Livre blanc technique

Présentation et architecture de la solution ECS

Résumé

Ce document fournit une présentation technique et conceptuelle d'ECS™, la plate-forme de stockage en mode objet software-defined à l'échelle du Cloud de Dell EMC™.

Février 2021

H14071.18

Révisions

Date	Description
Décembre 2015	version initiale
Mai 2016	Mise à jour pour 2.2.1
Septembre 2016	Mise à jour pour 3.0
Août 2017	Mise à jour pour 3.1
Mars 2018	Mise à jour pour 3.2
Septembre 2018	Mise à jour pour le matériel Gen 3
Février 2019	Mise à jour pour 3.3
Septembre 2019	Mise à jour pour 3.4
Février 2020	Modifications ECSDOC-628 mises à jour
Mai 2020	Mise à jour pour 3.5
Novembre 2020	Mise à jour pour 3.6
Février 2021	Mise à jour pour 3.6.1

Remerciements

Ce livre blanc a été conçu par les éléments suivants :

Auteur : Zhu, Jarvis

Table des matières

Synthèse

Les organisations ont besoin de solutions pour utiliser des services Cloud publics avec la même fiabilité et le même contrôle qu'une infrastructure Cloud privée. Dell EMC ECS est une plate-forme de stockage d'objet software-defined à l'échelle du Cloud prise en charge par le protocole IPv6 qui fournit des services de stockage S3, Atmos, CAS, Swift, NFSv3 et HDFS sur une plate-forme unique et moderne.

Avec ECS, les administrateurs peuvent gérer facilement l'infrastructure de stockage distribuée à l'échelle mondiale sous un espace de nommage mondial unique fournissant un accès au contenu en tout lieu. L'architecture hiérarchisée des composants de base d'ECS garantit la flexibilité et la résilience. Chaque couche peut être isolée et mise à l'échelle indépendamment des autres, avec une haute disponibilité.

La simplicité de l'accès à l'API RESTful pour les services de stockage est appréciée par les développeurs. L'utilisation de sémantiques HTTP comme GET et PUT simplifie la logique d'application requise par rapport aux opérations de fichiers traditionnelles mais familières basées sur des chemins. En outre, le système de stockage sous-jacent d'ECS est fortement cohérent, ce qui signifie qu'il peut garantir une réponse faisant autorité. Les applications nécessaires pour garantir la livraison de données faisant autorité sont en mesure de le faire sans logique de code complexe à l'aide d'ECS.

1 Introduction

Ce document fournit une présentation de la plate-forme de stockage en mode objet Dell EMC ECS. Il détaille l'architecture de conception d'ECS et les composants de base, tels que les mécanismes de protection des données et les services de stockage.

1.1 Public

Ce document s'adresse à toute personne souhaitant comprendre la valeur et l'architecture d'ECS. Son but est de présenter le contexte tout en fournissant des liens vers des informations complémentaires.

1.2 Périmètre

Ce document se concentre principalement sur l'architecture ECS. Il ne couvre pas les procédures d'installation, d'administration et de mise à niveau du logiciel ou du matériel ECS. Il ne couvre pas non plus les spécificités de l'utilisation et de la création d'applications avec les API ECS.

Ce document fait l'objet de mises à jour régulières qui coïncident généralement avec des versions majeures ou de nouvelles fonctionnalités.

2 Valeur d’ECS

ECS offre une valeur ajoutée significative aux entreprises et prestataires de services à la recherche d’une plate-forme conçue pour prendre en charge une croissance rapide des données. Les principaux avantages et fonctionnalités d’ECS qui permettent aux entreprises de gérer et de stocker globalement du contenu distribué à grande échelle sont les suivants :

- **Plate-forme à l’échelle du Cloud** - ECS est une plate-forme de stockage en mode objet à la fois pour les charges applicatives traditionnelles et celles de nouvelle génération. L’architecture software-defined hiérarchisée d’ECS offre une évolutivité sans limites. Les points clés des fonctionnalités sont les suivants :
  - Infrastructure d’objet distribuée à l’échelle mondiale
  - Évolutivité de plusieurs exaoctets sans limite de capacité sur le pool de stockage, le cluster ou l’environnement fédéré
  - Aucune limite au nombre d’objets dans un système, un espace de nommage ou un bucket
  - Efficacité des charges applicatives à la fois pour les fichiers de petite taille et les fichiers volumineux, sans limites de taille d’objet
- **Déploiement flexible** - ECS offre une flexibilité inégalée avec des fonctionnalités telles que :
  - Déploiement d’appliance
  - Déploiement logiciel uniquement avec prise en charge de matériel standard certifié ou personnalisé
  - Prise en charge multiprotocole : objet (S3, Swift, Atmos, CAS) et fichier (HDFS, NFSv3)
  - Plusieurs charges applicatives : applications modernes et archivage à long terme
  - Stockage secondaire pour la hiérarchisation sur le Cloud avec Data Domain et Isilon à l’aide de CloudPools
  - Stratégies de mise à niveau sans perturbation vers les modèles ECS de génération actuelle
- **Solution adaptée à l’entreprise** - ECS offre aux clients un meilleur contrôle de leurs ressources de données avec un stockage de niveau entreprise dans un système sécurisé et conforme doté de fonctionnalités telles que :
  - Chiffrement des données au repos (D@RE) avec rotation des clés et gestion des clés externe
  - Communication entre sites chiffrée
  - Désactive les ports 9101/9206 par défaut pour permettre aux organisations de respecter les politiques de conformité
  - Création de rapports, rétention des enregistrements basée sur des règles et des événements et renforcement de la plate-forme pour la conformité à la norme SEC 17a-4(f), y compris la gestion avancée de la rétention, comme la conservation en vue d’un litige et la gouvernance min/max
  - Conformité avec les directives pour le renforcement de la sécurité STIG (Security Technical Implementation Guide) de l’Agence des systèmes d’information de la Défense (DISA)
  - Authentification, autorisation et contrôles d’accès avec Active Directory et LDAP
  - Intégration avec l’infrastructure de surveillance et d’alerte (traps SNMP et SYSLOG)
  - Fonctionnalités d’entreprise améliorées (multitenancy, surveillance de la capacité et alertes)

- **Réduction du coût TCO** - ECS peut réduire considérablement le coût total de possession (TCO) par rapport au stockage traditionnel et au stockage dans le Cloud public. Il offre même un TCO inférieur à celui des bandes pour la rétention à long terme. Voici ses principales caractéristiques :

  - Espace de nommage global
  - Performances avec les fichiers de petite taille et de grande taille
  - Migration transparente de Centera
  - Compatibilité totale avec Atmos REST
  - Temps système de gestion réduit
  - Faible empreinte du datacenter
  - Taux élevé d’utilisation du stockage

La conception d’ECS est optimisée pour les exemples d’utilisation principaux suivants :

- **Applications modernes** : ECS a été conçu pour le développement moderne, notamment pour les applications Web, mobiles et Cloud nouvelle génération. Le développement d’applications est simplifié grâce à un stockage fortement cohérent. Outre qu’ils bénéficient d’un accès en lecture/écriture multi-utilisateur et multisite simultané, les développeurs n’ont pas besoin de réécrire leurs applications à mesure que la capacité d’ECS change et augmente.
- **Stockage secondaire** - ECS est utilisé en tant que stockage secondaire afin de libérer l’espace de stockage primaire occupé par les données rarement utilisées, tout en assurant un accès raisonnable à ces données. C’est le cas notamment pour les produits de hiérarchisation basés sur des règles tels que Data Domain Cloud Tier et Isilon CloudPools. GeoDrive, une application Windows, permet aux systèmes Windows d’accéder directement à ECS pour stocker les données.
- **Archivage géo-protégé** : ECS sert de Cloud sur site sécurisé et abordable pour l’archivage et la rétention à long terme. L’utilisation d’ECS en tant que niveau d’archivage peut réduire considérablement les capacités de stockage primaire. Afin de permettre une meilleure efficacité du stockage pour les cas d’utilisation d’archivage à froid, un schéma de codage d’effacement (EC) 10+2 est disponible en plus du schéma par défaut 12+4.
- **Référentiel de contenu à l’échelle mondiale** : les référentiels de contenu non structuré contenant des données telles que des images et des vidéos sont souvent stockés dans des systèmes de stockage coûteux, ce qui empêche les entreprises de gérer de manière rentable la prolifération des données. ECS permet de consolider plusieurs systèmes de stockage dans un référentiel de contenu unique, efficace et accessible dans le monde entier.
- **Stockage pour l’Internet of Things** : l’Internet of Things (IoT) offre de nouvelles opportunités de chiffre d’affaires aux entreprises capables de tirer parti de la valeur des données client. ECS propose une architecture IoT efficace pour la collecte des données non structurées à très grande échelle. Sans aucune limite quant au nombre d’objets, à la taille des objets ou aux métadonnées personnalisées, ECS est la plate-forme idéale pour stocker les données IoT. ECS peut également rationaliser certains flux de travail analytiques en permettant l’analyse directe des données sur la plate-forme ECS sans qu’il soit nécessaire d’exécuter de longs processus d’extraction, de transformation et de chargement (ETL). Les clusters Hadoop peuvent exécuter des requêtes à l’aide de données stockées dans ECS par une autre API de protocole, par exemple S3 ou NFS.
- **Référentiel de preuves de vidéosurveillance** : contrairement aux données IoT, les données de vidéosurveillance présentent un nombre de stockages en mode objet beaucoup plus petit, mais un encombrement beaucoup plus important par fichier. Si l’authenticité des données est importante, la rétention des données n’est pas aussi critique. ECS peut être une zone de réception à moindre coût ou un emplacement de stockage secondaire pour ces données. Le logiciel de gestion vidéo peut tirer le meilleur parti des puissantes fonctionnalités de métadonnées personnalisées pour le marquage des fichiers avec des détails importants comme l’emplacement de la caméra, les exigences de rétention et de protection des données. En outre, les métadonnées peuvent être utilisées pour définir le fichier à l’état en lecture seule, afin de maintenir une authenticité de la preuve dans le fichier.
- **Data Lakes et analytique** : les données et l’analytique constituent désormais un facteur de différenciation concurrentielle et une source majeure de génération de valeur pour les organisations. Toutefois, la transformation des données en ressources d’entreprise précieuses est un processus complexe qui peut facilement impliquer des dizaines de technologies, d’outils et d’environnements différents. ECS fournit un ensemble de services pour aider le client à collecter, stocker, gouverner et analyser les données à n’importe quelle échelle.

3 Architecture

ECS est conçu avec quelques principes de conception de base, tels que l'espace de nommage global à forte cohérence, la capacité de scale out, la sécurisation multitenancy et des performances supérieures pour les objets de petite taille et de grande taille. ECS s'appuie sur un système entièrement distribué suivant le principe des applications Cloud, où chaque fonction du système est construite sous la forme d'une couche indépendante. Avec cette conception, chaque couche est évolutive horizontalement sur tous les nœuds du système.
Les ressources sont réparties sur tous les nœuds afin d'améliorer la disponibilité et de partager la charge.

Cette section traitera en détail de l'architecture d'ECS ainsi que de la conception du logiciel et du matériel.

3.1 Présentation

ECS est déployé sur un ensemble de matériel standard qualifié ou sous la forme d'une appliance de stockage clé en main. Les principaux composants d'ECS sont les suivants :

- **Portail ECS et services de provisionnement** : interface utilisateur Web et CLI basées sur l'API pour le libre-service, l'automatisation, la création de rapport et la gestion des nœuds ECS. Cette couche gère également les services de gestion des licences, d'authentification, de multitenancy et de provisionnement, comme la création de l'espace de nommage.
- **Services de données** : services, outils et API pour la prise en charge de l'accès au système en mode fichier et objet.
- **Moteur de stockage** : service de base responsable du stockage et de l'extraction des données, de la gestion des transactions, de la protection et de la réplication des données en local et entre les sites.
- **Fabric** : service de clustering pour la gestion de l'intégrité, de la configuration et de la mise à niveau et des alertes.
- **Infrastructure** : SUSE Linux Enterprise Server 12 pour le système d'exploitation de base dans l'appliance clé en main ou systèmes d'exploitation Linux qualifiés pour une configuration avec du matériel standard.
- **Matériel** : appliance clé en main ou matériel standard qualifié.

La Figure 1 présente une vue graphique de ces couches qui sont décrites en détail dans les sections suivantes.

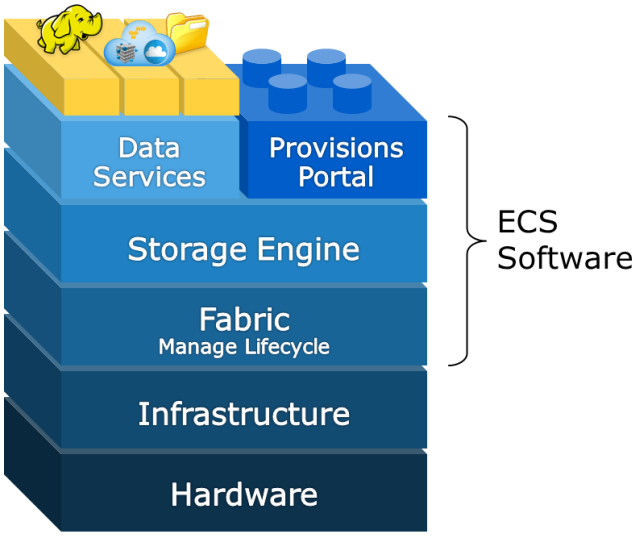

Figure 1 Couches d'architecture ECS

3.2 Portail ECS et services de provisionnement

Les administrateurs de stockage gèrent ECS à l'aide du portail ECS et des services de provisionnement. ECS fournit une interface utilisateur Web (WebUI) pour la gestion, l'octroi de licences et le provisionnement des nœuds ECS. Le portail possède des fonctionnalités complètes de création de rapport, incluant notamment :

- Utilisation de la capacité par site, pool de stockage, nœud et disque
- Surveillance des performances de la latence, du débit et de la progression de la réplication
- Informations de diagnostic, telles que l'état de restauration des nœuds et des disques

Le tableau de bord ECS fournit des informations générales sur l'intégrité et les performances au niveau du système. Cette vue unifiée améliore la visibilité globale du système. Les alertes notifient les utilisateurs des événements critiques, tels que les limites de capacité, les limites de quota, les défaillances des disques ou des nœuds ainsi que les défaillances logicielles. ECS fournit également une interface de ligne de commande pour l'installation, la mise à niveau et la surveillance d'ECS. L'accès aux nœuds pour l'utilisation de la ligne de commande s'effectue via SSH. La Figure 2 ci-dessous présente une capture d'écran du tableau de bord ECS.

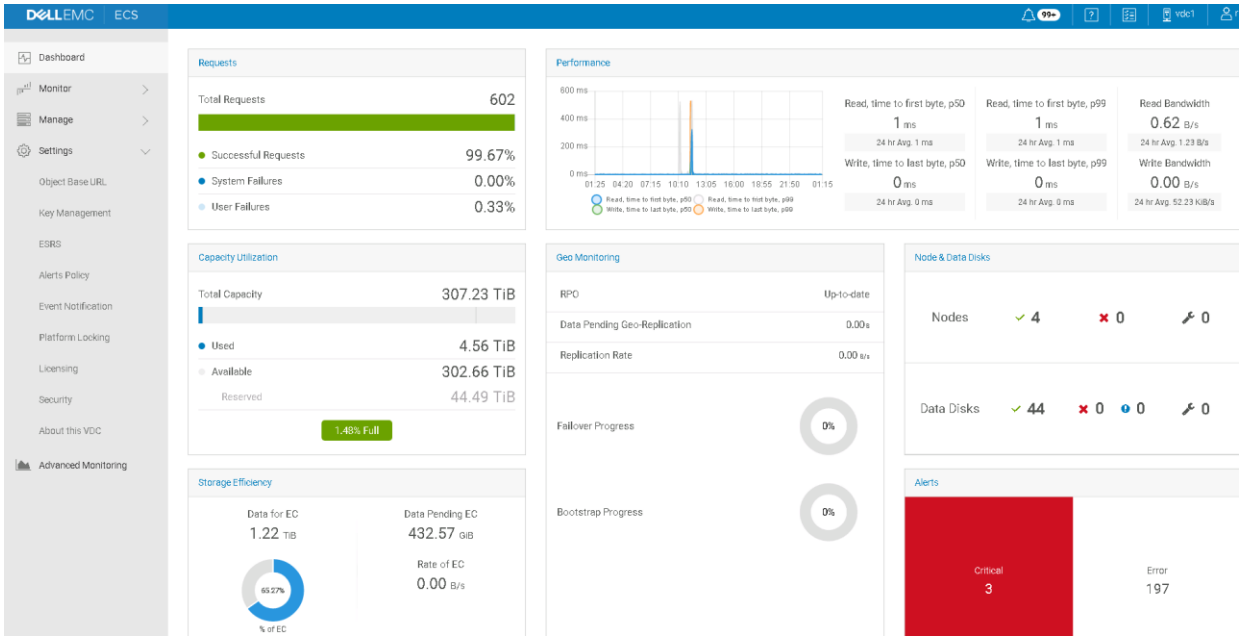

Figure 2 Tableau de bord de l'interface utilisateur Web d'ECS

Des rapports détaillés sur les performances sont disponibles dans l'interface utilisateur dans le dossier Advanced Monitoring. Les rapports s'affichent dans un tableau de bord Grafana. Des filtres permettent d'effectuer une recherche verticale dans les espaces de nommage, les protocoles ou les nœuds souhaités. La Figure 3 ci-dessous présente un exemple de rapport de performances du protocole S3.

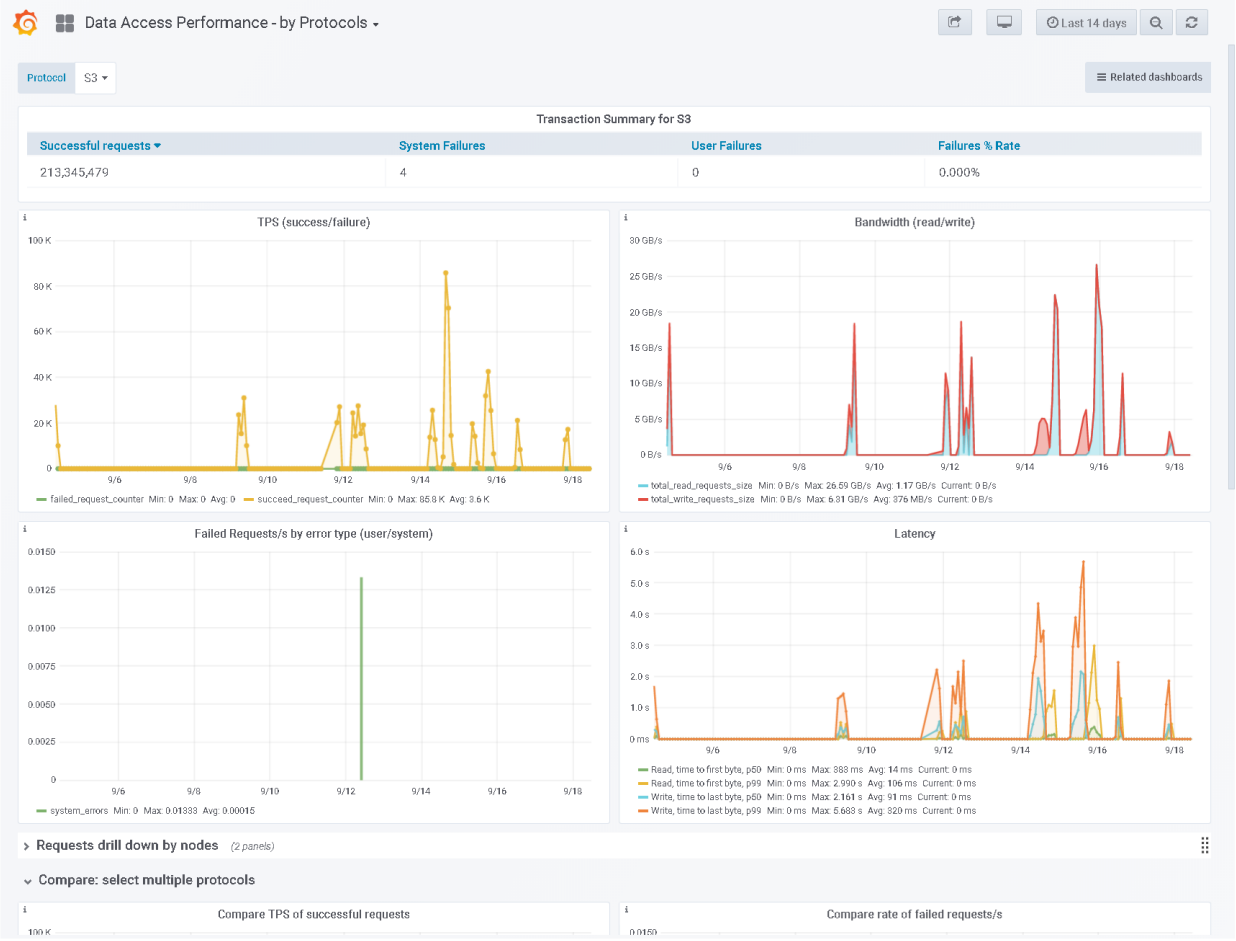

Figure 3 Visualisation de la surveillance avancée à l'aide de Grafana

ECS peut également être géré à l'aide des API RESTful. L'API de gestion permet aux utilisateurs d'administrer ECS à partir de leurs propres outils, scripts et applications nouvelles ou existantes. L'interface utilisateur Web d'ECS et les outils de ligne de commande sont créés à l'aide de l'ECS REST Management API.

