



## **SAN FC**

### Enterprise applications

NetApp  
May 19, 2024

# Sommaire

- SAN FC ..... 1
  - Alignement des LUN pour les E/S de la base de données Oracle ..... 1
  - Dimensionnement des LUN de la base de données Oracle et nombre de LUN ..... 2
  - Placement des LUN de la base de données Oracle ..... 3
  - Redimensionnement des LUN de la base de données Oracle et redimensionnement basé sur LVM ..... 4
  - Répartition LVM avec les bases de données Oracle ..... 5

# SAN FC

## Alignement des LUN pour les E/S de la base de données Oracle

L'alignement des LUN fait référence à l'optimisation des E/S par rapport à la disposition du système de fichiers sous-jacent.

Sur un système ONTAP, le stockage est organisé en unités de 4 Ko. Un bloc de 8 Ko de base de données ou de système de fichiers doit être mappé à exactement deux blocs de 4 Ko. Si une erreur de configuration de LUN déplace l'alignement de 1 Ko dans les deux sens, chaque bloc de 8 Ko existerait sur trois blocs de stockage de 4 Ko différents au lieu de deux. Cette configuration entraîne une augmentation de la latence et des E/S supplémentaires au sein du système de stockage.

L'alignement affecte également les architectures LVM. Si un volume physique au sein d'un groupe de volumes logiques est défini sur l'unité entière (aucune partition n'est créée), le premier bloc de 4 Ko de la LUN s'aligne sur le premier bloc de 4 Ko du système de stockage. Il s'agit d'un alignement correct. Des problèmes surviennent avec les partitions car elles déplacent l'emplacement de départ où le système d'exploitation utilise le LUN. Tant que le décalage est décalé en unités entières de 4 Ko, la LUN est alignée.

Dans les environnements Linux, créez des groupes de volumes logiques sur l'ensemble de l'unité de disque. Lorsqu'une partition est requise, vérifiez l'alignement en exécutant `fdisk -u` et vérifiez que le début de chaque partition est un multiple de huit. Cela signifie que la partition démarre à un multiple de huit secteurs de 512 octets, soit 4 Ko.

Voir également la discussion sur l'alignement des blocs de compression dans la section ["Efficacité"](#). Toute disposition alignée avec les limites des blocs de compression de 8 Ko est également alignée avec les limites de 4 Ko.

### Avertissements de mauvais alignement

La journalisation des opérations de reprise et des transactions de la base de données génère normalement des E/S non alignées qui peuvent entraîner des avertissements erronés concernant les LUN mal alignées sur ONTAP.

La journalisation effectue une écriture séquentielle du fichier journal avec des écritures de taille variable. Une opération d'écriture de journal qui ne s'aligne pas sur les limites de 4 Ko ne provoque généralement pas de problèmes de performances, car l'opération d'écriture de journal suivante termine le bloc. ONTAP est ainsi en mesure de traiter la quasi-totalité des écritures sous forme de blocs complets de 4 Ko, même si les données de blocs de 4 Ko ont été écrites dans deux opérations distinctes.

Vérifiez l'alignement à l'aide d'utilitaires tels que `sio` ou `dd` Qui peuvent générer des E/S à une taille de bloc définie. Les statistiques d'alignement des E/S sur le système de stockage peuvent être affichées à l'aide du `stats` commande. Voir ["Vérification de l'alignement WAFL"](#) pour en savoir plus.

L'alignement dans les environnements Solaris est plus compliqué. Reportez-vous à la section ["Configuration de l'hôte SAN ONTAP"](#) pour en savoir plus.

## Avertissement

Dans les environnements Solaris x86, prenez davantage soin de l'alignement approprié car la plupart des configurations comportent plusieurs couches de partitions. Les tranches de partition Solaris x86 existent généralement au-dessus d'une table de partition d'enregistrement d'amorçage maître standard.

