



データベース設定

Enterprise applications

NetApp
April 29, 2024

目次

データベース設定.....	1
Oracleデータベースのブロックサイズ	1
Oracleデータベースパラメータ：db_file_multiblock_read_count.....	1
Oracleデータベースパラメータ:filesystemio_options	2
Oracle Real Application Clusters (RAC) タイムアウト.....	3

データベース設定

Oracleデータベースのブロックサイズ

ONTAPでは内部的に可変ブロックサイズが使用されるため、Oracleデータベースには任意のブロックサイズを設定できます。ただし、ファイルシステムのブロックサイズがパフォーマンスに影響することがあり、場合によってはRedoブロックサイズを大きくすることでパフォーマンスが向上することがあります。

データファイルのブロックサイズ

OSによっては、ファイルシステムのブロックサイズを選択できます。Oracleデータファイルをサポートするファイルシステムでは、圧縮を使用する場合にブロックサイズを8KBにする必要があります。圧縮が不要な場合は、8KBまたは4KBのブロックサイズを使用できます。

512バイトのブロックを使用するファイルシステムにデータファイルが配置されていると、ファイルのミスアライメントが発生する可能性があります。LUNとファイルシステムはNetAppの推奨事項に基づいて適切にアライメントされていても、ファイルI/Oはミスアライメントされます。このようなミスアライメントが発生すると、原因で重大なパフォーマンス問題が発生します。

Redoログをサポートするファイルシステムでは、Redoブロックサイズの倍数のブロックサイズを使用する必要があります。そのためには、通常、RedoログファイルシステムとRedoログ自体の両方で512バイトのブロックサイズを使用する必要があります。

Redoブロックサイズ

Redo率が非常に高い場合は、少ない処理で効率よくI/Oを実行できるため、4KBのブロックサイズの方がパフォーマンスが向上する可能性があります。Redo率が50MBpsを超える場合は、4KBのブロックサイズをテストすることを検討してください。

一部のお客様では、ブロックサイズが512バイトのRedoログをブロックサイズが4KBのファイルシステムで使用し、非常に小さなトランザクションが大量に発生するという問題が報告されています。このパフォーマンスの問題は、複数の512バイトの変更を4KBの単一のファイルシステムブロックに適用する際のオーバーヘッドが原因で発生していましたが、ファイルシステムのブロックサイズを512バイトに変更することで解決されました。



* NetAppでは、関連するカスタマーサポートまたはプロフェッショナルサービス部門から指示があった場合、または正式な製品ドキュメントに基づく場合を除き、Redoブロックサイズを変更しないことを推奨しています*。

Oracleデータベースパラメータ : `db_file_multiblock_read_count`

。 `db_file_multiblock_read_count` パラメータは、シーケンシャルI/OでOracleが単一の処理として読み取るOracleデータベースブロックの最大数を制御します。

ただし、このパラメータは、すべての読み取り処理でOracleが読み取るブロック数には影響しません。また、

ランダムI/Oにも影響しません。影響を受けるのはシーケンシャルI/Oのブロックサイズだけです。

Oracleでは、このパラメータを未設定のままにしておくことを推奨しています。これにより、データベースソフトウェアは自動的に最適な値を設定できます。つまり、このパラメータは通常、I/Oサイズが1MBになる値に設定されます。たとえば、8KBブロックの1MB読み取りでは128ブロックを読み取る必要があるため、このパラメータのデフォルト値は128になります。

顧客サイトのNetAppで発生したデータベースパフォーマンスの問題のほとんどには、このパラメータの設定が誤っています。Oracleバージョン8および9では、この値を変更する正当な理由があります。そのため、パラメータが `init.ora` ファイル：データベースがOracle 10以降にアップグレードされたため。従来の設定である8または16をデフォルト値の128と比較すると、シーケンシャルI/Oのパフォーマンスが大幅に低下します。



* NetApp推奨*設定 `db_file_multiblock_read_count` パラメータがに存在してはなりません。 `init.ora` ファイル。NetAppでは、このパラメータを変更することでパフォーマンスが向上するという状況は発生していませんが、シーケンシャルI/Oのスループットに明らかな影響を及ぼすケースは少なくありません。