ECS prend en charge les serveurs de notification d'événements suivants, qui peuvent être définis à l'aide de l'interface utilisateur Web, de l'API ou de la CLI :

- Serveurs SNMP (Simple Network Management Protocol)
- Serveurs Syslog

Pour des informations plus détaillées sur la configuration des services de notification, consultez le guide *ECS Administrator's Guide*.

3.3 Services de données

Des méthodes d'objet et de fichier standard permettent d'accéder aux services de stockage ECS. Pour l'accès par les protocoles S3, Atmos et Swift, les API RESTful sur HTTP sont utilisées. Pour le stockage dédié aux contenus fixes (CAS), une méthode d'accès ou un SDK propriétaire sont utilisés. ECS prend en charge toutes les procédures NFSv3 en mode natif, à l'exception de LINK. Les buckets ECS sont désormais accessibles via S3a.

ECS fournit un accès multiprotocole grâce auquel les données reçues au moyen d'un protocole sont accessibles via d'autres protocoles. Cela signifie que les données peuvent être acquises via S3 et modifiées via NFSv3 ou Swift, ou vice versa. Cet accès multiprotocole présente quelques exceptions du fait de la sémantique du protocole et des représentations de la conception du protocole. Le Tableau 1 met en évidence les méthodes d'accès et les protocoles qui interagissent.

Tableau 1 Services de données ECS pris en charge et interopérabilité des protocoles

Protocoles		**Pris en charge**	**Interopérabilité**
Objet	S3	Fonctionnalités supplémentaires comme les mises à jour de la plage d'octets et les ACL enrichies	HDFS, NFS, Swift
	Atmos	Version 2.0	NFS (uniquement les objets basés sur un chemin d'accès et non ceux de style ID d'objet)
	Swift	API v2 et authentification Swift et Keystone v3	HDFS, NFS, S3
	CAS	SDK v3.1.544 ou versions supérieures	s.o.
Fichier	HDFS	Compatibilité de Hadoop 2.7	S3, NFS, Swift
	NFS	NFSv3	S3, Swift, HDFS, Atmos (uniquement les objets basés sur un chemin et non ceux de style ID d'objet)

Les services de données, également appelés « head services », sont chargés de traiter les demandes des clients, d'extraire les informations requises et de les transmettre au moteur de stockage pour traitement ultérieur. Tous les head services sont associés à un processus unique, *dataheadsvc,* qui s'exécute au sein de la couche d'infrastructure. Ce processus est encapsulé dans un conteneur Docker appelé *object-main,* qui s'exécute sur chaque nœud d'ECS. Pour plus d'informations, reportez-vous à la section *infrastructure* de ce document. Les exigences relatives aux ports de service des protocoles ECS, telles que le port 9020 pour la communication S3, sont disponibles dans la version la plus récente du guide *ECS Security Configuration Guide*.

3.3.1 Objet

ECS prend en charge les API S3, Atmos, Swift et CAS pour l'accès aux objets. À l'exception de CAS, les objets ou les données sont écrits, récupérés, mis à jour et supprimés via les appels HTTP ou HTTPS GET, POST, PUT, DELETE et HEAD. Pour CAS, des communications TCP standard et des appels et méthodes d'accès spécifiques sont utilisés.

ECS fournit une fonctionnalité de recherche de métadonnées d'objets à l'aide d'un langage de requête riche. Il s'agit d'une fonctionnalité puissante d'ECS qui permet aux clients d'objets S3 de rechercher des objets au sein de buckets à l'aide de métadonnées système et personnalisées. Bien qu'il soit possible d'effectuer une recherche à l'aide de n'importe quelles métadonnées, en la faisant porter sur les métadonnées qui ont été spécifiquement configurées pour être indexées dans un bucket, ECS peut renvoyer des requêtes plus rapidement, en particulier pour les buckets avec des milliards d'objets.

Jusqu'à trente champs de métadonnées définis par l'utilisateur peuvent être indexés par bucket. Les métadonnées sont spécifiées au moment de la création du bucket. La fonctionnalité de recherche de métadonnées peut être activée sur les buckets faisant l'objet d'un chiffrement côté serveur. Toutefois, tout attribut de métadonnées utilisateur indexé utilisé en tant que clé de recherche ne sera pas chiffré.

Remarque : L'écriture de données dans les buckets configurés pour indexer les métadonnées a un impact sur les performances. Cet impact sur les opérations s'accentue lorsque le nombre de champs indexés augmente. Ce point doit faire l'objet d'une attention particulière lors du choix de l'indexation des métadonnées dans un bucket et, le cas échéant, du nombre d'index à conserver.

Pour les objets CAS, l'API de requête CAS offre une fonction similaire de recherche d'objets sur la base des métadonnées correspondantes qui sont conservées. Cette fonction n'a pas besoin d'être activée explicitement.

Pour plus d'informations sur les API ECS et les API de recherche de métadonnées, consultez le *Guide d'accès aux données d'ECS* le plus récent. Pour les SDK Atmos et S3, consultez le site GitHub Dell EMC Data Services SDK ou Dell EMC ECS. Pour CAS, consultez le site de la communauté Centera. La communauté ECS donne accès à de nombreux exemples, ressources et fournit une assistance aux développeurs.

Les applications clientes telles que S3 Browser et Cyberduck offrent un moyen rapide de tester les données stockées dans ECS ou d'y accéder. ECS Test Drive est fourni librement par Dell EMC et permet d'accéder à un système ECS public à des fins de test et de développement. Après l'inscription à ECS Test Drive, des points de terminaison REST sont fournis avec les informations d'identification de chacun des protocoles d'objet. N'importe qui peut utiliser ECS Test Drive pour tester son application API S3.

Remarque : Seul le nombre de métadonnées pouvant être indexées par bucket est limité à 30 dans ECS. Il n'y a aucune limite au nombre total de métadonnées personnalisées stockées par objet, seul le nombre indexé pour la recherche rapide est limité.

3.3.2 HDFS

ECS peut stocker les données du système de fichiers Hadoop. En tant que système de fichiers compatible avec Hadoop, ECS permet aux organisations de créer de larges référentiels d'entreprise que l'analytique Hadoop peut consommer et traiter. Le service de données HDFS est compatible avec Apache Hadoop 2.7, avec la prise en charge des ACL à granularité fine et des attributs de système de fichiers étendus.

ECS a été validé et testé avec Hortonworks (HDP 2.7). ECS prend également en charge les services tels que YARN, MapReduce, Pig, Hive/Hiveserver2, HBase, Zookeeper, Flume, Spark et Sqoop.

3.3.2.1 Prise en charge de Hadoop S3A

ECS prend en charge le client Hadoop S3A pour le stockage des données Hadoop. S3A est un connecteur Open Source pour Hadoop, basé sur le SDK Amazon Web Services (AWS) officiel. Il a été créé pour résoudre les problèmes de mise à l'échelle du stockage et de coûts rencontrés par de nombreux administrateurs Hadoop avec HDFS. Hadoop S3A connecte les clusters Hadoop à une zone de stockage d'objets compatible S3 qui se trouve dans le Cloud public, le Cloud hybride ou sur site.

Remarque : La prise en charge de S3A est disponible sur Hadoop 2.7 ou version supérieure.

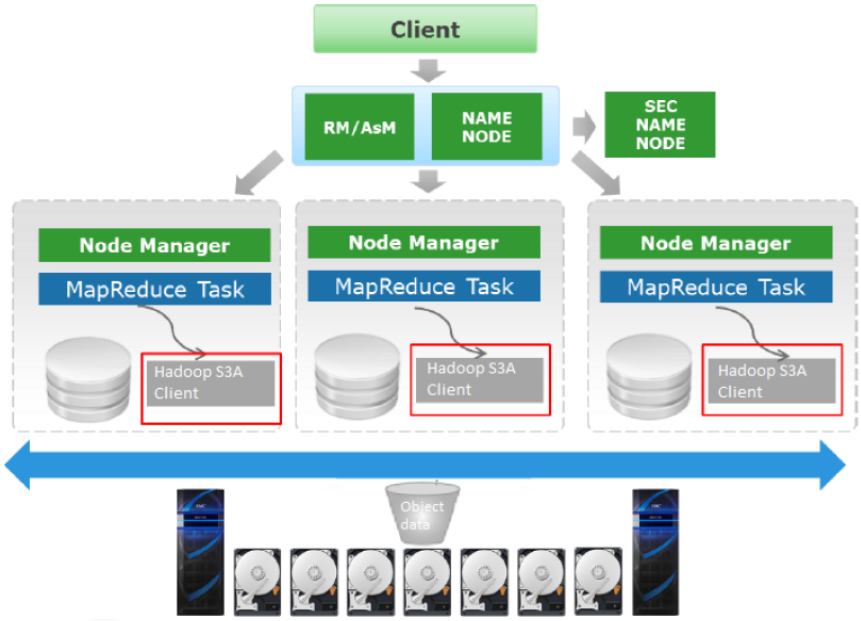

Figure 4 Hadoop et architecture ECS

Comme illustré par la Figure 4, lorsque le client configure le cluster Hadoop sur un HDFS traditionnel, sa configuration S3A indique aux données d'objet ECS d'effectuer toutes les activités HDFS. Sur chaque nœud Hadoop HDFS, les composants Hadoop traditionnels utilisent le client S3A Hadoop pour assurer l'activité HDFS.

Analyse de la configuration Hadoop à l'aide de la console de service ECS

La console de service ECS (SC) peut lire et interpréter vos paramètres de configuration Hadoop pour ce qui relève des connexions à ECS pour S3A. En outre, SC propose une fonction *Get_Hadoop_Config* qui lit la configuration du cluster Hadoop et vérifie les valeurs des paramètres S3A, s'assurant qu'ils ne contiennent pas de faute de frappe ni d'erreur. Contactez l'équipe de support ECS pour obtenir de l'aide sur l'installation de la console de service ECS.

Implémentation de Privacera avec Hadoop S3A

Cette solution est un fournisseur tiers qui a implémenté un agent Hadoop côté client et s'intègre avec Ambari pour la sécurité granulaire S3 (AWS et ECS). Bien que Privacera prenne en charge la distribution Cloudera d'Hadoop (CDH), Cloudera (un autre fournisseur tiers) ne prend pas en charge cette solution sur CDH.

Remarque : Les utilisateurs CDH doivent utiliser les services de sécurité ECS IAM. S'il vous faut un accès sécurisé à S3A sans utiliser ECS IAM, contactez l'équipe de support technique.

Pour plus d'informations sur la prise en charge de S3A, reportez-vous au guide *ECS Data Access Guide* le plus récent.

Sécurité d'Hadoop S3A

ECS IAM permet à l'administrateur Hadoop de configurer des règles d'accès pour contrôler l'accès aux données S3A Hadoop. Une fois les règles d'accès définies, les administrateurs Hadoop peuvent paramétrer deux types d'accès utilisateur :

- Utilisateurs/Groupes IAM
  - Créer des groupes IAM rattachés aux stratégies

  - Créer des utilisateurs IAM membres d'un groupe IAM

- Assertions SAML (utilisateurs fédérés)

  - Créer des rôles IAM rattachés aux stratégies
  - Configurer CrossTrustRelationship entre le fournisseur d'identité (AD FS) et ECS pour mapper les groupes AD aux rôles IAM

L'administrateur ECS et l'administrateur Hadoop doivent travailler ensemble pour prédéfinir des règles appropriées. Les exemples fictifs suivants décrivent trois types d'utilisateurs Hadoop pour lesquels nous allons créer des stratégies. En voici la liste :

- **Administrateur Hadoop** : effectue toutes les opérations, sauf la création et la suppression de buckets
- **Utilisateur Hadoop avancé** : effectue toutes les opérations, sauf la création de buckets, la suppression de buckets et la suppression d'objets
- **Utilisateur Hadoop en lecture seule** : liste et lit les objets uniquement

Pour plus d'informations sur ECS IAM, voir ECS IAM à la page 39.

3.3.2.2 Support clients ECS HDFS

ECS a été intégré avec Ambari, ce qui vous permet de déployer facilement le fichier jar du client ECS HDFS et de spécifier ECS HDFS comme système de fichiers par défaut dans un cluster Hadoop. Le fichier jar est installé sur chaque nœud au sein d'un cluster Hadoop participant. ECS offre des fonctions de stockage et de système de fichiers équivalentes à ce que les nœuds de noms et de données effectuent dans un déploiement Hadoop. ECS rationalise le flux de travail de Hadoop en éliminant la nécessité de migrer les données vers un DAS Hadoop local et/ou en créant un minimum de trois copies. La Figure 5 ci-dessous montre le fichier jar du client ECS HDFS installé sur chaque nœud de calcul Hadoop et le flux de communication général.

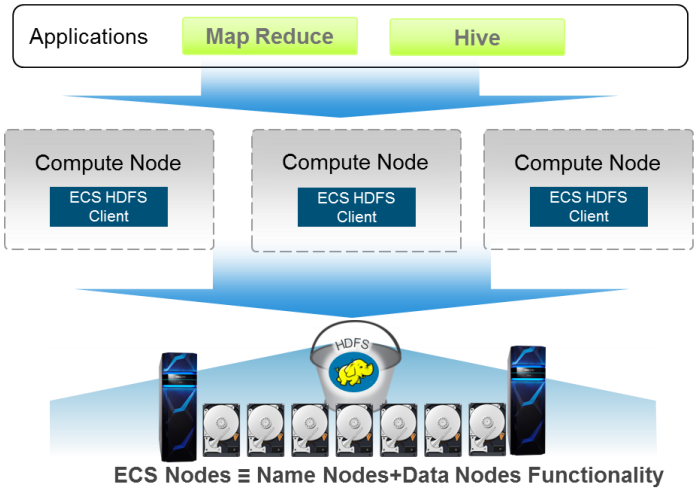

Figure 5 ECS servant de nœuds de noms et de données pour un cluster Hadoop

Les autres améliorations apportées à ECS pour HDFS incluent les éléments suivants :

- **Authentification d'utilisateur proxy** : emprunt d'identité pour Hive, HBase et Oozie.
- **Sécurité** : mise en œuvre de l'ACL côté serveur et ajout du superutilisateur et du groupe superutilisateur Hadoop, ainsi que du groupe par défaut sur les buckets.

3.3.3 NFS

ECS inclut la prise en charge native des fichiers avec NFSv3. Les principales fonctionnalités du service de données de fichier NFSv3 sont les suivantes :

- **Espace de nommage global** : accès aux fichiers à partir de n'importe quel nœud sur n'importe quel site.
- **Verrouillage global** : dans NFSv3, le verrouillage est **uniquement recommandé**. ECS prend en charge les implémentations de clients conformes qui autorisent les verrous partagés et exclusifs, basés sur des plages et obligatoires.
- **Accès multiprotocole** : accès aux données à l'aide de différentes méthodes de protocole.

Les exportations NFS, les autorisations et les mappages de groupes d'utilisateurs sont créés à l'aide de l'interface utilisateur Web ou de l'API. Les clients compatibles NFSv3 montent les exportations à l'aide des noms d'espace de nommage et de bucket. Voici un exemple de commande permettant de monter un bucket :

```
mount –t nfs –o vers=3 s3.dell.com:/namespace/bucket
```

Pour assurer la transparence du client lors d'une défaillance d'un nœud, un répartiteur de charge est recommandé pour ce flux de travail.

ECS a étroitement intégré les autres implémentations de serveur NFS, telles que *lockmgr*, *statd*, *nfsd* et *mountd*. Ces services ne dépendent donc pas de la couche d'infrastructure (système d'exploitation hôte) à gérer. La prise en charge de NFSv3 offre les fonctionnalités suivantes :

- Aucune limite de conception sur le nombre de fichiers ou de répertoires
- Taille d'écriture de fichier pouvant atteindre 16 To.
- Possibilité de mise à l'échelle sur un maximum de 8 sites avec un seul espace de nommage global (une seule exportation)
- Prise en charge de l'authentification Kerberos et AUTH_SYS

Les services de fichiers NFS traitent les demandes NFS provenant des clients ; toutefois, les données sont stockées sous forme d'objets dans ECS. Un descripteur de fichier NFS est mappé à un ID d'objet. Dans la mesure où le fichier est fondamentalement mappé à un objet, NFS dispose de fonctions comme le service de données en mode objet, ce qui inclut :

- Gestion des quotas au niveau du bucket
- Chiffrement au niveau de l'objet
- Write-Once-Read-Many (WORM) au niveau du bucket
  - Lors de la création du bucket, WORM est mis en œuvre à l'aide de la période de validation automatique.
  - WORM ne s'applique qu'aux buckets non conformes.

3.3.4 Connecteurs et passerelles

Plusieurs produits logiciels tiers ont la possibilité d'accéder au stockage en mode objet ECS. Des fournisseurs de logiciel indépendants (ISV) comme Panzura, Cteraet Syncplicity créent une couche de services qui offre un accès client au stockage en mode objet ECS via des protocoles traditionnels tels que SMB/CIFS, NFS et iSCSI. Les organisations peuvent également accéder aux données ou les télécharger dans le stockage ECS avec les produits Dell EMC suivants :

- **Isilon CloudPools** : hiérarchisation des données basée sur des règles dans ECS à partir du stockage Isilon.
- **Data Domain Cloud Tier** : hiérarchisation native automatisée des données dédupliquées vers ECS à partir de Data Domain pour une rétention à long terme. Data Domain Cloud Tier fournit une solution sécurisée et rentable pour chiffrer les données dans le Cloud avec une réduction de l'encombrement du stockage et de la bande passante réseau.
- **Geodrive** : service de stockage à base de fichiers stub ECS pour ordinateurs de bureau et serveurs Microsoft® Windows®.

3.4 Moteur de stockage

Le moteur de stockage est au cœur d'ECS. La couche du moteur de stockage contient les principaux composants responsables du traitement des demandes, ainsi que du stockage, de la récupération, de la protection et de la réplication des données.

Cette section décrit les principes de conception et la façon dont les données sont représentées et gérées en interne.

3.4.1 Services de stockage

Le moteur de stockage ECS inclut les services suivants, comme illustré sur la Figure 6.

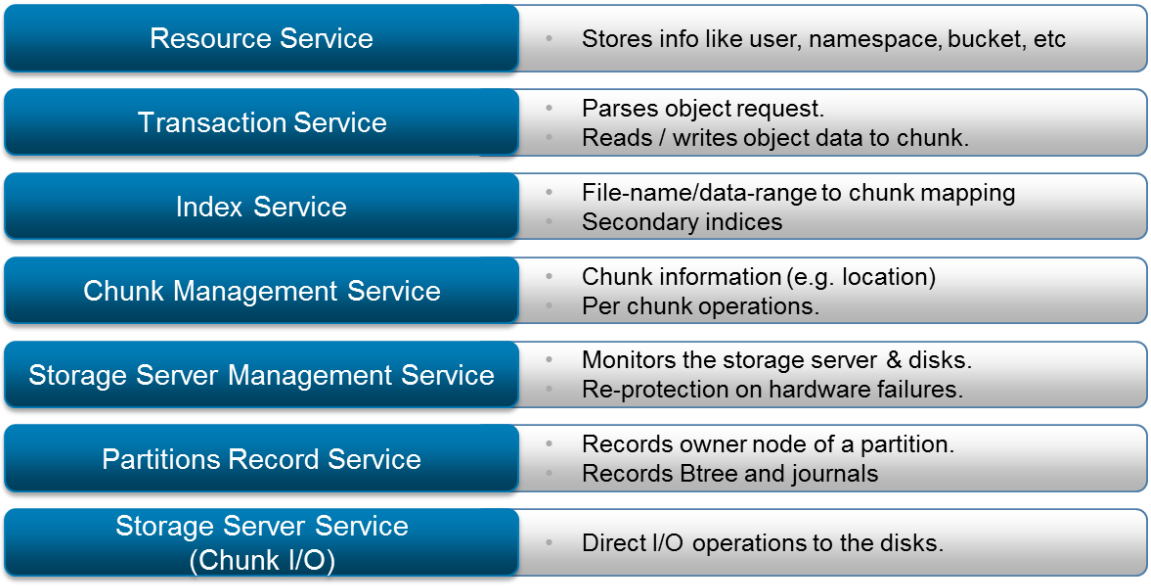

Figure 6 Services du moteur de stockage

Les services du moteur de stockage sont encapsulés dans un conteneur Docker. Celui-ci s'exécute sur chaque nœud ECS pour fournir un service distribué et partagé.