# Dimensionnement des LUN de la base de données Oracle et nombre de LUN

Il est essentiel de sélectionner la taille de LUN optimale et le nombre de LUN à utiliser pour optimiser les performances et la gestion des bases de données Oracle.

Une LUN est un objet virtualisé sur ONTAP qui existe sur tous les disques de l'agrégat d'hébergement. Par conséquent, les performances de la LUN ne sont pas affectées par sa taille, car la LUN exploite tout le potentiel de performance de l'agrégat, quelle que soit sa taille.

À titre de commodité, les clients peuvent souhaiter utiliser un LUN de taille spécifique. Par exemple, si une base de données est construite sur un groupe de disques LVM ou Oracle ASM composé de deux LUN de 1 To chacune, ce groupe de disques doit être développé par incréments de 1 To. Il peut être préférable de créer le groupe de disques à partir de huit LUN de 500 Go chacune, de sorte que le groupe de disques puisse être augmenté par incréments plus petits.

Il n'est pas recommandé d'établir une taille de LUN standard universelle, car cela peut compliquer la gestion. Par exemple, une taille de LUN standard de 100 Go peut fonctionner correctement lorsqu'une base de données ou un datastore se situe entre 1 et 2 To, mais qu'une base de données ou un datastore de 20 To nécessite 200 LUN. Cela signifie que les délais de redémarrage du serveur sont plus longs, que les différents utilisateurs doivent gérer davantage d'objets et que des produits tels que SnapCenter doivent effectuer des recherches sur de nombreux objets. L'utilisation d'un nombre inférieur de LUN de plus grande taille permet d'éviter de tels problèmes.

- Le nombre de LUN est plus important que la taille de LUN.
- La taille de LUN est principalement contrôlée par les exigences liées au nombre de LUN.
- Évitez de créer plus de LUN que nécessaire.

## Nombre de LUN

Contrairement à la taille de LUN, le nombre de LUN affecte les performances. La performance des applications dépend souvent de la capacité à réaliser des E/S parallèles via la couche SCSI. Ainsi, deux LUN offrent de meilleures performances qu'une seule LUN. L'utilisation d'un LVM tel que Veritas VxVM, Linux LVM2 ou Oracle ASM est la méthode la plus simple pour augmenter le parallélisme.

Les clients NetApp n'ont généralement pas eu l'avantage d'augmenter le nombre de LUN au-delà de seize. Toutefois, le test d'environnements 100 % SSD avec des E/S aléatoires très lourdes a permis d'améliorer encore jusqu'à 64 LUN.

**NetApp recommande** ce qui suit :



En général, de quatre à seize LUN suffisent pour prendre en charge les besoins en E/S d'une charge de travail de base de données donnée. Moins de quatre LUN peuvent créer des limites de performances en raison de limites dans les implémentations SCSI hôte.

# Placement des LUN de la base de données Oracle

Le placement optimal des LUN de base de données dans les volumes ONTAP dépend principalement de l'utilisation des différentes fonctionnalités ONTAP.

## Volumes

L'un des points de confusion les plus courants avec les nouveaux clients ONTAP est l'utilisation des volumes FlexVol, communément appelés simplement « volumes ».

Un volume n'est pas une LUN. Ces termes sont utilisés de façon synonymie avec de nombreux autres produits de fournisseurs, y compris les fournisseurs de cloud. Les volumes ONTAP sont des conteneurs de gestion simples. Ils ne fournissent pas les données en eux-mêmes, ni n'occupent l'espace. Il s'agit de conteneurs pour les fichiers et les LUN. Ils permettent d'améliorer et de simplifier la gestion, notamment à grande échelle.

## Volumes et LUN

Les LUN associées sont généralement situées en colocation dans un seul volume. Par exemple, une base de données qui nécessite 10 LUN doit généralement avoir les 10 LUN placées sur le même volume.



