

DDN[®]
STORAGE



**Big Data pour les sciences de la vie:
Accès aux données (performance,
flexibilité, sécurité et vie privée)**

2016/06

DataDirect Networks, Inc.

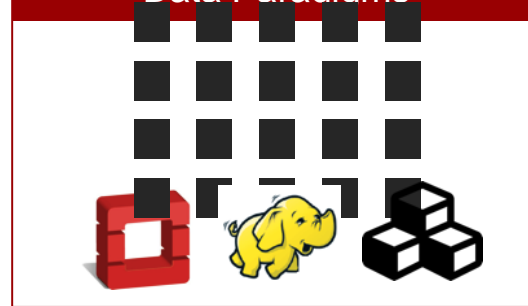
James Coomer jcoomer@ddn.com

Tendance dans le stockage haute performance

Continued Pressure to improve economics



Managing Scale and new Data Paradigms



Global Data Distribution



Managing performance for diverse workloads

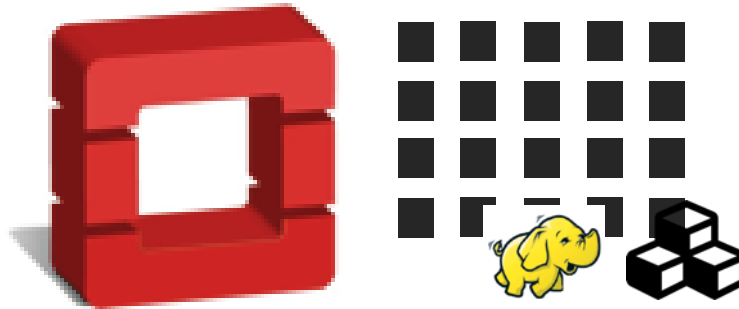


Sciences de la vie & stockage: tendances

Continued Pressure to improve economics



Managing Scale and new Data Paradigms



Global Data Distribution



Managing performance for diverse workloads

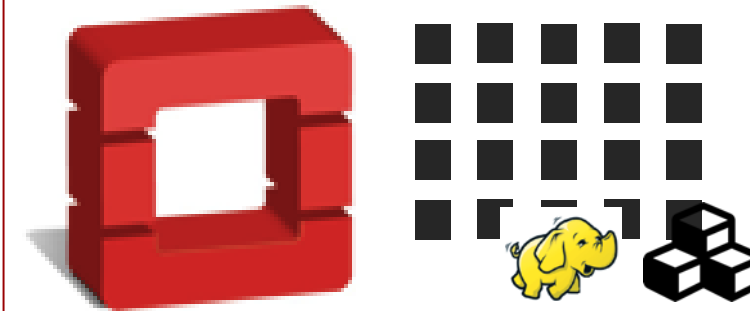


Sciences de la vie & stockage: tendances

Continued Pressure to improve economics



Managing Scale and new Data Paradigms



Global Data Distribution



Applying Security



Managing performance for diverse workloads

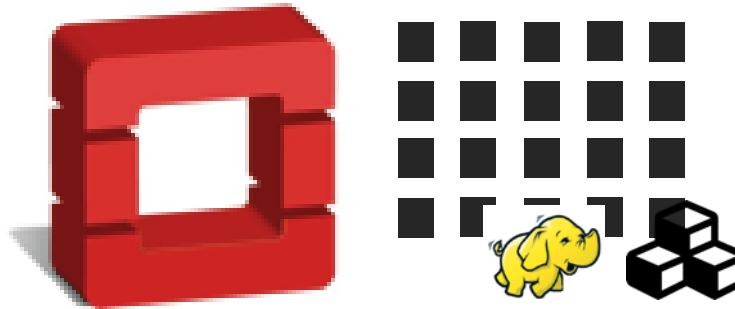


Sciences de la vie & stockage: tendances

Continued Pressure to improve economics



Managing Scale and new Data Paradigms



Global Data Distribution



Applying Security



Managing performance for diverse workloads

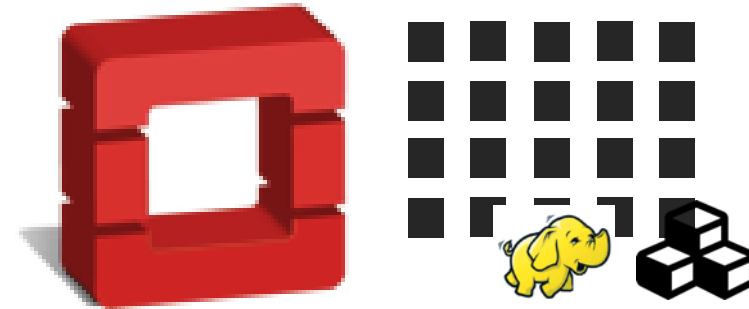


Les sciences de la vie représentent un défi pour le stockage aujourd'hui

Continued Pressure to improve economics



Managing Scale and new Data Paradigms



Global Data Distribution



Applying Security



Managing performance for diverse workloads



A propos de la Sécurité



Un peu de contexte

► La sécurité n'a pas été une priorité majeure des système de stockage HPC

- Conception orientée vers la Performance et la Scalabilité
- Pas de connexion directe à l'Internet et accès physique sécurisé

► Mais aujourd'hui, le stockage HPC N'EST PLUS limité à du scratch et aux répertoires des utilisateurs