Oracleデータベースパラメータ: `filesystemio_options`

Oracle初期化パラメータ `filesystemio_options` 非同期I/OとダイレクトI/Oの使用を制御します。

一般的な考え方に反して、非同期I/OとダイレクトI/Oは相互に排他的ではありません。NetAppでは、お客様の環境でこのパラメータの設定ミスが頻繁に発生し、この設定ミスが多くのパフォーマンス問題の直接的な原因となっていることを確認しています。

非同期I/Oとは、Oracle I/O処理を並行処理できることを意味します。さまざまなOSで非同期I/Oが使用可能になる前は、ユーザが多数のdbwriterプロセスを設定し、サーバプロセスの設定を変更していました。非同期I/Oでは、OS自体がデータベースソフトウェアに代わって効率的かつ並列的にI/Oを実行します。このプロセスによってデータがリスクにさらされることはなく、OracleのRedoロギングなどの重要な処理も同期的に実行されます。

ダイレクトI/OはOSのバッファキャッシュをバイパスします。UNIXシステムのI/Oは、通常、OSのバッファキャッシュを通過します。これは内部キャッシュを持たないアプリケーションでは便利ですが、OracleはSGA内に独自のバッファキャッシュを備えています。ほとんどの場合、OSのバッファキャッシュを使用するよりも、ダイレクトI/Oを有効にしてサーバRAMをSGAに割り当てる方が適しています。Oracle SGAはメモリをより効率的に使用します。さらに、I/OがOSバッファを通過すると、追加の処理が発生し、レイテンシが増加します。レイテンシの増加は、低レイテンシが重要な要件である書き込みI/Oの負荷が高い場合に特に顕著です。

オフシヨン `filesystemio_options` 次のとおりです。

- * `async`。* OracleはI/O要求をOSに送信して処理します。このプロセスにより、OracleはI/Oの完了を待たずに他の処理を実行できるため、I/Oの並列化が促進されます。
- * `directio`。* Oracleは、ホストOSキャッシュを介してI/Oをルーティングするのではなく、物理ファイルに対して直接I/Oを実行します。
- なし。Oracleは同期I/OとバッファI/Oを使用します。この構成では、共有サーバプロセスと専用サーバプロセスの選択、およびdbwriterの数がより重要になります。
- * `SETALL`。* Oracleは非同期I/OとダイレクトI/Oの両方を使用します。ほとんどすべての場合、`setall` 最適です。



。 `filesystemio_options` パラメータは、DNFS環境とASM環境では効果がありません。DNFSまたはASMを使用すると、自動的に非同期I/OとダイレクトI/Oの両方が使用されません。

一部のお客様では、特に以前のRed Hat Enterprise Linux 4 (RHEL4) リリースで、過去に非同期I/Oの問題が発生していました。インターネット上のいくつかの時代遅れのアドバイスは、時代遅れの情報のために非同期I/Oを避けることを提案しています。非同期I/Oは、現在のすべてのOSで安定しています。OSの既知のバグがない限り、無効にする理由はありません。

データベースでバッファI/Oが使用されている場合は、ダイレクトI/Oに切り替えてもSGAサイズの変更が必要になることがあります。バッファI/Oを無効にすると、ホストOSキャッシュがデータベースに提供するパフォーマンス上のメリットがなくなります。RAMをSGAに再度追加すると、この問題が解決します。最終的には、I/Oパフォーマンスの向上につながります。

RAMはOSのバッファキャッシュよりもOracle SGAに使用の方がほとんどですが、最適な値を特定できない場合もあります。たとえば、断続的にアクティブになるOracleインスタンスが多数あるデータベースサーバでは、SGAサイズが非常に小さいバッファI/Oを使用することを推奨します。この方法では、実行中のすべてのデータベースインスタンスが、空いているOSのRAMを柔軟に使用できます。これは非常にまれな状況ですが、一部のお客様のサイトで確認されています。



* NetApp推奨*設定 `filesystemio_options` 終了: ``setall``ただし、状況によっては、ホストのバッファキャッシュが失われた場合にOracle SGAの拡張が必要になることがあります。