- L'utilisation d'un rapport LUN/volumes de 1:1, c'est-à-dire une LUN par volume, n'est **pas** une bonne pratique formelle.
- À la place, les volumes doivent être considérés comme des conteneurs pour les charges de travail ou les datasets. Il peut y avoir une seule LUN par volume ou il peut y en avoir plusieurs. La bonne réponse dépend des exigences de gestion.
- La diffusion des LUN sur un nombre inutile de volumes peut entraîner une surcharge supplémentaire et des problèmes de planification pour des opérations telles que les opérations de snapshot, un nombre excessif d'objets affichés dans l'interface utilisateur et entraîner l'atteinte des limites de volume de la plate-forme avant que la limite de LUN ne soit atteinte.

## Volumes, LUN et snapshots

Les règles et planifications Snapshot sont placées sur le volume, et non sur la LUN. Un jeu de données composé de 10 LUN ne nécessite qu'une seule règle de snapshot lorsque ces LUN sont co-localisées dans le même volume.

En outre, la colocation de toutes les LUN associées à un jeu de données donné dans un seul volume permet d'effectuer des opérations de snapshot atomiques. Par exemple, une base de données résidant sur 10 LUN ou un environnement d'application VMware comprenant 10 systèmes d'exploitation différents peut être protégé comme un objet unique et cohérent si les LUN sous-jacentes sont tous placés sur un seul volume. S'ils sont placés sur des volumes différents, les snapshots peuvent être synchronisés à 100 %, même s'ils sont programmés en même temps.

Dans certains cas, il peut être nécessaire de diviser un jeu de LUN associé en deux volumes différents en raison des exigences de restauration. Par exemple, une base de données peut contenir quatre LUN pour les fichiers de données et deux LUN pour les journaux. Dans ce cas, un volume de fichiers de données avec 4 LUN et un volume de journaux avec 2 LUN peuvent être la meilleure option. La raison en est une capacité de restauration indépendante. Par exemple, le volume des fichiers de données peut être restauré de manière sélective à un état antérieur, ce qui signifie que les quatre LUN seraient rétablies à l'état du snapshot, tandis que le volume du journal contenant ses données stratégiques ne serait pas affecté.

## Volumes, LUN et SnapMirror

Les règles et opérations SnapMirror sont, tout comme les opérations Snapshot, exécutées sur le volume, et non sur la LUN.

La colocation de LUN associées dans un seul volume vous permet de créer une relation SnapMirror unique et de mettre à jour toutes les données qu'elle contient en une seule mise à jour. Comme pour les instantanés, la mise à jour sera également une opération atomique. La destination SnapMirror dispose d'une réplique instantanée unique des LUN source. Si les LUN ont été réparties sur plusieurs volumes, les répliques peuvent être cohérentes les unes avec les autres.

## Volumes, LUN et QoS

S'il est possible d'appliquer la QoS de manière sélective à chaque LUN, il est généralement plus facile de la configurer au niveau du volume. Par exemple, toutes les LUN utilisées par les invités dans un serveur ESX donné peuvent être placées sur un seul volume, puis une règle de qualité de service adaptative de ONTAP peut être appliquée. Vous obtenez ainsi une limite d'IOPS par To qui s'applique à toutes les LUN.

De même, si une base de données nécessitait 100 000 IOPS et occupait 10 LUN, il serait plus facile de définir une seule limite de 100 000 IOPS sur un seul volume que de définir 10 limites individuelles de 10 000 IOPS, une sur chaque LUN.

## Dispositions multi-volumes

Dans certains cas, la distribution de LUN sur plusieurs volumes peut être avantageuse. La principale raison est la répartition des contrôleurs. Par exemple, un système de stockage haute disponibilité peut héberger une base de données unique dans laquelle chaque contrôleur a besoin du potentiel de traitement et de mise en cache complet. Dans ce cas, la conception type consisterait à placer la moitié des LUN dans un seul volume sur le contrôleur 1 et l'autre moitié des LUN dans un seul volume sur le contrôleur 2.

