

# Architectures de stockage parallèle et distribué

(Tour d'horizon très rapide)

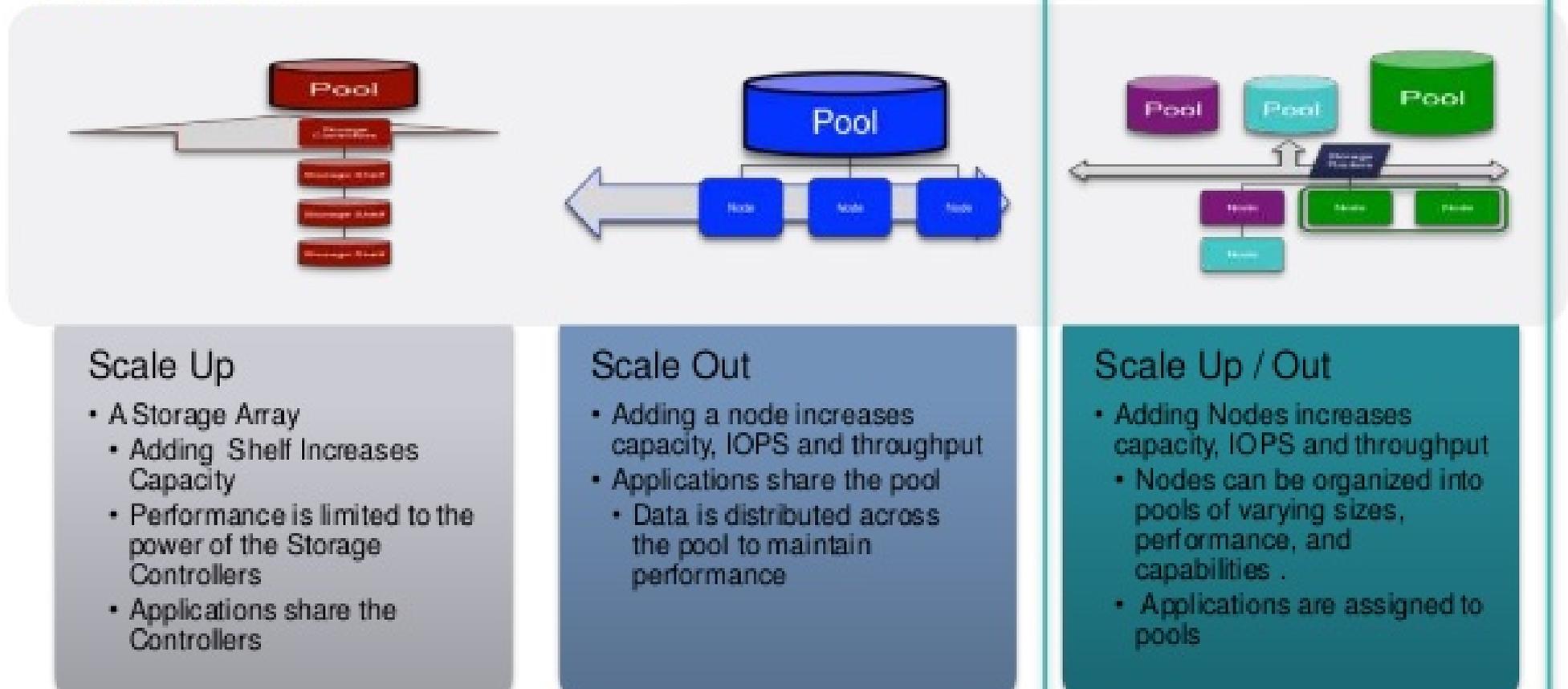
ANF UST4HPC, Fréjus 2018  
Loïc Tortay

- « Partage » : des composants de stockage
- Architectures sans partage (*shared nothing*) :
  - serveurs indépendants
  - disques « dans » un serveur : disponibilité des données liées au serveur ➔ éventuellement redondance des données sur d'autre(s) serveur(s)
- Avec partage :
  - réseau de stockage (SAN)
  - disques partagés par au moins deux serveurs : redondance d'accès aux données

- Réplication (« facile » mais coûteuse en espace)
- Codes correcteurs à effacement (*erasure coding*), y compris RAID (« difficiles » mais économes en espace)
- Entre composants
- Entre serveurs
- Redondance sous le contrôle de qui ou quoi ?
- Mélange des genres ?
- Pour la réplication (mais pas uniquement) : informations de topologie (*failure domain, protection domain, failure group, ...*)

- Pour architectures avec **et** sans partage

## The Basics



Cisco

- Infrastructure & accès par les clients

- Fichier : HTTP, iRods, dCache, système de fichiers distribué (NFS) ou parallèle (BeeGFS, GPFS, Lustre, etc.), ...
- Méta-données :
  - espace de nommage partagé, « réel » ou « simulé »
  - permissions & comptabilité (identités des utilisateurs)
- Souvent, redirection d'un serveur de méta-données vers les serveurs de données (AFS, iRods, dCache, NFS v4.1/pNFS, BeeGFS, GPFS, ...)

- « Périphérique disque » distant : RBD, iSCSI, NBD, etc.
- En général pour y mettre un système de fichiers (souvent local mais parfois parallèle)
- Peu/pas de méta-données au niveau du périphérique bloc :
  - au « dessous » : fournisseur du bloc
  - au « dessus » : application qui utilise le bloc

- Données dans des objets « indépendants » : pas d'espace de nommage (partagé)
- Standards OSD T10/SCSI
- Sens « moderne » : accès avec API de type REST (S3, etc.) ou pas/pas uniquement (CDMI)
- Objet à tous les niveaux :
  - matériel (DDN WOS, Seagate Kinetic, ...)
  - système de fichiers (Ceph, Lustre, Panasas, ...)
  - application de stockage (S3, GCS, ...)
  - application utilisateur (Dropbox, Facebook, ...)
- Attributs & permissions

- Accès transparent ou pas ?
- Modification de fichier/objet en place ou pas ?
- Exécution de binaire ou pas ?
- Copie locale + PUT/GET ?

- Espace temporaire/projet/home, disponibilité, sauvegarde ?
- Accès/distribution vers l'extérieur (rsync, sftp, http, Globus online, iRods, ...)

- Pas assez technique ou détaillé ?
  - voir présentation de Benoît Parrein à l'ANF Stockage distribué 2016 (diffusée par mail aux participants de cette ANF)
  - ou ma présentation « bases du stockage distribué pour les administrateurs systèmes », école IN2P3/CNRS Stockage distribué 2017