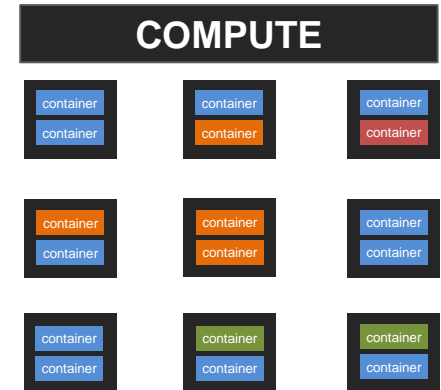
- Un même cluster sollicité pour plusieurs types d'utilisation
- Le matériel dédié n'est pas satisfaisant en terme coût/performance
- Les données sécurisées ne doivent être accessibles et visibles UNIQUEMENT aux personnes autorisées

► Authentification par utilisateur/nœud

► Capacité d'audit

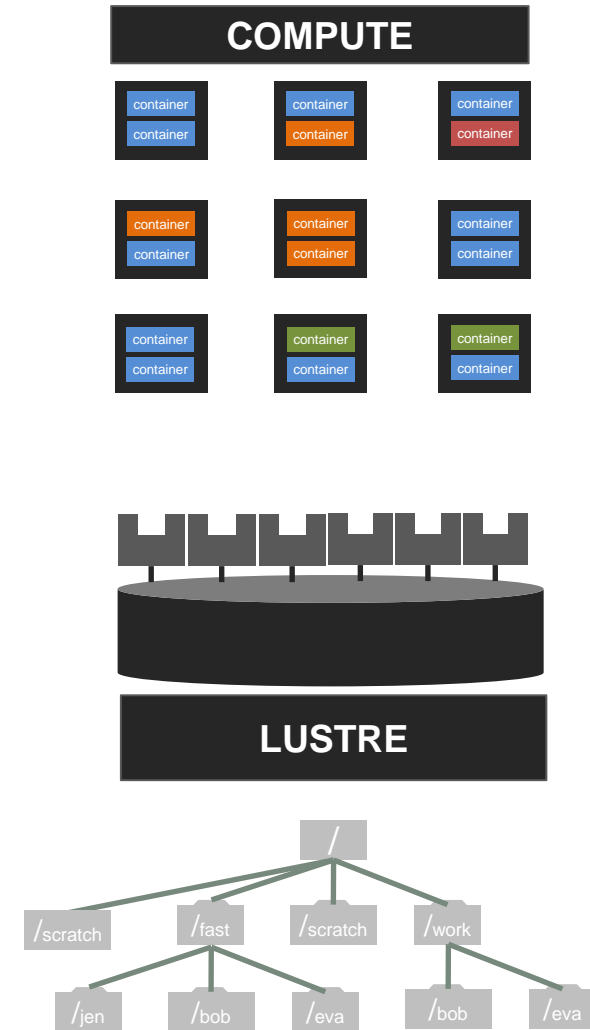
Isolement des données avec Docker

► La virtualiation dans les sciences de la vie apporte des capacités d'isolement (multi-tenancy)



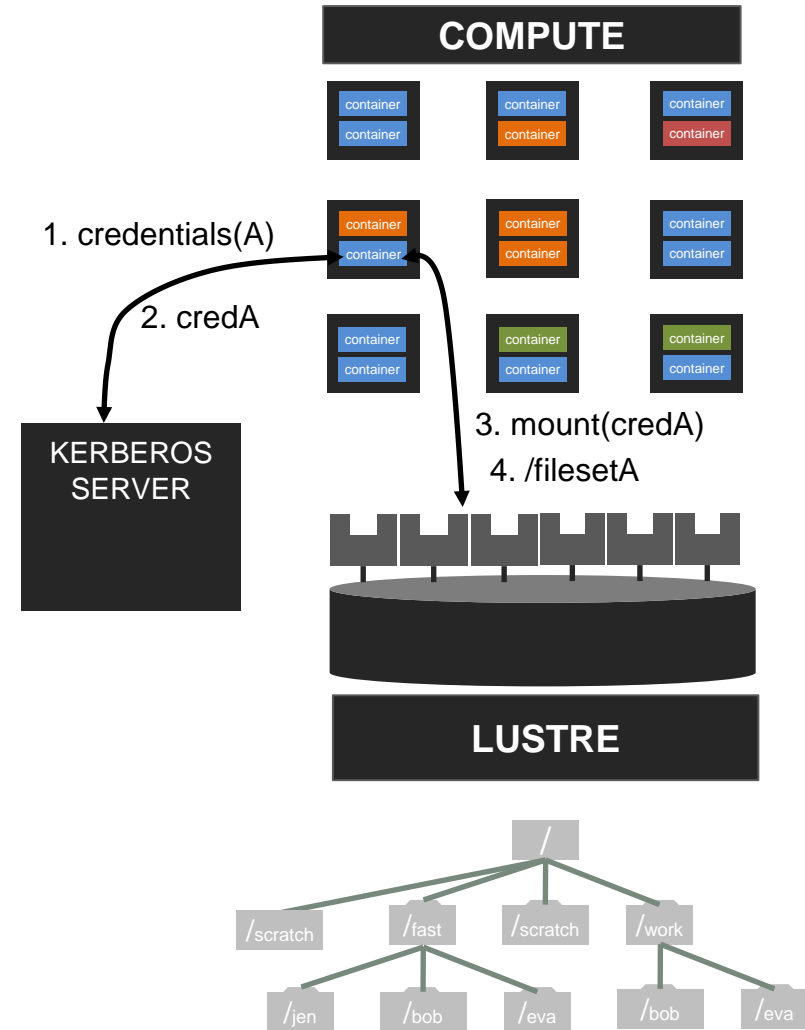
Isolement des données avec Docker

- ▶ The Lustre Filesystem provides good performance, but **breaks** the isolation
- ▶ all containers would see all data only being restricted by POSIX permissions/ACLs