De même, la répartition des contrôleurs peut être utilisée pour l'équilibrage de la charge. Un système haute disponibilité hébergeant 100 bases de données de 10 LUN chacune peut être conçu où chaque base de données reçoit un volume de 5 LUN sur chacun des deux contrôleurs. Il en résulte une charge symétrique garantie de chaque contrôleur au fur et à mesure que des bases de données supplémentaires sont provisionnées.

Cependant, aucun de ces exemples ne correspond à un ratio volume/LUN de 1:1. L'objectif reste d'optimiser la gestion en co-localisant les LUN associées dans les volumes.

Par exemple, la conteneurisation est un rapport LUN/volume 1:1. Chaque LUN peut représenter une seule charge de travail et doit être gérée individuellement. Dans ce cas, un rapport de 1:1 peut être optimal.

## Redimensionnement des LUN de la base de données Oracle et redimensionnement basé sur LVM

Lorsqu'un système de fichiers SAN a atteint sa limite de capacité, il existe deux options pour augmenter l'espace disponible :

- Augmentez la taille des LUN
- Ajoutez une LUN à un groupe de volumes existant et développez le volume logique contenu

Bien que le redimensionnement des LUN soit une option d'augmentation de la capacité, il est généralement

préférable d'utiliser un LVM, y compris Oracle ASM. L'une des principales raisons pour lesquelles les LVM existent est d'éviter la nécessité d'un redimensionnement des LUN. Avec une LVM, plusieurs LUN sont reliées entre elles dans un pool de stockage virtuel. Les volumes logiques extraits de ce pool sont gérés par le LVM et peuvent être facilement redimensionnés. Il est également possible d'éviter les points sensibles sur un disque en distribuant un volume logique donné à tous les LUN disponibles. Une migration transparente peut généralement être effectuée à l'aide du gestionnaire de volumes pour déplacer les extensions sous-jacentes d'un volume logique vers de nouvelles LUN.

## Répartition LVM avec les bases de données Oracle

La répartition des LVM consiste à distribuer les données entre plusieurs LUN. Les performances de nombreuses bases de données en sont ainsi considérablement améliorées.

Avant l'ère des disques Flash, la répartition était utilisée pour surmonter les limites de performances des disques rotatifs. Par exemple, si un système d'exploitation doit effectuer une opération de lecture de 1 Mo, la lecture de ce 1 Mo de données à partir d'un seul disque demande beaucoup de tête de lecture lorsque le transfert des 1 Mo est lent. Si ce 1 Mo de données a été réparti sur 8 LUN, le système d'exploitation pourrait exécuter huit opérations de lecture de 128 K en parallèle et réduire le temps nécessaire au transfert de 1 Mo.

