



# Configuration ONTAP

## Enterprise applications

NetApp  
May 09, 2024

# Sommaire

- Configuration ONTAP ..... 1
  - RAID et les bases de données Oracle ..... 1
  - Bases de données Oracle et gestion de la capacité de stockage ..... 1
  - Bases de données Oracle et machines virtuelles de stockage ..... 2
  - Gestion des performances des bases de données Oracle avec la QoS ONTAP ..... 3
  - Bases de données Oracle et fonctionnalités d'efficacité ONTAP ..... 5
  - Provisionnement fin avec les bases de données Oracle ..... 9
  - Basculer/basculer des bases de données Oracle et du contrôleur ONTAP ..... 11

# Configuration ONTAP

## RAID et les bases de données Oracle

RAID désigne l'utilisation de la redondance pour protéger les données contre la perte d'un disque.

Des questions se posent parfois au sujet des niveaux RAID dans la configuration du stockage NetApp utilisé pour les bases de données Oracle et d'autres applications d'entreprise. De nombreuses meilleures pratiques Oracle en matière de configuration de baie de stockage contiennent des avertissements concernant l'utilisation de la mise en miroir RAID et/ou l'évitement de certains types de RAID. Bien qu'elles soulèvent des points valides, ces sources ne s'appliquent pas au RAID 4 et aux technologies NetApp RAID DP et RAID-TEC utilisées dans ONTAP.

RAID 4, RAID 5, RAID 6, RAID DP et RAID-TEC utilisent tous la parité pour s'assurer qu'une panne de disque n'entraîne pas de perte de données. Ces options RAID offrent une meilleure utilisation du stockage que la mise en miroir, mais la plupart des implémentations RAID présentent des inconvénients pour les opérations d'écriture. La réalisation d'une opération d'écriture sur d'autres implémentations RAID peut nécessiter plusieurs lectures de disque pour régénérer les données de parité, un processus communément appelé la pénalité RAID.

Cependant, ONTAP n'entraîne pas cette pénalité RAID. Cela est dû à l'intégration de NetApp WAFL (Write Anywhere File Layout) à la couche RAID. Les opérations d'écriture sont fusionnées dans la mémoire RAM et préparées sous la forme d'une couche RAID complète, y compris la génération de la parité. ONTAP n'a pas besoin d'effectuer de lecture pour effectuer une écriture, ce qui signifie que ONTAP et WAFL évitent la pénalité RAID. Les performances des opérations stratégiques pour la latence, telles que la journalisation de reprise, sont assurées sans aucun obstacle. Les écritures aléatoires des fichiers de données n'entraînent aucune pénalité RAID résultant de la régénération de la parité.

En ce qui concerne la fiabilité statistique, même RAID DP offre une meilleure protection que la mise en miroir RAID. Le problème principal est la demande sur disques lors de la reconstruction RAID. Avec une configuration RAID en miroir, le risque de perte de données en cas de défaillance d'un disque pendant la reconstruction vers son partenaire dans la configuration RAID est bien plus grand que le risque de défaillance simultanée de trois disques dans une configuration RAID DP.

## Bases de données Oracle et gestion de la capacité de stockage

La gestion d'une base de données ou d'une autre application d'entreprise avec un stockage d'entreprise prévisible, gérable et haute performance requiert de l'espace libre sur les disques pour la gestion des données et des métadonnées. La quantité d'espace libre requise dépend du type de disque utilisé et des processus métier.

L'espace libre est défini comme tout espace qui n'est pas utilisé pour les données réelles et inclut l'espace non alloué sur l'agrégat lui-même et l'espace inutilisé au sein des volumes constitutifs. Le provisionnement fin doit également être envisagé. Par exemple, un volume peut contenir une LUN de 1 To, dont seulement 50 % sont utilisés par des données réelles. Dans un environnement à provisionnement fin, cet espace semble être consommé de 500 Go. Toutefois, dans un environnement entièrement provisionné, la capacité totale de 1 To semble être utilisée. Les 500 Go d'espace non alloué sont masqués. Cet espace n'est pas utilisé par les données réelles et doit donc être inclus dans le calcul de l'espace libre total.

Les recommandations de NetApp pour les systèmes de stockage utilisés pour les applications d'entreprise sont les suivantes :

## Des agrégats SSD, y compris les systèmes AFF



**NetApp recommande** un minimum de 10% d'espace libre. Cela inclut tout l'espace inutilisé, y compris l'espace libre au sein de l'agrégat ou d'un volume, ainsi que tout espace libre alloué en raison de l'utilisation du provisionnement complet, mais qui n'est pas utilisé par les données réelles. L'espace logique n'est pas important, la question est de savoir quelle quantité d'espace physique réellement disponible pour le stockage des données.

La recommandation de 10 % d'espace libre est très prudente. Les agrégats SSD peuvent prendre en charge des charges de travail à des niveaux d'utilisation encore plus élevés, sans affecter les performances. Cependant, à mesure que l'utilisation de l'agrégat augmente, le risque de manquer d'espace augmente également si l'utilisation n'est pas surveillée de près. De plus, même si vous utilisez un système à 99 % de capacité, les performances risquent d'être moins élevées, mais vous devrez probablement interrompre la gestion pour l'empêcher de se remplir complètement lors de la commande de matériel supplémentaire. L'acquisition et l'installation de disques supplémentaires peuvent prendre un certain temps.

