



NFS

Enterprise applications

NetApp
January 12, 2026

目次

NFS	1
概要	1
NFSハアクション	1
Linux NFSv3 TCPスロットテーブル	1
ADRとNFS	2
nfs-rootonlyおよびmount-rootonly	2
NFSエクスポートポリシー：superuserとsetuid	3
NFSv4 / 4.1構成	3
Oracle Direct NFS（dNFS）	5
Direct NFS	5
Direct NFSとホストファイルシステムへのアクセス	8
NFSのリースとロック	9
NFSv4のステートフル	9
NFSv4ロック	9
NFSv4リース	9
NFSv4猶予期間	11
リースタイムアウトと猶予期間	12
NFSキャッシング	13
NFSテンソウサイズ	13

NFS

概要

NetAppは30年以上にわたってエンタープライズクラスのNFSストレージを提供してきましたが、クラウドベースのインフラへの移行が進むにつれ、その使用がますます増えています。その理由は、シンプルさです。

NFSプロトコルには、要件が異なる複数のバージョンが含まれています。ONTAPを使用したNFSの完全な概要構成については、を参照してください。 ["TR-4067『NFS on ONTAP Best Practices』"](#)。次のセクションでは、より重要な要件と一般的なユーザーエラーについて説明します。

NFSハアシヨン

オペレーティングシステムNFSクライアントがNetAppでサポートされている必要があります。

- NFSv3は、NFSv3標準に準拠したOSでサポートされます。
- Oracle dNFSクライアントではNFSv3がサポートされています。
- NFSv4は、NFSv4標準に準拠するすべてのOSでサポートされます。
- NFSv4.1およびNFSv4.2では、特定のOSのサポートが必要です。を参照してください ["NetApp IMT"](#) サポートされているOSの場合。
- NFSv4.1でのOracle dNFSのサポートには、Oracle 12.2.0.2以降が必要です。



。 ["NetAppのサポートマトリックス"](#) NFSv3およびNFSv4の場合、特定のオペレーティングシステムは含まれません。RFCに準拠するすべてのOSが一般的にサポートされています。オンラインのIMTでNFSv3またはNFSv4のサポートを検索する場合は、該当するOSが表示されないため、特定のOSを選択しないでください。すべてのOSは、一般ポリシーで暗黙的にサポートされています。

Linux NFSv3 TCPスロットテーブル

TCPスロットテーブルは、NFSv3でホストバスアダプタ（HBA）のキュー深度に相当します。一度に未処理となることのできるNFS処理の数を制御します。デフォルト値は通常16ですが、最適なパフォーマンスを得るには小さすぎます。逆に、新しいLinuxカーネルでTCPスロットテーブルの上限をNFSサーバが要求でいっぱいになるレベルに自動的に引き上げることができるため、問題が発生します。

パフォーマンスを最適化し、パフォーマンスの問題を回避するには、TCPスロットテーブルを制御するカーネルパラメータを調整します。

を実行します `sysctl -a | grep tcp.*.slot_table` コマンドを実行し、次のパラメータを確認します。

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

すべてのLinuxシステムに `sunrpc.tcp_slot_table_entries` `ただし、次のようなものがあります。
``sunrpc.tcp_max_slot_table_entries`。どちらも128に設定する必要があります。



これらのパラメータを設定しないと、パフォーマンスに大きく影響する可能性があります。Linux OSが十分なI/Oを発行していないためにパフォーマンスが制限される場合もあります。一方では、Linux OSが問題で処理できる以上のI/Oを試行すると、I/Oレイテンシが増加します。

ADRとNFS

一部のお客様から、内のデータに対する過剰なI/Oが原因でパフォーマンスの問題が発生すると報告されています。ADR 場所。通常、この問題は、大量のパフォーマンスデータが蓄積されるまで発生しません。過剰なI/Oの理由は不明ですが、Oracleプロセスがターゲットディレクトリを繰り返しスキャンして変更を求めたことが原因と考えられます。

の取り外し `noac` および / または `actimeo=0` マウントオプションを使用すると、ホストOSのキャッシュを実行してストレージI/Oレベルを削減できます。



