Le stockage à l'IBMP

Vendu à DataCore ? 😇

JL Evrard - X/Stra - 5 novembre 2024

Une architecture essentiellement bipolaire

Vous avez dit "DataCore" ?

- La première brique tourne autour d’un système Software Defined Storage (SDS) "blocs" basé sur le logiciel DataCore **SANSymphony**
- La seconde brique repose sur un stockage pure objet qui utilise la solution DataCore **Swarm**
- La dernière brique consiste en une automatisation de passage de l’un à l’autre via le logiciel **FileFly** de chez… DataCore

Le bloc selon SANSymphony

Le SDS simple et efficace

- Un choix qui s’est imposé à l’époque (~2015) face à deux vagues :
  - Une numérisation de plus en plus forte des données scientifiques
  - La mise en place de systèmes virtualisés
- La volonté d’offrir un service 24/24 7/7 aux chercheurs avec un maximum de sécurité, mais avec le minimum de backup :
  - volume de données trop important : coût prohibitif pour les sauvegarder

Le bloc selon SANSymphony

Architecture "BioCore" - v1

Pare-feu 1

Protection actif/passif

Pare-feu 2

Commutateur 1

Commutateur 2

Switch Réseau 1 10 Gbps

Switch Réseau 2 10 Gbps

Serveur VMware 1

Cluster de virtualisation

Serveur VMware 2

Metacontrôleur 1

Metacontrôleur 2

Stockage 1 - 132 To

Liens Fiber Channel redondants

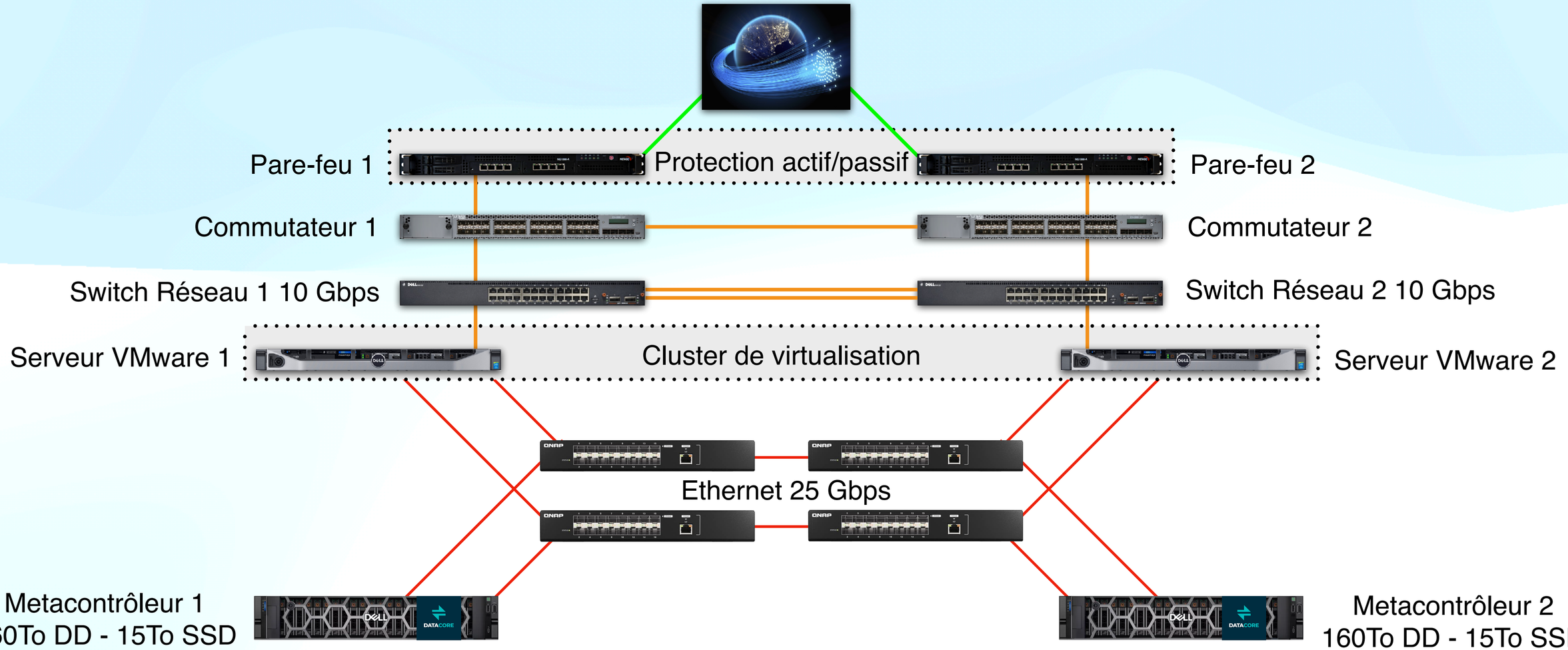
Stockage 2 - 132 To

iSCSI 1 10 Gbps - 128 To

iSCSI 2 10 Gbps - 128 To

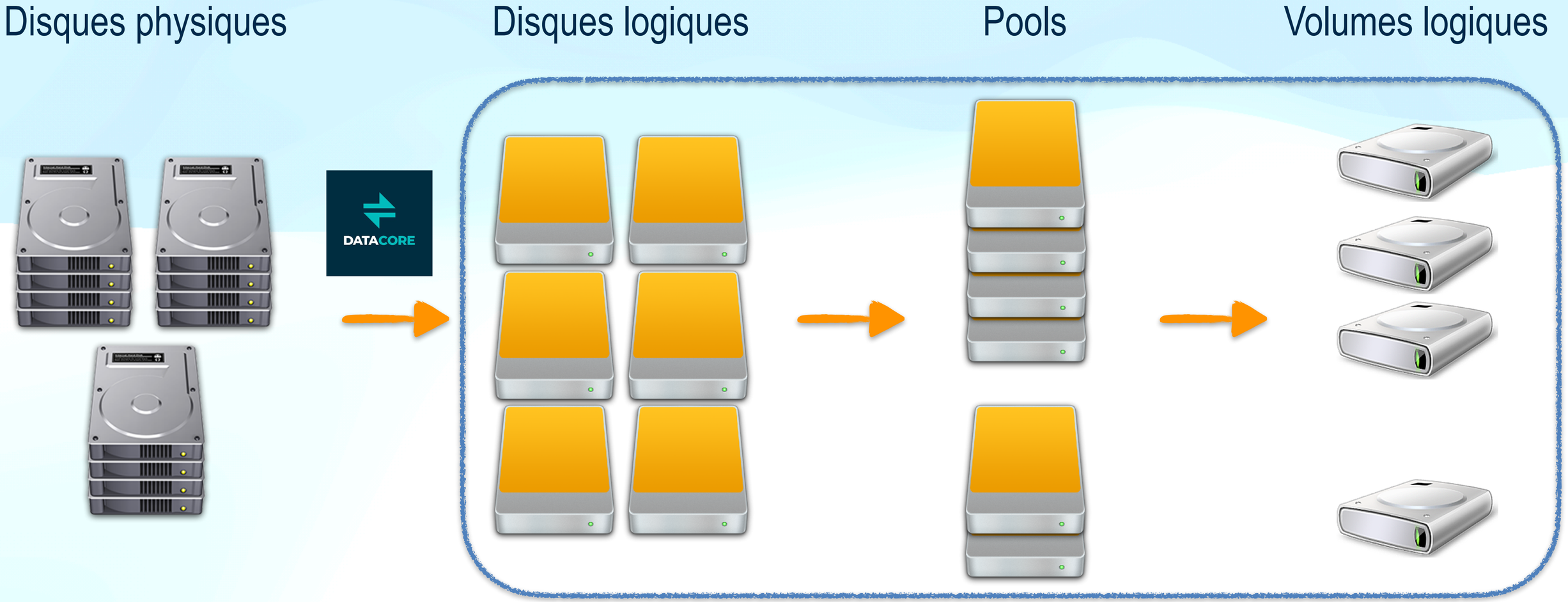
Le bloc selon SANSymphony

Architecture "BioCore" - v2 (fin 2024)



