



# **Configuration de stockage sous-jacente**

## **Enterprise applications**

NetApp  
May 19, 2024

# Sommaire

- Configuration de stockage sous-jacente ..... 1
  - Bases de données PostgreSQL avec systèmes de fichiers NFS ..... 1
  - PostgreSQL avec systèmes de fichiers SAN ..... 3

# Configuration de stockage sous-jacente

## Bases de données PostgreSQL avec systèmes de fichiers NFS

Les bases de données PostgreSQL peuvent être hébergées sur les systèmes de fichiers NFSv3 ou NFSv4. La meilleure option dépend de facteurs extérieurs à la base de données.

Par exemple, le comportement de verrouillage NFSv4 peut être préférable dans certains environnements en cluster. (Voir ["ici"](#) pour plus d'informations)

Dans le cas contraire, les fonctionnalités de la base de données doivent être proches des mêmes, y compris les performances. La seule exigence est l'utilisation du `hard` option de montage. Ceci est nécessaire pour garantir que les délais d'expiration ne produisent pas d'erreurs d'E/S irrécupérables.

Si NFSv4 est choisi en tant que protocole, NetApp recommande d'utiliser NFSv4.1. Certaines améliorations fonctionnelles du protocole NFSv4 dans NFSv4.1 améliorent la résilience sur NFSv4.0.

Utilisez les options de montage suivantes pour les charges de travail de base de données générales :

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsize=65536,wsiz=65536
```

Si des E/S séquentielles lourdes sont attendues, la taille du transfert NFS peut être augmentée comme décrit dans la section suivante.

### Tailles de transfert NFS

Par défaut, ONTAP limite la taille des E/S NFS à 64 Ko.

Les E/S aléatoires utilisent la plupart des applications et bases de données une taille de bloc bien inférieure à la taille maximale de 64 Ko. Les E/S de blocs volumineux sont généralement parallélisées de sorte que le maximum de 64 Ko ne limite pas non plus l'obtention d'une bande passante maximale.

Dans certains cas, le maximum de 64 000 charges de travail entraîne une limitation. En particulier, les opérations à thread unique, telles que les opérations de sauvegarde ou de restauration, ou encore les analyses de table complète de base de données s'exécutent plus rapidement et plus efficacement si la base de données peut exécuter moins d'E/S, mais plus volumineuses. La taille optimale de gestion des E/S pour ONTAP est de 256 Ko.

La taille maximale de transfert pour un SVM ONTAP donné peut être modifiée comme suit :

```
Cluster01::> set advanced
Warning: These advanced commands are potentially dangerous; use them only
when directed to do so by NetApp personnel.
Do you want to continue? {y|n}: y
Cluster01::*> nfs server modify -vserver vserver1 -tcp-max-xfer-size
262144
Cluster01::*>
```

### Avertissement

Ne réduisez jamais la taille de transfert maximale autorisée sur ONTAP en dessous de la valeur de rsize/wsize des systèmes de fichiers NFS actuellement montés. Cela peut provoquer des blocages ou même une corruption des données avec certains systèmes d'exploitation. Par exemple, si les clients NFS sont actuellement définis sur une taille rsize/wsize de 65536, la taille maximale du transfert ONTAP peut être ajustée entre 65536 et 1048576 sans effet car les clients eux-mêmes sont limités. Réduire la taille de transfert maximale en dessous de 65536 peut endommager la disponibilité ou les données.

Une fois la taille de transfert augmentée au niveau ONTAP, les options de montage suivantes sont utilisées :

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsiz=262144,wsiz=262144
```

## Tables d'emplacements TCP NFSv3

Si NFSv3 est utilisé avec Linux, il est essentiel de définir correctement les tables d'emplacements TCP.

Les tables d'emplacements TCP sont l'équivalent NFSv3 de la profondeur de file d'attente de l'adaptateur de bus hôte (HBA). Ces tableaux contrôlent le nombre d'opérations NFS qui peuvent être en attente à la fois. La valeur par défaut est généralement 16, un chiffre bien trop faible pour assurer des performances optimales. Le problème inverse se produit sur les noyaux Linux plus récents : la limite de la table des emplacements TCP augmente automatiquement par envoi de demandes, jusqu'à atteindre le niveau de saturation du serveur NFS.

Pour des performances optimales et pour éviter les problèmes de performances, ajustez les paramètres du noyau qui contrôlent les tables d'emplacements TCP.

