



Configuration de la base de données

Enterprise applications

NetApp
May 19, 2024

Sommaire

- Configuration de la base de données. 1
 - Tailles des blocs de base de données Oracle 1
 - Paramètres de la base de données Oracle : db_file_multibloc_read_count 1
 - Paramètres de la base de données Oracle : filesystemio_options 2
 - Délais d'expiration Oracle RAC (Real application clusters) 3

Configuration de la base de données

Tailles des blocs de base de données Oracle

ONTAP utilise en interne une taille de bloc variable, ce qui signifie que les bases de données Oracle peuvent être configurées avec n'importe quelle taille de bloc. Cependant, la taille des blocs du système de fichiers peut affecter les performances et, dans certains cas, une taille de bloc de reprise supérieure peut améliorer les performances.

Tailles des blocs de fichiers de données

Certains systèmes d'exploitation offrent un choix de tailles de blocs de système de fichiers. Pour les systèmes de fichiers prenant en charge les fichiers de données Oracle, la taille de bloc doit être de 8 Ko lorsque la compression est utilisée. Lorsque la compression n'est pas requise, vous pouvez utiliser une taille de bloc de 8 Ko ou 4 Ko.

Si un fichier de données est placé sur un système de fichiers avec un bloc de 512 octets, des fichiers mal alignés sont possibles. Il est possible que le LUN et le système de fichiers soient correctement alignés en fonction des recommandations de NetApp, mais les E/S de fichier sont mal alignées. Un tel mauvais alignement entraînerait de graves problèmes de performances.

Les systèmes de fichiers prenant en charge les journaux de reprise doivent utiliser une taille de bloc qui représente un multiple de la taille de bloc de reprise. Cela nécessite généralement que le système de fichiers redo log et le fichier redo log lui-même utilisent une taille de bloc de 512 octets.

Rétablir les tailles des blocs

Avec des taux de reprise très élevés, il est possible que des tailles de bloc de 4 Ko soient plus performantes, car les taux de reprise élevés permettent d'exécuter les E/S en moins d'opérations et de manière plus efficace. Si les taux de reprise sont supérieurs à 50 Mbit/s, envisagez de tester une taille de bloc de 4 Ko.

Quelques problèmes clients ont été identifiés avec les bases de données à l'aide de journaux de reprise avec une taille de bloc de 512 octets sur un système de fichiers d'une taille de bloc de 4 Ko et de nombreuses transactions très petites. La surcharge liée à l'application de plusieurs modifications de 512 octets à un seul bloc du système de fichiers de 4 Ko a entraîné des problèmes de performances qui ont été résolus en changeant le système de fichiers pour qu'il utilise une taille de bloc de 512 octets.



NetApp vous recommande de ne pas modifier la taille du bloc de reprise, sauf si un service client ou un service professionnel vous en informe ou si le changement est basé sur la documentation officielle du produit.

Paramètres de la base de données Oracle : `db_file_multibloc_read_count`

Le `db_file_multiblock_read_count` Paramètre contrôle le nombre maximal de blocs de base de données Oracle lus par Oracle au cours d'une opération, pendant les E/S séquentielles

Toutefois, ce paramètre n'affecte pas le nombre de blocs lus par Oracle au cours des opérations de lecture, ni

le nombre d'E/S aléatoires Seule la taille de bloc des E/S séquentielles est affectée.

Oracle recommande à l'utilisateur de ne pas définir ce paramètre. Cela permet au logiciel de base de données de définir automatiquement la valeur optimale. Cela signifie généralement que ce paramètre est défini sur une valeur qui produit une taille d'E/S de 1 Mo. Par exemple, une lecture de 1 Mo de blocs de 8 Ko nécessite la lecture de 128 blocs. La valeur par défaut de ce paramètre est donc 128.

La plupart des problèmes de performance de base de données observés par NetApp sur les sites des clients provenaient de paramètres incorrects. Des raisons valides ont été données pour modifier cette valeur avec les versions 8 et 9 d'Oracle. Par conséquent, le paramètre peut être présent sans le savoir dans `init.ora` Fichiers car la base de données a été mise à niveau vers Oracle 10 et versions ultérieures. La configuration héritée de 8 ou 16, par rapport à la valeur par défaut 128, nuit de manière significative aux performances d'E/S séquentielles.



NetApp recommande de régler le `db_file_multiblock_read_count` le paramètre ne doit pas être présent dans le `init.ora` fichier. NetApp n'a jamais observé d'amélioration des performances suite à la modification de ce paramètre, mais le débit d'E/S séquentielles subit une importante dégradation dans de nombreux cas.