3.4.2 Data

Les principaux types de données stockées dans ECS peuvent être résumés comme suit :

- **Données** : contenu stocké au niveau de l'application ou de l'utilisateur, par exemple une image. Le terme « données » est synonyme d'objet, fichier ou contenu. Les applications peuvent stocker un nombre illimité de métadonnées personnalisées avec chaque objet. Le moteur de stockage écrit les données et les métadonnées personnalisées fournies par l'application dans un référentiel logique. Les métadonnées personnalisées sont une fonctionnalité robuste des systèmes de stockage modernes qui fournissent des informations complémentaires ou permettent la catégorisation des données stockées. Les métadonnées personnalisées sont structurées en paires clé-valeur et fournies avec les demandes d'écriture.
- **Métadonnées système** : informations système et attributs relatifs aux données utilisateur et aux ressources système. Les métadonnées système peuvent être grossièrement classées comme suit :
  - **Identifiants et descripteurs** : ensemble d'attributs utilisés en interne pour identifier les objets et leurs versions. Les identifiants sont soit des ID numériques, soit des valeurs de hachage qui ne sont pas utilisées à l'extérieur du contexte du logiciel ECS. Les descripteurs définissent des informations telles que le type de codage.
  - **Clés de chiffrement sous forme chiffrée** : les clés de chiffrement des données sont considérées comme des métadonnées système. Elles sont stockées sous forme chiffrée au sein de la structure de la table de répertoire principale.
  - **Balises internes** : ensemble d'indicateurs permettant de déterminer si les mises à jour de la plage d'octets ou le chiffrement sont activés et de coordonner la mise en cache et la suppression.
  - **Informations sur l'emplacement** : ensemble d'attributs comprenant l'emplacement de l'index et des données, par exemple les décalages d'octets.
  - **Horodatages** : ensemble d'attributs qui assure le suivi de l'heure, par exemple pour la création ou la mise à jour d'un objet.
  - **Informations de configuration/tenancy** : espace de nommage et contrôle d'accès aux objets.

Les métadonnées de données et système sont écrites par *fragments* dans ECS. Un fragment ECS est un conteneur logique d'espace contigu de 128 Mo. Chaque fragment peut avoir des données provenant de différents objets, comme illustré ci-dessous sur la Figure 7. ECS utilise l'indexation pour effectuer le suivi de toutes les parties d'un objet susceptibles d'être réparties sur différents fragments et nœuds.

Les fragments sont écrits selon un modèle de type ajout seul. Le comportement de type ajout seul signifie que la demande de modification ou de mise à jour d'un objet existant émise par une application ne modifie pas ni ne supprime les données écrites précédemment dans un fragment, mais écrit plutôt les nouvelles modifications ou mises à jour dans un nouveau fragment. Par conséquent, aucun verrouillage n'est requis pour les E/S. De même, aucune invalidation du cache n'est nécessaire. La conception ajout seul simplifie également la gestion des versions de données. Les anciennes versions des données sont conservées dans les fragments précédents. Si la gestion des versions S3 est activée et si une version plus ancienne des données est requise, elle peut être récupérée ou restaurée vers une version précédente à l'aide de l'API REST S3.

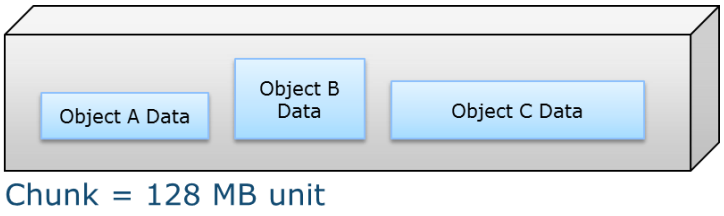

Figure 7 Fragment de 128 Mo stockant les données de trois objets

La section *Intégrité et protection des données* ci-dessous explique comment les données sont protégées au niveau des fragments.

3.4.3 Gestion des données

ECS utilise un ensemble de tables logiques pour stocker les informations relatives aux objets. Les paires clé-valeur sont finalement stockées sur le disque dans une arborescence B+ pour une indexation rapide des emplacements de données. En stockant la paire clé-valeur dans un arbre de recherche équilibré comme l'arborescence B+, l'emplacement des données et des métadonnées est accessible rapidement. ECS met en œuvre une arborescence de fusion à deux niveaux structurée à la manière d'un journal, dans laquelle cohabitent deux structures d'arbre : une arborescence de petite taille se trouve dans la mémoire (table de mémoire) et l'arborescence B+ principale réside sur le disque. La recherche des paires clé-valeur s'effectue d'abord dans la mémoire, puis, si nécessaire, au niveau de l'arborescence principale B+ sur le disque.
Les entrées de ces tables logiques sont d'abord enregistrées dans le journal, puis elles sont écrites sur des disques sous forme de fragments en miroir triple. Les journaux permettent de suivre les transactions qui n'ont pas encore été validées et inscrites dans l'arborescence B+. Après la consignation de chaque transaction dans un journal, la table en mémoire est mise à jour. Dès que la table en mémoire est saturée ou après un certain laps de temps, elle est fusionnée, triée ou vidée vers l'arborescence B+ sur le disque. Le nombre de fragments de journal utilisés par le système est négligeable par rapport aux fragments de l'arborescence B+. La Figure 8 illustre ce processus.

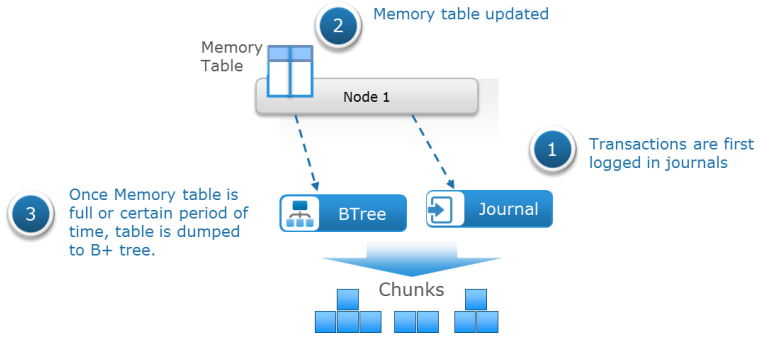

Figure 8 Table de mémoire vidée vers l'arborescence B+

Les informations stockées dans la table d'objets (OB) sont présentées ci-dessous dans le Tableau 2. La table OB contient les noms d'objets et l'emplacement de leur fragment selon un décalage et une longueur spécifiques dans ce fragment. Dans cette table, le nom de l'objet est la clé de l'index et la valeur est l'emplacement du fragment. La couche d'index dans le moteur de stockage est responsable du mappage entre nom d'objet et fragment.

Tableau 2 Entrées de la table d'objets

Object Name	Emplacement du fragment
ImgA	• C1:décalage:longueur
FileB	• C2:décalage:longueur • C3:décalage:longueur

La table de fragments (CT) enregistre l'emplacement de chaque fragment, comme indiqué dans le Tableau 3.

Tableau 3 Entrées de la table de fragments

ID de fragment	Emplacement
C1	• nœud1:disque1:fichier1:décalage1:longueur • nœud2:disque2:fichier1:décalage2:longueur • nœud3:disque2:fichier6:décalage:longueur

ECS a été conçu en tant que système distribué, de sorte que le stockage et l'accès aux données sont répartis sur tous les nœuds. Les tables utilisées pour gérer les données et les métadonnées des objets se remplissent avec le temps, à mesure que le stockage est utilisé et s'accroît. Les tables sont divisées en partitions et affectées à différents nœuds, chaque nœud devenant le propriétaire des partitions qu'il héberge pour chacune des tables. Par exemple, pour obtenir l'emplacement d'un fragment, la table d'enregistrements de partition (PR) est interrogée afin de déterminer le nœud propriétaire qui connaît l'emplacement du fragment. Une table PR de base est illustrée dans le Tableau 4 ci-dessous.

Tableau 4 Entrées de la table d'enregistrements de partition

ID de la partition	Propriétaire
P1	Nœud 1
P2	Nœud 2
P3	Nœud 3

En cas de panne d'un nœud, les autres nœuds prennent possession de ses partitions. Les partitions sont recréées par la lecture de la racine de l'arborescence B+ et la relecture des journaux stockés sur le disque. La Figure 9 illustre le basculement de la propriété de la partition.

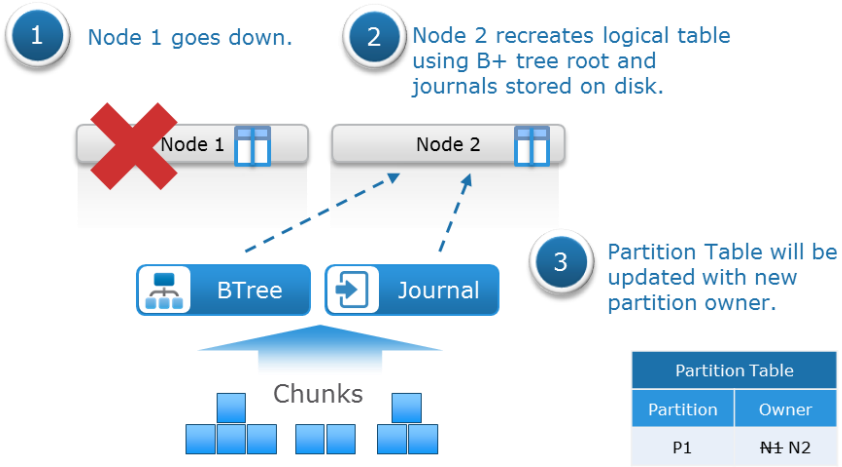

Figure 9 Basculement de la propriété de la partition

3.4.4 Flux de données

Les services de stockage sont disponibles à partir de n'importe quel nœud. Les données sont protégées par des segments EC distribués sur les disques, les nœuds et les racks. ECS exécute une somme de contrôle et stocke le résultat de chaque opération d'écriture. Si les premiers octets de données sont compressibles, ECS compresse les données. Pour les lectures, les données sont décompressées et la somme de contrôle stockée est validée. Voici un exemple de flux de données pour une écriture en cinq étapes :

1. Le client envoie une demande de création d'objet à un nœud.
2. Le nœud qui prend en charge la demande écrit les nouvelles données de l'objet dans un fragment repo (abréviation anglaise de référentiel).
3. Si l'opération d'écriture sur disque réussit, une transaction PR est effectuée pour obtenir le nom et l'emplacement du fragment.
4. Le propriétaire de la partition enregistre la transaction dans les entrées du journal.
5. Une fois la transaction enregistrée dans le journal, un accusé de réception est envoyé au client.

La Figure 10 ci-dessous présente un exemple de flux de données pour la lecture d'une architecture de disque dur de type Gen2 et EX300, EX500 et EX3000 :

1. Une demande de lecture d'objet est envoyée du client vers le nœud 1.
2. Le nœud 1 utilise une fonction de hachage sur le nom de l'objet pour déterminer quel nœud est propriétaire de la partition de la table logique où se trouvent ces informations d'objet. Dans cet exemple, le nœud 2 est le propriétaire, et c'est donc lui qui effectue une recherche dans les tables logiques pour obtenir l'emplacement du fragment. Dans certains cas, la recherche peut se produire sur deux nœuds différents, par exemple lorsque l'emplacement n'est pas mis en cache dans les tables logiques du nœud 2.
3. À l'étape précédente, l'emplacement du fragment est fourni au nœud 1 qui émet ensuite une demande de lecture de décalage d'octet sur le nœud contenant les données, le nœud 3 dans cet exemple, et qui envoie les données au nœud 1.
4. Le nœud 1 envoie les données au client demandeur.

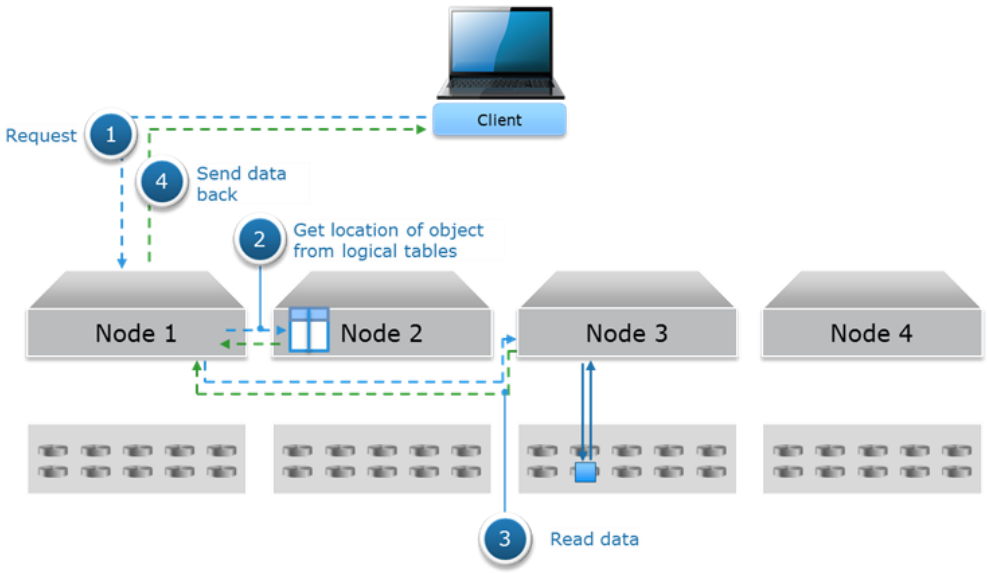

Figure 10 Flux de données de lecture pour l'architecture de disque dur

La Figure 11 ci-dessous présente un exemple de flux de données pour la lecture d'une architecture All-Flash de type EXF900 :

1. Une demande de lecture d'objet est envoyée du client vers le nœud 1.
2. Le nœud 1 utilise une fonction de hachage sur le nom de l'objet pour déterminer quel nœud est propriétaire de la partition de la table logique où se trouvent ces informations d'objet. Dans cet exemple, le nœud 2 est le propriétaire, et c'est donc lui qui effectue une recherche dans les tables logiques pour obtenir l'emplacement du fragment. Dans certains cas, la recherche peut se produire sur deux nœuds différents, par exemple lorsque l'emplacement n'est pas mis en cache dans les tables logiques du nœud 2.
3. À l'étape précédente, l'emplacement du fragment est fourni au nœud 1 qui lit ensuite directement les données du nœud 3.
4. Le nœud 1 envoie les données au client demandeur.

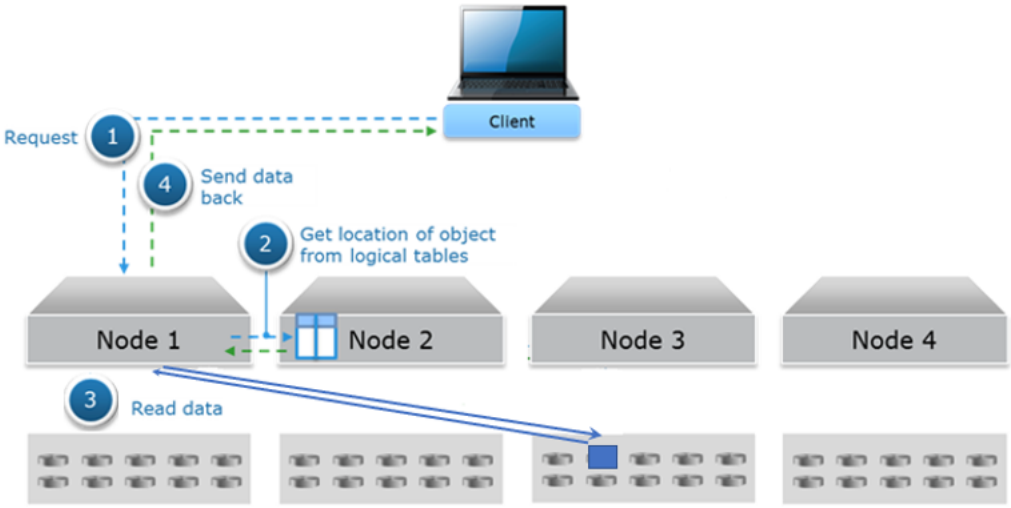

Figure 11 Flux de données de lecture pour une architecture All-Flash

Remarque : Dans une architecture All-Flash de type EXF900, chaque nœud peut lire les données d'un autre nœud directement, à l'inverse de l'architecture de disque dur dans laquelle chaque nœud ne peut lire que ses propres données.

3.4.5 Optimisation des écritures pour la taille du fichier

Pour les écritures de petite taille sur le stockage, ECS utilise une méthode appelée *box-carting* pour réduire l'impact sur les performances. Le box-carting regroupe en mémoire plusieurs écritures de petite taille, 2 Mo ou moins, et les écrit en une seule opération de disque. Le box-carting limite le nombre d'allers-retours au disque requis pour traiter les écritures individuelles.

Pour les écritures d'objets plus volumineux, les nœuds d'ECS peuvent traiter les demandes d'écriture pour le même objet simultanément et tirer parti des écritures simultanées sur plusieurs piles du cluster ECS. Ainsi, ECS peut acquérir et stocker efficacement des objets de petite taille et de grande taille.

3.4.6 Récupération d'espace

Dans le cadre de l'écriture de fragments en mode ajout seul, les données sont ajoutées ou mises à jour en gardant d'abord en place les données écrites d'origine, puis en créant de nouveaux segments de fragments qui peuvent ou non être inclus dans le conteneur de fragments de l'objet d'origine. L'avantage de la modification des données en mode ajout seul réside en un modèle d'accès aux données actif/actif qui n'est pas entravé par les problèmes de verrouillage des fichiers des systèmes de fichiers traditionnels. Dans ces conditions, lorsque les objets sont mis à jour ou supprimés, les données dans des fragments ne sont plus référencées ou deviennent inutiles. ECS a recours aux deux méthodes suivantes de nettoyage de la mémoire pour récupérer de l'espace à partir de fragments complets rejetés, ou de fragments contenant un mélange de fragments d'objets supprimés et non supprimés qui ne sont plus référencés :

- **Nettoyage de la mémoire normal** : lorsqu'un fragment entier est garbage, l'espace est récupéré.
- **Nettoyage de la mémoire partiel par fusion** : lorsqu'un fragment est à 2/3 garbage, ses parties valides sont fusionnées avec d'autres fragments partiellement remplis au sein d'un nouveau fragment, puis l'espace est récupéré.

Un nettoyage de la mémoire a également été appliqué à l'API d'accès aux services de données ECS CAS pour nettoyer les BLOB orphelins. Les BLOB orphelins, qui sont des BLOB non référencés identifiés dans les données CAS stockées sur ECS, sont éligibles pour la récupération d'espace via les méthodes normales de nettoyage de la mémoire.

3.4.7 Mise en cache des métadonnées SSD

Les métadonnées ECS sont stockées dans des arborescences B. Chaque arborescence B peut contenir des entrées dans la mémoire, dans les transactions de journal et sur le disque. Pour que le système dispose d'une vue d'ensemble d'une arborescence B en particulier, les trois emplacements sont interrogés, ce qui implique souvent plusieurs recherches sur le disque.

Afin de minimiser la latence des recherches de métadonnées, un mécanisme de cache de disque SSD est disponible en option dans ECS 3.5. Le cache contient les pages d'arborescence B récemment consultées. Cela signifie que les opérations de lecture sur les dernières arborescences B atteignent toujours le cache de disque SSD et évitent les déplacements vers les disques rotatifs.