* NetApp推奨* 配置しないこと ADR ファイルシステムノテエタ `noac` または `actimeo=0` パフォーマンスの問題が発生する可能性があるからです。分離 ADR データを別のマウントポイントに格納（必要な場合）

nfs-rootonlyおよびmount-rootonly

ONTAPには、という名前のNFSオプションがあります。 `nfs-rootonly` これにより、サーバが高ポートからのNFSトラフィック接続を受け入れるかどうかを制御されます。セキュリティ対策として、1024未満の送信元ポートを使用してTCP/IP接続を開くことができるのはrootユーザだけです。これは、このようなポートは通常、ユーザプロセスではなくOS用に予約されているためです。この制限により、NFSトラフィックが実際のオペレーティングシステムNFSクライアントからのものであり、NFSクライアントをエミュレートする悪意のあるプロセスではないことを確認できます。Oracle dNFSクライアントはユーザスペースドライバですが、このプロセスはrootとして実行されるため、通常は `nfs-rootonly`。接続は低ポートから行われます。

。 `mount-rootonly` オプションは環境NFSv3のみです。1024より大きいポートからRPCマウント呼び出しを受け入れるかどうかを制御します。dNFSを使用すると、クライアントは再びルートとして実行されるため、1024未満のポートを開くことができます。このパラメータは効果がありません。

NFSバージョン4.0以降でdNFSを使用して接続を開いているプロセスは、rootとして実行されないため、1024以上のポートが必要です。。 `nfs-rootonly` dNFSが接続を完了するには、パラメータを `disabled` に設定する必要があります。

状況 `nfs-rootonly` が有効になっている場合、マウントフェーズでdNFS接続を開いているときにハングします。 `sqlplus` の出力は次のようになります。

```
SQL>startup
ORACLE instance started.
Total System Global Area 4294963272 bytes
Fixed Size                  8904776 bytes
Variable Size               822083584 bytes
Database Buffers           3456106496 bytes
Redo Buffers                7868416 bytes
```

パラメータは次のように変更できます。

```
Cluster01::> nfs server modify -nfs-rootonly disabled
```



まれに、nfs-rootonlyとmount-rootonlyの両方をdisabledに変更しなければならないことがあります。サーバが非常に多くのTCP接続を管理している場合、1024未満のポートが使用できない可能性があります、OSはより高いポートを使用するように強制されます。接続を完了するには、これら2つのONTAPパラメータを変更する必要があります。

NFSエクスポートポリシー：superuserとsetuid

OracleバイナリがNFS共有に配置されている場合は、エクスポートポリシーにsuperuser権限とsetuid権限を含める必要があります。

ユーザホームディレクトリなどの汎用ファイルサービスに使用される共有NFSエクスポートでは、通常、rootユーザが引き下げられます。これは、ファイルシステムをマウントしたホスト上のrootユーザからの要求が、より低い権限を持つ別のユーザとして再マッピングされることを意味します。これは、特定のサーバ上のrootユーザが共有サーバ上のデータにアクセスできないようにすることで、データを保護するのに役立ちます。setuidビットは、共有環境ではセキュリティリスクになることもあります。setuidビットを使用すると、コマンドを呼び出すユーザとは別のユーザとしてプロセスを実行できます。たとえば、rootがsetuidビットを持つシェルスクリプトはrootとして実行されます。そのシェルスクリプトが他のユーザによって変更される可能性がある場合、root以外のユーザはスクリプトを更新することでrootとしてコマンドを問題できます。

Oracleバイナリには、rootが所有するsetuidビットを使用するファイルが含まれます。OracleバイナリがNFS共有にインストールされている場合は、エクスポートポリシーに適切なsuperuser権限とsetuid権限が含まれている必要があります。次の例では、ルールに allow-suid 許可します superuser (root) システム認証を使用したNFSクライアントのアクセス。

```
Cluster01::> export-policy rule show -vserver vserver1 -policyname orabin
-fields allow-suid,superuser
vserver  policyname ruleindex superuser allow-suid
-----
vserver1 orabin      1          sys      true
```

