GlusterFS over InfiniBand

Premiers résultats

EQ <emmanuel.quemener@ens-lyon.fr>

Plate-forme

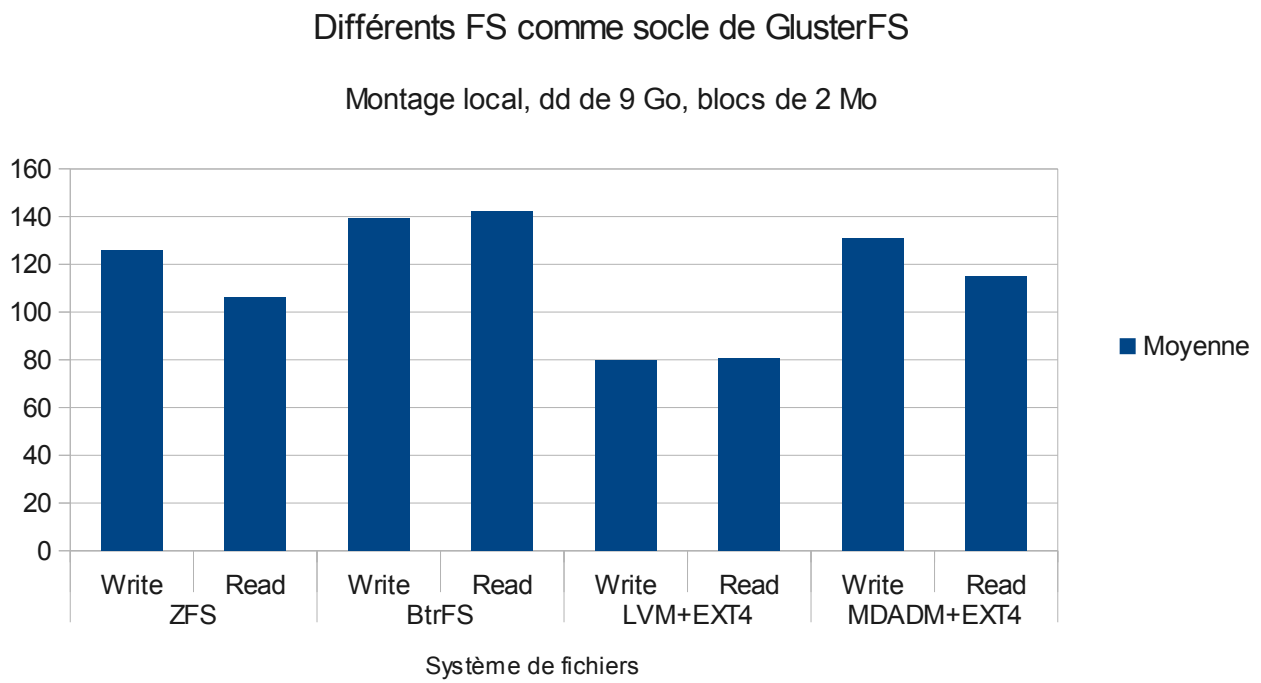
- OS : Debian Wheezy 7.0 (gelée)
- Les nœuds
  - 48 v22z avec 8 Go de RAM, 2 bicoeurs
  - 16 x41 avec 16 Go de RAM, 2 quadricoeurs
  - 8 v40z avec 16/32 Go de RAM, 4 monocoeurs
- Les serveurs
  - 1 R510 avec 12 SATA 3.5, 24 Go de RAM
  - 1 x41 avec 8 SATA 2.5, 16 Go de RAM
- Le réseau
  - Une interconnexion GE entre nœuds (sur 4 commutateurs)
  - Une matrice InfiniBand DDR (x41) et SDR (v22z et v40z)

Les Tests

- Sur les agrégations de disques
- Sur les interconnexions
- Sur la performance locale
- Sur la performance distante
- Sur le passage à l'échelle