Voici quelques points clés de la nouvelle mise en cache des métadonnées SSD :

- Amélioration de la latence de lecture et des TPS (transactions par seconde) pour les petits fichiers à l'échelle du système
- Un lecteur Flash de 960 Go par nœud
- Les nouveaux nœuds de fabrication incluent le disque SSD en option
- Les nœuds existants (Gen 3 et Gen 2) peuvent être mis à niveau via des kits de mise à niveau et une installation en libre-service
- Des disques SSD peuvent être ajoutés pendant qu'ECS est en ligne
- Amélioration pour les charges applicatives d'analytique de fichiers de petite taille qui nécessitent de lire rapidement des jeux de données volumineux
- Tous les nœuds d'un VDC doivent disposer de disques SSD pour que cette fonctionnalité soit disponible

Le fabric ECS détecte l'installation d'un kit SSD. Cela déclenche l'initialisation automatique du système et l'utilisation du nouveau disque. La Figure 12 indique que le cache SSD est activé.

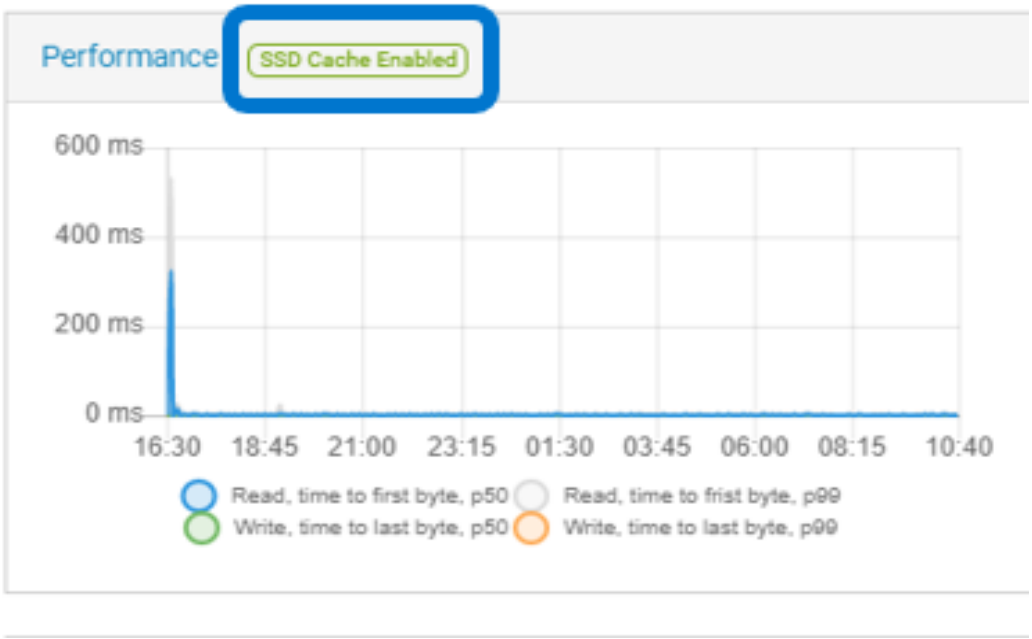

Figure 12 Cache SSD activé

La mise en cache des métadonnées SSD améliore les petites lectures et le référencement des buckets. Comme nous l'avons testé dans notre laboratoire, les performances de la liste s'améliorent de 50 % avec des objets de 10 Mo. Les performances de lecture s'améliorent de 35 % avec des objets de 10 Ko et de 70 % avec des objets de 100 Ko.

3.4.8 DVR du Cloud

ECS prend en charge la fonctionnalité d'enregistrement vidéo numérique (DVR) Cloud, qui répond aux exigences juridiques en matière de copyright des câbles et des entreprises de satellites. La condition est que chaque unité d'enregistrement mappée à un objet sur l'ECS doit être copiée un nombre de fois prédéterminé. Le nombre de copies prédéterminées est appelé Fanout. Le nombre prédéterminé de copies (facteur pyramidal) n'est pas vraiment une exigence de redondance ou de gain de performances, mais plutôt une exigence légale de copyright pour les entreprises de câble et de satellite. ECS prend en charge :

- La création d'un nombre à facteur pyramidal de copies de l'objet créé dans ECS
- L'autorisation de la lecture d'une copie spécifique
- L'autorisation de la suppression d'une copie spécifique
- L'autorisation de la suppression de toutes les copies
- L'autorisation de la copie d'une copie spécifique
- L'autorisation du référencement de toutes les copies
- L'autorisation du référencement des buckets pour les objets à facteur pyramidal

La fonction DVR Cloud peut être activée via la console de service. La première fois, vous devez activer la fonction DVR Cloud depuis la console de service. Après l'activation de la fonction DVR Cloud, elle est activée par défaut pour tous les nouveaux nœuds.

Exécutez la commande ci-dessous dans la console de service pour activer la fonction DVR Cloud :

```
service-console run Enable_CloudDVR
```

La fonctionnalité DVR Cloud prend en charge les API. Vous pouvez consulter le guide *ECS Data Access Guide* pour plus d'informations

3.5 Fabric

La couche Fabric fournit des fonctionnalités de clustering, d'intégrité des clusters, de gestion des logiciels, de gestion de la configuration, de mise à niveau et d'alerte. Elle est chargée de maintenir l'exécution des services et la gestion des ressources, telles que les disques, les conteneurs et le réseau. Elle assure le suivi des modifications apportées à l'environnement et réagit en conséquence, notamment pour détecter les défaillances. Elle fournit aussi des alertes relatives à l'intégrité du système. La couche Fabric est composé des éléments suivants :

- **Agent de nœud** : gère les ressources de l'hôte (disques, réseau, conteneurs, etc.) et les processus système.
- **Gestionnaire du cycle de vie** : gestion du cycle de vie des applications, qui comprend le démarrage des services, la restauration, la notification et la détection des pannes.
- **Gestionnaire de persistance** : coordonne et synchronise l'environnement distribué ECS.
- **Registre** : zone de stockage d'images Docker pour le logiciel ECS.
- **Bibliothèque d'événements** : contient l'ensemble des événements survenant sur le système.
- **Gestionnaire de matériel** : fournit des informations sur l'état, les événements et le provisionnement de la couche matérielle à des services de niveau supérieur. Ces services ont été intégrés pour la prise en charge de matériel générique.

3.5.1 Agent de nœud

L'agent de nœud est un agent léger écrit en Java qui s'exécute en mode natif sur tous les nœuds ECS.
Ses principales fonctions comprennent la gestion et le contrôle des ressources de l'hôte (conteneurs Docker, disques, pare-feu, réseau) et la surveillance des processus du système. Il peut s'agir, par exemple, de formater et de monter des disques, d'ouvrir les ports requis, de vérifier que tous les processus sont en cours d'exécution et de déterminer les interfaces réseau publiques et privées. Il comprend un flux d'événements qui fournit à un gestionnaire de cycle de vie des événements ordonnés pour indiquer les événements qui se produisent sur le système. Une CLI de Fabric est utile pour diagnostiquer les problèmes et examiner l'état général du système.

3.5.2 Gestionnaire du cycle de vie

Le gestionnaire du cycle de vie s'exécute sur un sous-ensemble de trois ou cinq nœuds et gère le cycle de vie des applications qui s'exécutent sur les nœuds. Chaque gestionnaire de cycle de vie est responsable du suivi de plusieurs nœuds. L'objectif principal de ce composant est de gérer l'intégralité du cycle de vie de l'application ECS, depuis le démarrage jusqu'au déploiement, en passant par la détection des pannes, la restauration, la notification et la migration. Il examine les flux de l'agent de nœud et pilote l'agent pour gérer la situation. Lorsqu'un nœud est en panne, il répond aux défaillances ou aux incohérences de l'état du nœud en restaurant le système à un état connu de bon fonctionnement. Si une instance du gestionnaire du cycle de vie est arrêtée, une autre prend sa place.

3.5.3 registre

Le registre contient les images Docker d'ECS utilisées lors de l'installation, de la mise à niveau et du remplacement de nœud. Un conteneur Docker appelé *fabric-registry* s'exécute sur un nœud au sein du rack ECS et contient le référentiel des images Docker d'ECS ainsi que les informations nécessaires pour les installations et les mises à niveau. Bien que le registre soit disponible sur un nœud à la fois, toutes les images Docker sont mises en cache localement sur chaque nœud, de sorte que tout nœud peut utiliser le registre.

3.5.4 Bibliothèque d'événements

La bibliothèque d'événements est utilisée au sein de la couche Fabric pour mettre à disposition les flux d'événements de cycle de vie et d'agent de nœud. Les événements générés par le système sont conservés de manière persistante dans la mémoire partagée et sur disque afin de fournir des informations historiques sur l'état et l'intégrité du système ECS. Ces flux d'événements ordonnés peuvent être utilisés pour restaurer le système à un état spécifique en relisant les événements ordonnés stockés. Des événements de nœud tels que démarré, arrêté ou dégradé en sont des exemples.

3.5.5 Gestionnaire de matériel

Le gestionnaire de matériel est intégré à l'agent de Fabric pour prendre en charge le matériel standard. Son objectif principal est de fournir des informations spécifiques sur l'état du matériel et sur les événements, ainsi que d'assurer le provisionnement de la couche matérielle vers des services de niveau supérieur au sein d'ECS.

3.6 Infrastructure

Les nœuds de l'appliance ECS exécutent actuellement SUSE Linux Enterprise Server 12 pour l'infrastructure. Pour le logiciel ECS déployé sur du matériel standard personnalisé, le système d'exploitation peut également être RedHat Enterprise Linux ou Coreos. Les déploiements personnalisés sont effectués via un processus formel de demande et de validation. Docker est installé sur l'infrastructure pour le déploiement des couches encapsulées d'ECS. Le logiciel ECS étant écrit en Java, la machine virtuelle Java (JVM) est installée dans le cadre de l'infrastructure.

3.6.1 Docker

ECS s'exécute sur le système d'exploitation en tant qu'application Java et est encapsulé dans plusieurs conteneurs Docker. Les conteneurs sont isolés mais partagent les ressources et le matériel du système d'exploitation sous-jacent. Certaines parties du logiciel ECS s'exécutent sur tous les nœuds, tandis que d'autres s'exécutent sur un ou plusieurs nœuds. Les composants qui s'exécutent dans un conteneur Docker sont les suivants :

- **object-main** : contient les ressources et les processus relatifs aux services de données, au moteur de stockage et aux services de portail et de provisionnement. S'exécute sur chaque nœud d'ECS.
- **fabric-lifecycle** :contient les processus, les informations et les ressources nécessaires à la surveillance au niveau du système, à la gestion de la configuration et à la gestion de l'intégrité. Un nombre impair d'instances de fabric-lifecycle est toujours en cours d'exécution. Par exemple, trois instances s'exécutent sur un système à quatre nœuds et cinq instances sur un système à huit nœuds.

- **fabric-zookeeper** : service centralisé pour la coordination et la synchronisation des processus distribués, des informations de configuration, des groupes et des services d'attribution de nom. Il est appelé gestionnaire de persistance et s'exécute sur un nombre impair de nœuds, par exemple cinq dans un système à huit nœuds.
- **fabric-registry** : registre des images Docker d'ECS. Une seule instance s'exécute par rack ECS.

Il existe d'autres processus et outils qui s'exécutent en dehors d'un conteneur Docker, notamment les outils de l'agent de nœud de Fabric et de la couche d'abstraction matérielle. La Figure 13 ci-dessous fournit un exemple de la façon dont les conteneurs ECS peuvent être exécutés sur un déploiement à huit nœuds.

Figure 13 Conteneurs Docker et agents dans un exemple de déploiement à huit nœuds

La Figure 14 présente la sortie de la ligne de commande de la commande *docker ps* sur un nœud, qui affiche les quatre conteneurs utilisés par ECS dans Docker. Une liste s'affiche avec tous les services associés à l'objet disponibles sur le système.

```
admin@hop-u300-11-pub-01:~> sudo docker ps
CONTAINER ID        IMAGE                                                COMMAND                  CREATED             STATUS
      PORTS               NAMES
7ba30ce42be2        ecs-monitoring/telegraf:3.5.0.0-825.b6b07cf9         "/entrypoint.sh "        5 weeks ago         Up 5 weeks
                          object-telegraf
e22513635cab        ecs-monitoring/grafana:3.5.0.0-825.b6b07cf9          "/entrypoint.sh "        5 weeks ago         Up 5 weeks
                          object-grafana
ee9db1ea40bc        emcvipr/object:3.5.0.0-120417.6a358e139f1            "/opt/vipr/boot/boot…"   5 weeks ago         Up 5 weeks
                          object-main
d11a7acd55e5        ecs-monitoring/throttler:3.5.0.0-825.b6b07cf9        "/entrypoint.sh "        5 weeks ago         Up 5 weeks
                          object-throttler
f94026797bb3        ecs-monitoring/fluxd:3.5.0.0-825.b6b07cf9            "/entrypoint.sh "        5 weeks ago         Up 5 weeks
                          object-fluxd
c7b8530a8bb9        caspian/fabric:3.5.0.0-4076.7d40a97                  "./boot.sh lifecycle"    5 weeks ago         Up 5 weeks
                          fabric-lifecycle
bffdf8836853        caspian/fabric-zookeeper:3.5.6.0-99.0354df7          "./boot.sh 1 1=169.2…"   5 weeks ago         Up 5 weeks
                          fabric-zookeeper
f4420f7f7d51        caspian/fabric-registry:2.3.1.0-68.10d1aca           "/opt/docker-registr…"   5 weeks ago         Up 5 weeks
                          fabric-registry
admin@hop-u300-11-pub-01:~> sudo dockobj
hop-u300-11-pub-01:/ # cd /opt/storageos/
hop-u300-11-pub-01:/opt/storageos # ls bin/*svc
bin/blobsvc     bin/coordinatorsvc  bin/eventsvc  bin/objcontrolsvc  bin/storagemanagementsvc
bin/cassvc      bin/dataheadsvc     bin/filesvc   bin/objheadsvc     bin/syssvc
bin/controlsvc  bin/ecsportalsvc    bin/hdfssvc   bin/resourcesvc    bin/transformsvc
```

Figure 14 Processus, ressources, outils et fichiers binaires dans le conteneur object-main

4 Modèles matériels de l’appliance

Des points d’entrée flexibles permettent à ECS d’être rapidement mis à l’échelle jusqu’à plusieurs pétaoctets et exaoctets de données. Avec un impact minime sur l’entreprise, une solution ECS peut évoluer de manière linéaire en matière de capacité et de performances, par l’ajout de nœuds et de disques.

Les modèles matériels de l’appliance ECS sont caractérisés par la génération de matériel. La gamme d’appliances de troisième génération, connue sous le nom de Gen 3 ou gamme EX, inclut trois modèles matériels. Cette section fournit un aperçu général de la série EX. Pour des informations complètes, consultez le *Guide du matériel des appliances ECS gamme EX*.

Pour plus d’informations sur le matériel de l’appliance ECS de première et deuxième génération, consultez le guide *Dell EMC ECS D- and U-Series Hardware Guide*.

4.1 Série EX

Les modèles d’appliance de la gamme EX sont basés sur des serveurs et des commutateurs Dell standard. Les offres de la gamme sont les suivantes :

- **EX300** : l’appliance EX300 offre une capacité brute de départ de 60 To. Il s’agit de la plate-forme de stockage idéale pour les applications Cloud natives et les initiatives de transformation numérique des clients. Les appliances EX300 se prêtent particulièrement bien à la modernisation des déploiements Centera. Plus important encore, l’appliance EX300 peut évoluer de manière rentable vers des capacités plus importantes. Elle fournit 12 disques par nœud et des options de disque de 1 To, 2 To, 4 To, 8 To et 16 To (toutes identiques dans le nœud)
- **EX500** : l’appliance EX500 est la dernière édition de l’appliance, conçue pour fournir un juste équilibre entre économie et densité. Avec des options de 12 ou 24 disques de 8 To, 12 To et 16 To (tous identiques dans le nœud). Les clusters s’étendent de 480 To à 6,1 Po par rack. Cette série constitue une solution polyvalente pour les entreprises de taille intermédiaire souhaitant prendre en charge des cas d’utilisation d’applications modernes et/ou d’archivage à long terme.
- **EX3000** : l’appliance EX3000 dispose d’une capacité maximale de 11,5 Po de stockage brut par rack, avec 30 à 90 disques par nœud, de 12 To ou 16 To, et peut évoluer vers plusieurs exaoctets sur plusieurs sites. Elle offre une solution de datacenter à la fois puissante et évolutive, idéale pour les charges applicatives avec des encombrements de données plus importants. Ces nœuds sont disponibles dans deux configurations différentes, appelées EX3000S et EX3000D. L’appliance EX3000S est un châssis à un nœud tandis que l’appliance EX3000D en contient deux. Ces nœuds haute densité offrent le remplacement à chaud des disques. Ils commencent avec un minimum de trente disques par nœud. C’est à partir de trente disques par nœud ECS que les gains de performances liés à l’ajout de disques diminuent. Avec un minimum de 30 disques ou plus dans chaque nœud, les attentes en matière de performances sont similaires sur tous les nœuds EX3000, quel que soit le nombre de disques.
- **EXF900** : l’appliance EXF900 est une solution de stockage d’objet All-Flash de nœuds hyperconvergés pour les déploiements ECS à faible latence et à E/S par seconde élevées. Avec des options de 12 ou 24 disques et de disques SSD NVMe de 3,84 To (le pilote de disque SSD NVMe de 7,68 To sera pris en charge lorsque le matériel sera disponible). Cette plate-forme commence avec une configuration minimale brute de 230 To et évolue jusqu’à 1,4 Po de capacité brute par rack. La Figure 15 montre un nœud EXF900.

Figure 15 Nœud EXF900

Remarque : La fonctionnalité de cache de lecture SSD ne s’applique pas à EXF900. Cloud DVR n’est pas pris en charge sur EXF900. Tech Refresh n’est pas pris en charge avec EXF900. EXF900 ne peut pas coexister avec un autre matériel non EXF900 dans un VDC. EXF900 ne peut pas coexister avec tout autre matériel non EXF900 dans GEO (tous les sites doivent être EXF900).

Les options de capacité de départ de la gamme EX permettent aux clients de démarrer un déploiement ECS présentant uniquement la capacité nécessaire, puis de l’étendre en toute simplicité, au gré de l’évolution de leur besoins. Pour plus d’informations sur les appliances de la sériee EX, consultez la notice technique *ECS Appliance Specification Sheet*. Elle décrit également les appliances précédentes Gen 2 des séries U et D.

Les mises à jour post-déploiement des nœuds de la gamme EX ne sont pas prises en charge.
Ces composants sont les suivants :

- Modification du processeur
- Ajustement de la capacité mémoire
- Mise à niveau de la taille du disque dur

4.2 Réseau de l’appliance

Depuis le lancement des appliances de la gamme EX, il est possible d’utiliser une paire redondante de commutateurs de gestion back-end dédiés. En passant au nouveau matériel de commutation d’appliance, ECS est désormais en mesure d’adopter un mode de commutation de configuration front-end et back-end.

Les appliances EX300, EX500 et EX3000 utilisent toutes le commutateur Dell EMC S5148F pour la paire de commutateurs front-end et la paire de commutateurs back-end. L’appliance EXF900 utilise le commutateur Dell EMC S5248F pour la paire de commutateurs front-end et la paire de commutateurs back-end, ainsi que le modèle S5232F pour le commutateur back-end d’agrégation. Notez que les clients ont la possibilité d’utiliser leurs propres commutateurs front-end au lieu des commutateurs Dell EMC.

4.2.1 S5148F : commutateurs publics front-end

Deux commutateurs Ethernet Dell EMC S5148F 25 GbE 1U en option peuvent être obtenus pour une connexion réseau, ou le client peut fournir sa propre paire HA de 10 GbE ou 25 GbE pour la connectivité front-end. Les commutateurs publics sont souvent appelés *hare* et *rabbit* ou simplement front-end.

Avertissement : Il est nécessaire de disposer de connexions du réseau du client aux deux commutateurs frontaux (Rabbit et Hare) afin de maintenir l’architecture à haute disponibilité de l’appliance ECS. Si le client décide de ne pas se connecter à son réseau avec la haute disponibilité requise, il n’y a aucune garantie de haute disponibilité des données pour l’utilisation de ce produit.