NFSv4 / 4.1構成

ほとんどのアプリケーションで、NFSv3とNFSv4の違いはほとんどありません。通常、アプリケーションI/O

は非常に単純なI/Oであり、NFSv4の高度な機能の一部からあまりメリットが得られません。上位バージョンのNFSは、データベースストレージから見ると「アップグレード」ではなく、機能を追加したNFSのバージョンとみなすべきです。たとえば、Kerberosプライバシーモード (krb5p) のエンドツーエンドのセキュリティが必要な場合は、NFSv4が必要です。



* NetAppでは* NFSv4の機能が必要な場合はNFSv4.1を使用することを推奨します。NFSv4.1では、一部のエッジにおける耐障害性を向上させるために、NFSv4プロトコルの機能がいくつか強化されています。

NFSv4への切り替えは、マウントオプションを単にvers=3からvers=4.1に変更するよりも複雑です。ONTAPを使用したNFSv4設定の詳細（OSの設定に関するガイダンスなど）については、を参照してください。"[TR-4067『NFS on ONTAP』のベストプラクティス](#)"。このTRの以降のセクションでは、NFSv4を使用するための基本的な要件の一部について説明します。

NFSv4ドメイン

NFSv4 / 4.1の設定について詳しくは本ドキュメントでは説明していませんが、よく発生する問題の1つとして、ドメインマッピングの不一致があります。sysadminから見ると、NFSファイルシステムは正常に動作しているように見えますが、アプリケーションからは特定のファイルに対する権限やsetuidに関するエラーが報告されます。場合によっては、管理者は、アプリケーションバイナリのアクセス許可が破損していると誤って判断し、実際の問題がドメイン名であったときにchownまたはchmodコマンドを実行しました。

ONTAP SVMでNFSv4ドメイン名が設定されます。

```
Cluster01::> nfs server show -fields v4-id-domain
vserver    v4-id-domain
-----
vserver1   my.lab
```

ホストのNFSv4ドメイン名は、/etc/idmap.cfg

```
[root@host1 etc]# head /etc/idmapd.conf
[General]
#Verbosity = 0
# The following should be set to the local NFSv4 domain name
# The default is the host's DNS domain name.
Domain = my.lab
```

ドメイン名が一致している必要があります。マッピングされていない場合は、次のようなマッピングエラーが表示されます。/var/log/messages：

```
Apr 12 11:43:08 host1 nfsidmap[16298]: nss_getpwnam: name 'root@my.lab'
does not map into domain 'default.com'
```

アプリケーションバイナリ（Oracleデータベースバイナリなど）には、rootが所有するsetuidビットのファイルが含まれています。つまり、NFSv4ドメイン名が一致していないとOracleの起動に失敗し、という名前のフ

ファイルの所有権または権限に関する警告が表示されます。 `oradism`` をクリックします。
``$ORACLE_HOME/bin` ディレクトリ。次のように表示されます。

```
[root@host1 etc]# ls -l /orabin/product/19.3.0.0/dbhome_1/bin/oradism
-rwsr-x--- 1 root oinstall 147848 Apr 17 2019
/orabin/product/19.3.0.0/dbhome_1/bin/oradism
```

所有権が `nobody` のファイルが表示される場合は、NFSv4ドメインのマッピングに問題がある可能性があります。

```
[root@host1 bin]# ls -l oradism
-rwsr-x--- 1 nobody oinstall 147848 Apr 17 2019 oradism
```

これを修正するには、`/etc/idmap.cfg` ファイルをONTAPのv4-id-domain設定に対して作成し、整合性を確保します。設定されていない場合は、必要な変更を行い、`nfsidmap -c`` をクリックし、変更が反映されるまでしばらく待ちます。これで、ファイル所有権が `root` として正しく認識されます。ユーザがを実行しようとした場合 ``chown root` NFSドメインの設定が修正される前に、このファイルで次のコマンドを実行する必要があります。 `chown root` をもう一度クリックします

Oracle Direct NFS (dNFS)