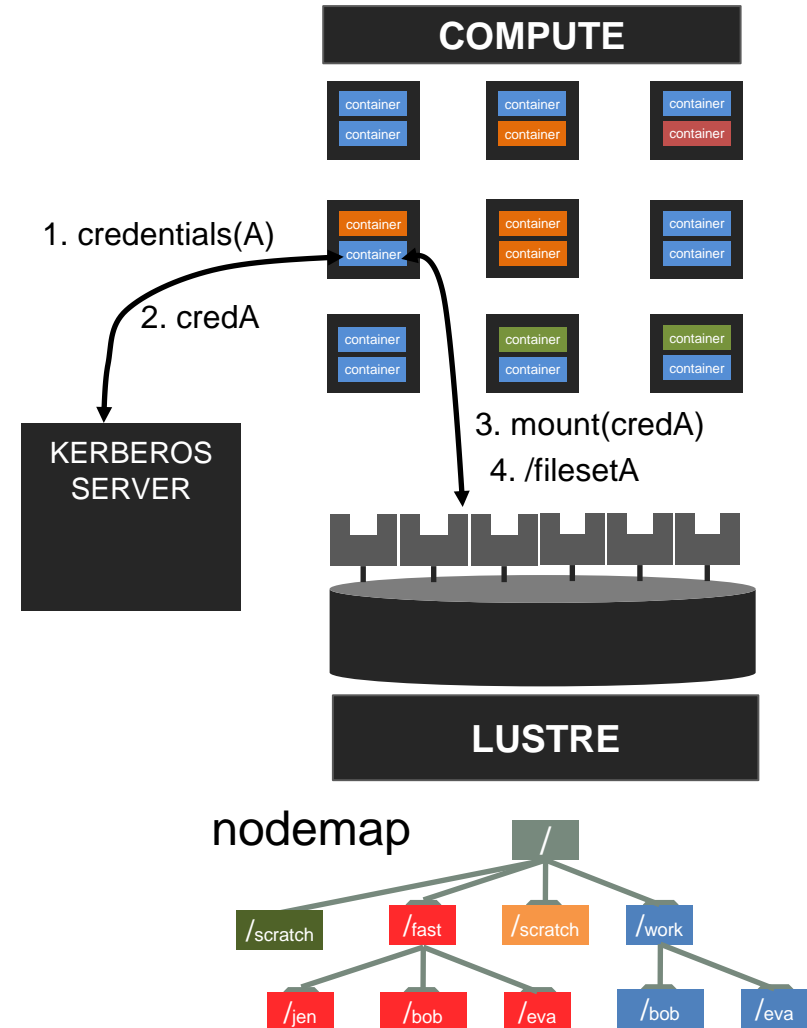
Isolement des données avec Docker

► Utilisation d'un Linux “Kerberisé” pour l'authentification des containers Dockers auprès du système de fichiers



Isolement des données avec Docker

- ▶ Deux éléments clef pour isoler les données :
- ▶ Notion de nodemap, associant des filesets avec NIDs (Lustre containers)
- ▶ Utilisation de montage de sous-répertoires



Performance des clients Lustre dans des containers



Les sciences de la vie représentent un défi pour le stockage aujourd'hui

Managing performance for diverse workloads



Comment relever ce défi ?

