

Stockage

ALBERT SHIH

Direction Informatique
Observatoire de Paris

14 juin 2019



Volume

- Explosion volume de données.

Volume

- Explosion volume de données.
- 2000 : 100Mbit/s, 9 Go, 1 Gbit/s, 320Mbit/s, N/A. 2019 : 1Gbits/s, 12To 10-100Gbit/s, 12Gbits/s, 50Mpx

Volume

- Explosion volume de données.
- 2000 : 100Mbit/s, 9 Go, 1 Gbit/s, 320Mbit/s, N/A. 2019 : 1Gbits/s, 12To 10-100Gbit/s, 12Gbits/s, 50Mpx
- 50 Po en 2016 pour le LHC..

Volume

- Explosion volume de données.
- 2000 : 100Mbit/s, 9 Go, 1 Gbit/s, 320Mbit/s, N/A. 2019 : 1Gbits/s, 12To 10-100Gbit/s, 12Gbits/s, 50Mpx
- 50 Po en 2016 pour le LHC..
- NenuFar, entre 1.5To ← 30To par jour, 4Po/an.

Volume

- Explosion volume de données.
- 2000 : 100Mbit/s, 9 Go, 1 Gbit/s, 320Mbit/s, N/A. 2019 : 1Gbits/s, 12To 10-100Gbit/s, 12Gbits/s, 50Mpx
- 50 Po en 2016 pour le LHC..
- NenuFar, entre 1.5To ← 30To par jour, 4Po/an.
- SKA (Square Kilometric Array) : Plusieurs To/s, 300–600 Po/an.

Brutes

- Nécessite haute/très hautes performances.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.
- Sécurité basse.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.
- Sécurité basse.
- Débit réseau.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.
- Sécurité basse.
- Débit réseau.
- Débit bus.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.
- Sécurité basse.
- Débit réseau.
- Débit bus.
- Parallélisation.

Brutes

- Nécessite haute/très hautes performances.
- Conservation très courte.
- Très près des traitements.
- Sécurité basse.
- Débit réseau.
- Débit bus.
- Parallélisation.
- Rangement.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.
- (Très) près des traitements.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.
- (Très) près des traitements.
- Conservation moyenne.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.
- (Très) près des traitements.
- Conservation moyenne.
- Utile ?.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.
- (Très) près des traitements.
- Conservation moyenne.
- Utile ?.
- Accès parallèle ou pas.

Traitement

- Nécessite très hautes performances (HPC)... ou pas.
- Volume très important.
- Sécurité moyenne ou pas.
- (Très) près des traitements.
- Conservation moyenne.
- Utile ?.
- Accès parallèle ou pas.
- Si possible fragmenter.

Types

Données traités

- Volume important.

Données traités

- Volume important.
- Sécurité élevé. ou pas

Données traités

- Volume important.
- Sécurité élevé. ou pas

Données traités

- Volume important.
- Sécurité élevé. ou pas
- Conservation longue.

Données traités

- Volume important.
- Sécurité élevé. ou pas
- Conservation longue.
- Rangement/Métadonnées.

Données traités

- Volume important.
- Sécurité élevé. ou pas
- Conservation longue.
- Rangement/Métadonnées.
- Migration données.

Données consolidées

- Volume important.

Données consolidées

- Volume important.
- Sécurité très élevé.

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.
- Conservation très long.

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.
- Conservation très long.
- Métadonnées, identifications, etc. . .

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.
- Conservation très long.
- Métadonnées, identifications, etc. . .
- Migration des supports.

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.
- Conservation très long.
- Métadonnées, identifications, etc. . .
- Migration des supports.
- Temps d'accès requis.

Données consolidées

- Volume important.
- Sécurité très élevé.
- Accessible à diffusions.
- Conservation très long.
- Métadonnées, identifications, etc. . .
- Migration des supports.
- Temps d'accès requis.
- Fréquence d'accès.

Types

Archivages

- Données froides.

Archivages

- Données froides.
- Peu/pas de modifications.

Archivages

- Données froides.
- Peu/pas de modifications.
- Peu/pas d'accès.

Archivages

- Données froides.
- Peu/pas de modifications.
- Peu/pas d'accès.
- Ranger (métadonnées).

Archivages

- Données froides.
- Peu/pas de modifications.
- Peu/pas d'accès.
- Ranger (métadonnées).
- Migration des supports

Archivages

- Données froides.
- Peu/pas de modifications.
- Peu/pas d'accès.
- Ranger (métadonnées).
- Migration des supports
- Temps d'accès requis.

Archivages

- Données froides.
- Peu/pas de modifications.
- Peu/pas d'accès.
- Ranger (métadonnées).
- Migration des supports
- Temps d'accès requis.
- Fréquence d'accès.

Support

- Bandes.

Support

- Bandes.
- Support optique.
- Disques Dur.

Support

- Bandes.
- Support optique.
- Disques Dur.
- SSD.

Support

- Bandes.
- Support optique.
- Disques Dur.
- SSD.
- ... le cloud.

Types

- Types : LTO (Linear Tape Open), IBM 3592 (TS 11xx), Oracle TS10000.

Types

- Types : LTO (Linear Tape Open), IBM 3592 (TS 11xx), Oracle TS10000.
- Capacité ×2 sur 2 ans.

