



# データベース設定

## Enterprise applications

NetApp  
May 09, 2024

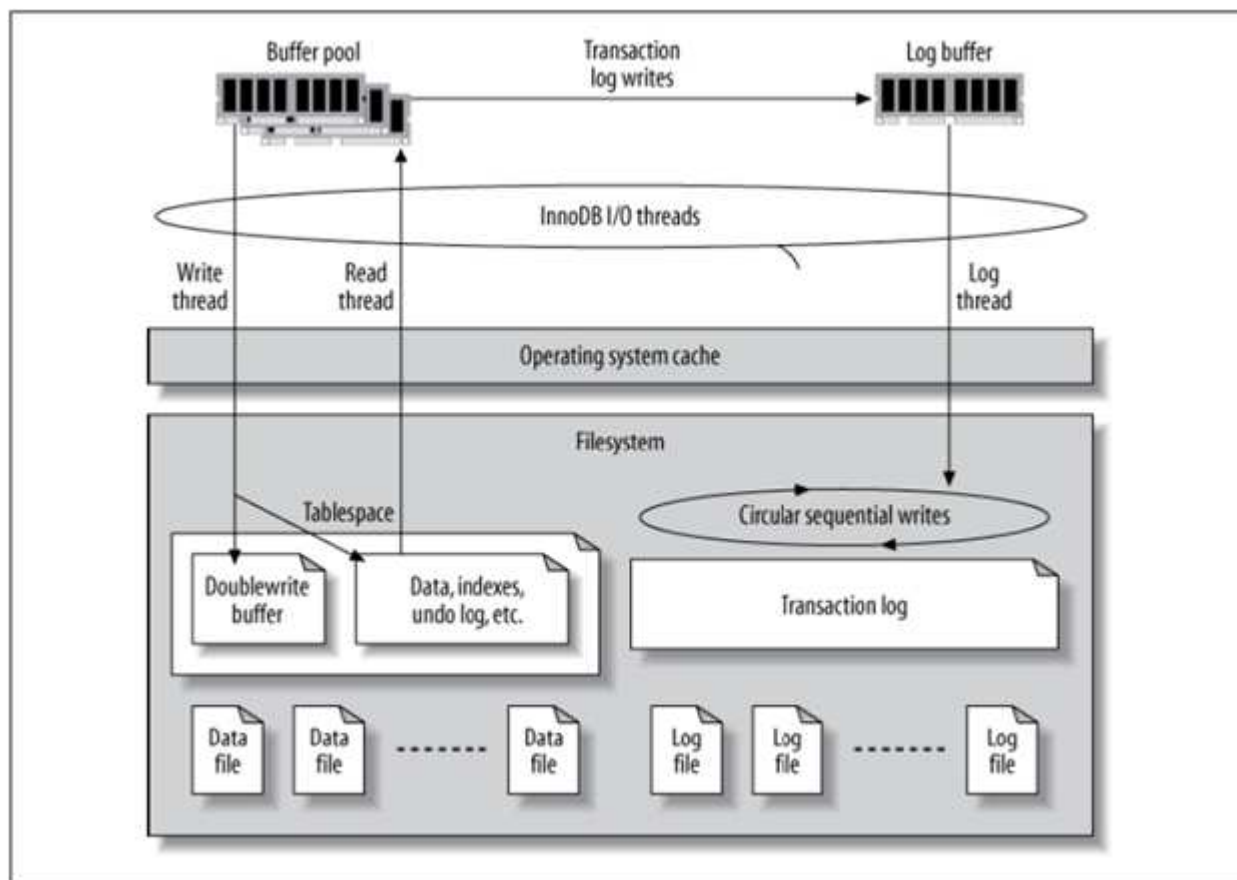
# 目次

データベース設定 .....	1
MySQLとInnoDB .....	1
MySQLセットイハラメエタ .....	3
innodb_log_file_size .....	4
innodb_flush_log_at_trx_commit .....	4
innodb_doublewrite .....	5
innodb_buffer_pool_size .....	5
innodb_flush_method .....	5
innodb_io_capacity .....	6
innodb_lru_scan_depth .....	7
open_file_limits .....	7