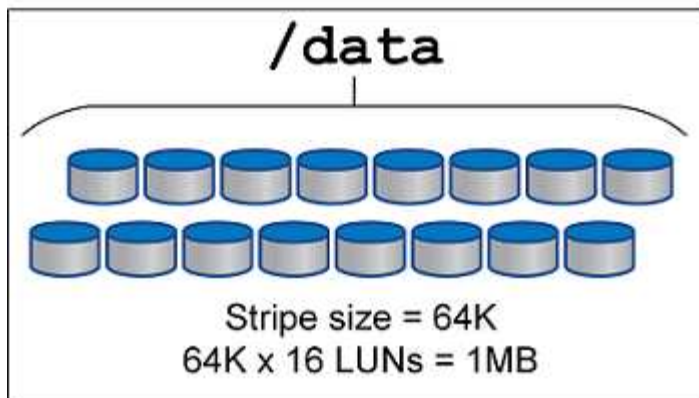
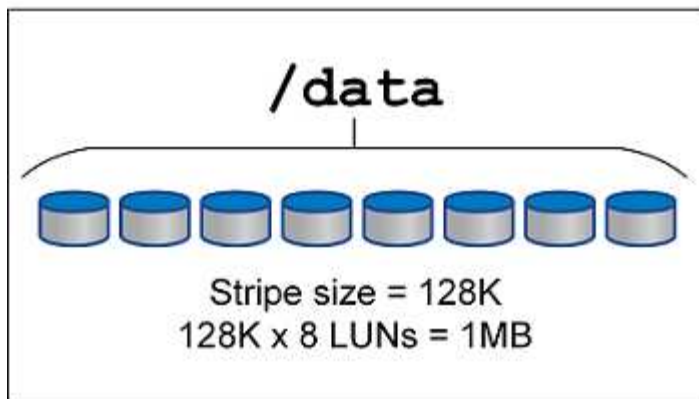
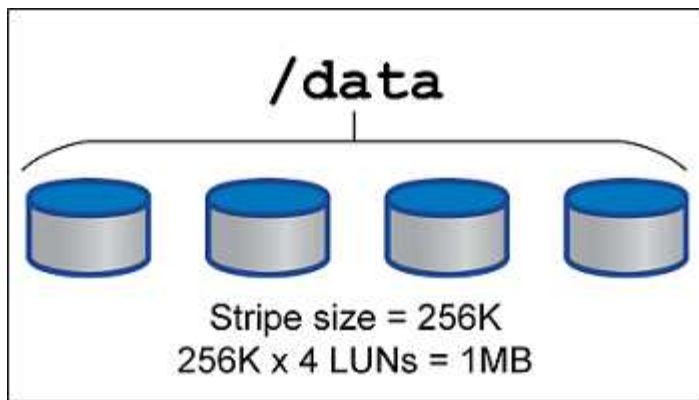
Le striping avec des disques rotatifs était plus difficile, car le modèle d'E/S devait être connu à l'avance. Si la répartition n'a pas été correctement réglée pour les véritables modèles d'E/S, les configurations à bandes risquent d'endommager les performances. Avec les bases de données Oracle, et en particulier les configurations 100 % Flash, le striping est beaucoup plus facile à configurer et a fait ses preuves pour améliorer considérablement les performances.

Par défaut, les gestionnaires de volumes logiques, tels que la bande Oracle ASM, ne le font pas pour le système d'exploitation natif LVM. Certaines lient plusieurs LUN ensemble en tant que périphérique concaténé. Résultat : des fichiers de données existent sur un seul périphérique LUN. Ceci provoque des points chauds. Les autres implémentations LVM prennent par défaut en charge les extensions distribuées. Cette méthode est similaire à la répartition, mais elle est plus grossière. Les LUN du groupe de volumes sont tranchées en grandes parties, appelées extensions et généralement mesurées en plusieurs mégaoctets. Ensuite, les volumes logiques sont distribués sur ces extensions. Il en résulte des E/S aléatoires sur un fichier qui doit être bien réparti entre les LUN, mais les opérations d'E/S séquentielles ne sont pas aussi efficaces qu'elles pourraient l'être.

Les E/S des applications exigeantes en performances sont presque toujours de (a) en unités de taille de bloc de base ou (b) d'un mégaoctet.

L'objectif principal d'une configuration à bandes est de s'assurer que les E/S de fichier unique peuvent être exécutées comme une seule unité, et que les E/S de plusieurs blocs, d'une taille de 1 Mo, peuvent être parallélisées de façon homogène sur toutes les LUN du volume réparti. Cela signifie que la taille de bande ne doit pas être inférieure à la taille du bloc de base de données, et que la taille de bande multipliée par le nombre de LUN doit être de 1 Mo.

La figure suivante présente trois options possibles pour le réglage de la taille et de la largeur des bandes. Le nombre de LUN est sélectionné pour répondre aux exigences de performances comme décrit ci-dessus, mais dans tous les cas, le total des données dans une seule bande est de 1 Mo.





## 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 DÉLICTELLE (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.