Types

- Types : LTO (Linear Tape Open), IBM 3592 (TS 11xx), Oracle TS10000.
- Capacité ×2 sur 2 ans.
- LTO-8 : 12To, LTO-9 : 24To, ..., LTO-12 : 192To.

Consommation énergétique

- Overland Neo 80

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000
 - 650W (Moyenne) & 440 kg

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000
 - 650W (Moyenne) & 440 kg
 - MD 1280 (×6) : 16800W & 780 kg

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000
 - 650W (Moyenne) & 440 kg
 - MD 1280 (×6) : 16800W & 780 kg
- SL 8500

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000
 - 650W (Moyenne) & 440 kg
 - MD 1280 (×6) : 16800W & 780 kg
- SL 8500
 - 2881 W Idle → 7157 W (Max), & 5.5t

Consommation énergétique

- Overland Neo 80
 - 200W (Max) & 60 kg
 - MD1280 : 2800 W & 130kg
- Overland Neo 8000
 - 650W (Moyenne) & 440 kg
 - MD 1280 (×6) : 16800W & 780 kg
- SL 8500
 - 2881 W Idle → 7157 W (Max), & 5.5t
 - MD1280 (×119) : 333200 W & 15t

Posix vs objet

- Accès Posix (ou presque) (DAS, NFS, etc. . .)
 - + Facilité utilisation
 - - Extensibilité
 - - - - Migration support

Posix vs objet

- Accès Posix (ou presque) (DAS, NFS, etc. . .)
 - + Facilité utilisation
 - - Extensibilité
 - - - - Migration support
- Accès type objet (Amazon S3, iRods, etc. . .)
 - +++ Extensibilité
 - +++ Migration support
 - - *complexité* d'utilisation.

- Stockage conçu pour le HPC.

- Stockage conçu pour le HPC.
- GPFS, PanFS.

- Stockage conçu pour le HPC.
- GPFS, PanFS.
- Lustre, BeegFS.

- Stockage conçu pour le HPC.
- GPFS, PanFS.
- Lustre, BeegFS.
- HDFS

FS distribué

- CephFS, GlusterFS

FS distribué

- CephFS, GlusterFS
- iRods, OpenIO, OneData, etc...

Le cloud

- Le cloud, c'est quoi ?

Le cloud

- Le cloud, c'est quoi ?
- Achat stockage.

Le cloud

- Le cloud, c'est quoi ?
- Achat stockage.
- Stockage objet

Le cloud

- Le cloud, c'est quoi ?
- Achat stockage.
- Stockage objet
- Attention coût.

Le cloud

- Le cloud, c'est quoi ?
- Achat stockage.
- Stockage objet
- Attention coût.
- Complexe.

StockageS
○○○○○○○

Bande magnétique
○○

Stockages
○○○○

Choix
●○○

Et moi ?
○

Le graal

Le graal



Le graal

Le graal

- ...
- je n'ai pas trouvé...

Besoin

Besoin

- Type.

Besoin

- Type.
- Vitesse.

Besoin

- Type.
- Vitesse.
- Volume.

Besoin

- Type.
- Vitesse.
- Volume.
- Temps de conservation.

Besoin

- Type.
- Vitesse.
- Volume.
- Temps de conservation.
- Taille des fichiers/Rangement.

Besoin

- Type.
- Vitesse.
- Volume.
- Temps de conservation.
- Taille des fichiers/Rangement.
- Utilisation.

Besoin

- Type.
- Vitesse.
- Volume.
- Temps de conservation.
- Taille des fichiers/Rangement.
- Utilisation.
- Migration.

Besoin

- Type.
- Vitesse.
- Volume.
- Temps de conservation.
- Taille des fichiers/Rangement.
- Utilisation.
- Migration.
- Niveau d'importance.

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).
- Financier.

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).
- Financier.
- Place/Refroidissement/Électricité.

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).
- Financier.
- Place/Refroidissement/Électricité.
- Dans la durée.

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).
- Financier.
- Place/Refroidissement/Électricité.
- Dans la durée.
- Discussion, et ... facturation.

Ressources

- Humains (Étude, benchmarks, prototype, mise en place, accompagnement, exploitation ...).
- Financier.
- Place/Refroidissement/Électricité.
- Dans la durée.
- Discussion, et ... facturation.
- PGD ou DMP.

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.
- DAS (ZFS) pour les données «standards» et backups (≈ 1 Po).

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.
- DAS (ZFS) pour les données «standards» et backups (≈ 1 Po).
- ActiveCircle (Propriétaire) pour archivage (≈ 500 To).

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.
- DAS (ZFS) pour les données «standards» et backups (≈ 1 Po).
- ActiveCircle (Propriétaire) pour archivage (≈ 500 To).
- Content ?

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.
- DAS (ZFS) pour les données «standards» et backups (≈ 1 Po).
- ActiveCircle (Propriétaire) pour archivage (≈ 500 To).
- Content ?
 - Évolution stockage distribué (iRods, OneData, etc. . .).

Et moi ?

- Beegfs pour le calcul parallèle (≈ 1 Po).
- DAS+NFS pour le calcul scalaire.
- DAS (ZFS) pour les données «standards» et backups (≈ 1 Po).
- ActiveCircle (Propriétaire) pour archivage (≈ 500 To).
- Content ?
 - Évolution stockage distribué (iRods, OneData, etc. . .).
 - Réflexion remplacement ActiveCircle.