Oracle Real Application Clusters (RAC) タイムアウト

Oracle RACは、クラスタの健全性を監視する複数のタイプの内部ハートビートプロセスを備えたクラスタウェア製品です。



の情報 "[MissCount](#)" セクションには、ネットワーク・ストレージを使用するOracle RAC環境に関する重要な情報が記載されています。多くの場合、RACクラスタがネットワーク・パスの変更やストレージのフェイルオーバー/スイッチオーバー操作に耐えられるように、デフォルトのOracle RAC設定を変更する必要があります。

ディスクタイムアウト

プライマリストレージ関連のRACパラメータは、`disktimeout`です。このパラメータは、投票ファイルI/Oが完了しなければならないしきい値を制御します。状況に応じて `disktimeout` パラメータの値を超えると、そのRACノードがクラスタから削除されます。このパラメータのデフォルトは200です。ストレージのテイクオーバーとギブバックの標準的な手順では、この値で十分です。

テイクオーバーやギブバックには多くの要素が影響するため、NetAppでは、RAC構成を本番環境に導入する前に徹底的にテストすることを強く推奨します。ストレージフェイルオーバーの完了に必要な時間に加えて、Link Aggregation Control Protocol (LACP; リンクアグリゲーション制御プロトコル) の変更が伝播されるまでの時間も長くなります。また、SANマルチパスソフトウェアはI/Oタイムアウトを検出し、代替パスで再試行する必要があります。データベースが非常にアクティブな場合は、投票ディスクI/Oが処理される前に、大量のI/Oをキューに入れて再試行する必要があります。

ストレージのテイクオーバーやギブバックを実際に実行できない場合は、データベースサーバでケーブルを取り外すテストを実行して影響をシミュレートできます。



- NetAppの推奨事項*：
- を終了します。 `disktimeout` パラメータを指定します。デフォルト値は200です。
- RAC構成は常に十分にテストしてください。

MissCount

。 `misscount` パラメータは通常、RACノード間のネットワークハートビートにのみ影響します。デフォルトは30秒です。Gridバイナリがストレージレイ上にある場合やOSのブートドライブがローカルでない場合は、このパラメータが重要になることがあります。これには、ブートドライブがFC SANに配置されたホスト、NFSブートOS、およびVMDKファイルなどの仮想データストアに配置されたブートドライブが含まれます。

ブートドライブへのアクセスがストレージのテイクオーバーやギブバックによって中断された場合、Gridバイナリの場所またはOS全体が一時的に停止する可能性があります。ONTAPがストレージ処理を完了するのに必要な時間、およびOSがパスを変更してI/Oを再開するのに必要な時間が、 `misscount` しきい値。そのため、ブートLUNまたはGridバイナリへの接続がリストアされたあと、ノードはただちに削除されます。ほとんどの場合、削除とその後のリブートは実行されますが、リブートの理由を示すログメッセージは表示されません。すべての構成に影響するわけではないので、RAC環境内のSANブート、NFSブート、またはデータストアベースのホストをテストして、ブートドライブへの通信が中断してもRACが安定した状態になるようにします。

ローカルでないブートドライブまたはローカルでないファイルシステムをホストしている場合 `grid` バイナリ、 `misscount` 一致するように変更する必要があります `disktimeout`。このパラメータを変更した場合は、さらにテストを行い、ノードのフェイルオーバー時間など、RACの動作への影響を特定します。



- NetAppの推奨事項*：
- そのままにします。 `misscount` 次のいずれかの条件が適用されない場合は、デフォルト値の30のパラメータを使用します。
 - `grid` バイナリが、ネットワークに接続されたドライブ（NFS、iSCSI、FC、データストアベースのドライブを含む）に配置されている。
 - OSがSANブートである。
- このような場合は、ネットワークの中断がOSへのアクセスに影響するか、 `GRID_HOME` ファイルシステム：このような中断によって原因Oracle RACデーモンが停止し、 `misscount`-ベースのタイムアウトおよび削除。タイムアウトのデフォルトは27秒です。これは `misscount` マイナス `reboottime`。このような場合、 `misscount 200` にして一致させる `disktimeout`。

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用权を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用权については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。