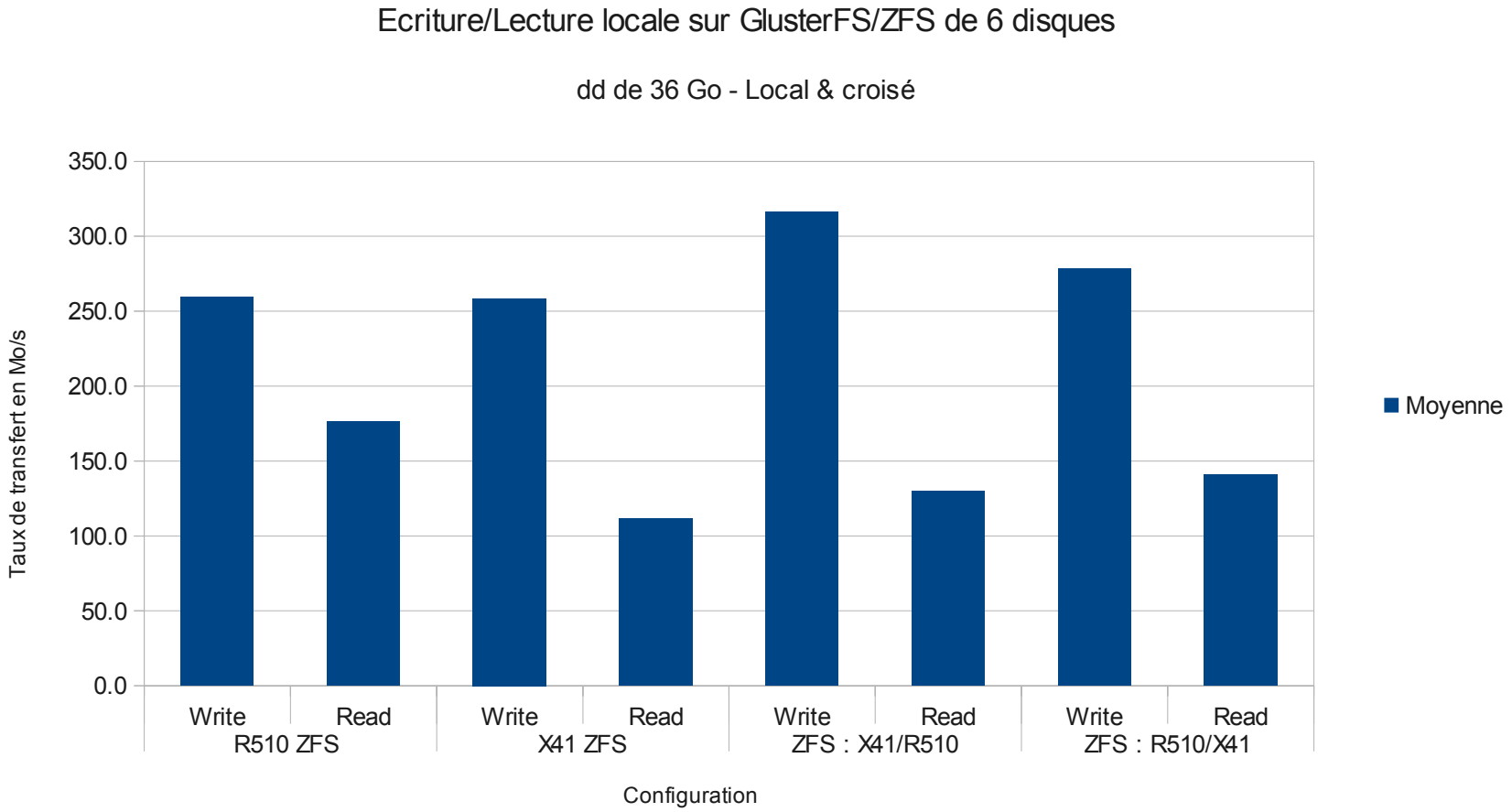
Agréger les disques, mais comment

- ZFS
- BtrFS
- LVM+Ext4
- MDADM+Ext4
- Sécurité des données
  - Type Raid5 (sauf BtrFS)