## Les agrégats HDD, y compris les agrégats Flash Pool



**NetApp recommande** un minimum de 15 % d'espace libre lorsque des disques rotatifs sont utilisés. Cela inclut tout l'espace inutilisé, y compris l'espace libre au sein de l'agrégat ou d'un volume, ainsi que tout espace libre alloué en raison de l'utilisation du provisionnement complet, mais qui n'est pas utilisé par les données réelles. Les performances seront affectées aux approches de la liberté d'expression de 10 %.

# Bases de données Oracle et machines virtuelles de stockage

La gestion du stockage des bases de données Oracle est centralisée sur un SVM (Storage Virtual machine)

Un SVM, connu sous le nom de vserver sur l'interface de ligne de commandes ONTAP, est une unité fonctionnelle de base du stockage. Il est utile de comparer un SVM à un invité sur un serveur VMware ESX.

Lors de l'installation initiale, ESX ne possède pas de fonctionnalités préconfigurées, telles que l'hébergement d'un système d'exploitation invité ou la prise en charge d'une application utilisateur. Il s'agit d'un conteneur vide jusqu'à ce qu'une machine virtuelle (VM) soit définie. ONTAP fonctionne de manière similaire : Lors de la première installation de ONTAP, aucune fonctionnalité de service des données n'est disponible tant qu'un SVM n'est pas créé. Pour configurer les services de données.

À l'instar des autres aspects de l'architecture de stockage, les meilleures options pour la conception des SVM et de l'interface logique (LIF) dépendent largement des exigences d'évolutivité et des besoins de l'entreprise.

## SVM

Il n'existe aucune bonne pratique officielle de provisionnement des SVM pour ONTAP. La bonne approche dépend des exigences en matière de gestion et de sécurité.

La plupart des clients utilisent un SVM principal pour la plupart de leurs besoins quotidiens, mais ils en créent

un petit pour des besoins particuliers. Par exemple, vous pouvez créer :

- SVM d'une base de données stratégique gérée par une équipe de spécialistes
- SVM pour un groupe de développement auquel un contrôle administratif complet a été attribué afin de pouvoir gérer leur propre stockage indépendamment
- SVM pour les données sensibles de l'entreprise, telles que les données de rapports financiers ou de ressources humaines, pour lesquelles l'équipe administrative doit être limitée

Dans un environnement de colocation, on peut attribuer à chaque locataire une SVM dédiée aux données. La limite du nombre de SVM et de LIF par cluster, paire HA et nœud dépend du protocole utilisé, du modèle de nœud et de la version de ONTAP. Consulter le "[NetApp Hardware Universe](#)" pour ces limites.

## Gestion des performances des bases de données Oracle avec la QoS ONTAP

Pour gérer efficacement et en toute sécurité plusieurs bases de données Oracle, il est nécessaire de disposer d'une stratégie de qualité de service efficace. C'est pourquoi les systèmes de stockage modernes offrent des performances toujours plus élevées.

Plus précisément, l'adoption croissante des systèmes de stockage 100 % Flash a permis de consolider les charges de travail. Les baies de stockage qui reposent sur des supports rotatifs ne prennent généralement en charge qu'un nombre limité de charges de travail exigeantes en E/S, car leurs capacités IOPS sont limitées par rapport aux anciens disques rotatifs. Une ou deux bases de données fortement actives satureraient les disques sous-jacents bien avant que les contrôleurs de stockage n'atteignent leurs limites. Cela a changé. Il est possible de saturer les contrôleurs de stockage les plus puissants, car le nombre de disques SSD requis est relativement faible. Cela signifie que vous pouvez exploiter pleinement les capacités des contrôleurs sans craindre un effondrement soudain des performances lors de pics de latence des supports rotatifs.

À titre d'exemple de référence, un simple système AFF A800 HA à deux nœuds est capable de traiter jusqu'à un million d'IOPS aléatoires avant que la latence ne dépasse la milliseconde. On pourrait s'attendre à ce que très peu de charges de travail atteignent de tels niveaux. L'utilisation optimale de cette baie AFF A800 implique l'hébergement de plusieurs workloads. Pour ce faire, la sécurité et la prévisibilité exigent des contrôles de QoS.

Il existe deux types de qualité de service (QoS) dans ONTAP : les IOPS et la bande passante. Les contrôles de QoS peuvent être appliqués aux SVM, volumes, LUN et fichiers.

### QoS des IOPS

Un contrôle de la QoS pour les IOPS est évidemment basé sur l'ensemble des IOPS d'une ressource donnée, mais il existe un certain nombre d'aspects de la QoS pour les IOPS qui peuvent ne pas être intuitifs. Au départ, quelques clients ont été surpris par l'augmentation apparente de la latence lorsqu'un seuil d'IOPS est atteint. L'augmentation de la latence est la conséquence naturelle de la limitation des IOPS. Logiquement, il fonctionne de la même manière qu'un système de jetons. Par exemple, si un volume donné contenant des fichiers de données dispose d'une limite de 10 000 IOPS, chaque E/S arrivant doit d'abord recevoir un jeton pour poursuivre le traitement. Tant que plus de 10 000 jetons n'ont pas été consommés en une seconde donnée, aucun retard n'est présent. Si les opérations d'E/S doivent attendre la réception de leur jeton, cet attente apparaît comme une latence supplémentaire. Plus une charge de travail est élevée, plus les E/S sont longues à attendre dans la file d'attente pour le traitement de son tour, ce qui apparaît comme une latence plus élevée.