Oracleデータベースでは、NFSを2つの方法で使用できます。

まず、オペレーティングシステムの一部であるネイティブのNFSクライアントを使用してマウントされたファイルシステムを使用できます。これはカーネルNFS (kNFS) と呼ばれることもあります。NFSファイルシステムは、NFSファイルシステムを使用する他のアプリケーションとまったく同じように、Oracleデータベースによってマウントされ、使用されます。

2つ目の方法はOracle Direct NFS (dNFS) です。これは、OracleデータベースソフトウェアにNFS標準を実装したものです。DBAによるOracleデータベースの設定や管理方法は変更されません。ストレージシステム自体に正しい設定があるかぎり、DBAチームやエンドユーザはdNFSを透過的に使用できなければなりません。

dNFS機能を有効にしたデータベースでは、通常のNFSファイルシステムが引き続きマウントされています。データベースが開くと、Oracleデータベースは一連のTCP/IPセッションを開き、NFS操作を直接実行します。

Direct NFS

OracleのDirect NFSの主なメリットは'ホストのNFSクライアントをバイパスして'NFSサーバ上で直接NFSファイル操作を実行することですこれを有効にするには、Oracle Disk Manager (ODM) ライブラリを変更する必要があります。このプロセスの手順については、Oracleのマニュアルを参照してください。

dNFSを使用すると、I/Oが最も効率的な方法で実行されるため、I/Oパフォーマンスが大幅に向上し、ホストとストレージシステムの負荷が軽減されます。

さらに、Oracle dNFSには、ネットワークインターフェイスのマルチパスとフォールトトレランス用の*オプション*が含まれています。たとえば、2つの10Gbインターフェイスをバインドして、20Gbの帯域幅を提供できます。一方のインターフェイスで障害が発生すると、もう一方のインターフェイスでI/Oが再試行されま

す。全体的な処理はFCマルチパスとほぼ同じです。マルチパスは、1Gbイーサネットが最も一般的な標準であった数年前には一般的でした。ほとんどのOracleワークロードには10Gb NICで十分ですが、必要に応じて10Gb NICをボンディングできます。

dNFSを使用する場合は、Oracleのドキュメント1495104.1に記載されているパッチをすべてインストールしておくことが重要です。パッチをインストールできない場合は、環境を評価して、そのドキュメントに記載されているバグが原因の問題ではないことを確認する必要があります。必要なパッチをインストールできないためにdNFSを使用できない場合があります。

dNFSでラウンドロビン方式の名前解決（DNS、DDNS、NISなど）を使用しないでください。これには、ONTAPで利用できるDNSロードバランシング機能も含まれます。dNFSを使用するOracleデータベースがあるホスト名をIPアドレスに解決した場合、以降の検索でもホスト名が変更されないようにする必要があります。その結果、Oracleデータベースがクラッシュし、データが破損する可能性があります。

dNFSノユウコウカ

Oracle dNFSは、dNFSライブラリを有効にする以外の設定は必要ありませんが（必要なコマンドについてはOracleのマニュアルを参照してください）、dNFSで接続を確立できない場合は、カーネルNFSクライアントにサイレントに戻すことができます。この場合、パフォーマンスに重大な影響を与える可能性があります。

NFSv4.Xで複数のインターフェイス間でdNFSの多重化を使用する場合、または暗号化を使用する場合は、`orandstab`ファイルを設定する必要があります。構文は非常に厳格です。ファイルに小さなエラーがあると、起動がハングしたり、`orandstab`ファイルがバイパスされたりする可能性があります。

本書の執筆時点では、最新バージョンのOracle DatabaseであるNFSv4.1では、dNFSマルチパスは機能しません。プロトコルとしてNFSv4.1を指定する`orandstab`ファイルでは、特定のエクスポートに対して1つのpathステートメントのみを使用できます。これは、ONTAPがclientIDランキングをサポートしていないためです。この制限を解決するOracle Databaseパッチは、将来提供される可能性があります。

dNFSが期待どおりに動作することを確認する唯一の方法は、`v$dnfs`テーブルを照会することです。

以下は、`/etc`にあるサンプル`orandstab`ファイルです。これは、`orandstab`ファイルを配置できる複数の場所の1つです。

```
[root@jfs11 trace]# cat /etc/orandstab
server: NFSv3test
path: jfs_svmdr-nfs1
path: jfs_svmdr-nfs2
export: /dbf mount: /oradata
export: /logs mount: /logs
nfs_version: NFSv3
```