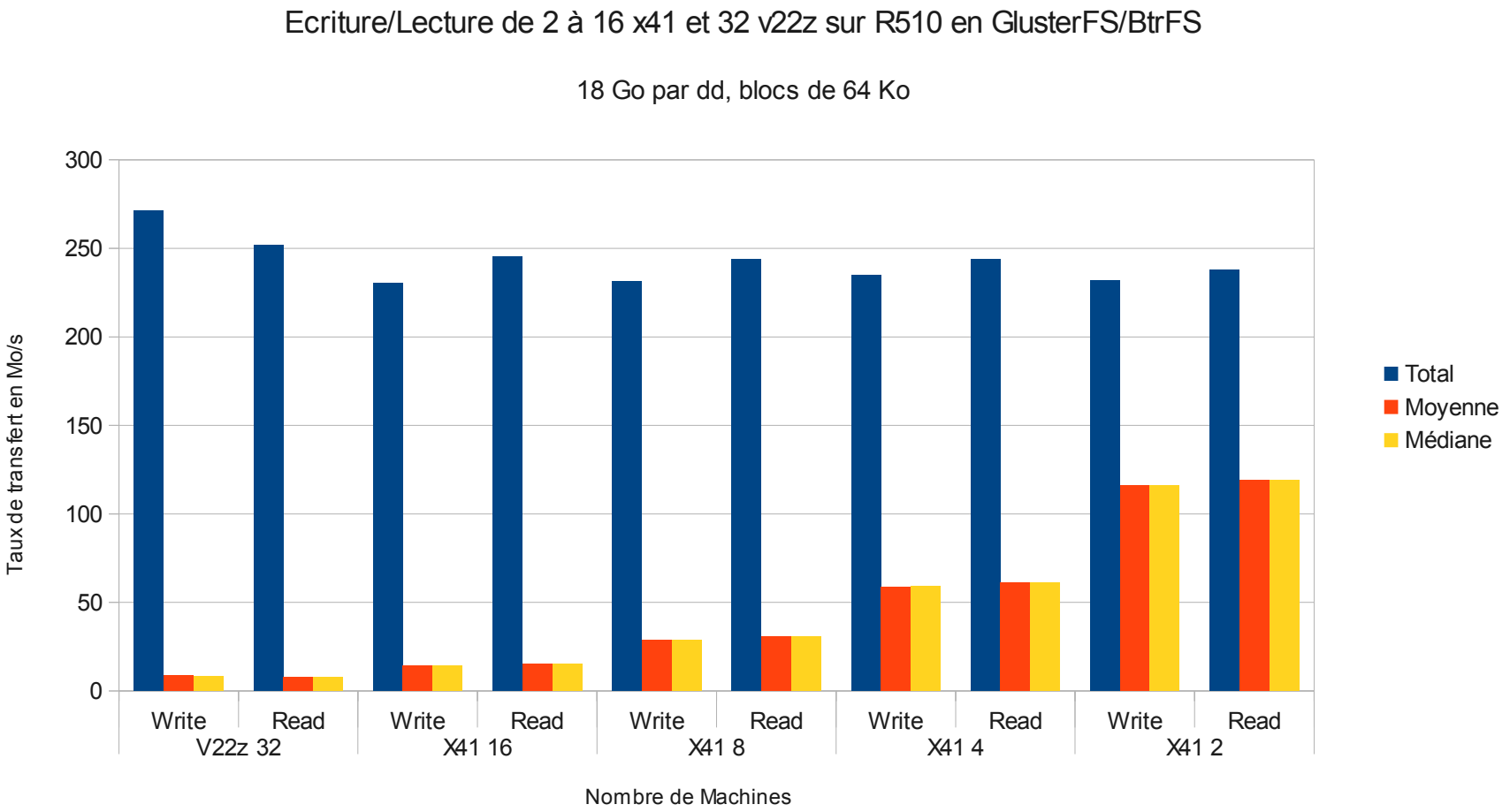
L'usage de serveurs récents

- 6 disques en ZFS mais 2 plates-formes différentes
- Montage local et croisé entre x41 et r510