Soyez prudent lorsque vous appliquez des contrôles QoS aux données des transactions de base de données/journaux de reprise. Alors que les demandes de performances liées à la journalisation de reprise sont généralement très élevées, bien inférieures à celles des fichiers de données, l'activité du journal de reprise est en rafales. L'E/S se produit en de brèves impulsions et une limite de QoS qui semble appropriée pour les niveaux d'E/S de reprise moyens peut être trop basse pour les exigences réelles. Cela peut entraîner de strictes limitations de performance en cas d'engagement de la QoS avec chaque pic de journal de reprise. En général, la journalisation des opérations de reprise et d'archivage ne doit pas être limitée par la QoS.

## QoS de la bande passante

Toutes les tailles d'E/S ne sont pas identiques. Par exemple, une base de données peut effectuer de nombreuses lectures de blocs de petite taille, ce qui entraînerait l'atteinte du seuil d'IOPS, mais il est également possible que les bases de données effectuent une analyse de table complète comprenant un très petit nombre de lectures de blocs volumineux, qui consomment une très grande quantité de bande passante, mais relativement peu d'IOPS.

De même, un environnement VMware peut générer un nombre très élevé d'IOPS aléatoires au démarrage, mais exécuter moins d'E/S, mais plus importantes, lors d'une sauvegarde externe.

Pour gérer efficacement les performances, les IOPS ou la bande passante doivent parfois être limitées, voire les deux.

## QoS minimale/garantie

De nombreux clients recherchent une solution incluant une QoS garantie, qui semble plus difficile à atteindre qu'elle ne le paraît et qui risque d'être très gaspillée. Par exemple, pour placer 10 bases de données avec une garantie de 10 000 IOPS, il est nécessaire de dimensionner un système dans le cas où les 10 bases de données s'exécutent simultanément à 10 000 IOPS, pour un total de 100 000.

La meilleure utilisation pour les contrôles QoS minimaux est de protéger les charges de travail stratégiques. Prenons l'exemple d'un contrôleur ONTAP avec un maximum de 500 000 IOPS et un mélange de charges de travail de production et de développement. Vous devez appliquer des règles de QoS maximales aux workloads de développement pour empêcher toute base de données de monopoliser le contrôleur. Vous appliqueriez ensuite des règles de QoS minimales aux charges de travail de production afin de vous assurer que les IOPS requises sont toujours disponibles, le cas échéant.

## La QoS adaptative

La QoS adaptative fait référence à la fonctionnalité ONTAP où la limite de QoS repose sur la capacité de l'objet de stockage. Elle est rarement utilisée avec les bases de données, car il n'existe généralement aucun lien entre la taille d'une base de données et ses exigences de performances. Les grandes bases de données peuvent être quasiment inertes, tandis que les bases de données plus petites peuvent être celles qui nécessitent le plus d'IOPS.

La QoS adaptative peut s'avérer très utile avec les datastores de virtualisation, car les exigences en IOPS de ces jeux de données ont tendance à être corrélées à la taille totale de la base de données. Un datastore plus récent contenant 1 To de fichiers VMDK devrait avoir besoin d'environ la moitié des performances pour un datastore de 2 To. La QoS adaptative vous permet d'augmenter automatiquement les limites de qualité de service lorsque le datastore est rempli de données.

# Bases de données Oracle et fonctionnalités d'efficacité ONTAP

Les fonctionnalités ONTAP d'optimisation de l'espace sont optimisées pour les bases de données Oracle. Dans la plupart des cas, la meilleure approche consiste à conserver les valeurs par défaut avec toutes les fonctionnalités d'efficacité activées.

Les fonctionnalités d'optimisation de l'espace, telles que la compression, la compaction et la déduplication, sont conçues pour augmenter la quantité de données logiques correspondant à un volume de stockage physique donné. Vous réduisez ainsi vos coûts et vos frais de gestion.

À un niveau élevé, la compression est un processus mathématique qui permet de détecter et d'encoder des modèles de données de manière à réduire les besoins en espace. En revanche, la déduplication détecte les blocs de données répétés et supprime les copies parasites. La compaction permet à plusieurs blocs logiques de données de partager le même bloc physique sur le support.



Reportez-vous aux sections ci-dessous sur le provisionnement fin pour une explication de l'interaction entre l'efficacité du stockage et la réservation fractionnaire.

## Compression

Avant la disponibilité des systèmes de stockage 100 % Flash, la compression basée sur les baies était d'une valeur limitée, car la plupart des charges de travail exigeantes en E/S nécessitaient un très grand nombre de piles pour obtenir une performance acceptable. Les systèmes de stockage contenaient invariablement beaucoup plus de capacité que nécessaire, ce qui a pour effet d'augmenter le nombre de disques. La situation a changé avec la montée du stockage Solid-State. Il n'est plus nécessaire de surprovisionner des disques uniquement pour obtenir de bonnes performances. L'espace disque d'un système de stockage peut être adapté aux besoins réels en termes de capacité.

La capacité accrue des disques SSD en termes d'IOPS permet presque toujours de réaliser des économies par rapport aux disques rotatifs. Toutefois, la compression peut réaliser davantage d'économies en augmentant la capacité effective des supports SSD.

Il existe plusieurs façons de compresser les données. De nombreuses bases de données incluent leurs propres fonctionnalités de compression, mais ce phénomène est rarement observé dans les environnements clients. La raison en est généralement la réduction des performances pour un **changement** de données compressées, plus avec certaines applications, il existe des coûts de licence élevés pour la compression au niveau de la base de données. Enfin, il y a les conséquences globales sur les performances des opérations des bases de données. Il est peu judicieux de payer un coût de licence par processeur élevé pour un processeur qui effectue la compression et la décompression des données plutôt que le véritable travail de base de données. Une meilleure option consiste à décharger la tâche de compression sur le système de stockage.