Le bloc selon SANSymphony

L’architecture logique



Le bloc selon SANSymphony

Le logiciel

- SANSymphony nécessite un OS Windows Server (ça fait peur, mais on s'y fait)
- Pour ce qui est du stockage utilisable : si Windows le voit, alors SANSymphony peut l'utiliser (USB, iSCSI, SATA, SAS, FC, NVMe…)
- Toute la gestion se fait en mode graphique à travers une console unique : objectivement, c'est très agréable
- L'intégration première était pour du VMware, mais à ce jour, tous les hyperviseurs du marché sont adressables (**Proxmox**, HyperV,…)

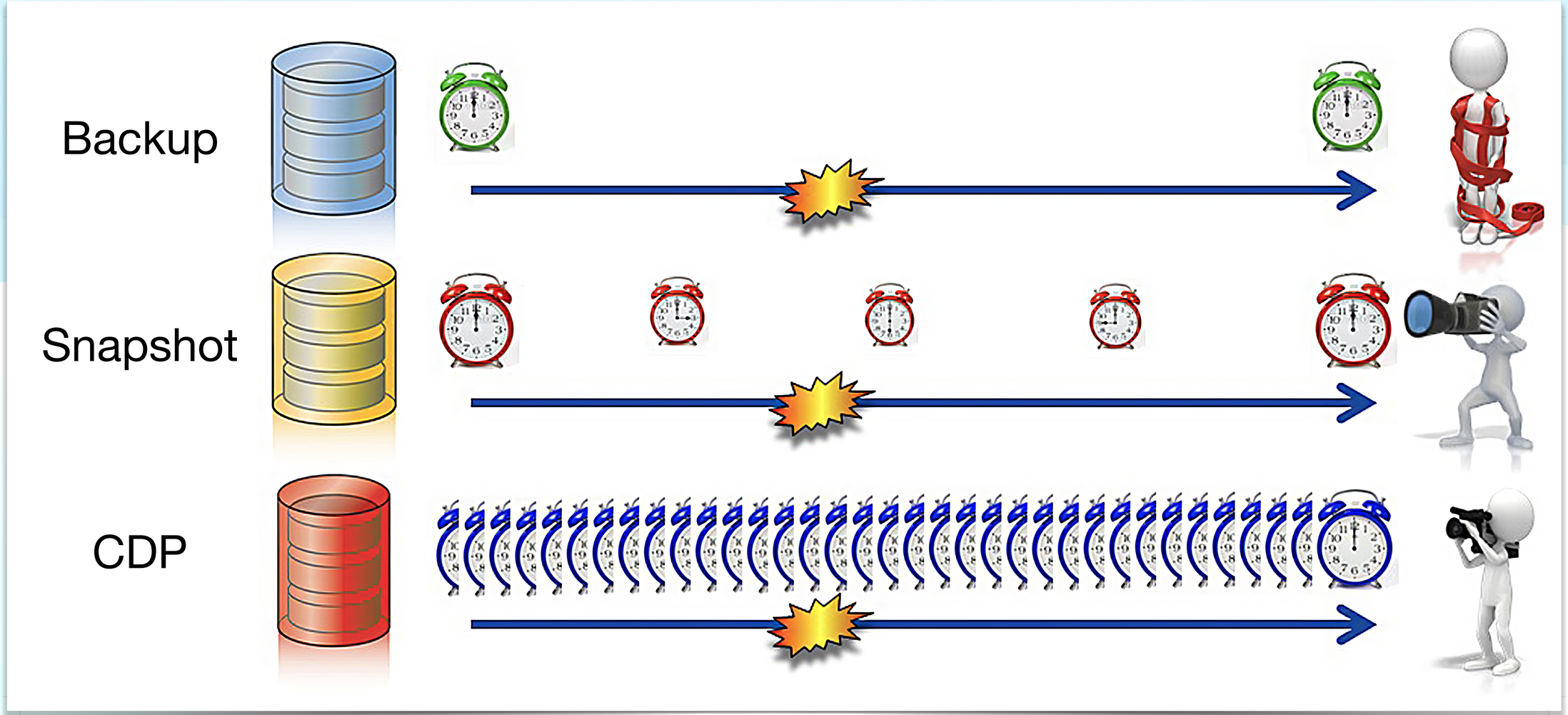
Le bloc selon SANSymphony

Le logiciel

- SANSymphony comporte de très nombreuses fonctionnalités :
  - Réplication synchrone (ou asynchrone)
  - Auto-tiering
  - CDP (Continuous Data Protection) - Snapshot
  - Déduplication
  - Compression
  - Optimisation de Capacité = déduplication & compression