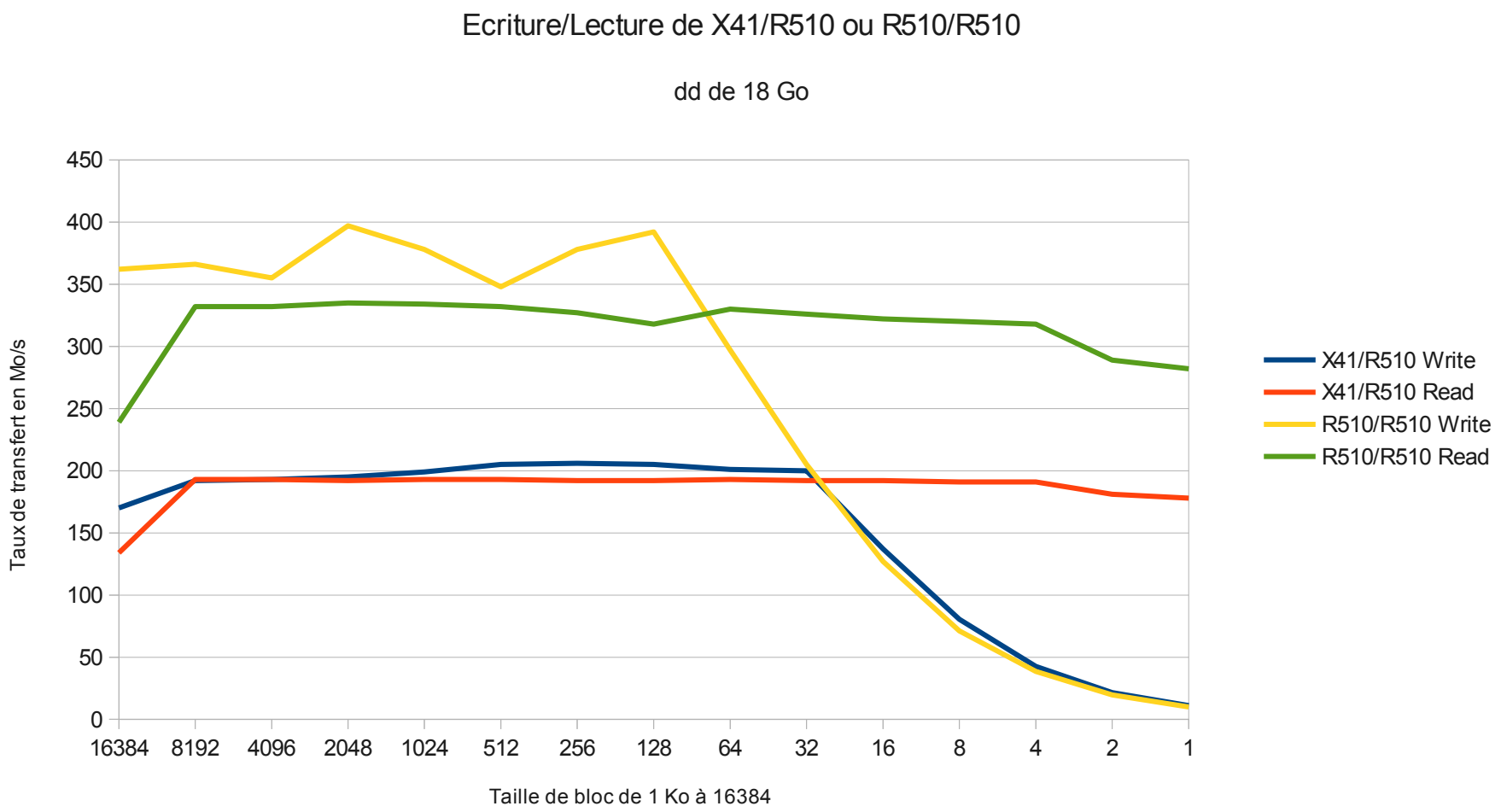
X nœuds sur 1 serveur

- De 1 à 32 nœuds sur un volume GlusterFS/BtrFS



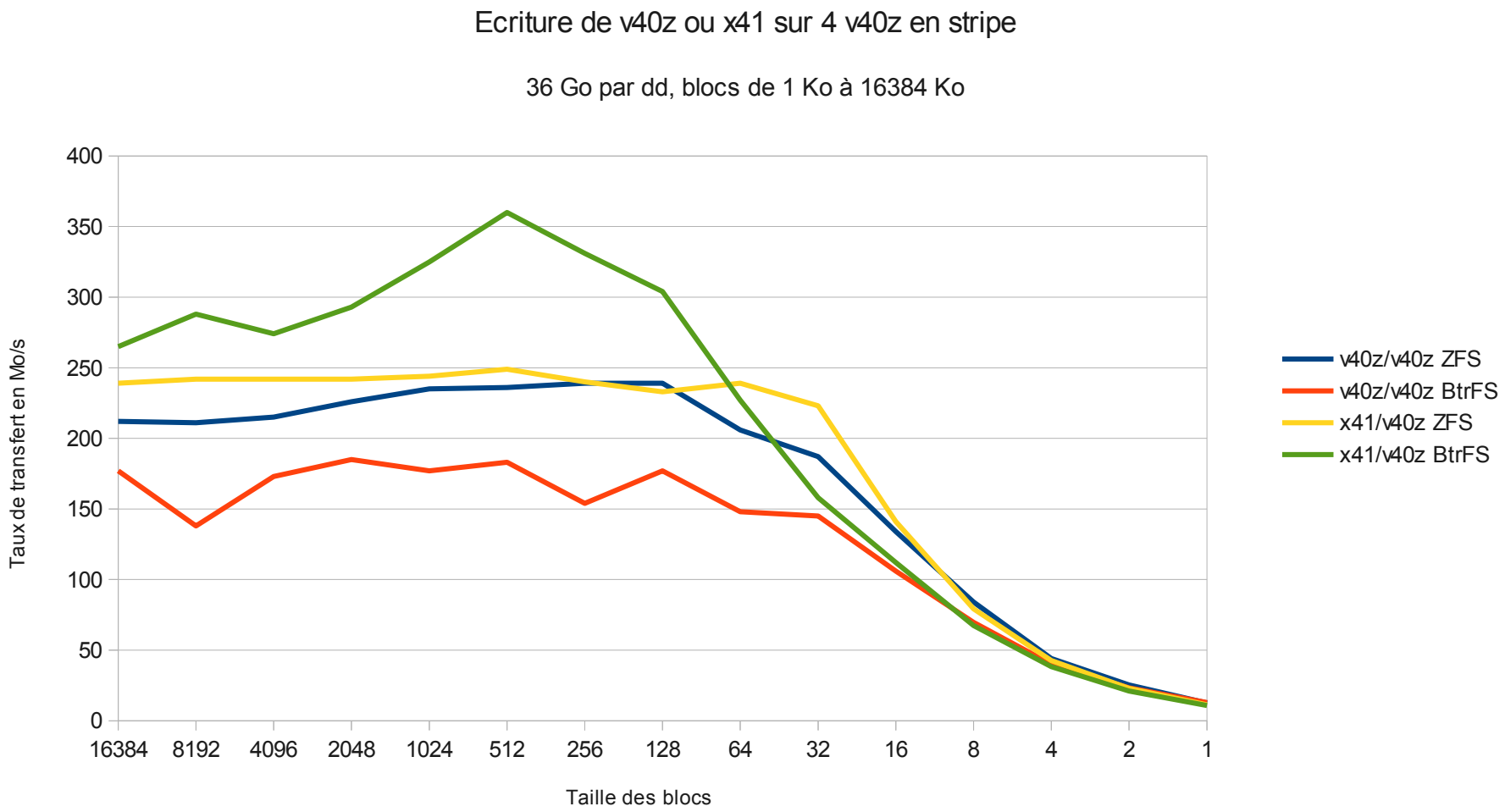