## Compression adaptative

La compression adaptative a été testée en profondeur avec des charges de travail exigeantes sans effet sur les performances, même dans un environnement 100 % Flash où la latence se mesure en microsecondes. Certains clients ont même signalé une augmentation des performances due à l'utilisation de la compression, car les données restent compressées dans le cache, augmentant ainsi la quantité de cache disponible dans un contrôleur.

ONTAP gère les blocs physiques dans des unités de 4 Ko. La compression adaptative utilise une taille de bloc

de compression par défaut de 8 Ko, ce qui signifie que les données sont compressées dans des unités de 8 Ko. La taille de bloc de 8 Ko la plus utilisée par les bases de données relationnelles est donc identique. Les algorithmes de compression deviennent plus efficaces avec la compression d'un volume croissant de données. Une taille de bloc de compression de 32 Ko serait plus compacte qu'une unité de bloc de compression de 8 Ko. Cela signifie que la compression adaptative utilisant une taille de bloc de 8 Ko par défaut entraîne des taux d'efficacité légèrement inférieurs, mais qu'une taille de bloc de compression inférieure présente également des avantages considérables. Les charges de travail de la base de données incluent une grande quantité d'activités de remplacement. Le remplacement d'un bloc de données de 32 Ko compressé de 8 Ko nécessite la lecture de l'intégralité des 32 Ko de données logiques, leur décompression, la mise à jour de la région de 8 Ko requise, la recompression, puis l'écriture de la totalité des 32 Ko sur les disques. Cette opération est très coûteuse pour un système de stockage. En effet, certaines baies de stockage concurrentes, basées sur des blocs de compression plus volumineux, affectent également considérablement les performances des charges de travail de la base de données.



La taille de bloc utilisée par la compression adaptative peut être augmentée jusqu'à 32 Ko. Cela peut améliorer l'efficacité du stockage et doit être envisagé pour les fichiers de repos tels que les journaux de transactions et les fichiers de sauvegarde lorsqu'une quantité importante de ces données est stockée sur la baie. Dans certains cas, les bases de données actives qui utilisent une taille de bloc de 16 ou 32 Ko peuvent également tirer parti de l'augmentation de la taille de bloc de la compression adaptative pour qu'elle corresponde. Consultez un représentant NetApp ou partenaire pour savoir si cette solution convient à votre charge de travail.



Les tailles de bloc de compression supérieures à 8 Ko ne doivent pas être utilisées avec la déduplication sur les destinations de sauvegarde en streaming. Les petites modifications apportées aux données sauvegardées affectent la fenêtre de compression de 32 Ko. Si la fenêtre change, les données compressées obtenues diffèrent dans l'ensemble du fichier. La déduplication a lieu après la compression, ce qui signifie que le moteur de déduplication voit chaque sauvegarde compressée différemment. Si la déduplication des sauvegardes en continu est nécessaire, seule une compression adaptative de bloc de 8 Ko doit être utilisée. Il est préférable d'utiliser la compression adaptative, car elle fonctionne à des blocs de taille réduite sans perturber l'efficacité de la déduplication. Pour des raisons similaires, la compression côté hôte interfère également avec l'efficacité de la déduplication.

## Alignement de compression

La compression adaptative dans un environnement de base de données nécessite un certain respect de l'alignement des blocs de compression. Cela ne préoccupe que les données soumises à des écrasements aléatoires de blocs très spécifiques. Cette approche est similaire à l'alignement global du système de fichiers, où le début d'un système de fichiers doit être aligné sur une limite de périphérique de 4 Ko et la taille de bloc d'un système de fichiers doit être un multiple de 4 Ko.

Par exemple, une écriture de 8 Ko dans un fichier est compressée uniquement si elle s'aligne sur une limite de 8 Ko dans le système de fichiers lui-même. Ce point signifie qu'il doit figurer sur le premier 8 Ko du fichier, le deuxième 8 Ko du fichier, etc. La manière la plus simple de garantir un alignement correct est d'utiliser le type de LUN correct, toute partition créée doit avoir un décalage par rapport au début du périphérique qui est un multiple de 8K, et utiliser une taille de bloc du système de fichiers qui est un multiple de la taille de bloc de la base de données.

Les données telles que les sauvegardes ou les journaux de transactions sont des opérations écrites de manière séquentielle sur plusieurs blocs, qui sont tous compressés. Par conséquent, il n'est pas nécessaire de considérer l'alignement. Le seul modèle d'E/S préoccupant est l'écrasement aléatoire des fichiers.



## Compaction

La compaction est une technologie qui améliore l'efficacité de la compression. Comme indiqué précédemment, la compression adaptative à elle seule permet d'économiser 2:1 au maximum, car elle se limite au stockage d'une E/S de 8 Ko dans un bloc WAFL de 4 Ko. Les méthodes de compression avec des blocs de taille supérieure améliorent l'efficacité. Cependant, elles ne conviennent pas aux données soumises à des remplacements de blocs de petite taille. La décompression d'unités de données de 32 Ko, la mise à jour d'une partie de 8 Ko, la recompression et l'écriture sur les disques entraînent une surcharge.