Paramètres de la base de données Oracle : `filesystemio_options`

Le paramètre d'initialisation Oracle `filesystemio_options` Contrôle l'utilisation des E/S asynchrones et directes

Contrairement à une idée reçue, ces deux types d'E/S ne s'excluent pas mutuellement. NetApp a observé que ce paramètre est souvent mal configuré dans les environnements des clients. Cette configuration incorrecte est la cause directe de nombreux problèmes de performances.

Les E/S asynchrones offrent la possibilité de paralléliser les opérations Oracle d'E/S. Avant la disponibilité des E/S asynchrones sur différents systèmes d'exploitation, les utilisateurs ont configuré de nombreux processus `dbwriter` et modifié la configuration du processus serveur. Avec les E/S asynchrones, le système d'exploitation lui-même exécute les E/S en parallèle pour le compte du logiciel de base de données. Ce processus ne présente aucun risque pour les données et les opérations critiques, telles que la journalisation de reprise Oracle, sont toujours exécutées de manière synchrone.

Les E/S directes contournent le cache du tampon du système d'exploitation. Sur un système UNIX, les E/S transitent normalement par le cache du tampon du système d'exploitation. Ceci est utile pour les applications qui ne maintiennent pas de cache interne, mais Oracle dispose de son propre cache de tampon dans la SGA. Dans la plupart des cas, il est préférable d'activer les E/S directes et d'allouer la RAM du serveur à la mémoire SGA plutôt que d'utiliser le cache du tampon du système d'exploitation. La SGA exploite la mémoire plus efficacement. En outre, lors de leur transit via le tampon du se, les E/S sont soumises à un traitement supplémentaire, ce qui augmente les latences. Cette augmentation est particulièrement visible lors des E/S intenses en écriture, pour lesquelles la faible latence est primordiale.

Les options pour `filesystemio_options` sont :

- **Async.** Oracle soumet des demandes d'E/S au système d'exploitation pour traitement. Ce qui lui permet d'effectuer d'autres tâches plutôt que d'attendre la fin des E/S et d'augmenter ainsi la parallélisation des E/S.
- **Directio.** Oracle effectue des E/S directement par rapport aux fichiers physiques plutôt que de router les E/S via le cache du système d'exploitation hôte.

- **None.** Oracle utilise des E/S synchrones et mises en tampon Dans cette configuration, le choix entre les processus serveur partagés et dédiés et le nombre de dbwriter est plus important.
- **Setall.** Oracle utilise des E/S asynchrones et directes Dans presque tous les cas, l'utilisation de `setall` est optimale.



Le `filesystemio_options` Ce paramètre n'a aucun effet dans les environnements dNFS et ASM. Dans ces environnements, les E/S asynchrones et directes sont automatiquement utilisées

Certains clients ont déjà rencontré des problèmes d'E/S asynchrones, notamment avec les versions précédentes de Red Hat Enterprise Linux 4 (RHEL4). Certains conseils obsolètes sur Internet suggèrent toujours d'éviter les E/S asynchrones en raison d'informations obsolètes. Les E/S asynchrones sont stables sur tous les systèmes d'exploitation actuels. Il n'y a aucune raison de le désactiver, en l'absence d'un bug connu avec le système d'exploitation.

Si une base de données utilise des E/S mises en tampon, un switch vers des E/S directes peut également justifier une modification de la taille de la mémoire SGA. La désactivation des E/S mises en tampon élimine le gain de performance fourni par le cache du se hôte pour la base de données. L'ajout de RAM à la SGA résout ce problème. Et devrait améliorer les performances nettes d'E/S.

Bien qu'il soit presque toujours préférable d'utiliser la RAM pour la SGA d'Oracle plutôt que pour le cache du tampon du système d'exploitation, il peut s'avérer impossible de déterminer ce qui est le plus avantageux. Par exemple, il est parfois préférable d'utiliser des E/S mises en tampon avec une mémoire SGA de très petite taille sur un serveur de base de données comportant de nombreuses instances Oracle actives par intermittence. Cette configuration permet à toutes les instances de base de données en cours d'exécution d'utiliser de manière flexible la RAM restante sur le système d'exploitation. Cette situation est très inhabituelle, mais elle a été observée sur certains sites clients.



NetApp recommande le réglage `filesystemio_options` à `setall`, Mais notez que dans certains cas, la perte du cache du tampon hôte peut nécessiter une augmentation de la SGA d'Oracle.