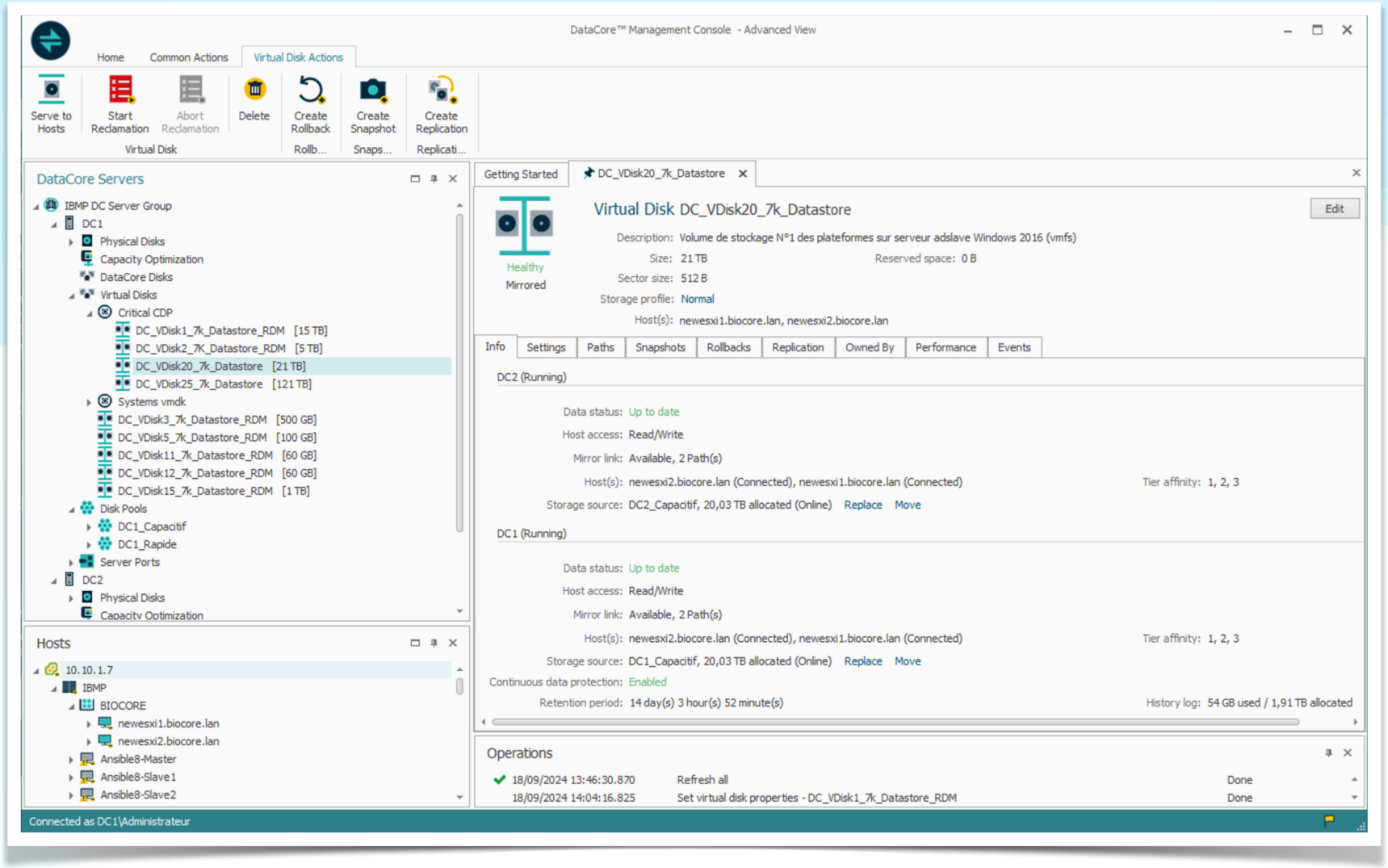
Le bloc selon SANSymphony

Le logiciel



Le bloc selon SANSymphony

Le logiciel



Nouveau design hyper-redondé

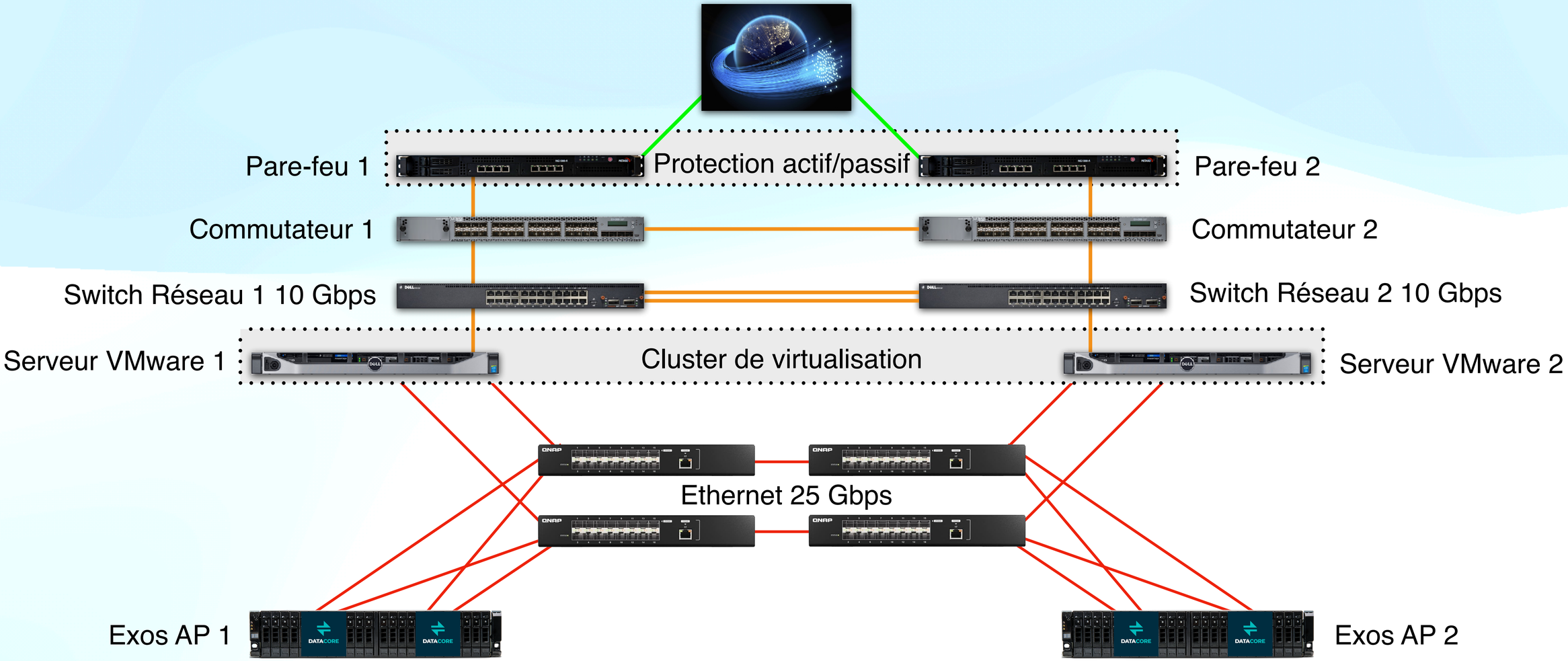
Seagate Exos AP (2U24)

Nouveau design hyper-redondé

Seagate Exos AP (2U24)

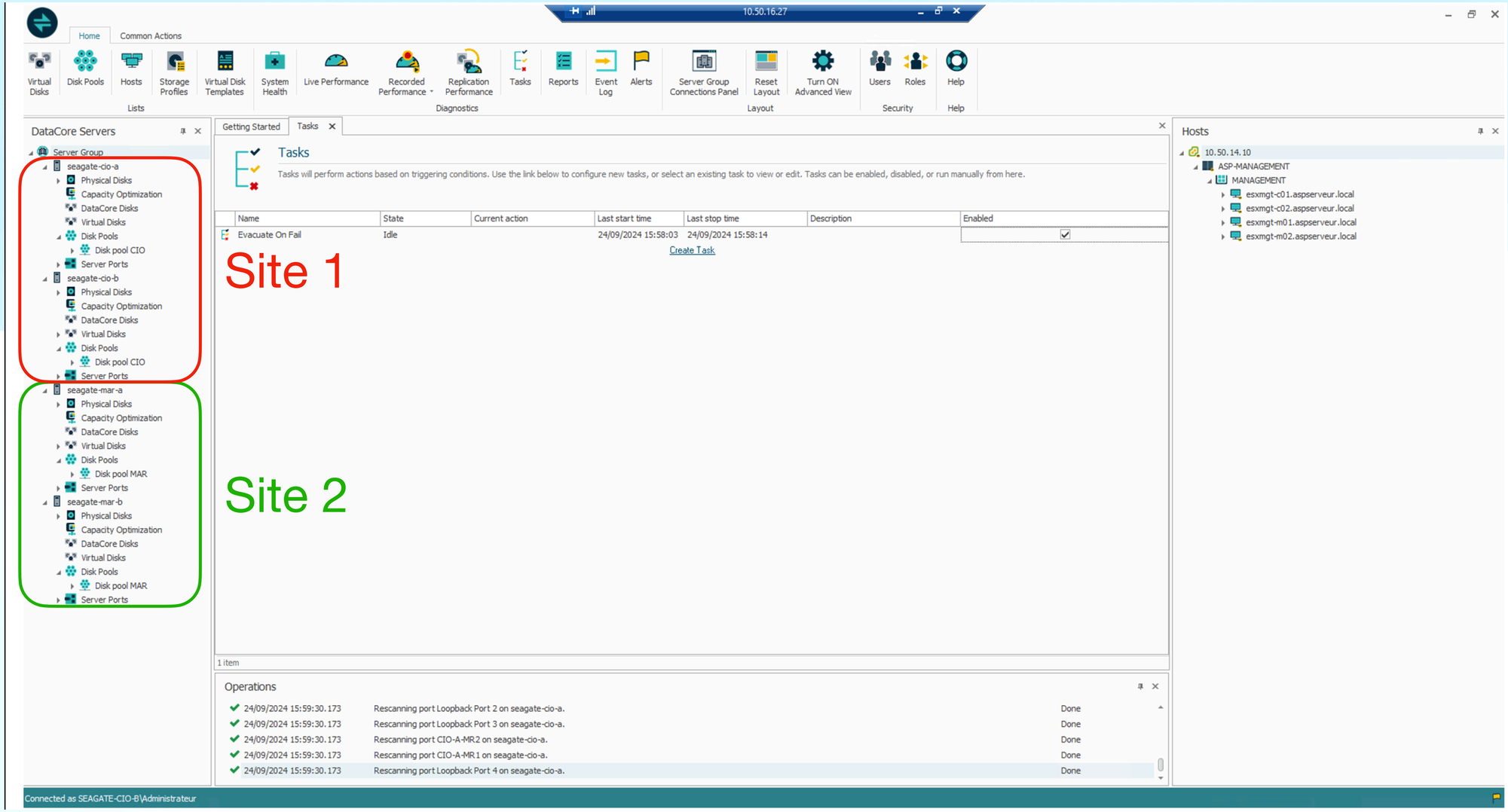
Nouveau design hyper-redondé

Seagate Exos AP (2U24)



Nouveau design hyper-redondé

Seagate Exos AP (2U24)