La compaction des données permet de stocker plusieurs blocs logiques dans des blocs physiques. Par exemple, une base de données avec des données fortement compressibles comme des blocs texte ou partiellement pleins peut être compressée de 8 Ko à 1 Ko. Sans compaction, 1 Ko de données occuperaient toujours un bloc complet de 4 Ko. La compaction des données à la volée permet de stocker 1 Ko de données compressées dans un espace physique de seulement 1 Ko, parallèlement à d'autres données compressées. Il ne s'agit pas d'une technologie de compression. Il s'agit simplement d'un moyen plus efficace d'allouer de l'espace sur les disques et, par conséquent, il ne doit pas créer d'effet détectable sur les performances.

Le degré d'économie obtenu varie. En général, les données déjà compressées ou chiffrées ne peuvent pas être compressées davantage et, par conséquent, la compaction de ces datasets ne peut pas être bénéfique. À contrario, les fichiers de données récemment initialisés ne contiennent qu'un petit peu plus que des métadonnées de bloc et des zéros compressent jusqu'à 80:1.

### Effacité du stockage sensible à la température

L'efficacité du stockage sensible à la température (TSSE) est disponible dans ONTAP 9.8 et versions ultérieures. Elle repose sur des cartes thermiques d'accès aux blocs pour identifier les blocs peu utilisés et les compresser avec une efficacité accrue.

## Déduplication

La déduplication permet de supprimer les tailles de bloc dupliquées d'un dataset. Par exemple, si le même bloc de 4 Ko existe dans 10 fichiers différents, la déduplication redirige ce bloc de 4 Ko au sein des 10 fichiers vers le même bloc physique de 4 Ko. Résultat : une amélioration de l'efficacité de ces données de 10:1.

Les données, telles que les LUN de démarrage invité VMware, se dédupliquent extrêmement bien, car elles sont constituées de plusieurs copies des mêmes fichiers du système d'exploitation. L'efficacité de 100:1 et plus ont été observées.

Certaines données ne contiennent pas de données dupliquées. Par exemple, un bloc Oracle contient un en-tête globalement unique à la base de données et une bande-annonce presque unique. Par conséquent, la déduplication d'une base de données Oracle permet rarement de réaliser plus de 1 % d'économies. La déduplication avec les bases de données MS SQL est légèrement meilleure, mais les métadonnées uniques au niveau des blocs restent une limitation.

Dans quelques cas, des économies d'espace allant jusqu'à 15 % ont été observées pour les bases de données de 16 Ko et les blocs volumineux. La bande de 4 Ko initiale de chaque bloc contient l'en-tête unique dans le monde, et le bloc de 4 Ko final contient la remorque presque unique. Les blocs internes sont candidats à la déduplication, bien que dans la pratique cela soit presque entièrement attribué à la déduplication des données mises à zéro.

De nombreuses baies concurrentes prétendent être capables de dédupliquer des bases de données en présumant qu'une base de données est copiée plusieurs fois. Il est également possible d'utiliser la déduplication NetApp, mais ONTAP offre une meilleure option : la technologie FlexClone de NetApp. Le résultat final est le même : plusieurs copies d'une base de données qui partagent la plupart des blocs physiques sous-jacents sont créées. L'utilisation de FlexClone est bien plus efficace que de prendre le temps

de copier les fichiers de base de données, puis de les dédupliquer. Il s'agit en effet de la non-duplication plutôt que de la déduplication, car un doublon n'est jamais créé à la première place.

## Efficacité et provisionnement fin

Les fonctions d'efficacité sont des formes de provisionnement fin. Par exemple, une LUN de 100 Go occupant un volume de 100 Go peut compresser à 50 Go. Aucune économie réelle n'est encore réalisée, car le volume est toujours de 100 Go. Le volume doit d'abord être réduit afin que l'espace économisé puisse être utilisé ailleurs sur le système. Si des modifications ultérieures de la LUN de 100 Go réduisent la taille des données compressibles, la LUN augmente et le volume pourrait se remplir.

Le provisionnement fin est fortement recommandé car il simplifie la gestion tout en améliorant la capacité exploitable avec les économies associées. La raison en est simple : les environnements de base de données comportent souvent beaucoup d'espace vide, un grand nombre de volumes et de LUN, ainsi que des données compressibles. Le provisionnement fin entraîne la réservation d'espace sur le stockage pour les volumes et les LUN au cas où un jour ils se traduiraient par une saturation de 100 % et contiendraient des données non compressibles à 100 %. Il est peu probable que cela se produise. Le provisionnement fin permet de récupérer et d'utiliser cet espace ailleurs. Il permet également de gérer la capacité en fonction du système de stockage lui-même, plutôt que de nombreux volumes et LUN plus petits.

Certains clients préfèrent utiliser le provisionnement lourd, soit pour des charges de travail spécifiques, soit généralement en fonction de pratiques opérationnelles et d'approvisionnement établies.

**Attention :** si un volume est configuré en mode lourd, il faut veiller à désactiver complètement toutes les fonctions d'efficacité de ce volume, y compris la décompression et la suppression de la déduplication à l'aide du `sis undo` commande. Le volume ne doit pas apparaître dans `volume efficiency show` sortie. Si c'est le cas, le volume est encore partiellement configuré pour les fonctions d'efficacité. Par conséquent, les garanties de remplacement fonctionnent différemment, ce qui augmente le risque que les dépassements de configuration entraînent un manque inattendu d'espace du volume, ce qui entraîne des erreurs d'E/S de la base de données.