La taille de bloc comme critère

- Un nœud ou un montage local sur R510



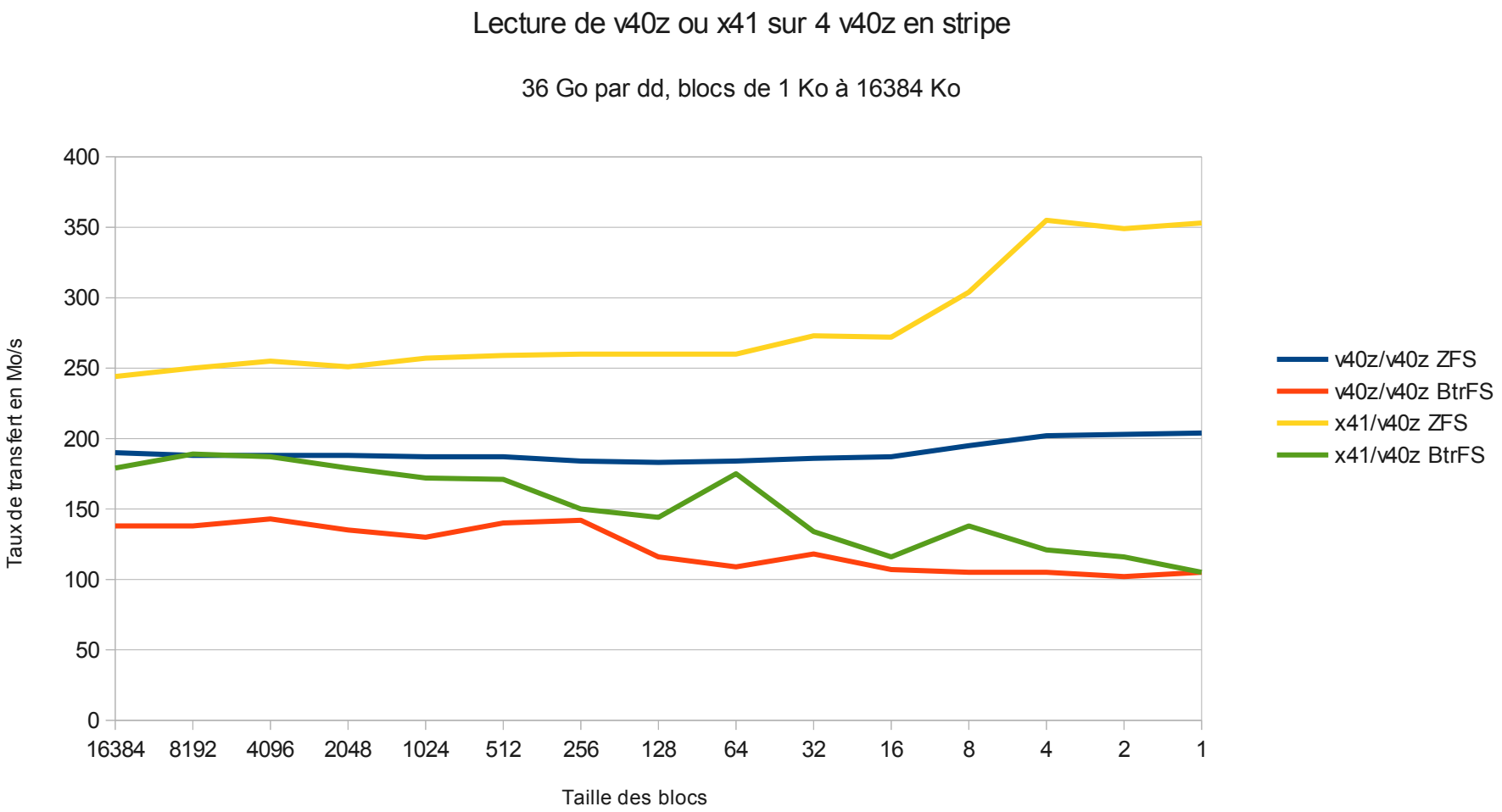
Le 4-Stripe : l'écriture

- Montage local ou distant par x41