L’objet selon Swarm

L’objet jusqu’au bout du disque

- Une solution trouvée "par hasard" :
  - ActiveScale (Western Digital)
    - Quantum - Passage en mode locatif…
      - Caringo Swarm - Racheté par DataCore…
- Une solution pensée objet depuis ses fondations : le disque est un objet qui contient des objets (triplication) ou des parties d’objets (Erasure Coding)
- Un OS propriétaire logé en RAM (CastorOS) - Un "essaim" de disques…

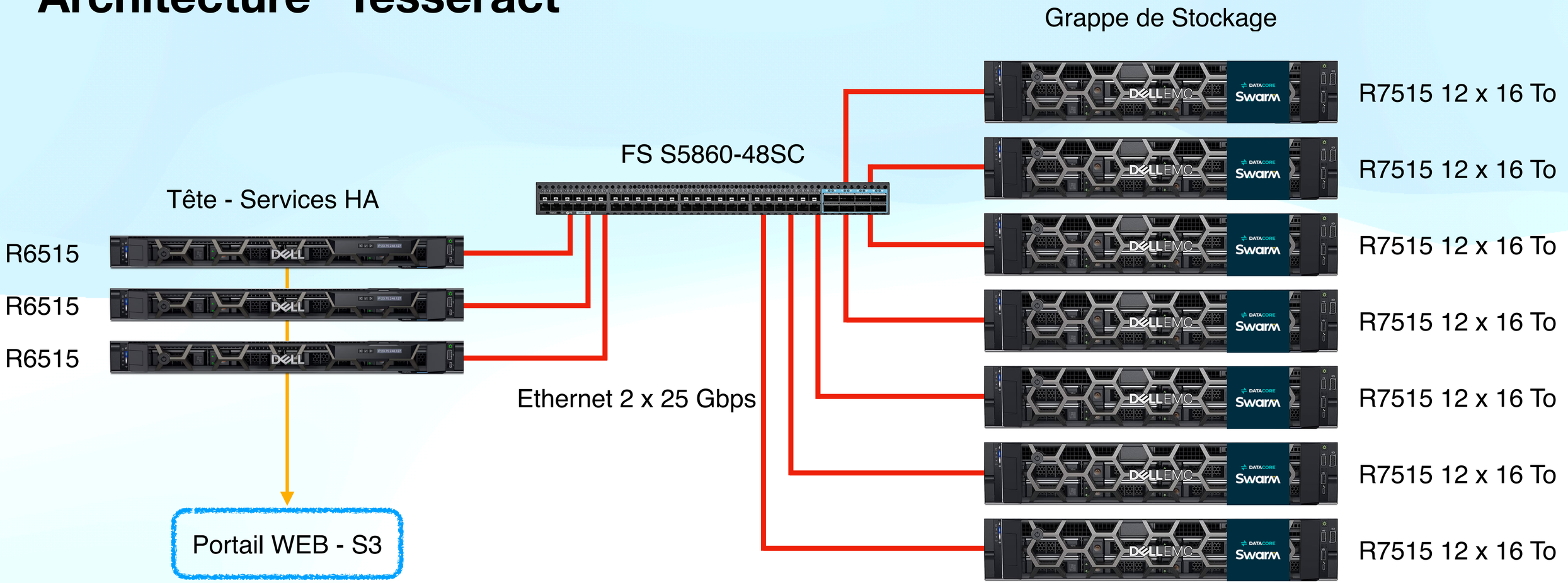
L’objet selon Swarm

La solution du pauvre…

- Besoin d’un "petit" S3 (10~100 To) ? SNS (Single Node Swarm) peut être une solution intéressante :
  - Une appliance capable d’être déployée en moins d’une heure
  - Installation en conteneurs automatisée
  - Possibilité de la raccrocher à un Swarm traditionnel
- Cas d’école : récupérer les données d’un matériel d’acquisition (et transférer vers le S3 "datacenter") sans se poser trop de questions…

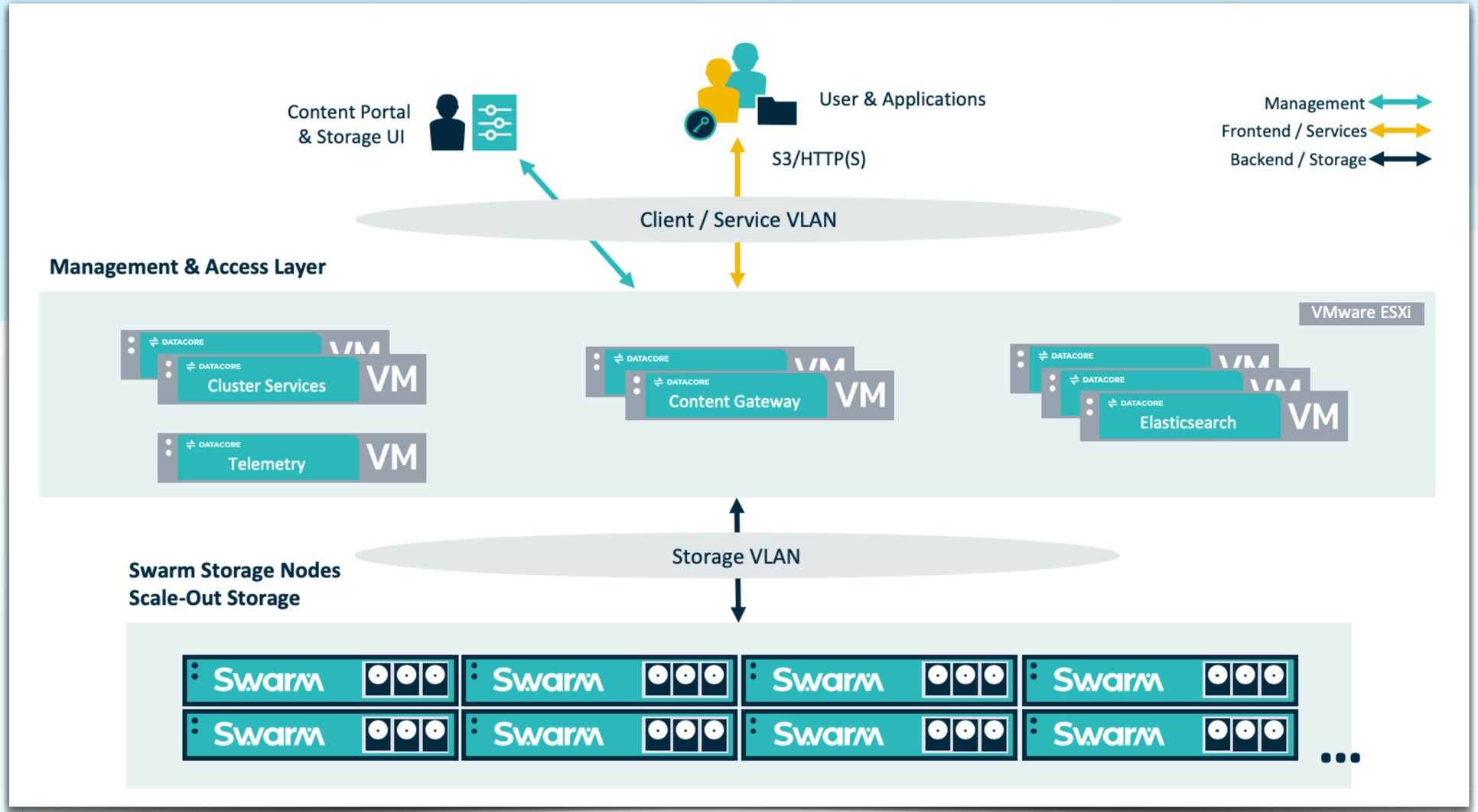
L’objet selon Swarm

Architecture "Tesseract"