Délais d'expiration Oracle RAC (Real application clusters)

Oracle RAC est un produit clusterware qui comporte plusieurs types de processus de pulsation internes qui contrôlent l'intégrité du cluster.



Les informations dans le "[misscount](#)" La section contient des informations essentielles pour les environnements RAC Oracle utilisant un stockage en réseau. Dans la plupart des cas, les paramètres RAC Oracle par défaut devront être modifiés pour garantir que le cluster RAC résiste aux modifications de chemin réseau et aux opérations de basculement/basculement du stockage.

disktimeout

Le paramètre RAC principal lié au stockage est `disktimeout`. Ce paramètre contrôle le seuil au sein duquel les E/S du fichier de vote doivent être terminées. Si le `disktimeout` Le paramètre est dépassé, puis le nœud RAC est supprimé du cluster. La valeur par défaut de ce paramètre est 200. Cette valeur doit être suffisante pour les procédures standard de Takeover et and Giveback du stockage.

NetApp recommande fortement de tester soigneusement les configurations RAC avant de les mettre en

production, car de nombreux facteurs affectent un basculement ou un rétablissement. Outre le temps nécessaire au basculement du stockage, la propagation des modifications du protocole LACP (Link Aggregation Control Protocol) nécessite également du temps supplémentaire. En outre, le logiciel de chemins d'accès multiples SAN doit détecter un délai d'expiration d'E/S et réessayer sur un autre chemin. Si une base de données est extrêmement active, une grande quantité d'E/S doit être mise en file d'attente et relancée avant le traitement des E/S du disque de vote.

En l'absence d'un basculement ou d'un retour de stockage réel, l'effet peut être simulé à l'aide de tests de câble Pull sur le serveur de base de données.

NetApp recommande ce qui suit :



- En quittant le `disktimeout` paramètre à la valeur par défaut de 200.
- Testez toujours soigneusement une configuration RAC.

misscount

Le `misscount` Le paramètre affecte normalement uniquement la pulsation réseau entre les nœuds RAC. La valeur par défaut est 30 secondes. Si les binaires de la grille se trouvent sur une matrice de stockage ou si le disque d'amorçage du système d'exploitation n'est pas local, ce paramètre peut devenir important. Cela inclut les hôtes avec des lecteurs de démarrage situés sur un SAN FC, les systèmes d'exploitation démarrés par NFS et les lecteurs de démarrage situés sur les datastores de virtualisation, tels qu'un fichier VMDK.

Si l'accès à un disque de démarrage est interrompu par un basculement ou un rétablissement du stockage, il est possible que l'emplacement binaire de la grille ou l'ensemble du système d'exploitation soit temporairement bloqué. Le temps nécessaire à ONTAP pour terminer l'opération de stockage et au système d'exploitation pour changer les chemins et reprendre les E/S peut être supérieur à `misscount` seuil. Par conséquent, un nœud est immédiatement supprimé une fois la connectivité à la LUN de démarrage ou aux binaires de la grille restaurée. Dans la plupart des cas, l'exclusion et le redémarrage qui s'ensuit se produisent sans message de journalisation indiquant la raison du redémarrage. Toutes les configurations ne sont pas affectées. Testez donc tout hôte de démarrage SAN, de démarrage NFS ou basé sur un datastore dans un environnement RAC afin que RAC reste stable si la communication avec le lecteur de démarrage est interrompue.

Dans le cas de lecteurs de démarrage non locaux ou d'un système de fichiers non local hébergeant `grid` binaires, le `misscount` devra être modifié pour correspondre `disktimeout`. Si ce paramètre est modifié, effectuez des tests supplémentaires pour identifier également les effets sur le comportement du RAC, tels que le temps de basculement du nœud.

NetApp recommande ce qui suit :

- Quittez le `misscount` paramètre à la valeur par défaut de 30, sauf si l'une des conditions suivantes s'applique :
 - `grid` Les fichiers binaires sont situés sur un disque connecté au réseau, y compris les disques basés sur NFS, iSCSI, FC et les datastores.
 - Le système d'exploitation est démarré sur un SAN.
- Dans de tels cas, évaluez l'effet des interruptions de réseau qui affectent l'accès au système d'exploitation ou `GRID_HOME` systèmes de fichiers. Dans certains cas, de telles interruptions provoquent le blocage des démons RAC Oracle, ce qui peut conduire à un `misscount`délai d'expiration et suppression basés sur`. Le délai par défaut est de 27 secondes, soit la valeur de ``misscount moins reboottime`. Dans de tels cas, augmenter `misscount` à 200 pour correspondre `disktimeout`.



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.