Ces commutateurs fournissent 48 ports de 25 GbE SFP28 et 6 ports de 100 GbE QSFP28. Informations complémentaires sur ces deux types de ports :

- SFP28 est une version améliorée de SFP+ :
  - Jusqu’à 16 Gbit/s pour SFP+, jusqu’à 28 Gbit/s pour SFP28
  - Même format
  - Compatibilité descendante avec les modules SFP+
- QSFP28 est une version améliorée de QSFP+ :
  - Jusqu’à 4 voies de 16 Gbit/s pour QSFP+, jusqu’à 4 voies de 28 Gbit/s pour QSFP28
    - QSFP+ : voies agrégées pour obtenir une connexion Ethernet de 40 Gbit/s
    - QSFP28 : voies agrégées pour obtenir une connexion Ethernet de 100 Gbit/s
  - Même format
  - Compatibilité descendante avec les modules QSFP+
  - Répartition possible en 4 voies individuelles de SFP28

Remarque : Deux câbles LAG 100 GbE sont fournis avec les commutateurs publics Dell EMC S5148F 25 GbE. Les organisations présentant leurs propres commutateurs publics doivent fournir les câbles LAG, SFP ou de raccordement externe nécessaires.

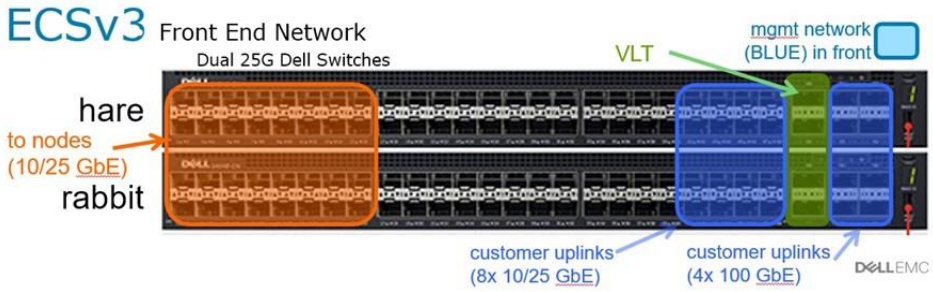

Figure 16 Conception et utilisation des ports de commutateur réseau front-end

La Figure 16 ci-dessus fournit une représentation visuelle de la façon dont les ports doivent être utilisés pour activer le trafic des nœuds ECS, ainsi que les ports de données sortantes du client. Il s’agit d’une norme standard valable pour l’ensemble des implémentations.

4.2.2 S5148F - commutateurs privés back-end

Les deux commutateurs Ethernet Dell EMC S5148F 25 GbE 1U requis avec 48 ports SFP 25 GbE et 6 ports 100 GbE de données sortantes sont inclus dans chaque rack ECS. Ils sont souvent appelés *fox* et *hound* ou simplement commutateurs back-end, et sont chargés du réseau de gestion. Dans les futures versions d’ECS, les commutateurs back-end fourniront également une séparation de réseau pour le trafic de réplication.
Le réseau privé sert principalement à la gestion et la console à distance, au démarrage PXE du gestionnaire d’installation et à l’activation de la gestion et du provisionnement à l’échelle du rack et du cluster.
La Figure 17 présente une vue de face de deux commutateurs Dell 25 GbE.

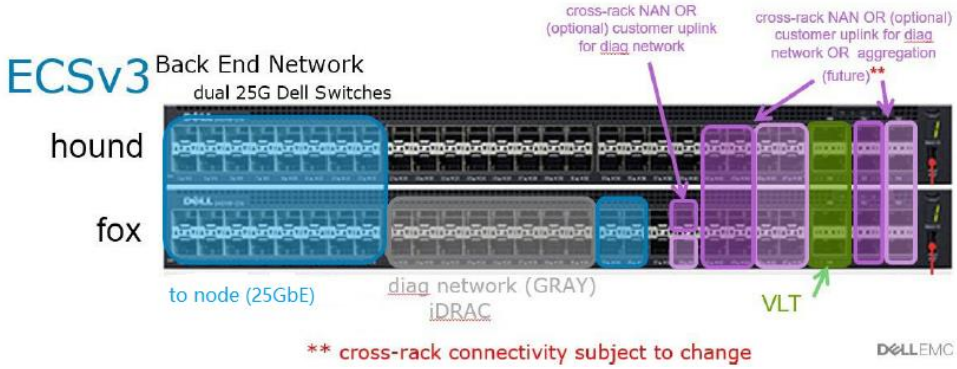

Figure 17 Conception et utilisation des ports de commutateur réseau back-end

Le schéma ci-dessus fournit une représentation visuelle de la façon dont les ports sont destinés à être utilisés pour activer le trafic de gestion ECS et les ports de diagnostic. Ces allocations de ports sont standard pour toutes les implémentations. Les ports pour de possibles utilisations futures sont notés en violet ; toutefois, cette utilisation est susceptible d’évoluer à l’avenir.

4.2.3 S5248F : commutateurs publics front-end

Dell EMC propose en option une paire HA de commutateurs frontaux 25 GbE S5248F pour la connexion du réseau client au rack. Ils disposent de deux câbles de jonction de liaisons virtuelles (QSFP28-DD) (VLT) de 200 GbE par paire de HA. Ces commutateurs sont appelés Hare et Rabbit. La Figure 18 offre une représentation visuelle de la façon dont les ports doivent être utilisés pour activer le trafic des nœuds ECS, ainsi que les ports de données sortantes du client.

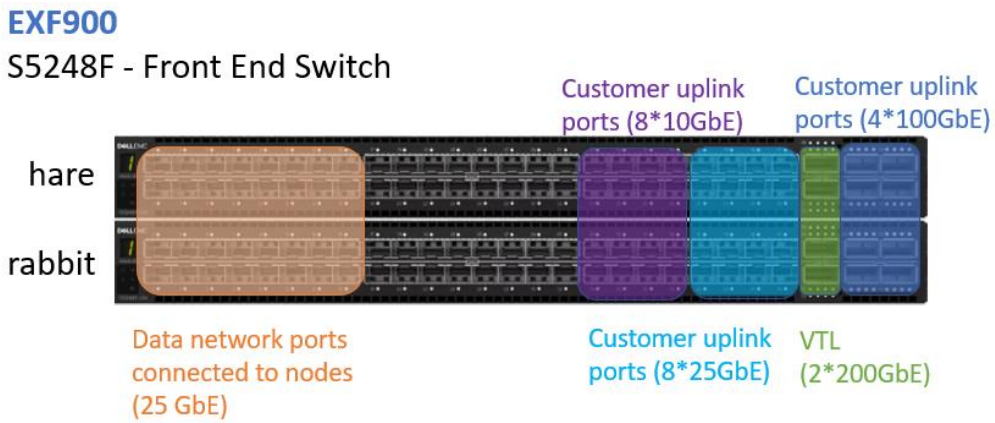

Figure 18 Conception et utilisation des ports de commutateur réseau front-end

4.2.4 S5248F - commutateurs privés back-end

Dell EMC fournit deux commutateurs back-end 25 GbE S5248F avec deux câbles VLT 200 GbE (QSFP28-DD). Ce commutateurs sont appelés Hound et Fox. Tous les câbles iDRAC des nœuds et toutes les connexions des câbles de gestion du commutateur front-end sont acheminés vers le commutateur Fox.
La Figure 19 fournit une représentation visuelle de la façon dont les ports sont destinés à être utilisés pour activer le trafic de gestion ECS et les ports de diagnostic. Ces allocations de ports sont standard pour toutes les implémentations.

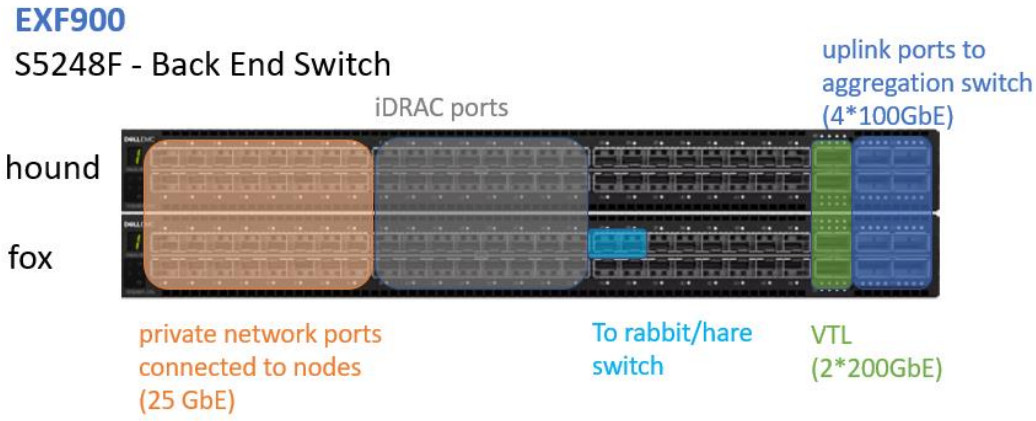

Figure 19 Conception et utilisation des ports de commutateur réseau back-end

4.2.5 S5232 - Commutateurs d’agrégation

Dell EMC fournit deux commutateurs d’agrégation back-end 100 GbE S5232F (AGG1 et AGG2) avec quatre câbles VLT 100 GbE. Ces commutateurs sont appelés Falcon et Eagle. Dans la Figure 20 ci-dessous, tous les ports étiquetés indiquent les désignations des ports. Cette configuration permet de connecter jusqu’à 7 racks de nœuds EXF900.

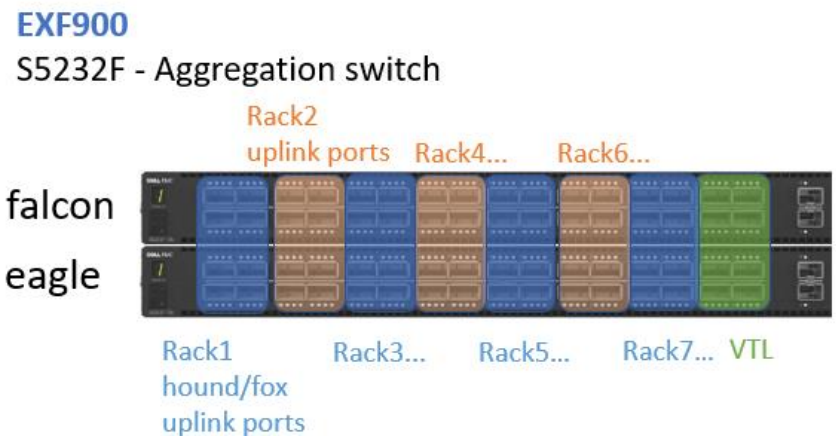

Figure 20 Conception et utilisation des ports de commutateur d’agrégation

Pour plus d’informations sur la mise en réseau et le câblage, reportez-vous au guide *ECS EX Series Hardware Guide.*

5 Séparation de réseau

ECS prend en charge la séparation des différents types de trafic réseau pour la sécurité et l'isolement des performances. Les types de trafic pouvant être séparés sont les suivants :

- Gestion
- Réplication
- Data

Il existe un mode de fonctionnement appelé *mode de séparation de réseau*. Dans ce mode, chaque nœud peut être configuré au niveau du système d'exploitation avec jusqu'à trois adresses IP, ou réseaux logiques, pour chacun des différents types de trafic. Cette fonctionnalité a été conçue pour offrir la possibilité de créer trois réseaux logiques distincts pour la gestion, la réplication et les données, ou de les combiner pour créer deux réseaux logiques, par exemple le trafic de gestion et de réplication dans un réseau logique et le trafic de données dans un autre réseau logique. Il est possible de configurer un second réseau de données logique pour le trafic CAS uniquement, ce qui permet de séparer ce dernier des autres types de trafic de données comme S3.

La mise en œuvre de la séparation de réseau par ECS exige que le trafic de chaque réseau logique soit associé à des services et des ports. Par exemple, les services du portail ECS communiquent via les ports 80 ou 443. Par conséquent, ces ports et services seront liés au réseau logique de gestion. Un deuxième réseau de données peut être configuré, mais pour le trafic CAS uniquement. Le Tableau 5 ci-dessous met en évidence les services associé à un type de réseau logique. Pour obtenir la liste complète des services associés aux ports, consultez la version la plus récente du guide *ECS Security Configuration Guide*.

Tableau 5 Mappage des services à un réseau logique

Services	Réseau logique	Identifier
WebUI et API, SSH, DNS, NTP, AD, SMTP	Gestion	public.mgmt
Données des clients	Data	public.data
	Données CAS uniquement	public.data2
Données de réplication	Réplication	public.repl
SRS (Dell EMC Secure Remote Services)	En fonction du réseau auquel la passerelle SRS est rattachée	public.data ou public.mgmt

Remarque : ECS 3.6 permet d'accéder aux données S3 sur les réseaux data (par défaut) et data2 (bien que S3 ne soit pas activé par défaut sur data2). Pour activer l'accès aux données S3 sur le réseau data2, vous devez disposer de public.data et contacter le support à distance ECS.

La séparation de réseau est réalisable de manière logique à l'aide de différentes adresses IP, virtuellement en utilisant différents VLAN ou physiquement en utilisant différents câbles. La commande *setrackinfo* est utilisée pour configurer les adresses IP et les réseaux VLAN. La configuration d'un réseau VLAN au niveau du commutateur ou du côté client est la responsabilité du client. Pour la séparation physique du réseau, les clients doivent soumettre une demande de qualification de produit (RPQ) en contactant Dell EMC Global Business Service. Pour plus d'informations sur la séparation de réseau, consultez le livre blanc *ECS Networking and Best Practices* qui fournit une vue générale de la séparation de réseau.

6 Security

La sécurité d'ECS est mise en œuvre au niveau de l'administration, du transport et des données. L'authentification de l'utilisateur et de l'administrateur est assurée via Active Directory, les méthodes LDAP, Keystone ou directement dans le portail ECS. La sécurité au niveau des données s'appuie sur le protocole HTTPS pour les données en mouvement et/ou le chiffrement côté serveur pour les données inactives.

6.1 Authentification

ECS prend en charge les méthodes d'authentification Active Directory, LDAP, Keystone et IAM pour permettre l'accès à la gestion et à la configuration d'ECS ; toutefois, il existe des limitations, comme indiqué dans le Tableau 6. Pour plus d'informations sur la sécurité, consultez le guide *ECS Security Configuration Guide* le plus récent.

Tableau 6 Méthodes d'authentification prises en charge

Méthode d'authentification	**Pris en charge**
Active Directory	• Les groupes AD sont pris en charge pour les utilisateurs de gestion. • Les groupes AD sont pris en charge pour les méthodes de provisionnement des utilisateurs d'objets à l'aide de clés en libre-service via l'API. • Le multidomaine est pris en charge.
LDAP	• Les utilisateurs de gestion peuvent s'authentifier individuellement via le protocole LDAP. • Les groupes LDAP ne sont pas pris en charge pour les utilisateurs de gestion. • LDAP est pris en charge pour les utilisateurs d'objets (clés en libre-service via l'API). • Le multidomaine est pris en charge.
Keystone	• Les règles RBAC ne sont pas encore prises en charge. • Les tokens non délimités ne sont pas pris en charge. • Il n'y a pas non plus de prise en charge de plusieurs serveurs Keystone par système ECS.
IAM	• Assure la fédération des identités et l'authentification unique (SSO) selon les normes SAML 2.0 • Disponible via le protocole S3 uniquement

6.2 Authentification des services de données

L'accès aux objets à l'aide d'API RESTful est sécurisé via HTTPS (TLS v1.2). Les demandes entrantes sont authentifiées à l'aide de méthodes définies telles que HBAC (Hash-based Message Authentication Code), Kerberos ou l'authentification par token. Le Tableau 7 ci-dessous présente les différentes méthodes utilisées pour chaque protocole.

Tableau 7 Authentification des services de données

Protocoles		**Méthodes d'authentification**
Objet	S3	V2 (HMAC-SHA1), V4 (HMAC-SHA256)
	Swift	Token - Keystone v2 et v3 (délimité, UUID, tokens PKI), SWAuth v1
	Atmos	HMAC-SHA1
	CAS	Fichier de clé secrète PEA
Fichier	HDFS	Kerberos
	NFS	Kerberos, AUTH_SYS

6.3 Chiffrement des données au repos (D@RE)

Les exigences de conformité régissent souvent l'utilisation du chiffrement pour protéger les données écrites sur les disques. Dans ECS, le chiffrement peut être activé au niveau de l'espace de nommage et du bucket. Les principales caractéristiques d'ECS D@RE sont les suivantes :

- Chiffrement à faible interaction au repos : activation facile et configuration simple
- CIPHER (AES-256 CTR) utilisés
- Chiffrement de clé publique RSA d'une longueur de 2 048 bits
- Prise en charge de la gestion de clés externe (EKM) au niveau du cluster :
  - Gemalto SafeNet
  - IBM Security Key Lifecycle Manager
- Rotation des clés
- Prise en charge de la sémantique de chiffrement S3 à l'aide d'en-têtes HTTP comme *x-amz-server-side-encryption*
- Conformité FIPS 140-2 aux normes de sécurité cryptographique du gouvernement américain

Remarque : Le mode FIPS 140-2 applique l'utilisation d'algorithmes approuvés uniquement dans D@RE. La conformité FIPS 140-2 concerne uniquement le module D@RE, pas l'ensemble du produit ECS.

ECS utilise une hiérarchie de clés pour chiffrer et déchiffrer les données. Le gestionnaire de clés natif stocke une clé privée commune à tous les nœuds pour déchiffrer la clé principale. Avec la configuration EKM, la clé principale est fournie par l'EKM. Les clés fournies par l'EKM résident dans la mémoire uniquement sur ECS. Elles ne sont jamais stockées dans un stockage persistant au sein d'ECS.

Dans un environnement géorépliqué, lorsqu'un nouveau système ECS rejoint une fédération existante, la clé principale est extraite à l'aide de la clé publique-privée du système existant et chiffrée à l'aide de la nouvelle paire de clés publique-privée générée à partir du nouveau système ayant rejoint la fédération. La clé

principale est dès lors globale et connue des deux systèmes de la fédération. Lors de l'utilisation d'EKM, tous les systèmes fédérés récupèrent la clé principale à partir du système de gestion des clés.

6.3.1 Rotation des clés

ECS prend en charge la modification des clés de chiffrement. Cela peut être effectué régulièrement afin de limiter la quantité de données protégées par un ensemble spécifique de clés de chiffrement de clé (clé KEK) ou en réponse à une faille ou une attaque potentielle. Un enregistrement de rotation de clé KEK est utilisé conjointement avec d'autres clés parentes pour créer des clés d'encapsulation virtuelles qui protègent les clés de chiffrement des données (clé DEK) et les KEK d'espace de nommage.

Les clés de rotation sont générées de manière native ou fournies et gérées par un EKM. ECS utilise la clé de rotation actuelle pour créer des clés d'encapsulation virtuelles afin de protéger les clés DEK ou les clés KEK, que la gestion des clés s'effectue de manière native ou externe.

Au cours des opérations d'écriture, ECS englobe les clés DEK générées de manière aléatoire à l'aide d'une clé d'encapsulation virtuelle créée au moyen du bucket et de la clé de rotation active.

Dans le cadre de la rotation des clés, ECS encapsule à nouveau tous les enregistrements KEK d'espace de nommage avec une nouvelle clé KEK principale virtuelle créée à partir d'une nouvelle clé de rotation, le contexte secret associé et la clé principale active. Cela permet de sécuriser l'accès aux données protégées par les précédentes clés de rotation.

L'utilisation d'un EKM affecte le chemin d'accès en lecture/écriture pour les objets chiffrés. La rotation des clés permet d'avoir une protection supplémentaire des données en utilisant des clés d'encapsulation virtuelles pour clés DEK et clés KEK d'espace de nommage. Les clés d'encapsulation virtuelles ne sont pas persistantes et proviennent de deux hiérarchies indépendantes de clés persistantes. Avec EKM, la clé de rotation n'est pas stockée dans ECS, ce qui augmente la sécurité des données. Nous ajoutons principalement de nouveaux enregistrements KEK et mettons à jour les ID actifs, mais ne supprimons jamais quoi que ce soit.

Les points supplémentaires à prendre en compte concernant la rotation des clés dans ECS sont les suivants :

- Le processus de rotation des clés modifie uniquement la clé de rotation actuelle. Les clés principales, de l'espace de nommage et du bucket existantes ne changent pas au cours du processus de rotation des clés.
- La rotation des clés au niveau de l'espace de nommage ou du bucket n'est pas prise en charge. En revanche, le champ d'application de la rotation est au niveau du cluster. Par conséquent, tous les nouveaux objets chiffrés du système sont concernés.
- Les données existantes ne font pas l'objet d'un nouveau chiffrement à cause de la rotation des clés.
- ECS ne prend pas en charge la rotation des clés lors des pannes.
  - TSO pendant la rotation : tâche de rotation des clés interrompue jusqu'à ce que le système ne soit plus en TSO.
  - Panne de site permanente (PSO) en cours. ECS doit sortir d'une PSO pour que la rotation des clés soit activée. Si une PSO se produit lors de la rotation, la rotation échoue immédiatement.
- Le chiffrement du bucket n'est pas nécessaire pour le chiffrement des objets via S3.
- Les métadonnées d'objets client indexées utilisées en tant que clés de recherche ne sont pas chiffrées.