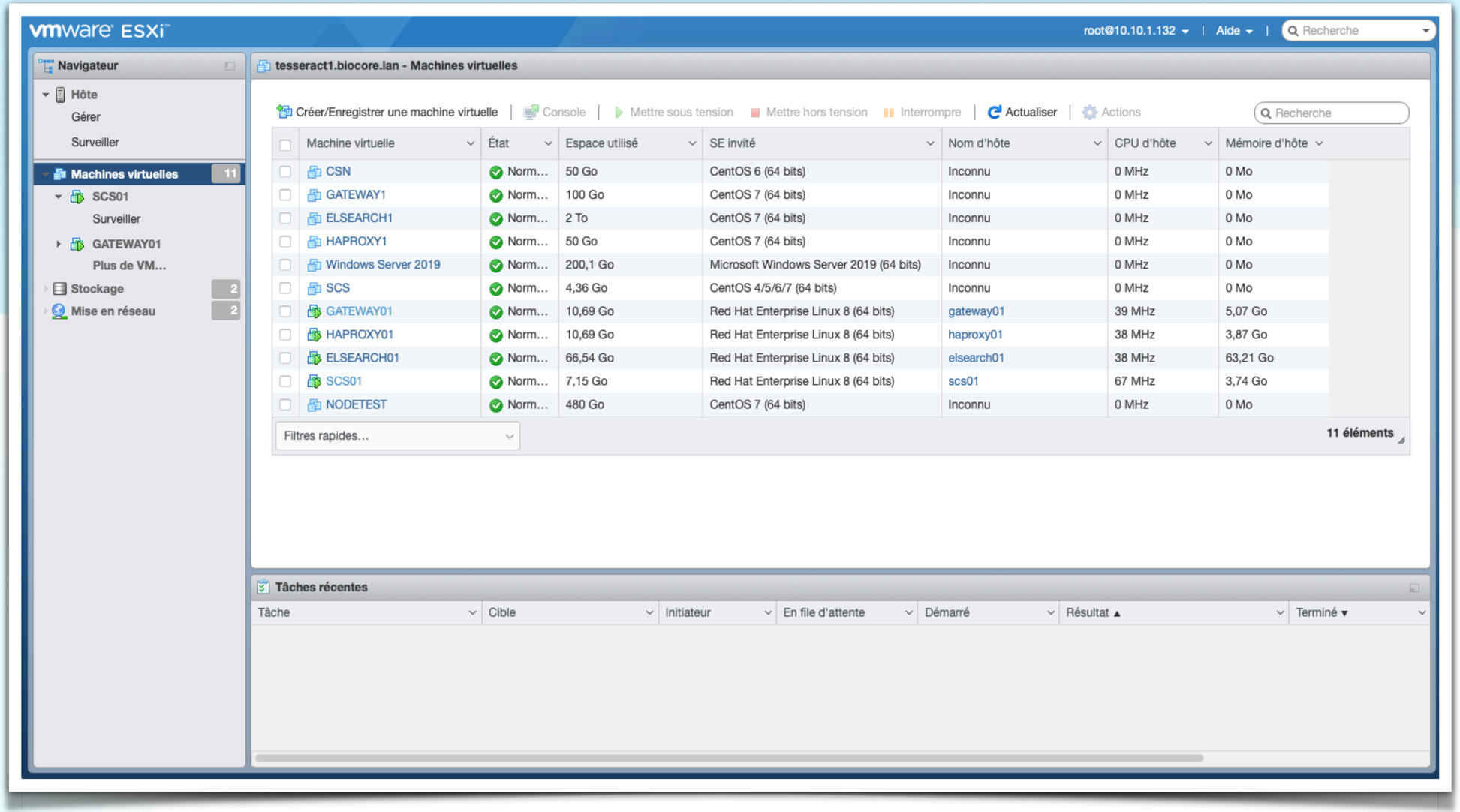
L’objet selon Swarm

Architecture logique



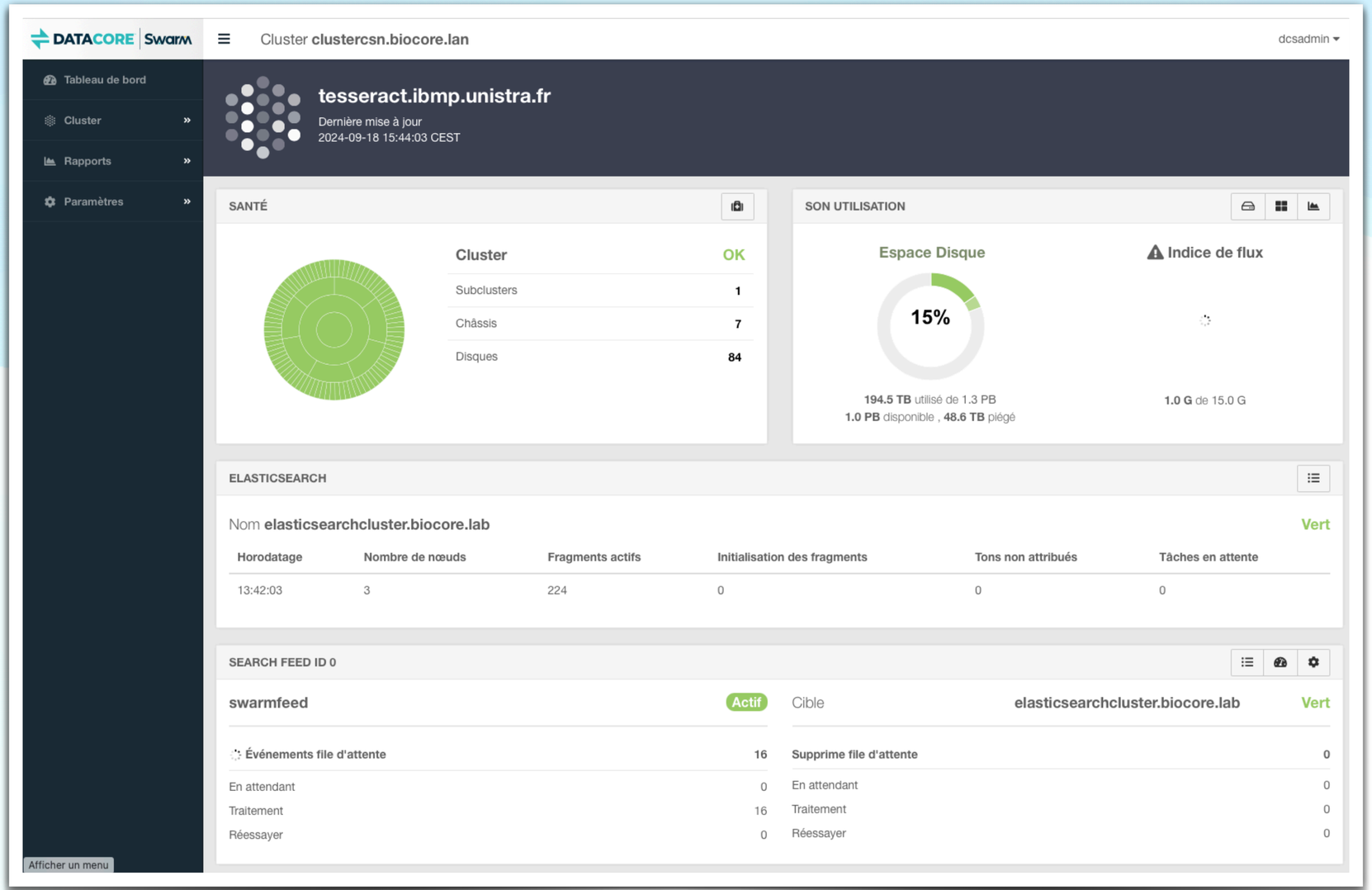
L’objet selon Swarm

Architecture logique



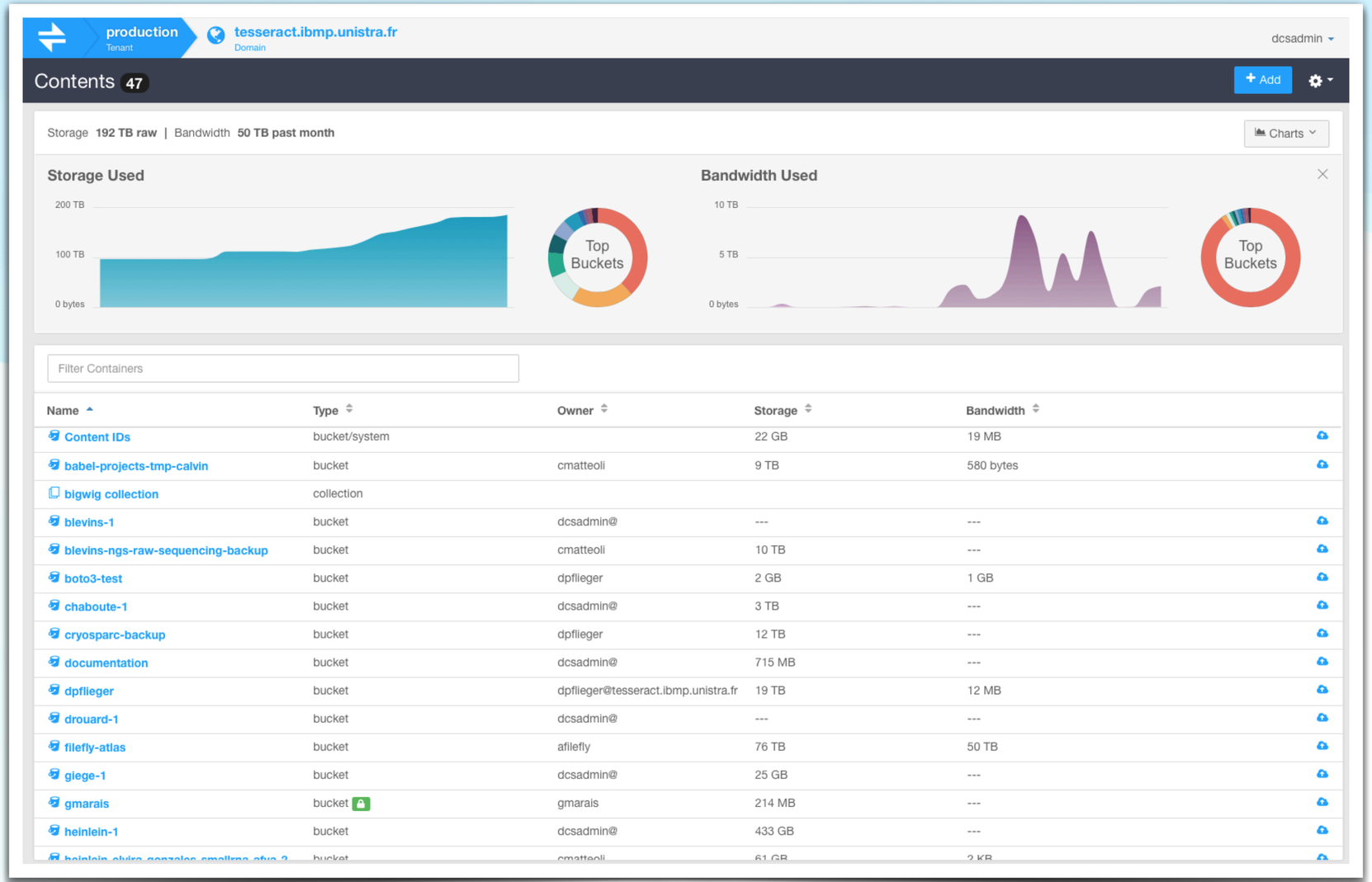
L’objet selon Swarm

Portails WEB



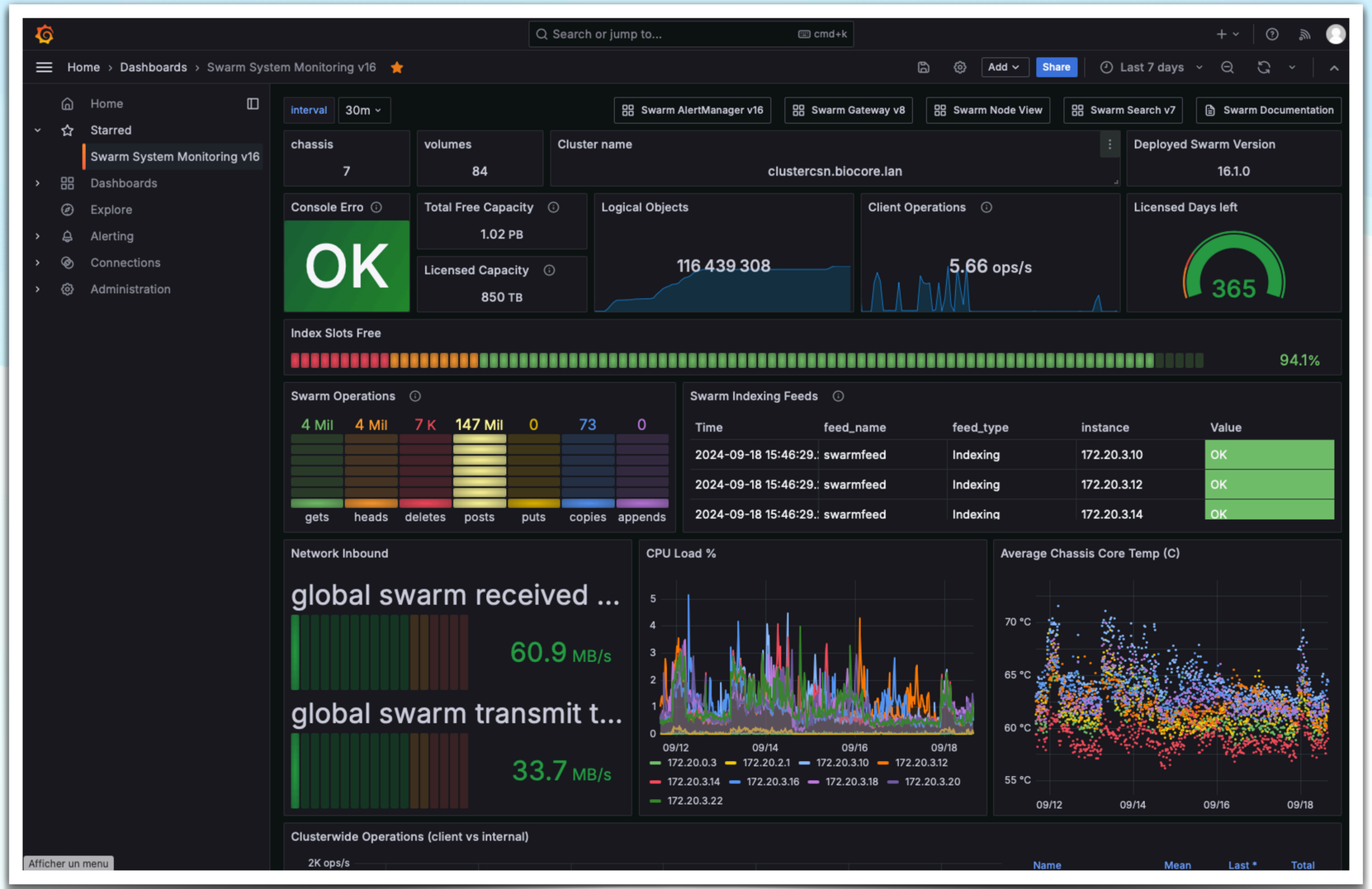
L’objet selon Swarm

Portails WEB



L’objet selon Swarm

Portails WEB



L’objet selon Swarm

Pour les utilisateurs

- Pas forcément "évident" de changer sa façon de faire
- Pour les "gros" utilisateurs :
  - La ligne de commande AWS
- Pour les personnels "lambda" :
  - Mountain Duck
- Idéalement un portail WEB dédié qui reste à faire…
- Et si on automatisait certains choses ?