▶ Petites E/S

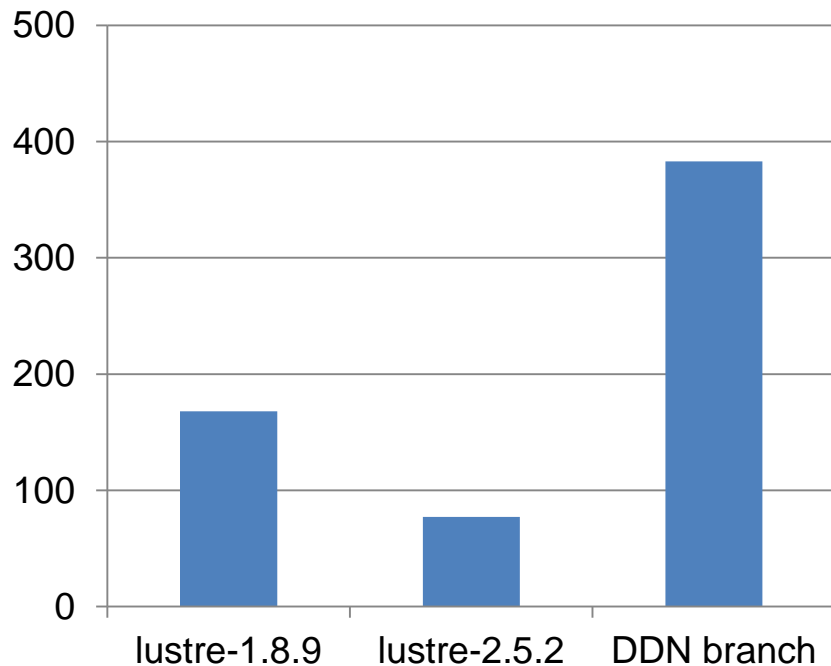
▶ Petits Fichiers

▶ mmap() E/S

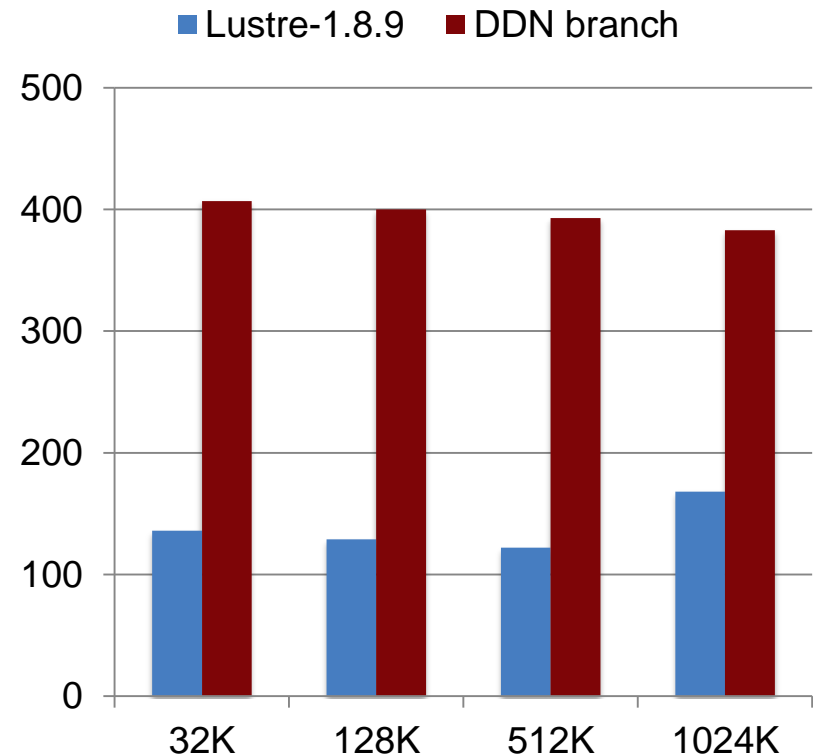
- Souvent suite à l'utilisation de Java
- GPFS peut se révéler sous-optimal dans ces conditions:
"La structure du noyau Linux implique pour GPFS un volume additionnel d'opérations de synchronisation afin de maintenir la sémantique de mmap. Ces opérations nuisent aux performances."

Exemple d'appel E/S mmap()

mmap() Read Performance (1MB Block Size)



mmap() Read Performance (32K to 1 MB Block Size)