最初に、指定したファイルシステムでdNFSが動作していることを確認します。


```
SQL> select dirname,nfsversion from v$dnfs_servers;
```

```
DIRNAME
```

```
-----
```

```
NFSVERSION
```

```
-----
```

```
/logs
```

```
NFSv3.0
```

```
/dbf
```

```
NFSv3.0
```

この出力は、これら2つのファイルシステムでdNFSが使用されていることを示していますが、`orantstab`が動作していることを意味するわけではない*。エラーが発生していた場合、dNFSはホストのNFSファイルシステムを自動検出しますが、このコマンドからも同じ出力が表示されることがあります。

マルチパスは次のようにチェックできます。

```
SQL> select svrname,path,ch_id from v$dnfs_channels;
```

```
SVRNAME
```

```
-----
```

```
PATH
```

```
-----
```

```
CH_ID
```

```
-----
```

```
NFSv3test
```

```
jfs_svmdr-nfs1
```

```
0
```

```
NFSv3test
```

```
jfs_svmdr-nfs2
```

```
1
```

```
SVRNAME
```

```
-----
```

```
PATH
```

```
-----
```

```
CH_ID
```

```
-----
```

```
NFSv3test
```

```
jfs_svmdr-nfs1
```

```
0
```

```

NFSv3test
jfs_svmdr-nfs2

[output truncated]

SVRNAME
-----
PATH
-----
      CH_ID
-----
NFSv3test
jfs_svmdr-nfs2
      1

NFSv3test
jfs_svmdr-nfs1
      0

SVRNAME
-----
PATH
-----
      CH_ID
-----

NFSv3test
jfs_svmdr-nfs2
      1

66 rows selected.

```

これらはdNFSが使用している接続です。SVRNAMEエントリごとに2つのパスとチャネルが表示されます。これは、マルチパスが機能していること、つまりoranfstabファイルが認識されて処理されたことを意味します。

Direct NFSとホストファイルシステムへのアクセス

アプリケーションやユーザのアクティビティが、ホストにマウントされた参照可能なファイルシステムに依存している場合、dNFSを使用すると原因の問題が発生することがあります。これは、dNFSクライアントがホストOSの帯域外でファイルシステムにアクセスするためです。dNFSクライアントは、OSが認識されていなくてもファイルの作成、削除、および変更を行うことができます。

シングルインスタンスデータベースのマウントオプションを使用すると、ファイルおよびディレクトリの属性のキャッシュが有効になり、ディレクトリの内容もキャッシュされます。そのため、dNFSでファイルが作成される可能性があり、OSがディレクトリの内容を再読み取りしてファイルがユーザに表示されるまでに少し時間がかかります。これは通常は問題になりませんが、まれに、SAP BR * Toolsなどのユーティリティで問題

が発生することがあります。この場合は、マウントオプションをOracle RACの推奨事項に変更して、問題に対処してください。この変更により、ホストのキャッシュがすべて無効になります。

マウントオプションを変更するのは、(a) dNFSが使用されていて、(b) 問題がファイルが参照可能になるまでの遅延が原因で発生した場合のみにしてください。dNFSを使用していない場合は、シングルインスタンスデータベースでOracle RACマウントオプションを使用すると、パフォーマンスが低下します。



通常とは異なる結果になる可能性のあるLinux固有のdNFSの問題については、の["Linux NFSのマウントオプション"](#)注を参照してください `nosharecache`。

NFSのリースとロック

NFSv3はステートレスです。つまり、NFSサーバ（ONTAP）は、どのファイルシステムがマウントされているのか、誰がどのロックが実際に有効であるのかを追跡しません。

ONTAPにはマウントの試行を記録する機能がいくつかあります。そのため、どのクライアントがデータにアクセスしている可能性があるかを把握できます。また、アドバイザリロックが存在する可能性がありますが、その情報が100%完了する保証はありません。NFSクライアントの状態の追跡はNFSv3標準には含まれていないため、この処理を完了できません。