Consultez la version la plus récente du guide *ECS Security Configuration Guide* pour plus d'informations sur D@RE, EKM et la rotation des clés.

6.4 ECS IAM

ECS Identity and Access Management (IAM) vous permet de contrôler et sécuriser les accès aux ressources ECS S3. Cette fonctionnalité garantit que chaque demande d'accès à une ressource ECS est identifiée, authentifiée et autorisée. ECS IAM permet à l'administrateur d'ajouter des utilisateurs, des rôles et des groupes. L'administrateur peut également limiter les accès en ajoutant des politiques aux entités ECS IAM.

Remarque : ECS IAM est destiné à être utilisé avec S3 uniquement. Il n'est pas activé pour les buckets CAS ou les buckets installés en tant que système de fichiers.

ECS IAM comprend les composants suivants :

- **Gestion des comptes** : vous permet de gérer les identités IAM au sein de chaque espace de nommage, telles que les utilisateurs, les groupes et les rôles.
- **Gestion des accès** : les accès sont gérés en créant des règles et en les rattachant à des identités ou des ressources IAM.
- **Fédération des identités :** les identités sont établies et authentifiées par SAML (Security Assertion Markup Language). Une fois les identités établies, vous obtiendrez des informations d'identification temporaires avec le service de jeton de sécurité pour accéder à la ressource.
- **Service de jeton de sécurité** : vous permet de demander des informations d'identification temporaires pour l'accès inter-comptes aux ressources, ainsi que pour les utilisateurs qui sont authentifiés à l'aide de l'authentification SAML d'un fournisseur d'identité d'entreprise ou d'un service d'annuaire.

En utilisant IAM, vous pouvez contrôler qui est authentifié et autorisé à utiliser les ressources ECS en créant et en gérant :

- **Les utilisateurs** : un utilisateur IAM représente une personne ou une application de l'espace de nommage qui peut interagir avec les ressources d'ECS.
- **Les groupes** : un groupe IAM est un ensemble d'utilisateurs IAM. Utilisez les groupes pour spécifier des autorisations pour un ensemble d'utilisateurs IAM.
- **Les rôles** : le rôle IAM est une identité qui peut être prise en charge par toute personne ayant besoin de ce rôle. Un rôle est similaire à un utilisateur : il s'agit d'une identité dotée de politiques d'autorisation qui déterminent ce que l'identité peut et ne peut pas faire
- **Les règles** : une règle IAM est un document au format JSON qui définit les autorisations d'un rôle. Attribuez et rattachez des règles aux utilisateurs IAM, aux groupes IAM et aux rôles IAM.
- **Le fournisseur SAML** : SAML est une norme ouverte pour l'échange de données d'authentification et d'autorisation entre un fournisseur d'identités et un prestataire de services. Le fournisseur SAML dans ECS est utilisé pour établir une relation de confiance entre ECS et un fournisseur d'identité (IdP) compatible SAML

Un compte ECS IAM est attribué à chaque système ECS. Ce compte prend en charge plusieurs espaces de nommage et possède des entités IAM associées qui sont définies dans son espace de nommage.

- Les espaces de nommage individuels prennent en charge la gestion du compte à l'aide des entités ECS IAM telles que les utilisateurs, les rôles et les groupes.
- Règles, autorisations, listes de contrôle d'accès (ACL) associées aux entités ECS IAM et prise en charge des ressources ECS S3 pour la gestion des accès aux fonctions d'ECS IAM.
- ECS IAM prend en charge l'accès inter-compte avec SAML (Security Assertion Markup Language) et les rôles.
- ECS IAM prend en charge la clé d'accès Amazon Web Services (AWS) pour accéder à IAM et S3 dans ECS.

Pour plus d'informations sur ECS IAM, reportez-vous au guide *ECS Security Guide* le plus récent.

6.5 Balisage des objets

Le balisage des objets permet de catégoriser les objets en attribuant des balises aux objets individuels. Un seul objet peut être associé à plusieurs balises, ce qui permet une catégorisation multidimensionnelle.

Une balise peut décrire des données sensibles comme un dossier médical, ou vous pouvez baliser un objet en l'associant à un certain produit qui pourra être catégorisé confidentiel. Le balisage est une sous-ressource d'un objet qui a un cycle de vie et est intégré aux opérations d'objet. Vous pouvez ajouter des balises aux nouveaux objets lorsque vous les téléchargez ou ajouter des balises à des objets existants. Il est possible d'utiliser des balises pour libeller des objets contenant des données confidentielles, telles que des informations d'identification personnelle (PII) ou des informations médicales protégées (PHI). Les balises ne doivent pas contenir d'informations confidentielles, car elles peuvent être consultées sans autorisation de lecture sur un objet.

6.5.1 Informations supplémentaires sur le balisage des objets

Cette section fournit des informations sur le balisage des objets dans IAM, le balisage des objets avec des politiques de bucket, la gestion du balisage des objets durant une TSO ou une PSO et le balisage des objets pendant la gestion du cycle de vie. Voici d'autres éléments à prendre en compte :

- Balisage des objets dans IAM
  - La fonction clé du balisage des objets en tant que système de catégorisation est son intégration avec des politiques IAM. Elle permet à l'administrateur de configurer des autorisations utilisateur spécifiques. Par exemple, l'administrateur peut ajouter une politique qui permet à tous les utilisateurs d'accéder à des objets dotés d'une balise spécifique ou de configurer et d'octroyer des autorisations aux utilisateurs qui peuvent gérer les balises sur des objets spécifiques. L'autre aspect clé du balisage des objets est la façon dont les balises sont conservées. Ce point est important, car il a un impact direct sur divers aspects du système.
- Balisage des objets avec des politiques de bucket
  - Le balisage d'objet vous permet de catégoriser les objets, mais également d'intégrer les balises à différentes politiques. La politique de gestion du cycle de vie vous permet de procéder à une configuration au niveau du bucket. Les versions antérieures d'ECS prennent en charge l'expiration, l'annulation des téléchargements incomplets et la suppression d'un marqueur de suppression de balisage d'objet ayant expiré. Le filtre peut inclure plusieurs conditions, par exemple une condition basée sur une balise. Chaque balise de la condition du filtre doit correspondre à la clé et à la valeur.
- Balisage des objets durant une TSO ou une PSO
  - Comme le balisage des objets est un jeu d'entrées dans les métadonnées du système, aucune manipulation particulière n'est requise lors d'une TSO ou PSO. Même s'il existe une limite définie du nombre de balises pouvant être associées à chaque objet, la taille occupée par les métadonnées du système avec le balisage des objets reste largement dans les limites de la mémoire.

- Balisage des objets pendant la gestion du cycle de vie
  - Le balisage des objets fait partie des métadonnées du système et est traité en même temps que les métadonnées du système au cours de la gestion du cycle de vie. La logique d'expiration et le Lifecycle Delete Scanner nécessitent de comprendre les politiques basées sur des balises. Les balises d'objets permettent une gestion fine du cycle de vie des objets dans le cadre de laquelle vous pouvez spécifier un filtre basé sur les balises, en plus d'un préfixe de nom clé, dans une politique de cycle de vie.

Consultez la version la plus récente du guide *ECS Security Configuration Guide*pour plus d'informations sur le balisage des objets sur ECS.

7 Intégrité et protection des données

Pour l'intégrité des données, ECS utilise des sommes de contrôle. Les sommes de contrôle sont créées lors des opérations d'écriture et sont stockées avec les données. Lors d'opérations de lecture, les sommes de contrôle sont calculées et comparées à la version stockée. Une tâche en arrière-plan vérifie de manière proactive les informations de la somme de contrôle.

Pour la protection des données, ECS utilise la mise en miroir triple pour les fragments de journal et des schémas EC séparés pour les fragments *repo* (données du référentiel utilisateur) et *btree* (arborescence B+).

Le codage d'effacement améliore la protection des données en cas de panne de disque, de nœud et de rack de manière efficace pour le stockage par rapport aux schémas de protection classiques. Le moteur de stockage ECS met en œuvre la correction des erreurs Reed Solomon à l'aide de deux schémas :

- 12+4 (par défaut) : le fragment est divisé en 12 segments de données. 4 segments de codage (parité) sont créés.
- 10+2 (archives inactives) : le fragment est divisé en 10 segments de données. 2 segments de codage sont créés.

Avec la valeur par défaut de 12+4, les 16 segments qui en résultent sont répartis entre les nœuds du site local. Les segments de données et de codage de chaque fragment sont répartis équitablement entre les nœuds du cluster. Par exemple, avec 8 nœuds, chaque nœud reçoit 2 segments (sur 16 au total).
Le moteur de stockage peut reconstruire un fragment à partir de 12 segments quelconques sur les 16.

ECS requiert un minimum de six nœuds pour l'option d'archives inactives, dans lequel un schéma de 10+2 au lieu de 12+4 est utilisé. EC s'arrête lorsque le nombre de nœuds descend au-dessous du minimum requis pour le schéma EC.

Lorsqu'un fragment est saturé ou après une période définie, il est scellé, la parité est calculée et les segments de codage sont écrits sur les disques de l'ensemble du domaine de défaillance. Les données de fragments sont conservées sous la forme d'une copie unique composée de 16 segments (12 de données, 4 de code) dispersées sur l'ensemble du cluster. ECS n'utilise les segments de code pour la reconstruction du fragment qu'en cas de défaillance.

Lorsque l'infrastructure sous-jacente d'un VDC change au niveau du nœud ou du rack, les couches de Fabric détectent la modification et déclenchent un analyseur de rééquilibrage comme tâche d'arrière-plan. L'analyseur calcule la meilleure disposition des segments EC sur les domaines de défaillance pour chaque fragment à l'aide de la nouvelle topologie. Si la nouvelle disposition offre une meilleure protection que la disposition existante, ECS répartit les segments EC dans une tâche en arrière-plan. Cette tâche a un impact minime sur les performances du système ; toutefois, il y aura une augmentation du trafic entre les nœuds lors du rééquilibrage. Un équilibrage des partitions de la table logique sur les nouveaux nœuds se produit également, et les fragments de journal et d'arborescence B+ nouvellement créés sont désormais alloués de manière égale sur les anciens et les nouveaux nœuds. La redistribution améliore la protection en local en tirant parti de toutes les ressources au sein de l'infrastructure.

Remarque : Il est recommandé de ne pas attendre que la plate-forme de stockage soit entièrement saturée avant d'ajouter des disques ou des nœuds. Un seuil de taux d'utilisation du stockage raisonnable est de 70 % en tenant compte du taux d'acquisition quotidien et du temps prévu pour la commande, la livraison et l'intégration des disques/nœuds ajoutés.

7.1 Conformité

Pour respecter les exigences de conformité de l'entreprise et du secteur (norme SEC 17a-4(f)) pour le stockage des données, ECS a implémenté les éléments suivants :

- **Renforcement de la plate-forme** : le renforcement s'attaque aux failles de sécurité dans ECS, avec notamment le verrouillage de la plate-forme pour désactiver l'accès aux nœuds ou aux clusters, la fermeture de tous les ports non essentiels (p. ex. *ftpd, sshd*), la journalisation complète des audits pour les commandes sudo et la prise en charge de SRS (Dell EMC Secure Remote Services) pour arrêter l'accès distant aux nœuds.
- **Création de rapport sur la conformité** : un agent système signale l'état de conformité du système, par exemple *Good* indique une conformité ou *Bad* indique une non-conformité.
- **Règles et rétention des enregistrements basée sur des politiques** : possibilité de limiter les modifications apportées aux enregistrements ou aux données en cours de rétention à l'aide de politiques, de périodes et de règles.
- **Gestion avancée de la rétention (ARM)** : pour répondre aux exigences de conformité de Centera, un ensemble de règles de rétention a été défini pour CAS uniquement.
  - **Rétention basée sur des événements** : active des périodes de rétention qui commencent lorsqu'un événement spécifié se produit.
  - **Conservation en vue d'un litige** : permet une prévention temporaire de la suppression des données soumises à une action légale.
  - **Gouverneur min./max.** : paramètre par bucket définissant la période de rétention minimale et maximale par défaut.

La conformité est activée au niveau de l'espace de nommage. Les périodes de rétention sont configurées au niveau du bucket. Les exigences de conformité certifient la plate-forme ; pour cette raison, la fonctionnalité de conformité est uniquement disponible pour ECS s'exécutant sur le matériel de l'appliance. Pour plus d'informations sur l'activation et la configuration de la conformité dans ECS, consultez le guide *ECS Data Access Guide*actuel et le guide *ECS Administrator's Guide* le plus récent.

8 Déploiement

ECS peut être déployé sous forme d'instance sur un ou plusieurs sites. Les blocs de construction d'un déploiement ECS sont les suivants :

- **Datacenter virtuel (VDC)** : il s'agit d'un cluster, généralement aussi appelé site ou zone géographique distincte, composé d'un ensemble d'infrastructures ECS gérées par une seule instance de Fabric.
- **Pool de stockage (SP)** : les SP peuvent être considérés comme un sous-ensemble de nœuds et leur stockage associé appartenant à un VDC. Un nœud ne peut appartenir qu'à un seul SP. EC est défini au niveau du SP avec un schéma de 12+4 ou 10+2. Un SP peut être utilisé en tant qu'outil pour séparer physiquement les données entre les clients ou les groupes de clients qui accèdent au stockage sur ECS.
- **Groupe de réplication (RG)** : les RG définissent où le contenu des SP est protégé et les emplacements à partir desquels les données sont accessibles. Un RG avec un seul site est parfois appelé RG local. Les données sont toujours protégées en local car elles sont écrites en prévision des pannes de disque, de nœud et de rack. Les RG avec deux sites ou plus sont parfois appelés RG globaux. Les RG globaux s'étendent sur jusqu'à 8 VDC et protègent contre les pannes de disque, de nœud, de rack et de site. Un VDC peut appartenir à plusieurs RG.
- **Espace de nommage** : un espace de nommage est conceptuellement similaire à un tenant dans ECS. Une des caractéristiques clés d'un espace de nommage tient au fait que ses utilisateurs ne peuvent pas accéder aux objets d'un autre espace de nommage.
- **Buckets** : es buckets sont des conteneurs d'objets créés dans un espace de nommage et peuvent parfois être considérés comme des conteneurs logiques pour les sous-clients. Dans S3, les conteneurs sont appelés buckets, et ce terme a été adopté dans ECS. Dans Atmos, l'équivalent d'un bucket est un sous-tenant ; dans Swift, il s'agit d'un conteneur et dans CAS, d'un pool CAS. Les buckets sont des ressources globales d'ECS. Chaque bucket est créé dans un espace de nommage, et chaque espace de nommage est créé dans un RG.

ECS s'appuie sur les systèmes d'infrastructure suivants :

- **DNS** : (obligatoire) recherches directes et inversées requises pour chaque nœud ECS.
- **NTP** : (obligatoire) serveur Network Time Protocol.
- **SMTP** : (en option) serveur Simple Mail Transfer Protocol permettant d'envoyer des alertes et des rapports.
- **DHCP** : (en option) obligatoire si vous attribuez des adresses IP via DHCP.
- **Fournisseurs d'authentification** : (en option) les administrateurs ECS peuvent être authentifiés à l'aide de groupes Active Directory et LDAP. Les utilisateurs d'objets peuvent être authentifiés à l'aide de Keystone. Les fournisseurs d'authentification ne sont pas obligatoires pour ECS. ECS dispose d'une fonctionnalité intégrée de gestion locale des utilisateurs ; sachez cependant que les utilisateurs créés en local ne sont pas répliqués entre les VDC.
- **Répartiteur de charge** : (obligatoire si le workflow l'exige, en option dans le cas contraire) la charge client doit être répartie sur les nœuds afin d'utiliser efficacement toutes les ressources disponibles dans le système. Si une appliance ou un service d'équilibrage de charge dédié est nécessaire pour gérer la charge sur les nœuds ECS, cette fonctionnalité doit être considérée comme obligatoire. Les développeurs qui écrivent des applications à l'aide du SDK ECS S3 peuvent tirer parti de la fonctionnalité intégrée de répartiteur de charge. Les répartiteurs de charge sophistiqués peuvent prendre en compte des facteurs supplémentaires, tels que la charge signalée par un serveur, les temps de réponse, l'état up/down, le nombre de connexions actives et l'emplacement géographique. Le client est responsable de la gestion du trafic client et de la détermination des exigences en

matière d'accès. Quelle que soit la méthode utilisée, il existe des options de base qui sont généralement considérées, comme l'allocation manuelle d'adresses IP, la permutation circulaire DNS, l'équilibrage de charge côté client, les appliances de répartiteur de charge et les répartiteurs de charge géographiques. Voici une brève description de chacune de ces méthodes :

- **Allocation manuelle d'adresses IP** : les adresses IP sont attribuées manuellement aux applications. Cela n'est généralement pas recommandé, car la charge ne sera peut-être pas distribuée ou la tolérance de panne ne sera pas assurée.
- **Permutation circulaire DNS** : une entrée DNS est créée qui inclut toutes les adresses IP de nœud. Les clients interrogent le DNS pour résoudre les noms de domaine complets des services ECS et reçoivent une réponse avec les adresses IP d'un nœud aléatoire. Cela peut permettre une certaine pseudo répartition de charge. Il est possible que cette méthode n'offre pas de tolérance de panne, car une intervention manuelle est souvent nécessaire pour supprimer les adresses IP des nœuds défaillants du DNS. Vous pouvez rencontrer des problèmes de durée de vie (TTL) avec cette méthode. Certaines implémentations de serveur DNS peuvent mettre en cache des recherches DNS pendant un certain temps afin que les clients qui se connectent dans un délai proche puissent le faire à la même adresse IP, ce qui réduit la quantité de distribution de charge vers les nœuds de données. Il n'est pas recommandé d'utiliser le DNS pour distribuer le trafic dans le cadre d'une permutation circulaire.
- **Équilibrage de charge** : les répartiteurs de charge constituent l'approche la plus courante pour la distribution de la charge client. Les clients peuvent envoyer le trafic vers un répartiteur de charge qui le reçoit et le transfère vers un nœud ECS sain. L'état de la connexion ou des bilans de santé proactifs permettent de vérifier la disponibilité de chaque nœud pour traiter les demandes. Les nœuds indisponibles ne sont pas utilisés tant qu'ils ne font pas l'objet d'un bilan de santé réussi. Il est possible de décharger le traitement SSL, qui sollicite très largement le processeur, pour libérer ces ressources sur ECS.
- **Équilibrage géographique de la charge** : l'équilibrage géographique de la charge s'appuie sur le DNS pour acheminer les recherches vers une appliance comme Riverbed SteelApp, qui utilise Geo-IP ou un autre mécanisme pour déterminer le meilleur site vers lequel acheminer le client.

8.1 Déploiement sur site unique

Lors du déploiement initial sur un site unique ou un seul cluster, les nœuds sont d'abord ajoutés à un SP. Les SP sont des conteneurs logiques de nœuds physiques. La configuration du SP consiste à sélectionner le nombre minimal requis de nœuds disponibles et à choisir le schéma EC par défaut 12+4 ou d'archives inactives 10+2. Des niveaux d'alerte critiques peuvent être définis au cours de la configuration du SP au départ et ultérieurement. Toutefois, il est impossible de modifier le schéma EC après l'initialisation du SP. Le premier SP créé est désigné comme le SP du système et est utilisé pour stocker les métadonnées de ce dernier. Il est impossible de supprimer le SP système.

Comme illustré sur la Figure 21, les clusters contiennent généralement un ou deux SP, un pour chaque schéma EC, mais si une organisation requiert une séparation physique des données, des SP supplémentaires sont utilisés pour mettre en œuvre des limites.