## Meilleures pratiques en matière d'efficacité

Recommandation NetApp :

### AFF par défaut

Les volumes créés sur ONTAP et exécutés sur un système AFF 100 % Flash sont à allocation dynamique, avec l'activation de toutes les fonctionnalités d'efficacité à la volée. Bien que les bases de données ne bénéficient généralement pas de la déduplication et puissent inclure des données non compressibles, les paramètres par défaut conviennent néanmoins à la plupart des charges de travail. ONTAP est conçu pour traiter efficacement tous les types de données et de modèles d'E/S, qu'ils entraînent ou non des économies. Les valeurs par défaut ne doivent être modifiées que si les raisons sont parfaitement comprises et si un écart est bénéfique.

### Recommandations générales

- Si les volumes et/ou les LUN ne sont pas à provisionnement fin, vous devez désactiver tous les paramètres d'efficacité car l'utilisation de ces fonctionnalités n'offre aucune économie et la combinaison du provisionnement lourd et de l'optimisation de l'espace peut provoquer des comportements inattendus, notamment des erreurs de manque d'espace.
- Si les données ne sont pas sujettes à des écrasements, par exemple avec des sauvegardes ou des journaux de transactions de base de données, vous pouvez atteindre une meilleure efficacité en activant TSSE avec une période de refroidissement faible.

- Certains fichiers peuvent contenir une quantité importante de données non compressibles, par exemple lorsque la compression est déjà activée au niveau de l'application, les fichiers sont cryptés. Si l'un de ces scénarios est vrai, envisagez de désactiver la compression pour permettre un fonctionnement plus efficace sur d'autres volumes contenant des données compressibles.
- N'utilisez pas la compression et la déduplication de 32 Ko pour les sauvegardes de bases de données. Voir la section [Compression adaptative](#) pour plus d'informations.

## Provisionnement fin avec les bases de données Oracle

Le provisionnement fin pour une base de données Oracle nécessite une planification minutieuse, car il en résulte une configuration d'espace sur un système de stockage qui n'est pas nécessairement physiquement disponible. Cela vaut vraiment le coup, car une fois correctement effectué, il en résulte des économies considérables et des améliorations en termes de gestion.

Le provisionnement fin, de nombreuses formes, fait partie intégrante de nombreuses fonctionnalités offertes par ONTAP à l'environnement applicatif d'entreprise. Le provisionnement fin est également étroitement lié aux technologies d'efficacité pour la même raison : les fonctionnalités d'efficacité permettent de stocker davantage de données logiques que ce qui existe techniquement sur le système de stockage.

La plupart des snapshots impliquent un provisionnement fin. Par exemple, une base de données classique de 10 To sur un système de stockage NetApp compte environ 30 jours de copies Snapshot. Cet arrangement donne lieu à environ 10 To de données visibles dans le système de fichiers actif et 300 To dédiés aux snapshots. La capacité totale de stockage de 310 To réside généralement dans un espace d'environ 12 To à 15 To. La base de données active consomme 10 To et les 300 To de données restantes ne nécessitent que 2 à 5 To d'espace, car seules les modifications apportées aux données d'origine sont stockées.

Le clonage est également un exemple de provisionnement fin. Un client NetApp majeur a créé 40 clones d'une base de données de 80 To à utiliser pour le développement. Si les 40 développeurs qui utilisent ces clones surécrivent chaque bloc dans chaque fichier de données, plus de 3,2 po de stockage seraient nécessaires. En pratique, le chiffre d'affaires est faible et l'espace collectif requis est proche de 40 To, car seules les modifications sont stockées sur les disques.

### Gestion de l'espace

Le provisionnement fin d'un environnement applicatif doit être extrêmement prudent, car les taux de modification des données peuvent augmenter de manière inattendue. Par exemple, la consommation d'espace due aux snapshots peut augmenter rapidement si les tables de base de données sont réindexées ou si des correctifs à grande échelle sont appliqués aux invités VMware. Une sauvegarde mal placée peut écrire une grande quantité de données dans un délai très court. Enfin, il peut être difficile de restaurer certaines applications si un système de fichiers manque d'espace de façon inattendue.

Avec une configuration soignée de, ces risques peuvent être maîtrisés `volume-autogrow` et `snapshot-autodelete` règles. Comme leurs noms l'indiquent, ces options permettent de créer des règles qui effacent automatiquement l'espace consommé par les snapshots ou augmentent un volume pour prendre en charge des données supplémentaires. De nombreuses options sont disponibles et les besoins varient selon les clients.

Voir la "[documentation sur la gestion du stockage logique](#)" pour une discussion complète de ces fonctionnalités.

## Réservations fractionnaires

La réserve fractionnaire fait référence au comportement d'une LUN dans un volume en ce qui concerne l'efficacité de l'espace. Lorsque l'option `fractional-reserve` est défini sur 100 %, toutes les données du volume peuvent connaître un taux de rotation de 100 % avec n'importe quel modèle de données, sans épuiser l'espace sur le volume.

Par exemple, prenons l'exemple d'une base de données située sur une seule LUN de 250 Go dans un volume de 1 To. La création d'un snapshot entraînerait immédiatement la réservation d'un espace supplémentaire de 250 Go dans le volume, garantissant ainsi que l'espace disponible sur le volume ne serait pas insuffisant pour quelque raison que ce soit. L'utilisation de réserves fractionnaires est généralement inutile car il est très peu probable que chaque octet du volume de base de données ait besoin d'être écrasé. Il n'y a aucune raison de réserver de l'espace pour un événement qui ne se produit jamais. Cependant, si un client ne peut pas surveiller la consommation d'espace dans un système de stockage et doit être certain que l'espace ne sera jamais épuisé, des réservations fractionnaires de 100 % seront nécessaires pour utiliser les snapshots.