NFSv4のステートフル

一方、NFSv4はステートフルです。NFSv4サーバは、どのクライアントが使用しているファイルシステム、どのファイルが存在するか、どのファイルやファイル領域がロックされているかなどを追跡します。つまり、状態のデータを最新の状態に保つためには、NFSv4サーバ間で定期的な通信が必要です。

NFSサーバによって管理されている最も重要な状態は、NFSv4ロックとNFSv4リースであり、これらは非常に密接に関連しています。それぞれがそれ自体でどのように機能し、それらが互いにどのように関連しているかを理解する必要があります。

NFSv4ロック

NFSv3では、ロックは推奨されます。NFSクライアントは、「ロックされた」ファイルを変更または削除できます。NFSv3のロックは自動的に期限切れになるわけではなく、削除する必要があります。これは問題を引き起こします。たとえば、クラスタ化されたアプリケーションでNFSv3ロックを作成していて、いずれかのノードで障害が発生した場合は、どうすればよいですか？残りのノードでアプリケーションをコーディングしてロックを解除できますが、それが安全であることをどのようにして確認できますか？「failed」ノードは動作しているが、クラスタの残りのノードと通信していない可能性があります。

NFSv4では、ロックの期間に制限があります。ロックを保持しているクライアントがNFSv4サーバにチェックインし続けるかぎり、他のクライアントはこれらのロックを取得できません。クライアントがNFSv4へのチェックインに失敗すると、最終的にサーバによってロックが取り消され、他のクライアントはロックを要求および取得できます。

NFSv4リース

NFSv4ロックはNFSv4リースに関連付けられます。NFSv4クライアントがNFSv4サーバとの接続を確立すると、リースを取得します。クライアントがロックを取得した場合（ロックにはさまざまな種類があります）、ロックはリースに関連付けられます。

このリースには定義済みのタイムアウトがあります。デフォルトでは、ONTAPはタイムアウト値を30秒に設定します。

```
Cluster01::*> nfs server show -vserver vserver1 -fields v4-lease-seconds

vserver    v4-lease-seconds
-----
vserver1   30
```

つまり、NFSv4クライアントはリースを更新するために、30秒ごとにNFSv4サーバにチェックインする必要があります。

リースはすべてのアクティビティによって自動的に更新されるため、クライアントが作業を行っている場合は追加操作を実行する必要はありません。アプリケーションが静かになり、実際の作業を行っていない場合は、代わりに一種のキープアライブ操作(シーケンスと呼ばれる)を実行する必要があります。それは基本的に「私はまだここにいる、私のリースを更新してください」と言っているだけです。

***Question:** What happens if you lose network connectivity for 31 seconds?
NFSv3はステートレスです。クライアントからの通信を期待していません。NFSv4はステートフルであり、リース期間が経過するとリースが期限切れになり、ロックが取り消され、ロックされたファイルが他のクライアントから利用可能になります。

NFSv3では、ネットワークケーブルを移動したり、ネットワークスイッチをリブートしたり、設定を変更したりすることができ、問題が発生しないように十分に確認することができます。アプリケーションは通常、ネットワーク接続が再び機能するのを辛抱強く待つだけです。

NFSv4では、作業を完了するまでに30秒かかります（ONTAP内でそのパラメータの値を増やした場合を除く）。それを超えると、リースはタイムアウトになります。通常、この結果、アプリケーションがクラッシュします。

たとえば、Oracleデータベースを使用していて、リースタイムアウトを超えるネットワーク接続（「ネットワークパーティション」と呼ばれることもあります）が失われると、データベースがクラッシュします。

これが発生した場合のOracleアラート・ログの出力例を次に示します

```
2022-10-11T15:52:55.206231-04:00
Errors in file /orabin/diag/rdbms/ntap/NTAP/trace/NTAP_ckpt_25444.trc:
ORA-00202: control file: '/redo0/NTAP/ctrl/control01.ctl'
ORA-27072: File I/O error
Linux-x86_64 Error: 5: Input/output error
Additional information: 4
Additional information: 1
Additional information: 4294967295
2022-10-11T15:52:59.842508-04:00
Errors in file /orabin/diag/rdbms/ntap/NTAP/trace/NTAP_ckpt_25444.trc:
ORA-00206: error in writing (block 3, # blocks 1) of control file
ORA-00202: control file: '/redo1/NTAP/ctrl/control02.ctl'
ORA-27061: waiting for async I/Os failed
```

syslogを確認すると、次のエラーのいくつかが表示されます。

```
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
```