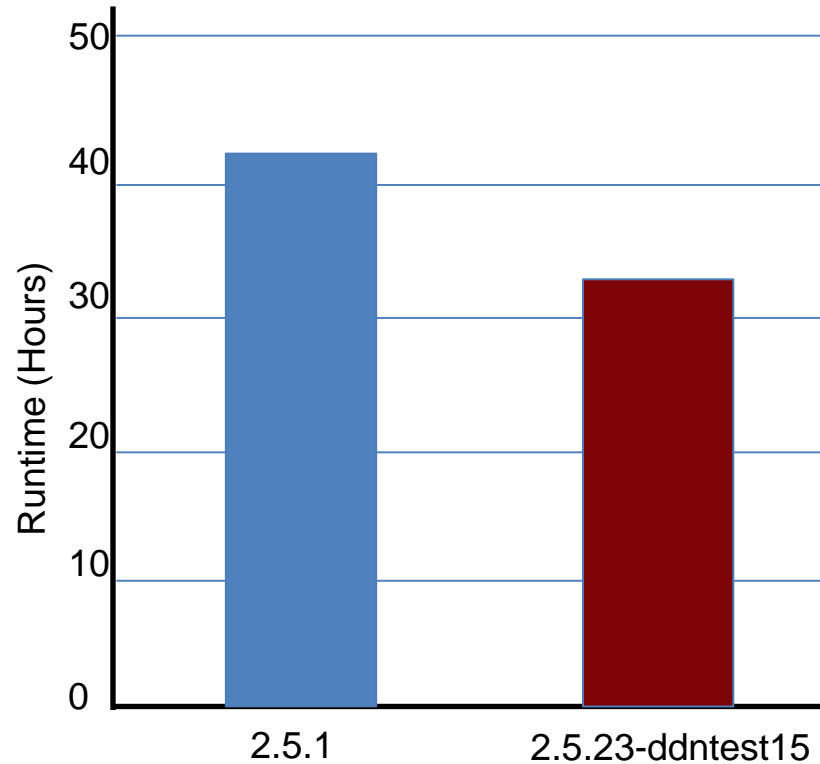


Sanger Institute Benchmarks

Samtools
20% plus rapide
grace aux
optimisations
Lustre de DDN

Lustre 2.5 Client Performance

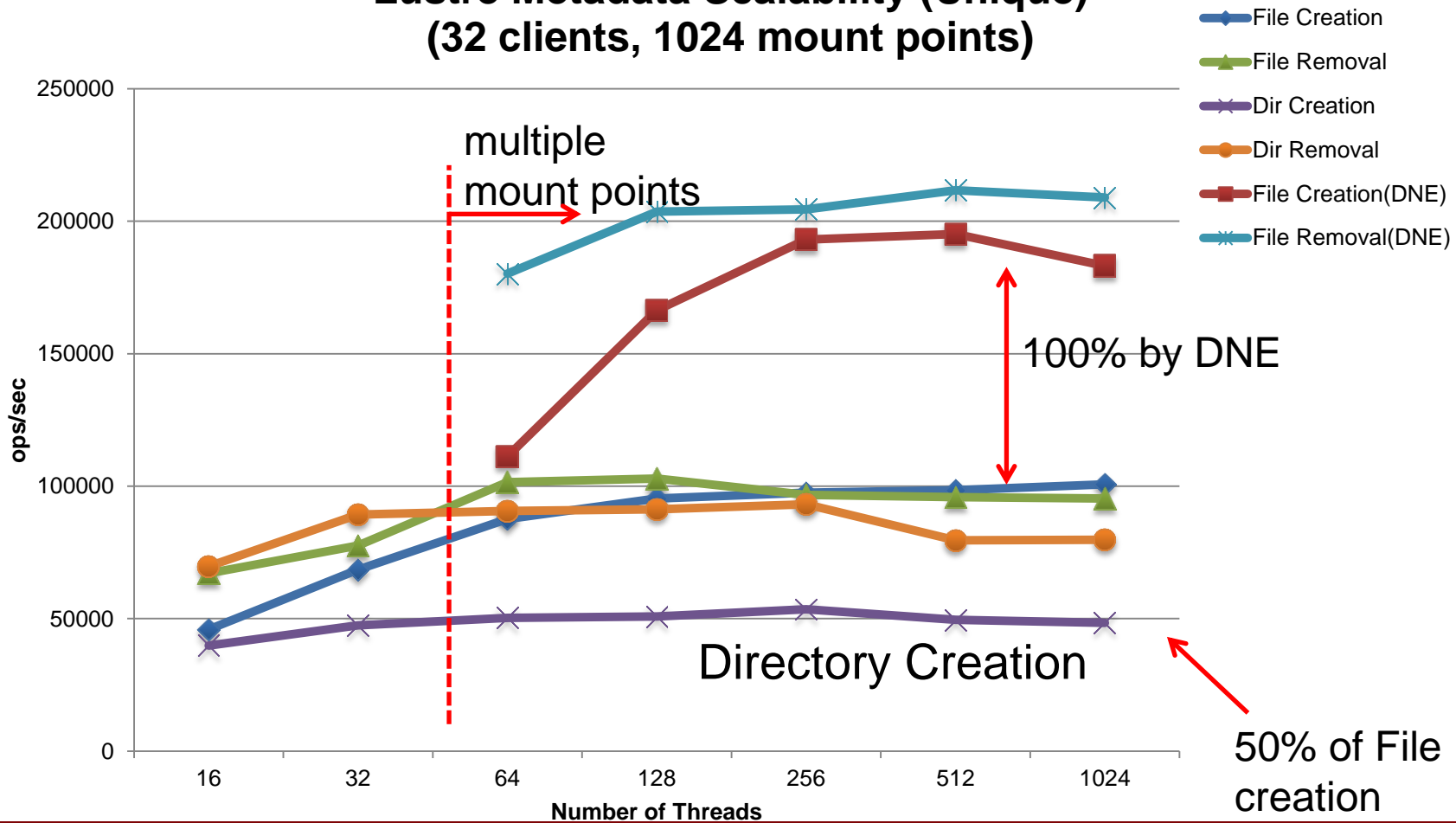
Human Genetics samtools workflow



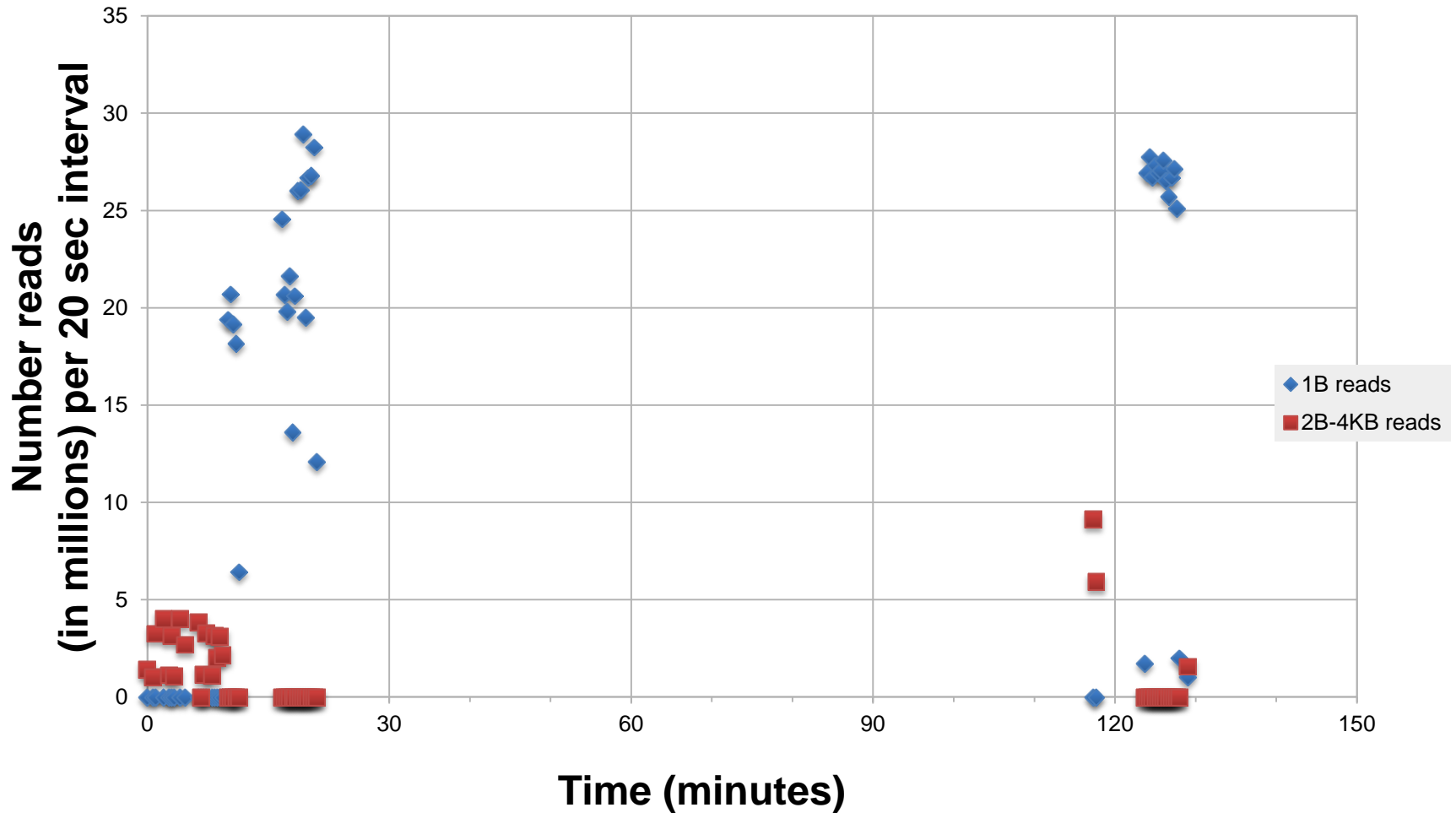
Lustre Metadata Performance