## Compression et déduplication

La compression et la déduplication sont deux formes de provisionnement fin. Par exemple, une empreinte des données de 50 To peut être compressée jusqu'à 30 To, ce qui permet d'économiser 20 To. Pour que la compression offre tous les avantages, il faut utiliser quelques 20 To pour d'autres données ou acheter le système de stockage avec moins de 50 To. Il en résulte une quantité de données stockées supérieure à ce qui n'est techniquement disponible sur le système de stockage. Du point de vue des données, il y a 50 To de données, même si celles-ci ne occupent que 30 To sur les disques.

Il est toujours possible que la compressibilité d'un dataset change, ce qui entraîne une consommation accrue de l'espace réel. Cette augmentation de la consommation signifie que la compression doit être gérée comme avec les autres formes de provisionnement fin en termes de surveillance et d'utilisation `volume-autogrow` et `snapshot-autodelete`.

La compression et la déduplication sont présentées plus en détail dans la section [xref:./oracle/efficiency.html](#)

## Compression et réservations fractionnaires

La compression est une forme d'allocation dynamique. Les réservations fractionnaires affectent l'utilisation de la compression, avec une remarque importante ; l'espace est réservé avant la création du snapshot. Normalement, la réserve fractionnaire n'est importante que si un instantané existe. S'il n'y a pas de snapshot, la réserve fractionnaire n'est pas importante. Ce n'est pas le cas avec la compression. Si une LUN est créée sur un volume avec compression, ONTAP conserve l'espace nécessaire pour prendre en charge un snapshot. Ce comportement peut être déroutant pendant la configuration, mais il est normal.

Prenons l'exemple d'un volume de 10 Go avec une LUN de 5 Go compressée à 2,5 Go sans copie Snapshot. Prenez en compte ces deux scénarios :

- La réserve fractionnaire = 100 entraîne une utilisation de 7,5 Go
- La réserve fractionnaire = 0 entraîne une utilisation de 2,5 Go

Le premier scénario comprend 2,5 Go de consommation d'espace pour les données actuelles et 5 Go d'espace pour représenter 100 % de chiffre d'affaires de la source en prévision de l'utilisation des snapshots. Le deuxième scénario ne réserve pas d'espace supplémentaire.

Bien que cette situation puisse sembler confuse, il est peu probable qu'elle soit rencontrée dans la pratique. La compression implique un provisionnement fin et le provisionnement fin dans un environnement LUN nécessite des réservations fractionnaires. Il est toujours possible d'écraser des données compressées par un élément

non compressible, ce qui signifie qu'un volume doit être à provisionnement fin pour la compression, pour réaliser des économies.

**NetApp recommande** les configurations de réserve suivantes :



- Réglez `fractional-reserve` à 0 lorsque la surveillance de la capacité de base est en place avec `volume-autogrow` et `snapshot-autodelete`.
- Réglez `fractional-reserve` à 100 s'il n'y a pas de capacité de surveillance ou s'il est impossible d'évacuer l'espace en quelque circonstance que ce soit.

## Allocation d'espace libre et d'espace LVM

L'efficacité du provisionnement fin des LUN actives dans un environnement de système de fichiers peut être perdue au fil du temps suite à la suppression des données. À moins que les données supprimées ne soient écrasées par des zéros (voir également "[ASMRU](#)") Ou l'espace est libéré avec la récupération d'espace TRIM/UNMAP, les données « effacées » occupent de plus en plus d'espace non alloué dans le système de fichiers. En outre, l'utilisation du provisionnement fin des LUN actives est limitée dans de nombreux environnements de base de données, car les fichiers de données sont initialisés sur leur taille complète au moment de la création.

Une planification minutieuse de la configuration de LVM peut améliorer l'efficacité et réduire les besoins en provisionnement du stockage et en redimensionnement des LUN. Lorsqu'un LVM tel que Veritas VxVM ou Oracle ASM est utilisé, les LUN sous-jacentes sont divisés en extensions qui ne sont utilisées que lorsque cela est nécessaire. Par exemple, si un dataset commence à 2 To mais peut atteindre 10 To au fil du temps, ce dataset peut être placé sur 10 To de LUN à provisionnement fin organisées dans un groupe de disques LVM. Elle occupant seulement 2 To d'espace au moment de la création et réclaire uniquement de l'espace supplémentaire, dans la mesure où les extensions sont allouées pour prendre en charge la croissance du volume des données. Ce processus est sûr tant que l'espace est surveillé.

## Basculement/basculement des bases de données Oracle et du contrôleur ONTAP

Il est nécessaire de bien comprendre les fonctions de basculement et de basculement du stockage pour s'assurer que les opérations de la base de données Oracle ne sont pas interrompues par ces opérations. En outre, les arguments utilisés par les opérations de basculement et de basculement peuvent affecter l'intégrité des données en cas d'utilisation incorrecte.