ログメッセージは通常、アプリケーションがフリーズする以外に、問題の最初の兆候です。通常、ネットワークの停止中は何も表示されません。これは、NFSファイルシステムにアクセスしようとするプロセスとOS自体がブロックされるためです。

エラーは、ネットワークが再び動作可能になると表示されます。上記の例では、接続が再確立されると、OSはロックの再取得を試みましたが、遅すぎました。リースが期限切れになり、ロックが削除されました。その結果、エラーがOracleレイヤまで伝播し、アラートログにメッセージが記録されます。これらのパターンは、データベースのバージョンと構成によって異なる場合があります。

要約すると、NFSv3はネットワークの中断は許容されますが、NFSv4はより機密性が高く、リース期間が定義されます。

30秒のタイムアウトが許容されない場合はどうなりますか。スイッチが再起動されたり、ケーブルが再配置されたりする動的に変化するネットワークを管理していて、その結果、時々ネットワークが中断される場合はどうなりますか。リース期間を延長することもできますが、その場合はNFSv4猶予期間の説明が必要です。

NFSv4猶予期間

NFSv3サーバをリブートすると、ほぼ瞬時にIOを処理できるようになります。それはクライアントの状態を維持することではありませんでした。そのため、ONTAPのテイクオーバー処理はほぼ瞬時に実行されることがよくあります。コントローラがデータの提供を開始する準備ができた時点で、ネットワークにARPを送信し、トポロジの変更を通知します。通常、クライアントはこれをほぼ瞬時に検出し、データの流れを再開します。

ただし、NFSv4では一時停止が発生します。これは、NFSv4がどのように機能するかの一部にすぎません。



次のセクションは、ONTAP 9.15.1の時点で最新のものです。リースとロックの動作、チューニングオプションはバージョンごとに変更される可能性があります。NFSv4リース/ロックのタイムアウトを調整する必要がある場合は、NetAppサポートで最新情報を確認してください。

NFSv4サーバは、リース、ロック、および誰がどのデータを使用しているかを追跡する必要があります。NFSサーバがパニック状態になってリブートされた場合、または一時的に電力が失われた場合、またはメンテナンス作業中に再起動された場合は、リース/ロックなどのクライアント情報が失われます。サーバは、処理を再開する前に、どのクライアントがどのデータを使用しているかを把握する必要があります。ここで猶予期間が入ります。

NFSv4サーバの電源が突然再投入された場合。再起動すると、IOを再開しようとするクライアントは、基本的に「リース/ロック情報が失われました。ロックを再登録しますか？」これが猶予期間の始まりですONTAPではデフォルトで45秒です。

```
Cluster01::> nfs server show -vserver vserver1 -fields v4-grace-seconds

vserver    v4-grace-seconds
-----
vserver1   45
```

その結果、再起動後、すべてのクライアントがリースとロックを再要求する間、コントローラはIOを一時停止します。猶予期間が終了すると、サーバはIO処理を再開します。

この猶予期間は、ネットワークインターフェイスの変更時のリース再生を制御しますが、ストレージフェイルオーバー時のリース再生を制御する2つ目の猶予期間があります `locking.grace_lease_seconds`。これはノードレベルのオプションです。

```
cluster01::> node run [node names or *] options
locking.grace_lease_seconds
```

たとえば、LIFのフェイルオーバーを頻繁に実行する必要があり、猶予期間を短縮する必要がある場合は、を変更します `v4-grace-seconds`。コントローラフェイルオーバー時のIO再開時間を短縮するには、を変更する必要があります `locking.grace_lease_seconds`。