Extensibilité des MDSs (Répertoire unique)

Lustre Metadata Scalability (Unique) (32 clients, 1024 mount points)

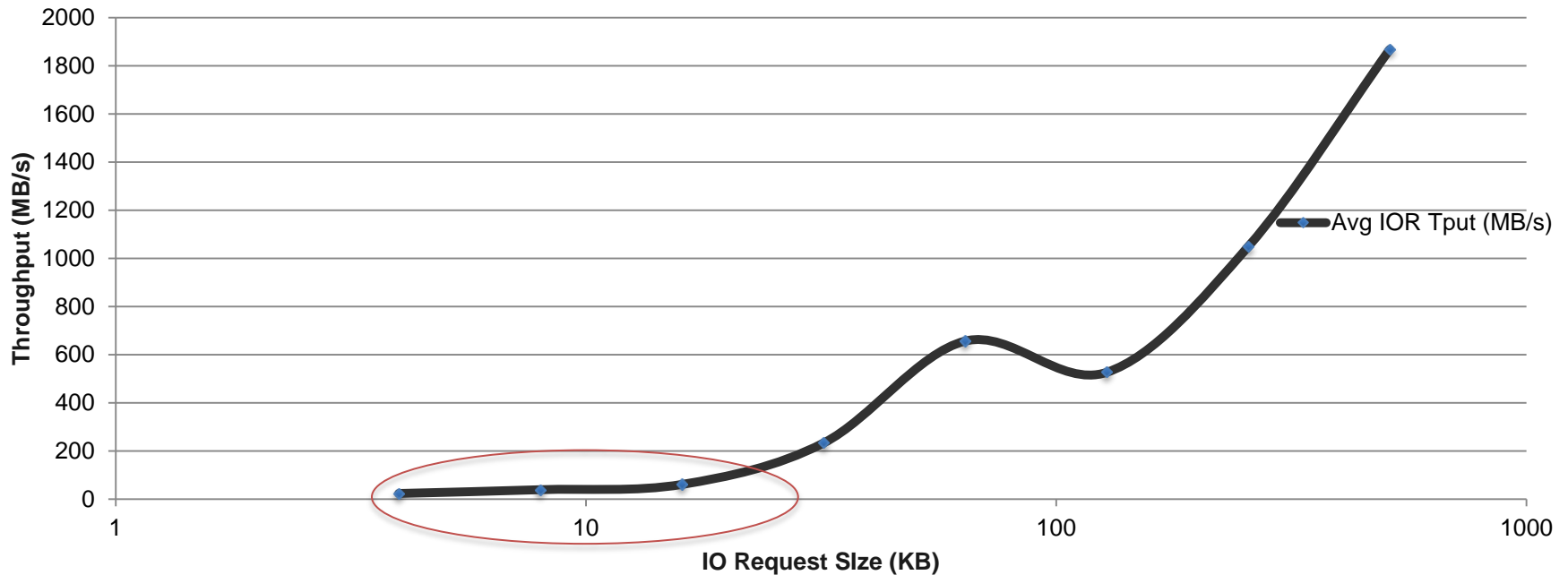


Le problème des petites E/S



Le déploiement de SSDs sous le PFS n'est pas une réponse satisfaisante

Parallel Filesystem on IME Demo Cluster
14x1U servers with SSDs
(50GB/s available)