- Dans des conditions normales, les écritures entrantes sur un contrôleur donné sont mises en miroir de manière synchrone sur son partenaire. Dans un environnement NetApp MetroCluster, les écritures sont également mises en miroir sur un contrôleur distant. Tant qu'une écriture n'est pas stockée sur un support non volatile dans tous les emplacements, elle n'est pas validée par l'application hôte.
- Le support qui stocke les données d'écriture est appelé mémoire non volatile ou NVMEM. Elle est également parfois appelée mémoire NVRAM, et peut être considérée comme un cache d'écriture, même si elle fonctionne comme un journal. En fonctionnement normal, les données de NVMEM ne sont pas lues ; elles sont uniquement utilisées pour protéger les données en cas de défaillance logicielle ou matérielle. Lors de l'écriture des données sur les disques, les données sont transférées de la mémoire RAM du système, et non de NVMEM.
- Lors d'une opération de basculement, un nœud d'une paire haute disponibilité reprend les opérations de son partenaire. Un basculement est quasiment identique, mais s'applique aux configurations MetroCluster

dans lesquelles un nœud distant prend le relais par rapport à un nœud local.

Lors des opérations de maintenance de routine, un basculement du stockage ou un basculement doivent être transparents, sauf en cas de brève pause potentielle dans les opérations en cas de changement des chemins réseau. La mise en réseau peut toutefois être complexe et il est facile d'y faire des erreurs. NetApp recommande donc de tester minutieusement les opérations de basculement et de basculement avant de mettre en production un système de stockage. C'est la seule façon de s'assurer que tous les chemins réseau sont correctement configurés. Dans un environnement SAN, vérifiez soigneusement le résultat de la commande `sanlun lun show -p` pour vous assurer que tous les chemins principaux et secondaires attendus sont disponibles.

Il convient de faire attention lors d'un basculement forcé ou d'un basculement forcé. Forcer une modification de la configuration du stockage avec ces options signifie que l'état du contrôleur propriétaire des disques est ignoré et que le nœud alternatif prend le contrôle des disques. Une force de basculement incorrecte peut entraîner une perte ou une corruption des données. En effet, un basculement forcé ou un basculement forcé peut rejeter le contenu de la NVMEM. Une fois le basculement ou le basculement effectué, la perte de ces données signifie que les données stockées sur les disques peuvent revenir à un état plus ancien du point de vue de la base de données.

Un basculement forcé avec une paire haute disponibilité normale devrait rarement être nécessaire. Dans la plupart des scénarios de défaillance, un nœud s'arrête et informe le partenaire qu'un basculement automatique a lieu. Il existe certains cas à la périphérie, par exemple une panne de déploiement où l'interconnexion entre les nœuds est perdue puis un contrôleur est perdu, dans lequel un basculement forcé est nécessaire. Dans ce cas, la mise en miroir entre les nœuds est perdue avant la panne du contrôleur, ce qui signifie que le contrôleur survivant n'aurait plus de copie des écritures en cours. Le basculement doit ensuite être forcé, ce qui signifie que des données peuvent être perdues.

La même logique s'applique à un basculement MetroCluster. Dans des conditions normales, le basculement est presque transparent. Toutefois, un incident peut entraîner une perte de connectivité entre le site survivant et le site de reprise sur incident. Du point de vue du site survivant, le problème ne pourrait être rien de plus qu'une interruption de la connectivité entre les sites, et le site d'origine pourrait encore traiter les données. Si un nœud ne peut pas vérifier l'état du contrôleur principal, seul un basculement forcé est possible.

**NetApp recommande** de prendre les précautions suivantes :



- Veillez à ne pas forcer accidentellement un basculement ou un basculement. En règle générale, il n'est pas nécessaire de forcer et le fait de forcer la modification peut entraîner la perte de données.
- Si un basculement ou un basculement forcé s'avère nécessaire, assurez-vous que les applications sont arrêtées, que tous les systèmes de fichiers sont démontés et que les groupes de volumes LVM (Logical Volume Manager) sont proposés en mode Variyoffed. Les groupes de disques ASM doivent être démontés.
- En cas de basculement forcé du MetroCluster, vous pouvez isoler le nœud défaillant de toutes les ressources de stockage restantes. Pour plus d'informations, consultez le Guide de gestion et de reprise sur incident de MetroCluster correspondant à la version appropriée de ONTAP.

## MetroCluster et plusieurs agrégats

MetroCluster est une technologie de réplication synchrone qui passe en mode asynchrone en cas d'interruption de la connectivité. Cette demande est la plus courante de la part des clients, car une réplication synchrone garantie signifie que l'interruption de la connectivité du site entraîne un blocage complet des E/S de la base de données, ce qui la met hors service.

Avec MetroCluster, les agrégats sont rapidement resynchronisés une fois la connectivité restaurée. Contrairement à d'autres technologies de stockage, MetroCluster ne devrait jamais nécessiter de mise en miroir complète après une panne de site. Seules les modifications delta doivent être expédiées.

Dans les jeux de données qui couvrent les agrégats, le risque est faible de nécessiter des étapes supplémentaires de restauration des données en cas de sinistre en cas de déploiement. En particulier, si (a) la connectivité entre les sites est interrompue, (b) la connectivité est restaurée, (c) les agrégats atteignent un état dans lequel certains sont synchronisés et d'autres ne le sont pas, puis (d) le site primaire est perdu, le site survivant dans lequel les agrégats ne sont pas synchronisés. Dans ce cas, une partie du dataset est synchronisée et il est impossible d'ouvrir des applications, des bases de données ou des datastores sans restauration. Si un dataset compte plusieurs agrégats, NetApp recommande vivement d'utiliser des sauvegardes basées sur des snapshots avec l'un des nombreux outils disponibles pour vérifier la restauration rapide dans ce scénario inhabituel.

## Informations sur le copyright

Copyright © 2024 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTEUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

## Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.