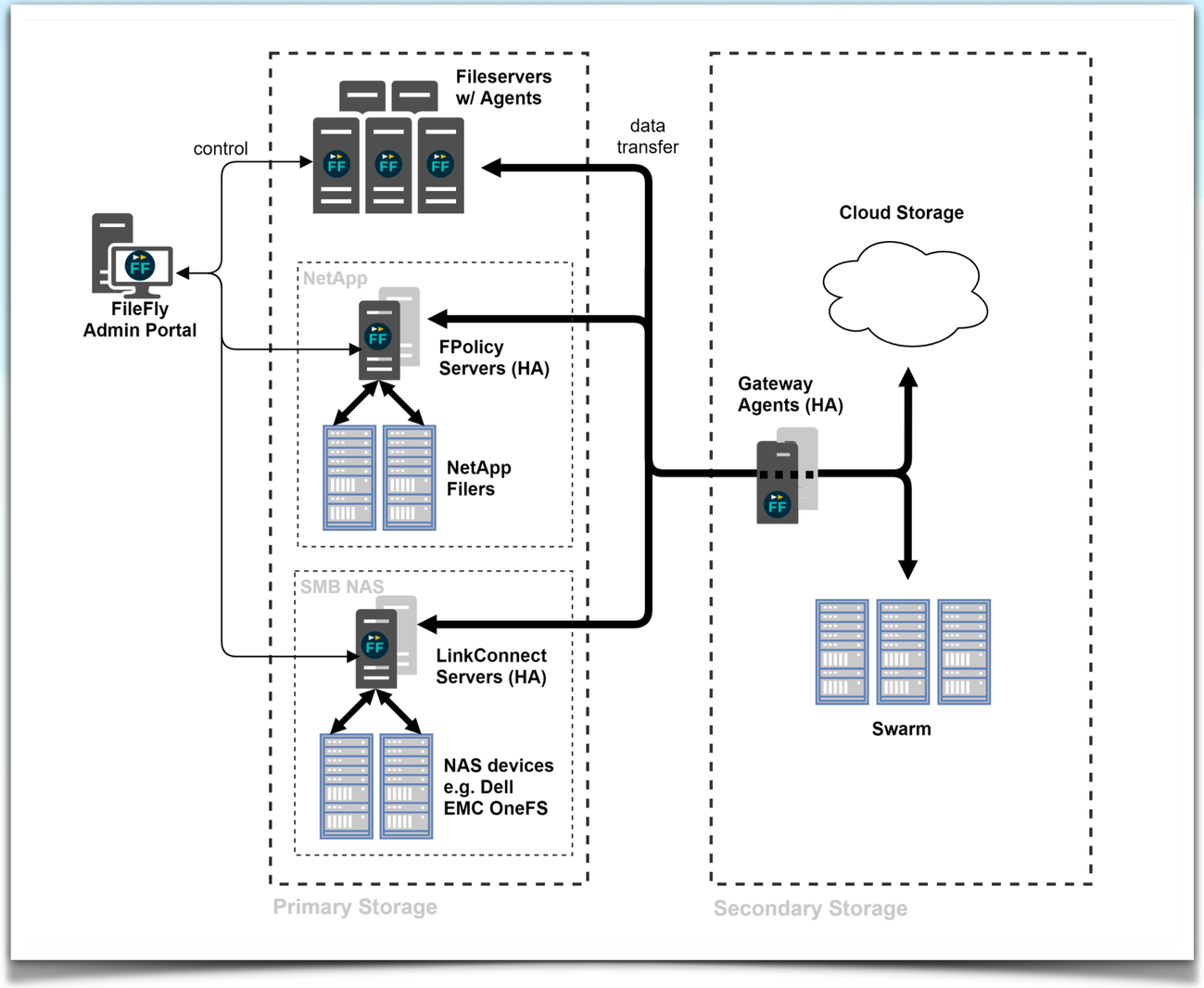
La solution FileFly

Le meilleur des deux mondes ?

- Il est très compliqué de forcer les chercheurs à trier les données chaudes et froides
- Par ailleurs, des données "froides" peuvent redevenir "chaudes"
- **FileFly** permet de déplacer automatiquement des données d’un stockage bloc (SMB) vers un stockage objet selon des critères temporels
- Le déplacement est totalement transparent pour l’utilisateur : les fichiers sont remplacés par des "alias" qui seront automatiquement regarnis s’il sont demandés

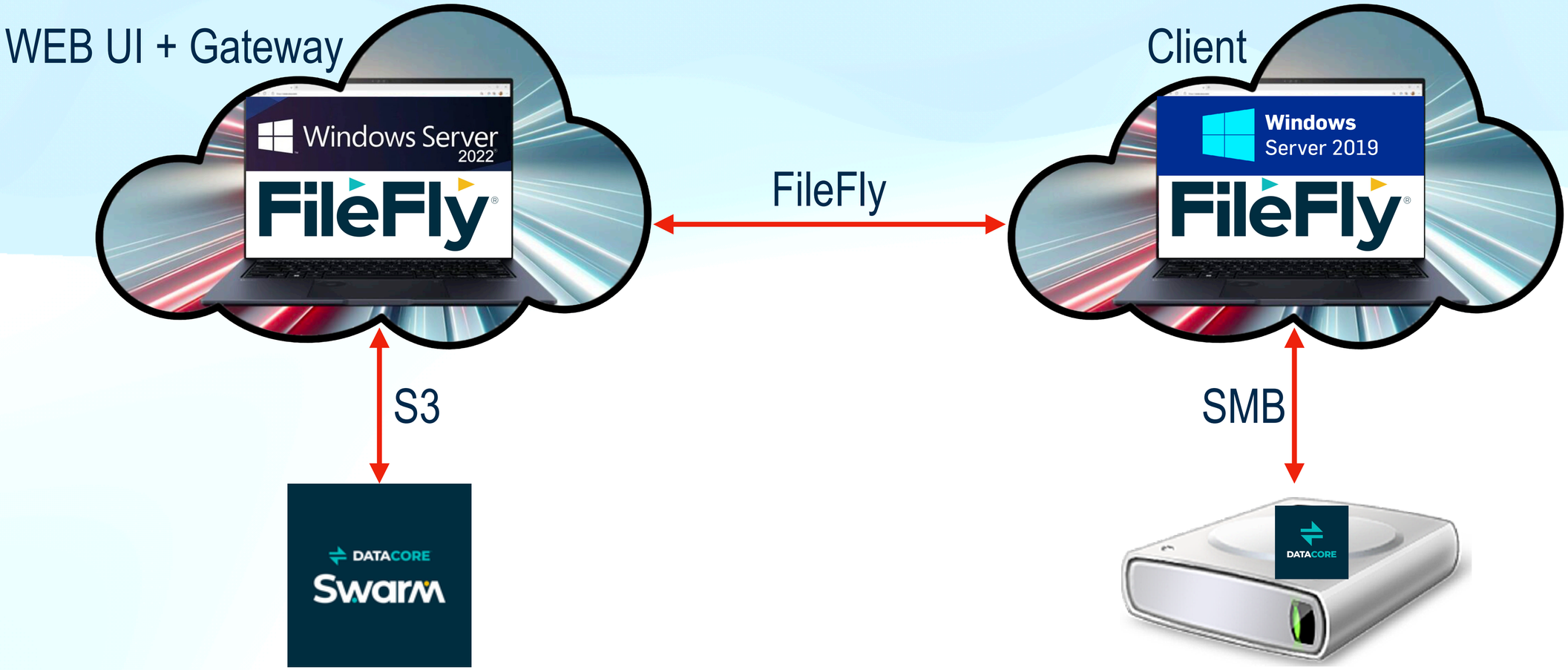
La solution FileFly

Le schéma logique (complet)