これらの値は、リスクと結果を十分に理解した上で慎重に変更してください。NFSv4.Xでのフェイルオーバー処理や移行処理でのIOの一時停止を完全に回避することはできません。ロック、リース、猶予期間はNFS RFCの一部です。多くのお客様には、フェイルオーバーにかかる時間が短縮されるため、NFSv3を使用することを推奨します。

リースタイムアウトと猶予期間

猶予期間とリース期間が接続されます。前述したように、デフォルトのリースタイムアウトは30秒です。つまり、NFSv4クライアントは少なくとも30秒ごとにサーバにチェックインする必要があります。そうしないと、リースとロックが失われます。この猶予期間はNFSサーバがリース/ロックデータを再構築できるようにするためのもので、デフォルトは45秒です。猶予期間はリース期間よりも長くする必要があります。これにより、リースを30秒以上更新するように設計されたNFSクライアント環境では、再起動後にサーバにチェッ

クインできます。猶予期間を45秒に設定することで、少なくとも30秒ごとにリースを更新することを期待するすべてのクライアントが確実に更新する機会を得ることができます。

30秒のタイムアウトが許容されない場合は、リース期間を延長することもできます。

60秒のネットワーク停止に耐えるためにリースタイムアウトを60秒に延長する場合は、猶予期間も延長する必要があります。つまり、コントローラフェイルオーバー中にIOが一時停止する時間が長くなります。

これは通常は問題ではありません。一般的なユーザはONTAPコントローラを年に1~2回更新するだけで、ハードウェア障害による計画外フェイルオーバーは非常にまれです。また、ネットワークに60秒のネットワーク停止が発生する可能性があり、リースタイムアウトを60秒にする必要がある場合は、まれにストレージシステムのフェイルオーバーに異議を唱えず、61秒の一時停止も発生する可能性があります。ネットワークが60秒以上頻繁に一時停止していることをすでに認識しています。

NFSキャッシング

次のマウントオプションが存在すると、ホストのキャッシュが無効になります。

```
cio, actimeo=0, noac, forcedirectio
```

これらの設定は、ソフトウェアのインストール、パッチ適用、およびバックアップ/リストアの処理速度に重大な悪影響を及ぼす可能性があります。場合によっては、特にクラスタ化されたアプリケーションでは、クラスタ内のすべてのノードにキャッシュの一貫性を提供するため、これらのオプションが必然的に必要になることがあります。それ以外の場合、顧客はこれらのパラメータを誤って使用し、結果は不要な性能の損傷です。

多くのお客様は、アプリケーションバイナリのインストール時やパッチ適用時に、これらのマウントオプションを一時的に削除しています。インストールまたはパッチ適用プロセス中にターゲットディレクトリを他のプロセスがアクティブに使用していないことをユーザーが確認した場合は、この削除を安全に実行できます。

NFSテンソウサイズ

ONTAPでは、デフォルトでNFS I/Oサイズが64Kに制限されています。

ほとんどのアプリケーションとデータベースでランダムI/Oを実行すると、ブロックサイズがはるかに小さくなり、最大64Kよりもはるかに小さくなります。ラージブロックI/Oは通常並列処理されるため、最大64Kも最大帯域幅の確保に制限されるわけではありません。

一部のワークロードでは、最大64Kに制限があります。特に、バックアップ/リカバリ処理やデータベースのフルテーブルスキャンなどのシングルスレッド処理は、実行回数が少なくとも大容量のI/Oを実行できるのであれば、より高速かつ効率的に実行できます。ONTAPに最適なI/O処理サイズは256Kです。

特定のONTAP SVMの最大転送サイズは、次のように変更できます。

```
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::*>
```



ONTAPで許容される最大転送サイズを、現在マウントされているNFSファイルシステムのrsizewsizeの値より小さくしないでください。これにより、一部のオペレーティングシステムでハングしたり、データが破損したりする可能性があります。たとえば、NFSクライアントのrsizewsizeが65536に設定されている場合は、クライアント自体が制限されているため、ONTAPの最大転送サイズを65536~1048576の間で調整しても効果はありません。最大転送サイズを65536未満に縮小すると、可用性やデータが損傷する可能性があります。

著作権に関する情報

Copyright © 2026 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。