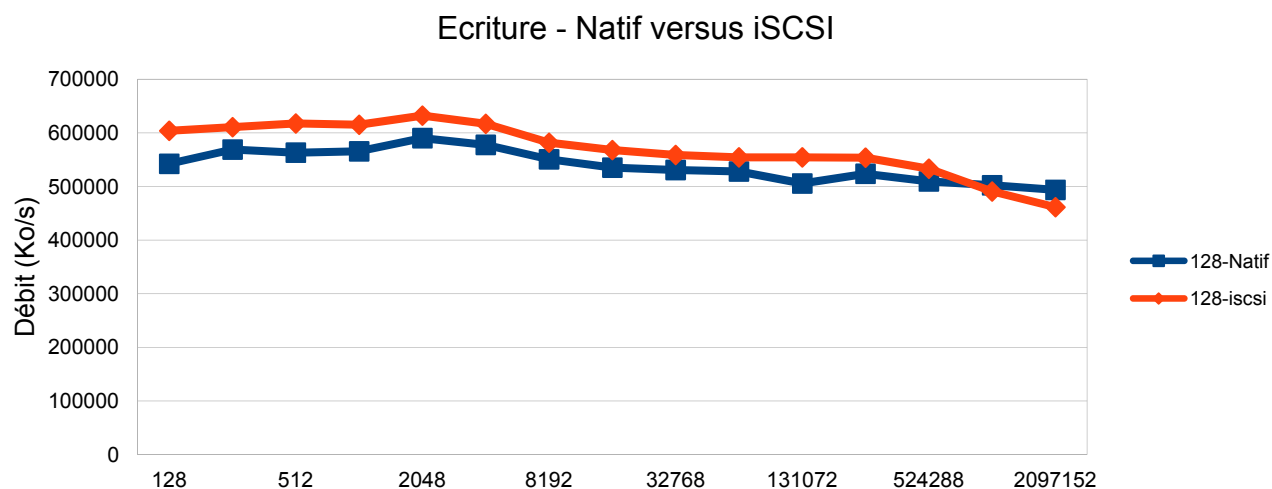


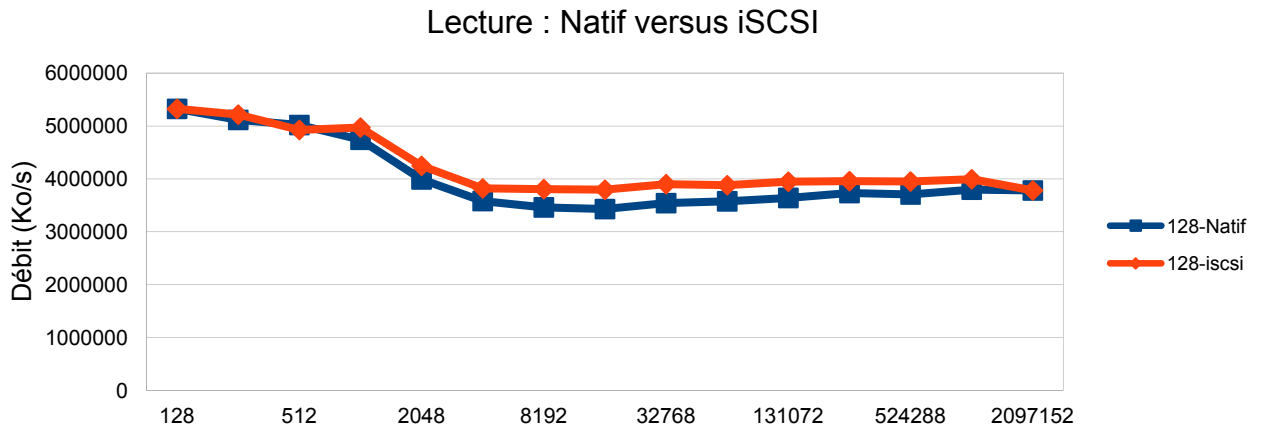
En Abscisse, la taille des fichiers, jusqu'à 2 Go.

La seconde série de mesure est effectuée depuis un serveur de fichier linux initiator iSCSI d'un des volumes de 11 To, après fabrication d'une partition pour l'ensemble, et création de son système de fichier type ext3.



Comparaison des performances entre les deux mesures, par une extraction des données ci-avant :





Pour contrer les mécanismes de cache de Linux, ci-dessous des tests basiques sur des fichiers plus conséquents :

```
[root@test essai]# time dd if=/dev/zero of=t8 bs=1M count=18000
18000+0 enregistrements lus
18000+0 enregistrements écrits
18874368000 octets (19 GB) copiés, 151,163 seconde, 125 MB/s
real    2m33.615s
user    0m0.021s
sys     0m47.209s

[root@test essai]# time dd if=/dev/zero of=t9 bs=1M count=36000
36000+0 enregistrements lus
36000+0 enregistrements écrits
37748736000 octets (38 GB) copiés, 304,039 seconde, 124 MB/s
real    5m4.492s
user    0m0.043s
sys     1m34.356s
```

Mesure au niveau de l'interface réseau :

Elle montre que le débit atteint est de 992 Mbps, soit 124 Moctets par seconde, constant sur 5 minutes de transfert. Les trames géantes - supérieures à 1500 octets - sont activées.

```
Last 300 seconds input:  15021 packets/sec 124032442 bytes/sec 99%
Last 300 seconds output:  7403 packets/sec  557621 bytes/sec 1%
```

5 - Mise en œuvre avec un système linux :

Exemple pour le volume « DR13 » sur un initiator dont l'adresse IP est 10.25.1.3

```
iscsiadm -m discovery -t sendtargets -p 10.25.1.2:3260
```

```

iscsiadm -m node
    10.25.1.2:3260,1 iqn.2011-01.fr.cnrs.montp:deleg
iscsiadm --mode node --targetname iqn.2011-01.fr.cnrs.montp:deleg -p
10.25.1.2:3260 --login
iscsiadm -m session
    tcp: [1] 10.25.1.2:3260,1 iqn.2011-01.fr.cnrs.montp:deleg
fdisk -l
    [...]
    Disque /dev/sdb: 12094.6 Go, 12094627905536 octets
    255 heads, 63 sectors/track, 1470421 cylinders
    Unités = cylindres de 16065 * 512 = 8225280 octets
    [...]
parted /dev/sdb
    mklabel gpt
    mkpart ext3 0 -1
    quit
mkfs.ext3 -j /dev/sdb1
tune2fs -m 0 /dev/sdb1
df -k
    [...]
    /dev/sdb1      11625822076 161952  11625660124 1%   /mnt/essai

```

6 - Exploitation :

Signaler tout problème ou poser vos questions par l'interface « extranet » de la Délégation :

https://extranet.dr13.cnrs.fr/appels/index.php?q=demande&service=1&id_cat=76

A noter qu'en cas de programmation de coupure électrique sur le local hébergeant ce stockage, le serveur sera coupé la veille en horaire ouvré.

Ne pas diffuser ce document sans autorisation.