Exécutez le `sysctl -a | grep tcp.*.slot_table` et observez les paramètres suivants :

```
# sysctl -a | grep tcp.*.slot_table
sunrpc.tcp_max_slot_table_entries = 128
sunrpc.tcp_slot_table_entries = 128
```

Tous les systèmes Linux doivent inclure `sunrpc.tcp_slot_table_entries`, mais seulement certains incluent `sunrpc.tcp_max_slot_table_entries`. Ils doivent tous deux être réglés sur 128.

## Avertissement

Si vous ne définissez pas ces paramètres, vous risquez d'avoir des effets importants sur les performances. Dans certains cas, les performances sont limitées car le système d'exploitation linux n'émet pas suffisamment d'E/S. Dans d'autres cas, les latences d'E/S augmentent à mesure que le système d'exploitation linux tente d'émettre plus d'E/S que ce qui peut être traité.

# PostgreSQL avec systèmes de fichiers SAN

Les bases de données PostgreSQL avec SAN sont généralement hébergées sur des systèmes de fichiers xfs, mais d'autres peuvent être utilisées si elles sont prises en charge par le fournisseur du système d'exploitation

Même si un seul LUN peut généralement prendre en charge jusqu'à 100 000 IOPS, les bases de données exigeantes en E/S nécessitent généralement l'utilisation de LVM avec répartition.

## Marquage LVM

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.

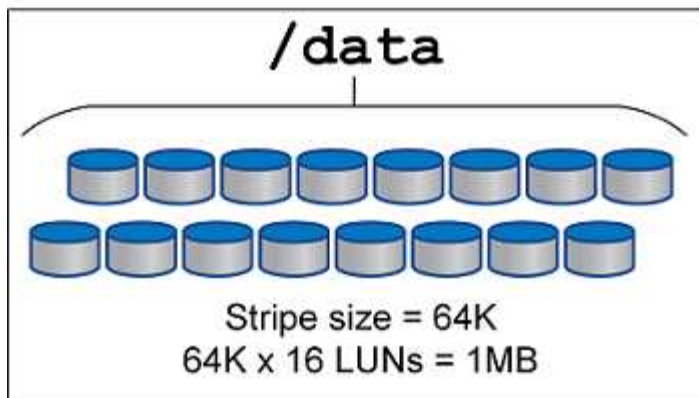
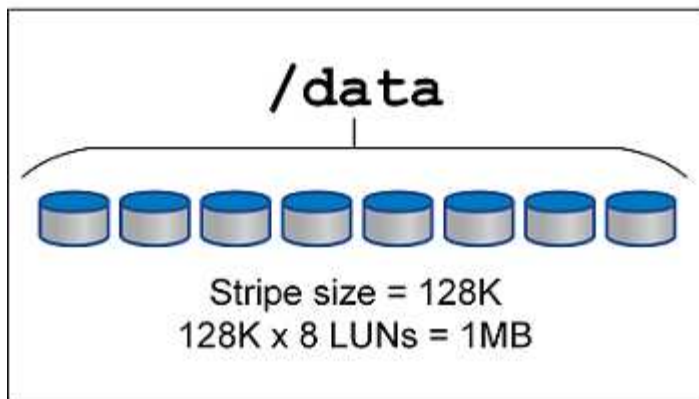
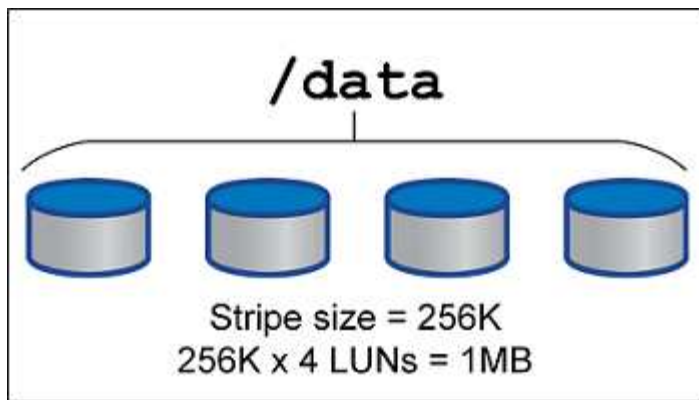
Le striping avec des disques rotatifs était plus difficile, car le modèle d'E/S devait être connu à l'avance. Si la répartition n'a pas été correctement réglée pour les véritables modèles d'E/S, les configurations à bandes risquent d'endommager les performances. Avec les bases de données Oracle, et en 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.