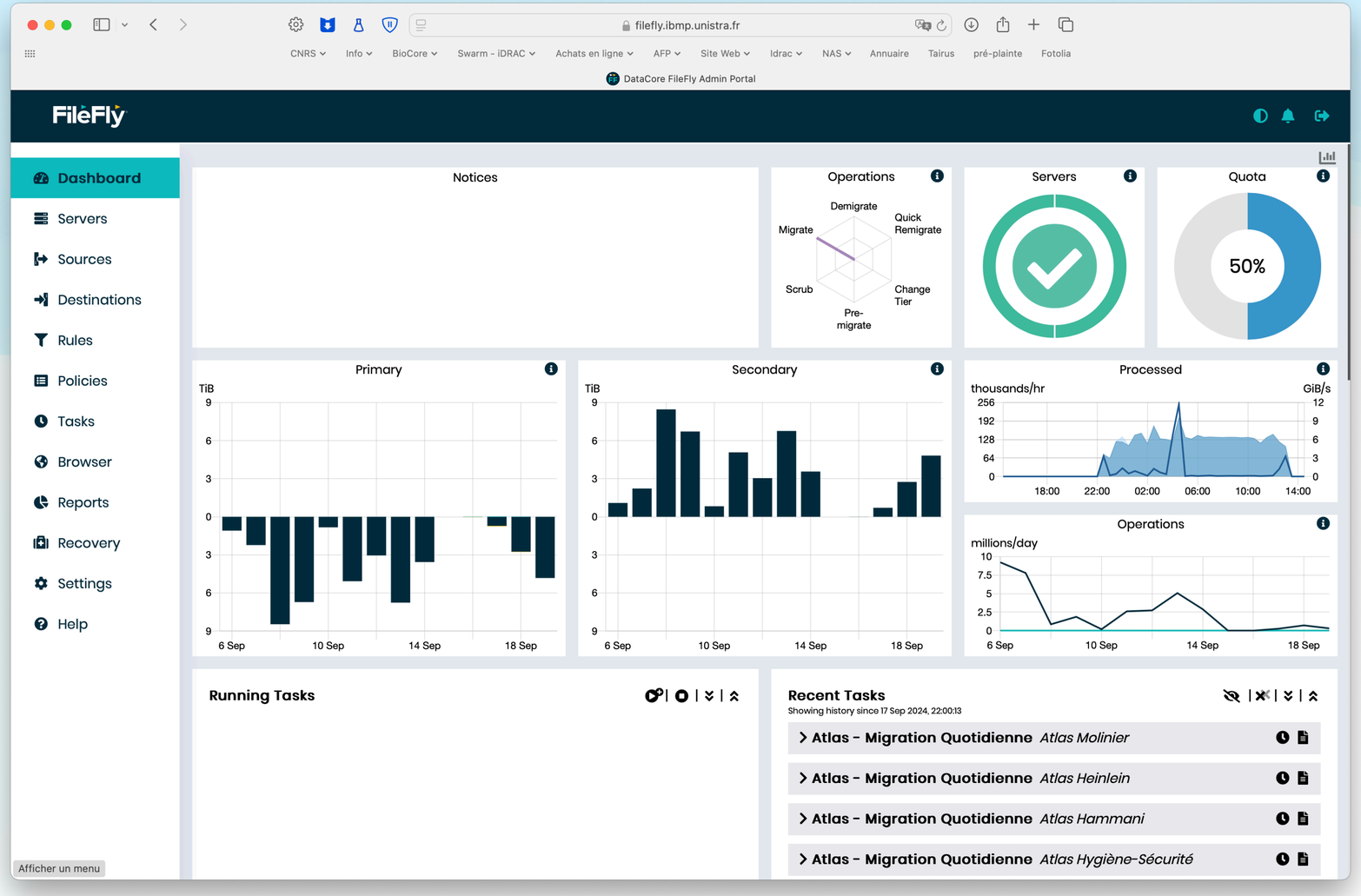
La solution FileFly

L’implémentation simplifiée IBMP



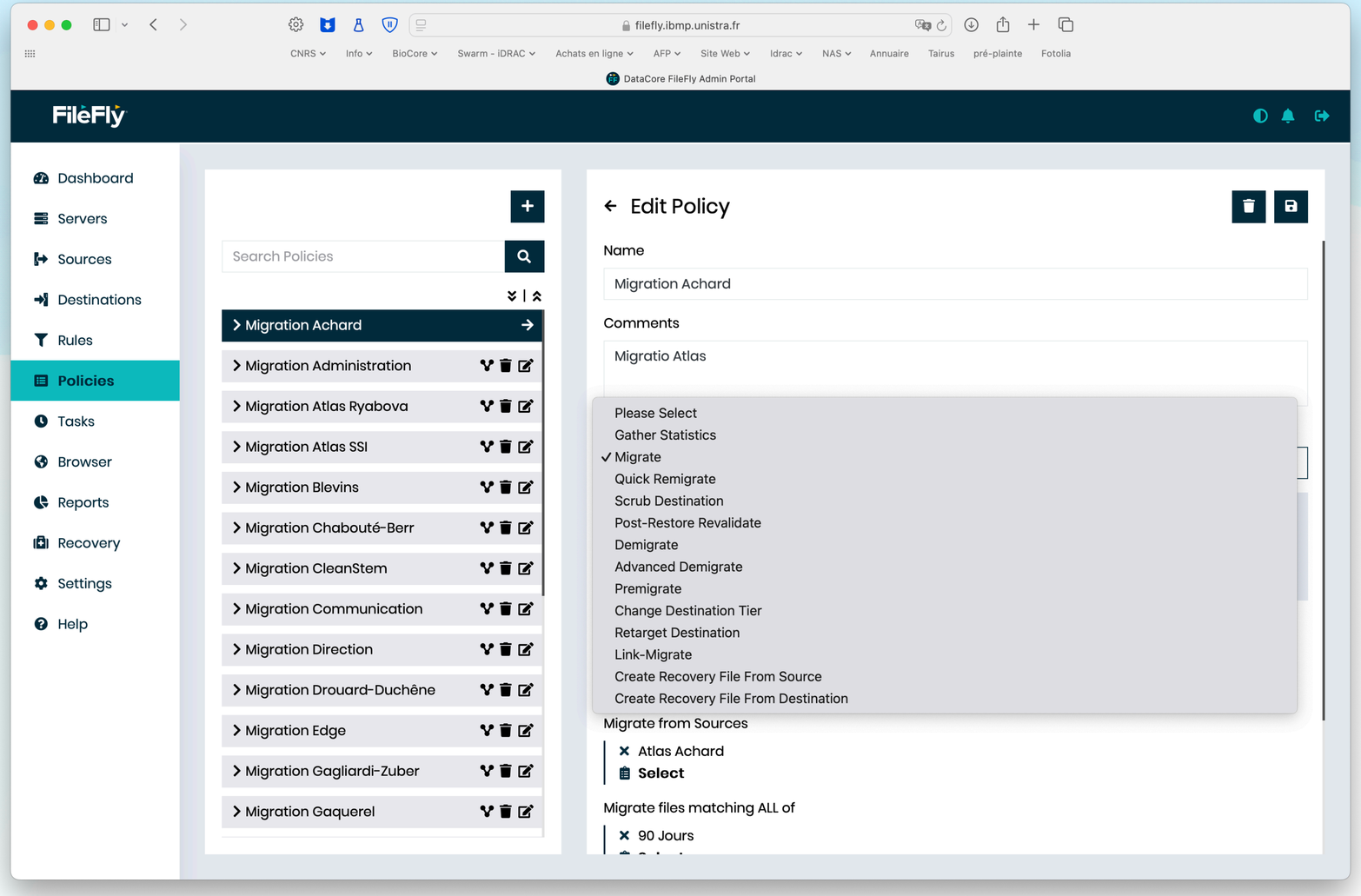
La solution FileFly

Une interface de gestion simple



La solution FileFly

Une interface de gestion simple



Le stockage à l’IBMP

Le bilan…

- Certes un investissement, mais la politique des licences "à vie" est un grand plus face aux velléités des acteurs du marché, et au final pas si cher…
- Des systèmes hyper-fiables qui nous ont sauvé la mise face à des pannes matérielles sévères (baie RAID qui part en sucette par exemple)
- Une simplicité au quotidien qui fait que la mobilisation humaine reste minime : pas besoin d’avoir une personne dédiée aux systèmes
- L’automatisation SMB <-> S3 est intéressante pour "désaturer" le système le plus coûteux
- La démocratisation du stockage objet reste un chantier à conduire…

Des questions ?

…ou une démo ?