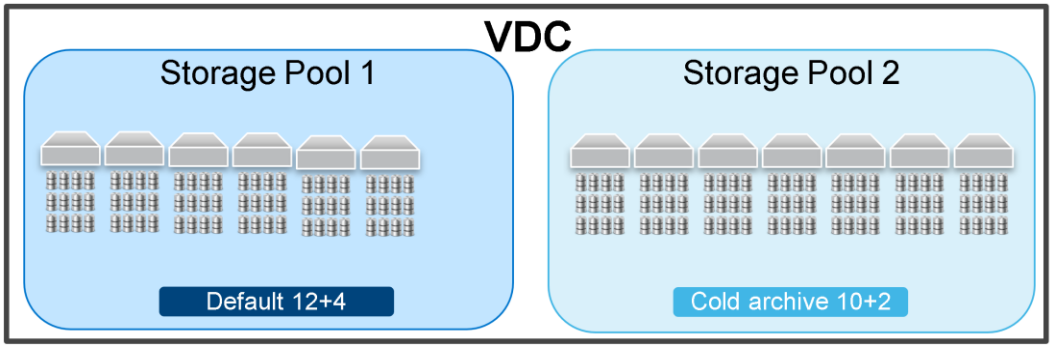

Figure 21 VDC avec deux pools de stockage, chacun configuré avec un schéma EC différent

Après l'initialisation du premier SP, un VDC peut être créé. La configuration du VDC implique la spécification des points de terminaison de gestion et de réplication. Sachez que même si l'initialisation du SP du système est obligatoire avant la création du VDC, la configuration du VDC n'entraîne pas l'affectation des SP, mais plutôt des adresses IP des nœuds.

Après la création d'un VDC, les RG sont configurés. Les RG sont des ressources globales avec une configuration incluant la désignation d'au moins un VDC, lui-même dans la configuration d'un site unique ou initial, ainsi que l'un des SP du VDC. Un RG avec un seul membre VDC protège les données localement au niveau du disque, du nœud et du rack. La section suivante développe les RG pour inclure les déploiements multisites.

Les espaces de nommage sont des ressources globales créées et affectées à un RG. C'est au niveau des espaces de nommage que sont définis les politiques de rétention, les quotas, les administrateurs de conformité et d'espace de nommage. L'accès au cours d'une panne (Access During Outage, ADO), décrit dans la section suivante, peut être configuré au niveau de l'espace de nommage. En règle générale, les tenants sont organisés au niveau de l'espace de nommage. Les tenants peuvent être une instance d'application ou une équipe, un utilisateur, un groupe professionnel ou tout autre regroupement adapté à l'organisation.

Les buckets sont des ressources globales qui peuvent englober plusieurs sites. La création d'un bucket implique son attribution à un espace de nommage et à un RG. C'est au niveau du bucket que sont activés l'état de propriété et l'accès aux fichiers ou à CAS. La Figure 22 ci-dessous illustre un SP dans un VDC avec un espace de nommage contenant deux buckets.

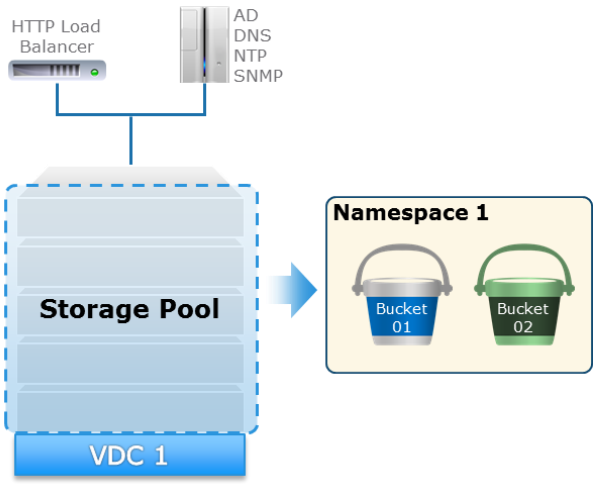

Figure 22 Exemple de déploiement à site unique

8.2 Déploiement sur plusieurs sites

Un déploiement multisite, également appelé environnement fédéré ou ECS fédéré, peut s'étendre sur un maximum de huit VDC. Les données sont répliquées dans ECS au niveau des fragments. Les nœuds qui participent à un RG envoient leurs données locales de manière asynchrone à un ou à tous les autres sites. Les données sont chiffrées à l'aide de l'algorithme AES256 avant d'être envoyées sur le réseau WAN via HTTP. Les principaux avantages reconnus de la fédération de plusieurs VDC sont les suivants :

- Consolidation des efforts de gestion de plusieurs VDC dans une ressource logique unique
- Protection au niveau du site, en plus de la protection en local au niveau des nœuds, des disques et des racks
- Accès distribué géographiquement à l'espace de stockage, de manière fortement cohérente et active partout

Cette section sur le déploiement multisite décrit les fonctionnalités spécifiques d'une solution ECS fédérée, telles que :

- **Cohérence des données** : par défaut, ECS fournit un service de stockage fortement cohérent.
- **Groupes de réplication** : conteneurs globaux servant à définir les limites de protection et d'accès.
- **Géo-mise en cache** : optimisation pour les workflows d'accès à un site distant dans les déploiements multisite.
- **ADO** : comportement de l'accès client lors d'une panne de site temporaire (TSO).

8.2.1 Cohérence des données

ECS est un système fortement cohérent qui utilise la propriété pour conserver une version faisant autorité de chaque espace de nommage, bucket et objet. La propriété est attribuée au VDC dans lequel l'espace de nommage, le bucket ou l'objet est créé. Par exemple, si un espace de nommage NS1 est créé sur VDC1, VDC1 possède NS1 et est responsable de la gestion de la version autorisée des buckets au sein de NS1. Si un bucket B1 est créé sur VDC2 à l'intérieur de NS1, VDC2 possède B1 et est responsable de la gestion de la version autorisée du contenu du bucket, ainsi que de chaque VDC propriétaire de l'objet. De même, si un objet O1 est créé dans B1 sur VDC3, VDC3 possède O1 et est responsable de la gestion de la version autorisée de O1 et des métadonnées associées.

La résilience de la protection des données multisite se fait au prix d'une augmentation de la surcharge de protection du stockage et de la consommation de bande passante sur le réseau WAN. Des requêtes d'index sont requises lors de l'accès à un objet ou de sa mise à jour à partir d'un site qui n'est pas propriétaire de l'objet. De même, des recherches d'index sur le réseau WAN sont également requises pour récupérer des informations, telles qu'une liste faisant autorité des buckets d'un espace de nommage ou des objets d'un bucket, détenues par un site distant.

Comprendre comment ECS utilise la propriété pour assurer un suivi fiable des données au niveau des espaces de nommage, des buckets et des objets aide les administrateurs et les propriétaires d'applications à prendre des décisions en matière d'accès lors de la configuration de leur environnement.

8.2.2 Groupe de réplication actif

Lors de la création d'un RG, un paramètre *Replicate to All Sites* est disponible ; il est désactivé par défaut ou peut être activé pour permettre l'utilisation de cette fonctionnalité. La réplication des données sur tous les sites signifie que les données écrites individuellement sur chaque VDC sont répliquées sur tous les autres VDC membres du RG. Par exemple, une instance d'ECS fédérée de type X-nombre-de-sites avec un RG actif configuré pour répliquer les données sur tous les sites donne lieu à une surcharge de protection X fois supérieure, ou X * 1,33 (1,2 avec le schéma EC d'archives inactives) au total pour la protection des données. La réplication sur tous les sites peut être pertinente, en particulier pour les petits jeux de données dans lesquels un accès local est important. Si vous désactivez cette option, toutes les données écrites dans chaque VDC sont répliquées vers un autre VDC. Le site primaire, là où l'objet est créé, et le site de stockage de la copie de réplication, protègent chacun les données en local à l'aide du schéma EC alloué au SP local. Autrement dit, seules les données d'origine sont répliquées sur le réseau WAN et non les segments de codage EC associés.

Les données stockées dans un RG actif sont accessibles aux clients via n'importe quel VDC membre du RG. La Figure 23 ci-dessous montre un exemple d'ECS fédéré créé à l'aide de VDC1, VDC2 et VDC3. Deux RG sont affichés : RG1 possède un seul membre, VDC1, et RG2 possède les trois VDC en tant que membres. Trois buckets sont affichés : B1, B2 et B3.

Dans cet exemple, les clients qui accèdent :

- au VDC1 ont accès à tous les buckets ;
- au VDC2 et au VDC3 ont uniquement accès aux buckets B2 et B3.

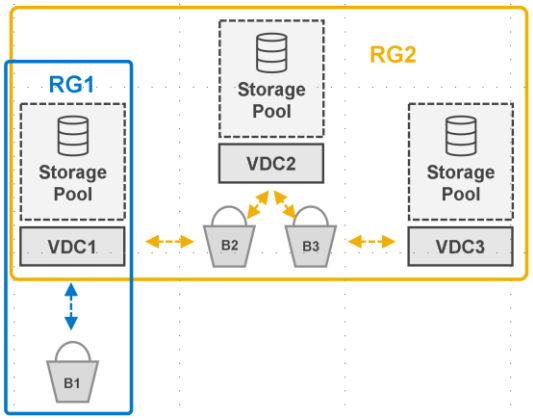

Figure 23 Accès au niveau du bucket par site avec des groupes de réplication à un ou plusieurs sites

8.2.3 Groupe de réplication passif

Un RG passif dispose de trois VDC membres. Deux des VDC sont désignés comme actifs et sont accessibles aux clients. Le troisième VDC est désigné comme passif et est utilisé en tant que cible de réplication uniquement. Le site passif est utilisé à des fins de restauration uniquement et ne permet pas un accès client direct. Les avantages de la réplication géo-passive sont les suivants :

- Réduction de la surcharge de protection du stockage en augmentant le potentiel des opérations XOR
- Contrôle au niveau de l'administrateur de l'emplacement utilisé pour le stockage avec réplication uniquement

Figure 24 montre un exemple de configuration géo-passive, où VDC 1 et VDC 2 sont des sites primaires (sources) qui répliquent tous deux leurs données (fragments) vers la cible de réplication, VDC 3.

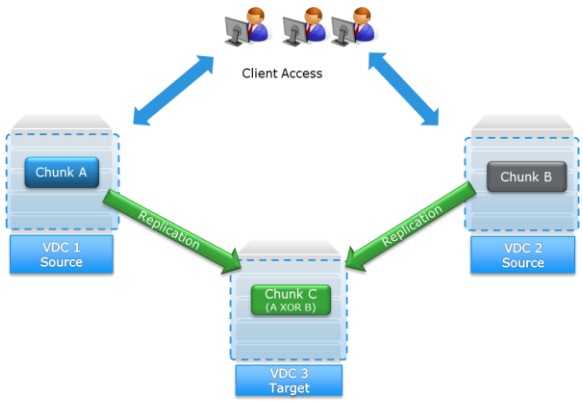

Figure 24 Chemins d'accès client et de réplication pour le groupe de réplication géo-passive

L'accès multisite à des données fortement cohérentes s'effectue à l'aide de la propriété des objets, du bucket et de l'espace de nommage sur l'ensemble des sites membres du RG. Les requêtes d'index intersites sur le réseau WAN sont nécessaires lorsque l'accès à l'API provient d'un VDC qui ne dispose pas de la ou des constructions logiques requises. Les recherches WAN permettent de déterminer la version autorisée des données. Par conséquent, si un objet créé sur Site 1 est lu à partir de Site 2, une recherche WAN est requise pour interroger le VDC propriétaire de l'objet, Site 1, afin de vérifier si les données de l'objet qui ont été répliquées sur Site 2 correspondent à la version la plus récente des données. Si Site 2 ne dispose pas de la

dernière version, il extrait les données nécessaires à partir de Site 1 ; dans le cas contraire, il utilise les données précédemment répliquées. Cela est illustré à la Figure 25 ci-dessous.

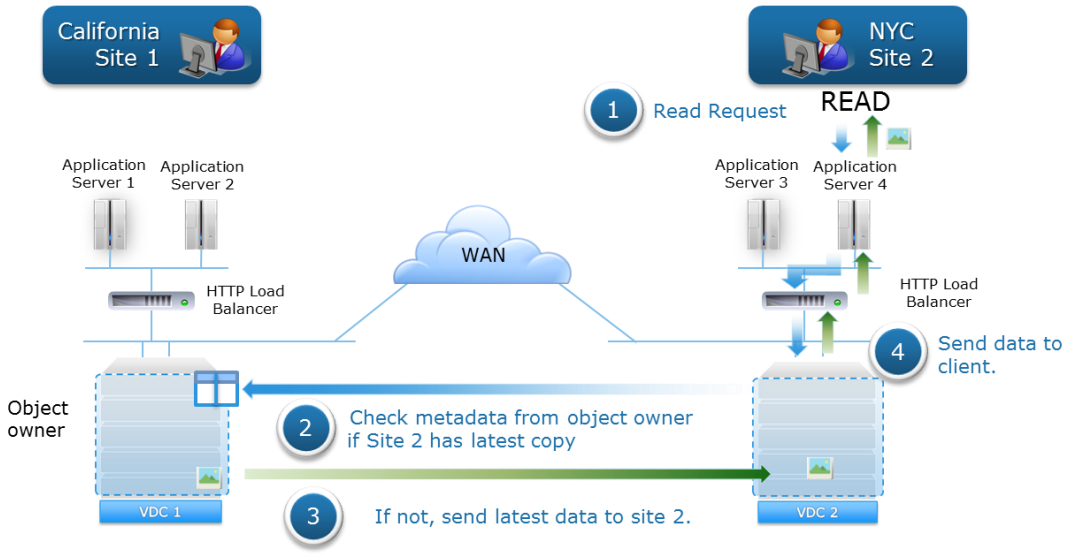

Figure 25 La demande de lecture du VDC non-propriétaire déclenche une recherche WAN sur le VDC propriétaire de l'objet

Le flux de données des écritures dans un environnement géorépliqué sur lequel deux sites mettent à jour le même objet est illustré sur la Figure 26. Dans cet exemple, Site 1 a initialement créé l'objet et en est propriétaire. Le codage d'effacement a été appliqué à l'objet, et les transactions de journal associées ont été écrites sur le disque sur Site 1. Le flux de données pour une mise à jour de l'objet reçue sur Site 2 est le suivant :

1. Site 2 écrit d'abord les données localement.
2. Site 2 met à jour de manière synchrone les métadonnées (écriture de journal) avec le propriétaire de l'objet, Site 1, et attend l'accusé de réception de la mise à jour des métadonnées à partir de Site 1.
3. Site 1 reconnaît l'écriture de métadonnées sur Site 2.
4. Site 2 reconnaît l'écriture sur le client.

Remarque : Site 2 réplique de manière asynchrone les données de Site 1, le site propriétaire de l'objet, selon le processus habituel. Si les données doivent être servies à partir de Site 1 avant d'être répliquées depuis Site 2, Site 1 récupère les données directement à partir de Site 2.

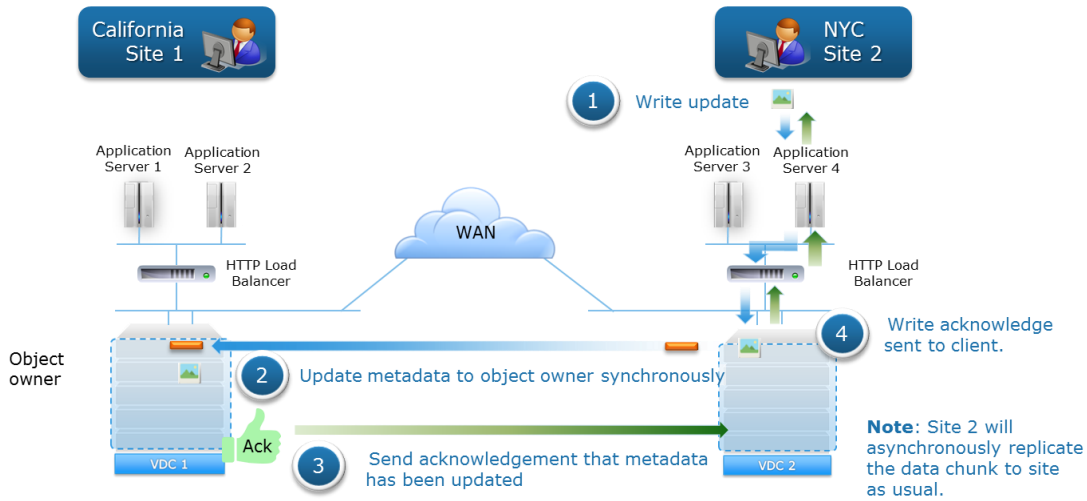

Figure 26 Mise à jour du même flux de données d'objets dans un environnement géorépliqué

Dans les scénarios de lecture et d'écriture au sein d'un environnement géorépliqué, il y a un temps de latence associé à la lecture et à la mise à jour des métadonnées et à la récupération des données à partir du site propriétaire de l'objet.

Remarque : Pour ECS 3.4 et versions ultérieures, vous pouvez supprimer un VDC d'un groupe de réplication (RG) dans une fédération contenant plusieurs VDC sans affecter les VDC ou les autres groupes de réplication associés au VDC. Le retrait d'un VDC du RG ne lance plus de PSO (panne de site permanente). Le retrait d'un VDC du RG lance la restauration.

Pour plus d'informations sur les groupes de réplication, reportez-vous au guide *ECS Administrator Guide* le plus récent.

8.2.4 Géo-mise en cache des données distantes

ECS optimise les temps de réponse pour accéder aux données stockées sur les sites distants en mettant en cache localement les objets lus sur le réseau WAN. Cela peut être utile pour les modèles d'accès multisites, dans lesquels les données sont souvent extraites d'un site distant ou non-propriétaire. Prenons un environnement géorépliqué avec trois sites, VDC1, VDC2 et VDC3, où un objet est écrit dans VDC1 et la copie de réplication de l'objet est stockée sur VDC2. Dans ce scénario, pour traiter une demande de lecture reçue sur VDC3, pour l'objet créé sur VDC1 et répliqué sur VDC2, les données d'objet doivent être envoyées à VDC3 à partir de VDC1 ou VDC2. La géo-mise en cache des données distantes fréquemment utilisées permet de réduire les temps de réponse. Un algorithme LRU (Least Recently Used) est utilisé pour la mise en cache. La taille du géo-cache est ajustée lorsqu'une infrastructure matérielle (disques, nœuds ou racks, par exemple) est ajoutée à un SP géorépliqué.

8.2.5 Comportement lors d'une panne de site

Une panne de site temporaire (TSO) fait généralement référence à une défaillance de la connectivité WAN ou d'un site entier, par exemple lors d'une catastrophe naturelle. ECS utilise des mécanismes de pulsation pour détecter et gérer les pannes de site temporaires. L'accès client et la disponibilité des opérations API au niveau de l'espace de nommage, des buckets et des objet au cours d'une TSO sont régis par les options ADO suivantes définies au niveau de l'espace de nommage et du bucket :

- **Off (par défaut)** : une cohérence élevée est maintenue au cours d'une panne temporaire.
- **On** : un accès cohérent à terme est autorisé lors d'une panne de site temporaire.

La cohérence des données lors d'une TSO est mise en œuvre au niveau du bucket. La configuration est définie au niveau de l'espace de nommage, qui définit le paramètre ADO par défaut lors de la création d'un nouveau bucket et peut être remplacé à la création d'un autre. En d'autres termes, la TSO peut être configurée pour certains buckets et non pour d'autres.

8.2.5.1 Access During Outage (ADO) non activé

Par défaut, ADO n'est pas activé et une forte cohérence est maintenue. Toutes les demandes de l'API client pour lesquelles des données autorisées d'espace de nommage, de bucket ou d'objet sont requises, mais temporairement indisponibles, échouent. Les opérations de lecture, de création, de mise à jour et de suppression sur les objets, ainsi que les listes de buckets qui ne sont pas détenus par un site en ligne, échouent. En outre, les opérations de création et de modification du bucket, de l'utilisateur et de l'espace de nommage échouent également.

Comme mentionné précédemment, le site initial propriétaire du bucket, de l'espace de nommage et d'un objet est le site sur lequel la ressource a été créée pour la première fois. Au cours d'une TSO, certaines opérations peuvent échouer si le site propriétaire de la ressource n'est pas accessible. Les points clés des opérations autorisées ou non autorisées au cours d'une panne de site temporaire sont les suivants :

- La création, la suppression et la mise à jour des buckets, des espaces de nommage, des utilisateurs d'objets, des fournisseurs d'authentification, des RG et des mappages de groupe et d'utilisateur NFS ne sont pas autorisés à partir d'un site.
- La liste des buckets au sein d'un espace de nommage est autorisée si le site propriétaire de l'espace de nommage est disponible.

HDFS/NFS active les buckets qui sont détenus par le site inaccessible en lecture seule.

8.2.5.2 ADO activé

Dans un bucket pour lequel ADO est activé, le service de stockage fournit des réponses cohérentes à terme au cours d'une TSO. Dans ce scénario, les opérations de lecture et, le cas échéant, d'écriture à partir d'un site secondaire (non-propriétaire) sont acceptées et honorées. En outre, une écriture sur un site secondaire au cours d'une TSO oblige le site secondaire à prendre possession de l'objet. Cela permet à chaque VDC de continuer à lire et à écrire des objets à partir de buckets dans un espace de nommage partagé. Enfin, la nouvelle version de l'objet devient la version autorisée de l'objet lors de la réconciliation post-TSO, même si une autre application met à jour l'objet sur le VDC propriétaire.