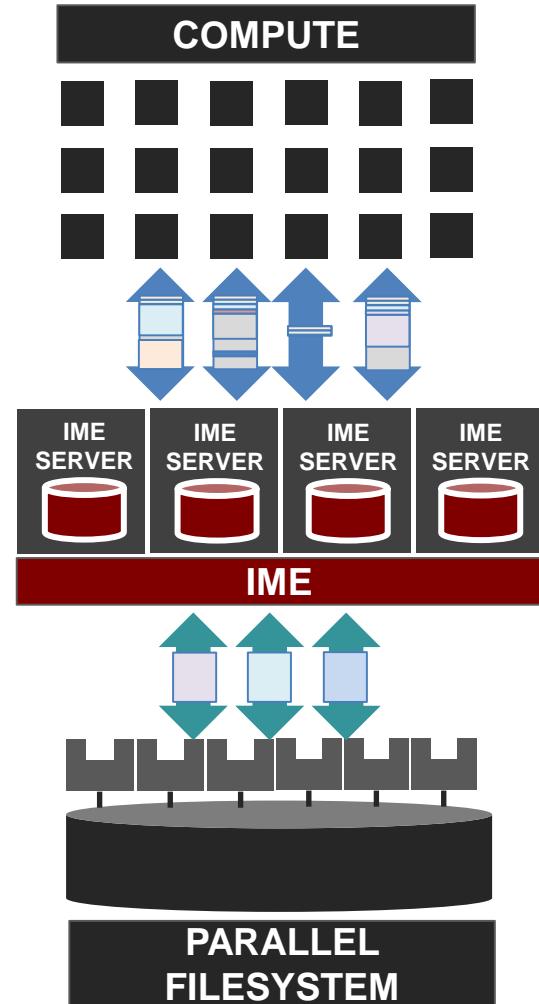
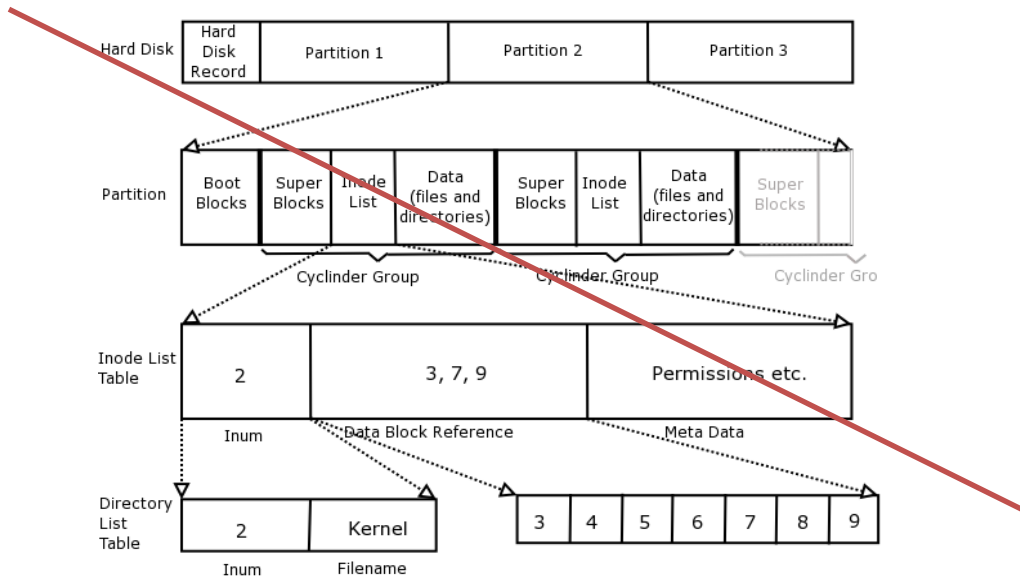


DDN | IME

Application I/O Workflow

- ▶ Insertion d'une strate additionnelle de NVMs distribués entre les applications et le PFS
- ▶ Architecture Client-Server
- ▶ Transparent pour les utilisateurs
- ▶ Transparent pour les applications
- ▶ Ré-organisation des données avant la reversement des données sur le PFS

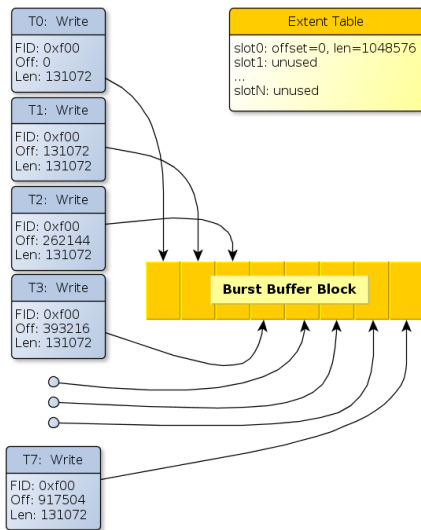
UNIX File System Layout



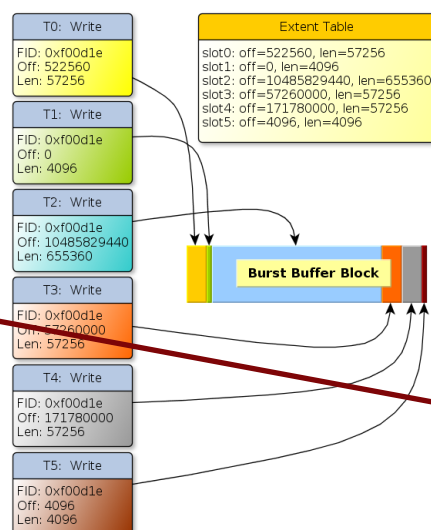
Structuration en Log dans IME

Consider two different application I/O patterns (write)