Le 4-stripe : la lecture

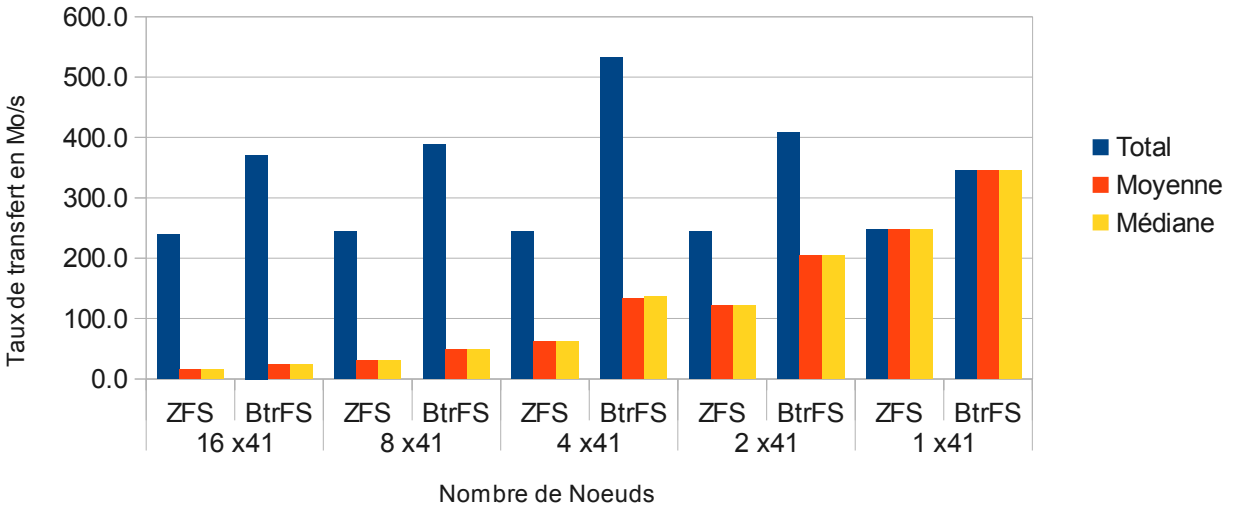
- Montage local ou distant par x41



Noeuds récents sur 4-stripe

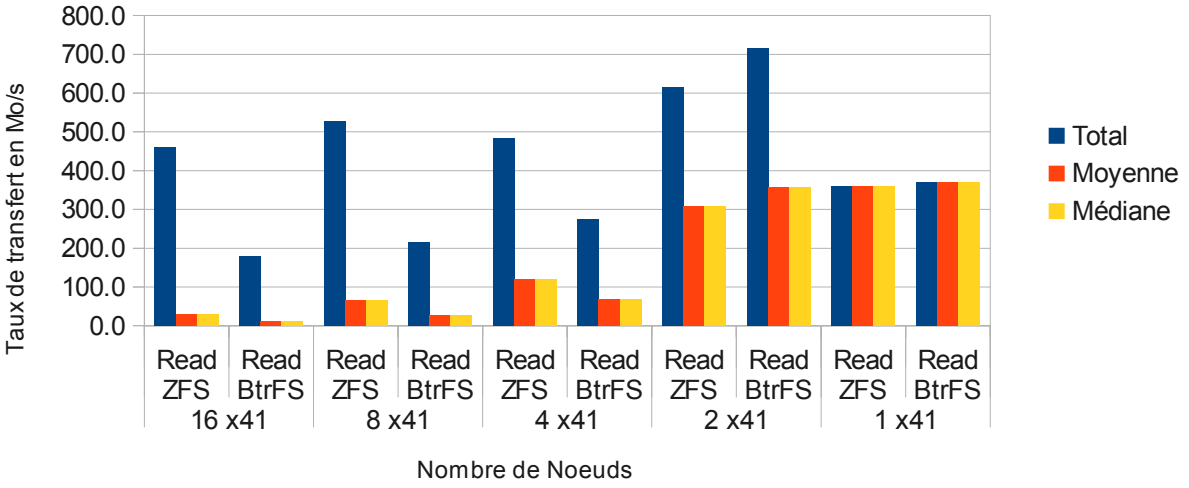