Bien qu'un grand nombre d'opérations d'objet se poursuivent au cours d'une panne de réseau, certaines opérations ne sont pas autorisées, telles que la création de buckets, d'espaces de nommage ou d'utilisateurs. Lorsque la connectivité réseau entre deux VDC est restaurée, le mécanisme de pulsation détecte automatiquement la connectivité, restaure le service et résout les incohérences des objets à partir des deux sites. Si le même objet est mis à jour sur le VDC A et le VDC B, la copie sur le VDC non-propriétaire est la copie faisant autorité. Par conséquent, si un objet détenu par le VDC B est mis à jour à la fois sur le VDC A et le VDC B lors de la synchronisation, la copie sur le VDC A est celle qui fait autorité et qui est conservée, tandis que l'autre copie est non référencée et disponible pour la récupération d'espace.

Lorsque plus de deux VDC font partie d'un RG, et si la connectivité réseau est interrompue entre un VDC et les deux autres, les opérations d'écriture/mise à jour/propriété se poursuivent comme elles le feraient sur deux VDC, mais le processus pour répondre aux demandes de lecture est plus complexe, ainsi que le montre la description ci-dessous.

Si une application demande un objet appartenant à un VDC qui n'est pas accessible, ECS envoie la demande au VDC ayant la copie secondaire de l'objet. Cependant, la copie du site secondaire peut avoir subi une réduction des données qui correspond à une opération de type XOR entre deux jeux de données différents, ce qui produit un nouveau jeu de données. Pour cette raison, le VDC du site secondaire doit récupérer au préalable les fragments de l'objet inclus dans l'opération XOR d'origine et il doit appliquer une opération XOR à ces fragments à l'aide de la copie de restauration. Cette opération renvoie le contenu du fragment initialement stocké sur le VDC en panne. Les fragments de l'objet restauré peuvent être réassemblés et renvoyés. Lorsque les fragments sont reconstruits, ils sont également mis en cache afin que le VDC puisse répondre plus rapidement aux demandes ultérieures. Il faut savoir que la reconstruction prend du temps. Plus il existe de VDC dans un RG, plus le nombre de fragments devant être extraits à partir d'autres VDC est élevé. Par conséquent, la reconstruction de l'objet prend plus de temps.

Lorsqu'un sinistre se produit, un VDC tout entier peut devenir irrécupérable. ECS traite le VDC irrécupérable comme une panne de site temporaire. Si la panne est permanente, l'administrateur système doit basculer définitivement le VDC à partir de la fédération pour déclencher un processus de basculement, ce qui démarre une nouvelle protection et une resynchronisation des objets stockés sur le VDC en panne. Les tâches de restauration s'exécutent en arrière-plan. Vous pouvez vérifier la progression de la restauration dans le portail ECS.

Une option de bucket supplémentaire pour ADO *en lecture seule (RO)* est disponible, ce qui garantit que la propriété de l'objet n'est jamais modifiée et supprime le risque de conflits causés par les mises à jour d'objets sur les sites en échec et en ligne au cours d'une panne de site temporaire. L'inconvénient de la propriété RO ADO est que, lors d'une panne de site temporaire, aucun nouvel objet ne peut être créé et aucun objet existant dans le bucket ne peut être mis à jour avant que tous les sites ne soient de nouveau en ligne. L'option RO ADO n'est disponible que lors de la création du bucket. Elle ne peut pas être modifiée par la suite. Cette option est désactivée par défaut.

Tableau 8 Tolérance de panne sur plusieurs sites

Modèle d'échec	Tolérance
Environnement géorépliqué	Jusqu'à une panne de site

8.3 Tolérance de panne

ECS est conçu pour tolérer diverses pannes d'équipement à l'aide d'un certain nombre de domaines de défaillance. La gamme de conditions de défaillance couvre différents scénarios, notamment :

- Défaillance d'un seul disque dur dans un seul nœud
- Défaillance de plusieurs disques durs dans un seul nœud
- Défaillance d'un seul disque dur dans plusieurs nœuds
- Défaillance de plusieurs disques durs dans plusieurs nœuds
- Défaillance d'un seul nœud
- Défaillance de plusieurs nœuds
- Perte de communication vers un VDC répliqué
- Perte d'un VDC entier répliqué

Dans le cas d'une configuration à un site, à deux sites ou géorépliquée, l'impact de la défaillance dépend de la quantité et du type des composants concernés. Toutefois, à chaque niveau, ECS fournit des mécanismes de défense contre l'impact des échecs de composant. La plupart de ces mécanismes ont déjà été abordés

dans ce livre blanc, mais sont présentés ici et sur la Figure 27 afin de montrer comment ils sont appliqués à la solution. Ces composants sont les suivants :

- Disk failure
  - Les segments EC ou les copies de réplica du même fragment ne sont pas stockés sur le même disque.
  - Calcul de la somme de contrôle lors des opérations d'écriture et de lecture.
  - Vérification en arrière-plan des sommes de contrôle par le vérificateur de cohérence.
- Node failure
  - Distribution des segments ou des copies de réplica d'un fragment de manière égale sur l'ensemble des nœuds d'un VDC.
  - Le Fabric ECS veille à ce que les services soient en cours d'exécution et gère des ressources telles que les disques et le réseau.
  - Les enregistrements de partition et les tables sont protégés par le basculement de la propriété de la partition de nœud à nœud.
- Défaillance du rack dans le VDC
  - Distribution des segments ou des copies de réplica d'un fragment de manière égale dans un VDC.
  - Une instance de registre de Fabric s'exécute dans chaque rack et peut être redémarrée sur n'importe quel autre nœud du même rack si le nœud tombe en panne.

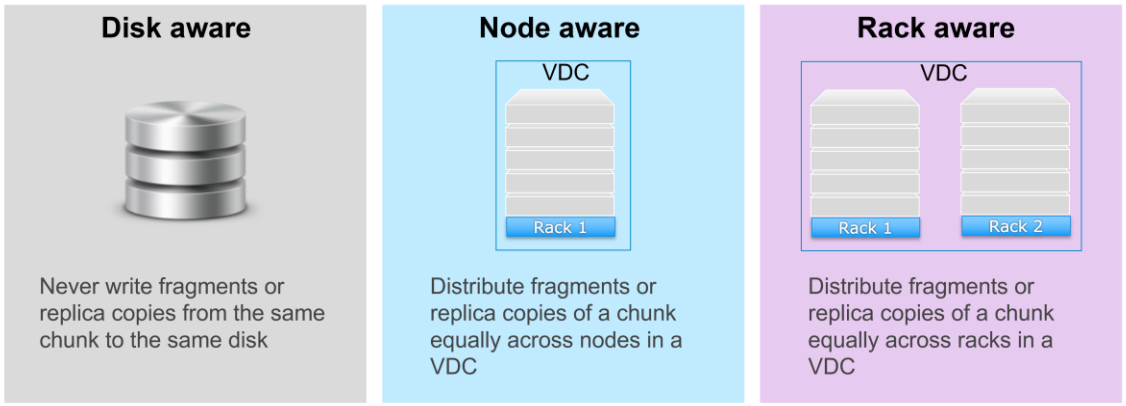

Figure 27 Mécanismes de protection au niveau des disques, des nœuds et des racks

Le tableau suivant définit le type et le nombre d'échecs de composant visés par chaque schéma EC par rapport à la configuration de base du rack. Le Tableau 9 souligne l'importance de prendre en compte l'impact des domaines de défaillance de protection sur la disponibilité globale des données et des services en ce qui concerne le nombre de nœuds requis pour chaque schéma EC.

Tableau 9 Protection par code d'effacement sur l'ensemble des domaines de défaillance

Schéma EC	Nombre de nœuds dans le VDC	Nombre de fragments par nœud	Protection des données EC contre...
12+4 Par défaut	5 ou moins	4	• Perte de quatre disques max. ou • Perte d'un nœud
	6 ou 7	3	• Perte de quatre disques max. ou • Perte d'un nœud et d'un disque d'un second nœud
	8 ou plus	2	• Perte de quatre disques max. ou • Perte de deux nœuds ou • Perte d'un nœud et de deux disques
	16 ou plus	1	• Perte de quatre nœuds ou • Perte de trois nœuds et des disques d'un nœud supplémentaire ou • Perte de deux nœuds et des disques de deux nœuds différents ou • Perte d'un nœud et des disques d'au maximum trois nœuds différents ou • Perte de quatre disques de quatre nœuds différents
10+2 Cold Storage	11 ou moins	2	• Perte de deux disques max. ou • Perte d'un nœud
	12 ou plus	1	• Perte de n'importe quel nombre de disques de deux nœuds différents ou • Perte de deux nœuds

8.4 Automatisation du remplacement de disque

Avec ECS 3.5 et versions ultérieures, les clients peuvent faire remplacer les disques défaillants par les services Dell EMC à l'aide du portail ECS (interface utilisateur Web) au workflow intuitif. Cette fonctionnalité permet :

- La résolution en interne des pannes de disque
- L'accélération du délai de réparation
- La flexibilité opérationnelle et la réduction du coût TCO

La page de maintenance du portail ECS fournit une visibilité administrateur sur tous les disques de chaque nœud. Lorsqu'un disque tombe en panne, le système lance automatiquement la récupération. Tous les types de ressources sur le disque sont restaurés et, quand le disque est prêt à être retiré du nœud, le portail ECS affiche un bouton Remplacer, comme illustré par la Figure 28.

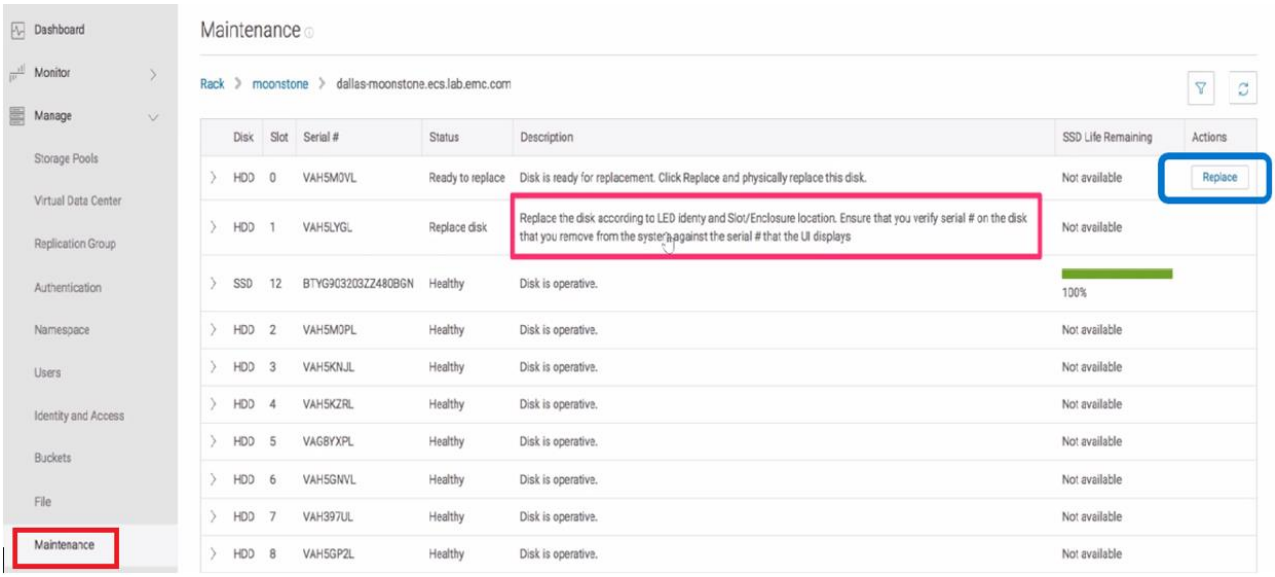

Figure 28 Automatisation du remplacement de disque

Remarque : Il n'est possible de remplacer qu'un seul disque à la fois. Cela permet d'éviter de remplacer le mauvais disque.

8.5 Actualisation des technologies

L'actualisation des technologies est un engagement direct de Dell EMC Professional Services, disponible à partir d'ECS 3.5, qui retire des clusters ECS sans interruption les nœuds matériels les plus anciens à l'aide de la fonctionnalité logicielle intégrée. Cette opération efficace est peu gourmande en ressources et peut être régulée avec précision. Le temps système précédemment associé à la mise hors service du matériel ECS en est ainsi réduit.

L'actualisation des technologies se divise en trois phases :

- **Extension de nœud** : ajout de nœuds Gen 3 au cluster existant
- **Migration des ressources** : déplacement de toutes les ressources sur les nœuds existants vers des nœuds Gen 3
- **Évacuation des nœuds** : nettoyage des anciens nœuds et suppression sur le cluster

Les Professional Services doivent être impliqués lors de la maintenance d'actualisation des technologies. Pour plus d'informations sur le renouvellement des technologies, reportez-vous à la version la plus récente du guide *ECS Tech Refresh Guide*.

9 Surcharge de protection du stockage

Chaque VDC membre d'un RG est responsable de sa propre protection EC des données au niveau local. Autrement dit, les données sont répliquées, mais pas les segments de codage associés. Bien que le schéma EC soit plus efficace sur le plan du stockage que d'autres formes de protection, telles que la mise en miroir d'un lecteur par copie intégrale, il entraîne une surcharge inhérente au stockage au niveau local. Toutefois, lorsqu'il est nécessaire que les copies secondaires soient répliquées hors site et que tous les sites aient accès aux données quand un site unique devient indisponible, les coûts de stockage deviennent plus importants qu'avec des méthodes traditionnelles de protection par copie des données de site à site. Cela est particulièrement vrai lorsque des données uniques sont réparties sur au moins trois sites.

ECS fournit un mécanisme qui permet d'accroître l'efficacité de la surcharge de protection du stockage en cas de fédération d'au moins trois sites. Dans un environnement répliqué à deux VDC, ECS réplique les fragments du VDC principal ou propriétaire sur un site distant afin d'assurer la haute disponibilité et la résilience. Il n'existe aucun moyen d'éviter la surcharge de 100 % de la protection par copie complète des données dans un déploiement ECS fédéré sur deux sites.

Considérons maintenant trois VDC dans un environnement multisite, VDC1, VDC2 et VDC3, où chaque VDC dispose de données uniques répliquées depuis chacun des autres VDC. VDC2 et VDC3 peuvent envoyer une copie de leurs données à VDC1 à des fins de protection. VDC1 aurait donc ses propres données d'origine, ainsi que les données répliquées à partir de VDC2 et VDC3. Cela signifie que VDC1 stockerait trois fois la quantité de données écrites sur son propre site.

Dans ce cas, ECS peut effectuer une opération XOR sur les données VDC2 et VDC3 stockées en local sur VDC1. Cette opération mathématique compare des quantités égales de données uniques, des fragments, et génère un résultat dans un nouveau fragment qui contient suffisamment de caractéristiques des deux fragments de données d'origine afin de permettre la restauration de l'un des deux ensembles d'origine. Ainsi, si auparavant il y avait trois ensembles uniques de fragments de données stockés sur VDC1, consommant trois fois la capacité disponible, il n'y en a désormais que deux : l'ensemble de données local d'origine et les copies de protection réduites par XOR.

Dans ce scénario, si VDC3 devient indisponible, ECS peut reconstruire les fragments de données VDC3 à l'aide de copies de fragment rappelées à partir de VDC2 et des données (C1 $\oplus$ C2) de VDC3 stockées localement sur VDC1. Ce principe s'applique à l'ensemble des trois sites faisant partie du RG et dépend de la présence de jeux de données uniques sur chacun des trois VDC. La Figure 29 montre un calcul XOR avec deux sites répliqués sur un troisième site.

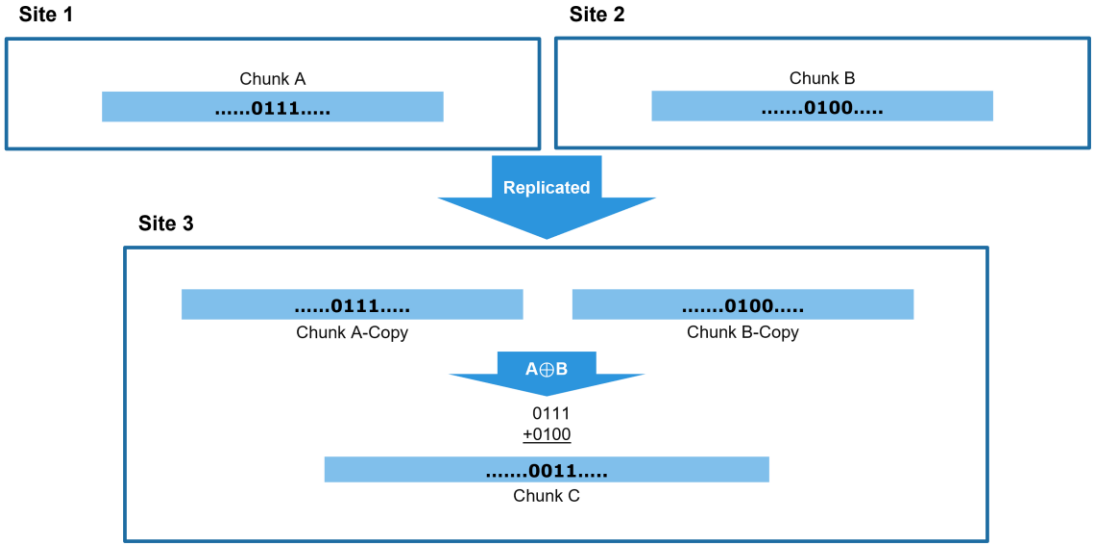

Figure 29 Efficacité de la protection des données avec XOR

Si les contrats de niveau de service d'une entreprise nécessitent une vitesse d'accès en lecture optimale même en cas de panne totale d'un site, l'option de réplication sur tous les sites oblige ECS à revenir à un stockage des copies intégrales des données répliquées sur tous les sites. Sans surprise, cela augmente les coûts de stockage proportionnellement au nombre de VDC faisant partie du RG. Par conséquent, une configuration à trois sites reviendrait à un triplement de la surcharge de protection du stockage. Le paramètre Replicate to All Sites est disponible à la création d'un RG et ne peut pas être activé et désactivé librement.

Au fur et à mesure que le nombre de sites fédérés augmente, l'optimisation XOR est plus efficace pour réduire la surcharge de protection du stockage causée par la réplication. Le Tableau 10 fournit des informations sur la surcharge de protection du stockage en fonction du nombre de sites pour le schéma EC normal de 12+4 et pour le schéma EC d'archives inactives de 10+2, illustrant la façon dont ECS peut améliorer l'efficacité du stockage au fur et à mesure que le nombre de sites liés augmente.

Remarque : Pour réduire la surcharge pour des données répliquées sur trois sites, et jusqu'à huit sites, les données uniques doivent être écrites relativement équitablement sur chaque site. En écrivant les données de manière égale sur l'ensemble des sites, chaque site aura un nombre similaire de fragments de réplica. Le nombre de fragments de réplica sur chaque site conduit à un nombre similaire d'opérations XOR pouvant se produire sur chaque site. L'efficacité maximale du stockage multisite est obtenue en réduisant le nombre maximal de fragments de réplica stockés à l'aide de XOR.

Tableau 10 Surcharge de protection du stockage

Nombre de sites dans le RG	EC 12+4	EC 10+2
1	1,33	1.2
2	2,67	2,4
3	2,00	1,8
4	1,77	1,6
5	1,67	1,5
6	1,60	1,44
7	1,55	1,40
8 (nombre max. de sites dans le RG)	1,52	1,37

10 Conclusion

Les organisations sont confrontées à des volumes de données et de stockage toujours plus élevés, en particulier dans l'espace de Cloud public. L'architecture scale-out et géodistribuée d'ECS fournit une plate-forme Cloud sur site qui peut évoluer jusqu'à plusieurs exaoctets de données avec un *coût total de possession* sensiblement inférieur à celui du stockage dans le Cloud public. ECS est une solution idéale en raison de sa polyvalence, de son hyperévolutivité, de ses fonctionnalités puissantes et de l'utilisation de matériel générique.

A Support technique et ressources

Dell.com/support propose des services et un support éprouvés répondant aux besoins des clients.

Des documents et vidéos techniques sur le stockage offrent aux clients l'expertise nécessaire pour tirer pleinement parti des plates-formes de stockage Dell EMC.