Séquentiel



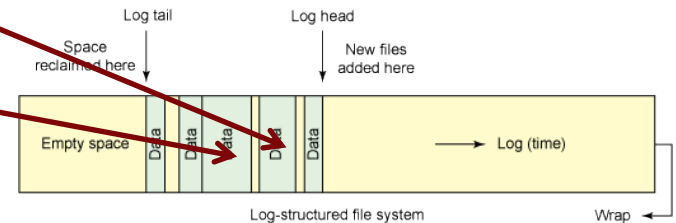
Non-séquentiel



'Burst buffer blocks' (BBB) sont les buffers générés au niveau du client

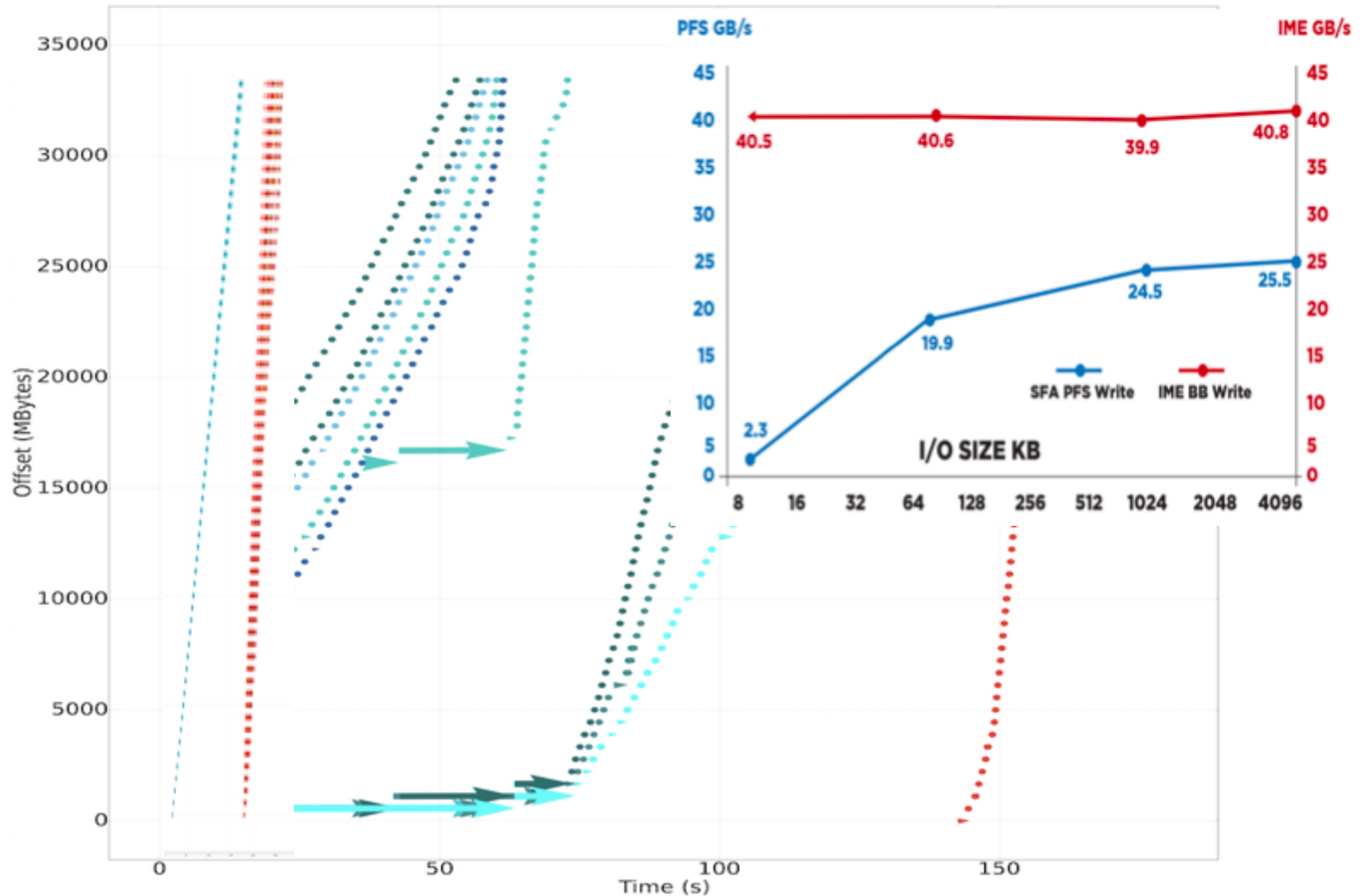
Notez que les contenus des BBB peuvent être alignés ou non.

La même méthode de stockage est utilisée pour les deux BBB et ce en dépit de la différence qualitative de contenu



IME vs Systèmes de fichiers parallèles :

Performance pour un fichier partagé (SSF)



Thank You!

Keep in touch with us



Team-jpsales@ddn.com



@ddn_limitless



company/datadirect-networks



102-0081
東京都千代田区四番町6-2
東急番町ビル 8F



[TEL:03-3261-9101](tel:03-3261-9101)
FAX : 03-3261-9140