Ecriture de 1 à 16 x41 sur 4 v40z en stripe

36 Go par dd, blocs de 128 Ko



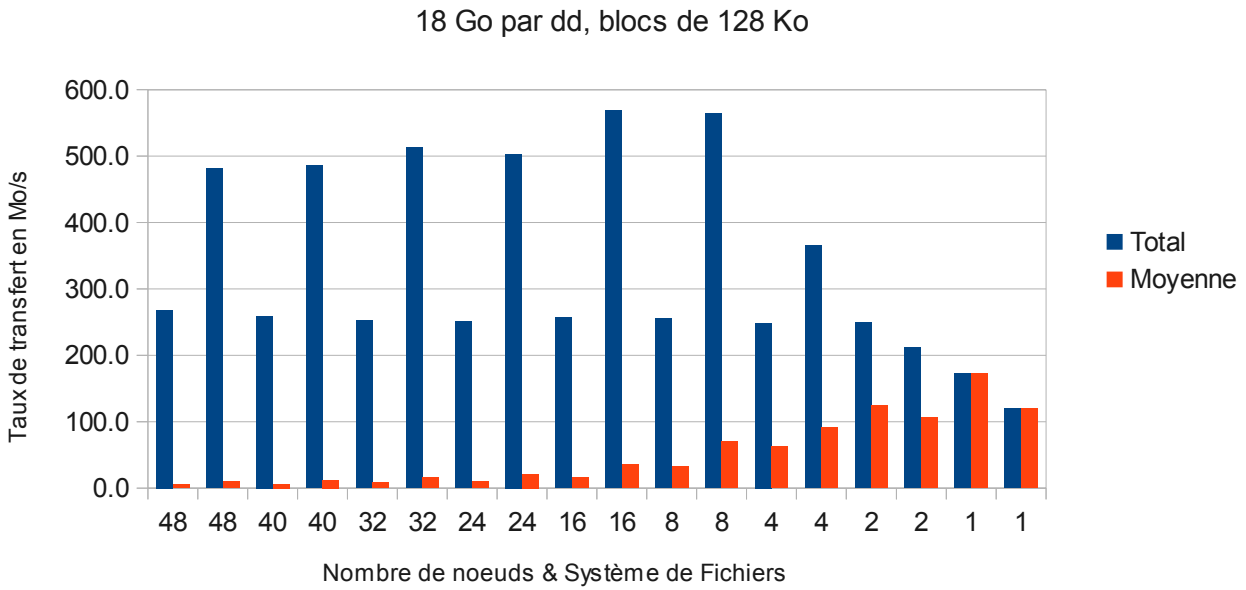
Lecture de 1 à 16 x41 sur 4 v40z en stripe

36 Go par dd, blocs de 128 Ko

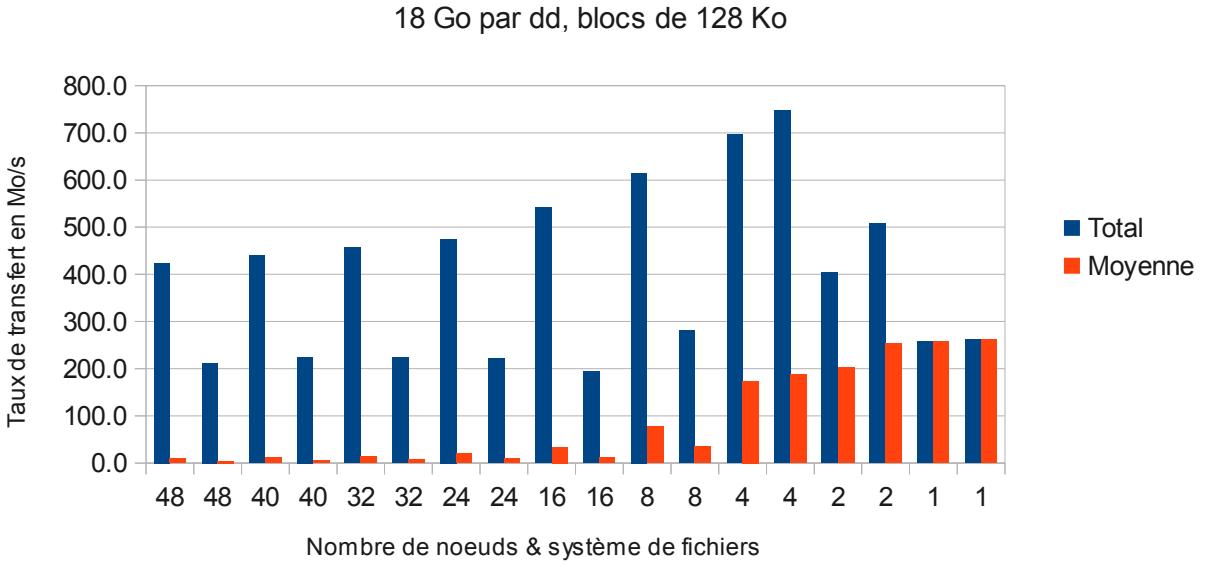


Noeuds anciens sur 4-stripe

Ecriture de 1 à 48 v22z sur 4 v40z en stripe



Lecture de 1 à 48 v22z sur v40z en stripe



Conclusions préliminaires

- FS : BtrFS ou ZFS comme choix (avant-gardistes...)
- Pertinence des tests
  - Pire des scénarii : écriture // des nœuds sur le volume
- Bonne scalabilité
  - Pas d'effondrement de performances
- Sur stockage unique récent : > 200 Mo/s
- Sur stockage unique ancien : > 120 Mo/s
- Sur stockage « stripe » ancien : > 400 Mo/s
  - Sur BtrFS en écriture & sur ZFS en lecture