# データベース設定

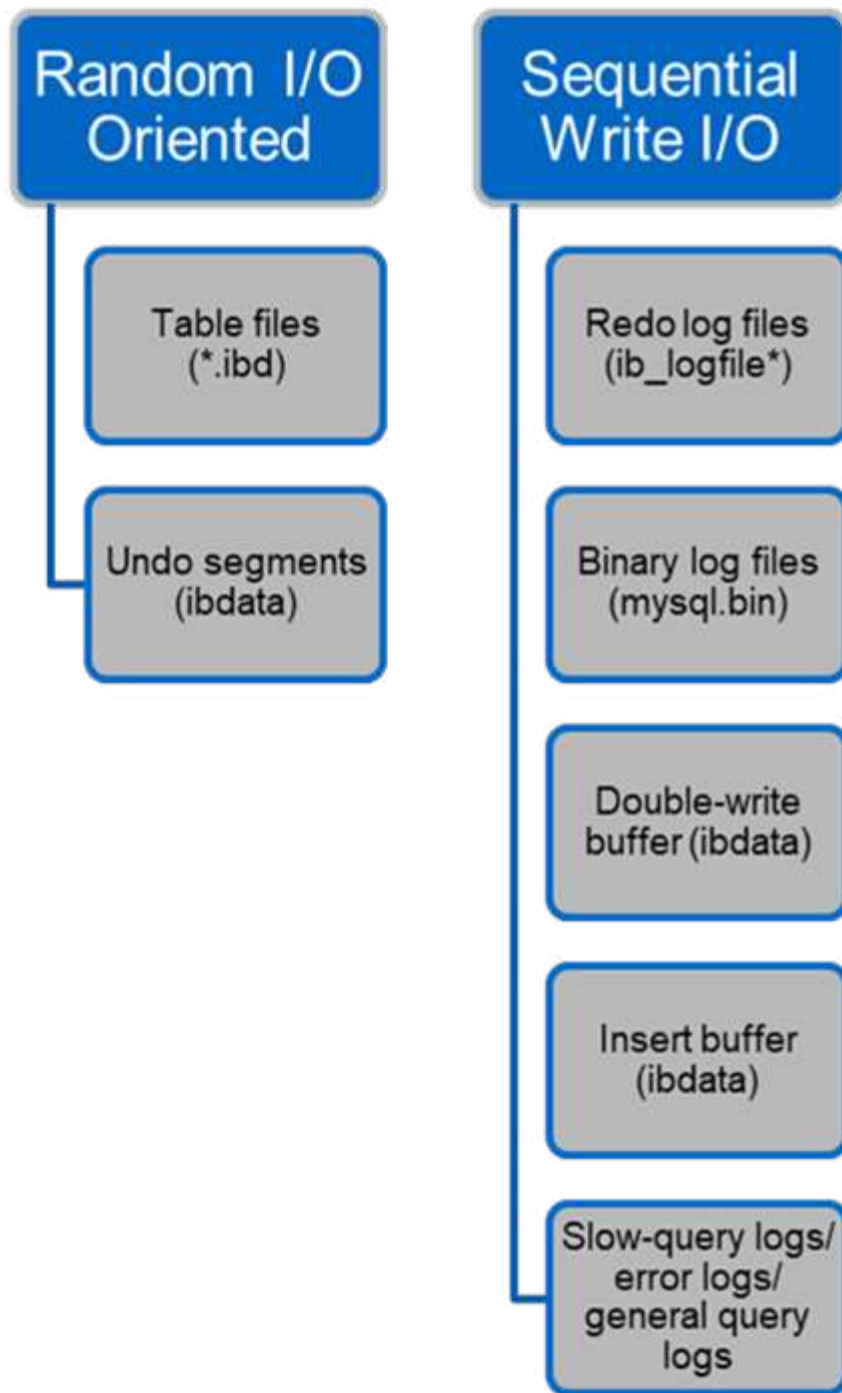
## MySQLとInnoDB

InnoDBはストレージとMySQLサーバの中間層として機能し、データをドライブに格納します。



MySQL I/Oは次の2つのタイプに分類されます。

- ランダムファイルI/O
- シーケンシャルファイルI/O



データファイルはランダムに読み取りおよび上書きされるため、IOPSが高くなります。そのため、SSDストレージを推奨します。

REDOログファイルとバイナリログファイルはトランザクションログです。それらはシーケンシャルに書き込まれるので、書き込みキャッシュを備えたHDDで優れたパフォーマンスを得ることができます。リカバリ時にシーケンシャルリードが発生しますが、ログファイルのサイズは通常データファイルより小さく、シーケンシャルリードはランダムリード（データファイルで発生）よりも高速であるため、パフォーマンスの問題が発生することはほとんどありません。

ダブル書き込みバッファはInnoDBの特別な機能です。InnoDBは、最初にフラッシュされたページをダブルライトバッファに書き込み、次にページをデータファイル上の正しい位置に書き込みます。このプロセスによ

り、ページの破損が防止されます。二重書き込みバッファがない場合、ドライブへの書き込みプロセス中に電源障害が発生すると、ページが破損する可能性があります。ダブル書き込みバッファへの書き込みはシーケンシャルであるため、HDD向けに高度に最適化されています。リカバリ時にシーケンシャルリードが発生します。

ONTAP NVRAMはすでに書き込み保護を提供しているため、ダブル書き込みバッファは必要ありません。MySQLにはパラメータがあります。 `skip_innodb_doublewrite`、ダブルライトバッファをディセーブルにします。この機能により、パフォーマンスが大幅に向上します。

挿入バッファはInnoDBの特別な機能でもあります。一意でないセカンダリインデックスブロックがメモリ内にない場合、InnoDBはエントリを挿入バッファに挿入して、ランダムなI/O操作を回避します。定期的に、挿入バッファはデータベース内のセカンダリインデックスツリーにマージされます。挿入バッファは、I/O要求を同じブロックにマージすることでI/O処理数を削減します。ランダムI/O処理はシーケンシャルです。また、インサートバッファはHDD用に高度に最適化されています。シーケンシャルライトと読み取りは、どちらも通常運用時に発生します。

元に戻すセグメントは、ランダムI/Oです。Multi-Version Concurrency (MVCC) を保証するために、InnoDBは元に戻すセグメントに古いイメージを登録する必要があります。元に戻すセグメントから以前の画像を読み取るには、ランダムな読み取りが必要です。繰り返し実行可能な読み取りで長時間のトランザクション（mysqldump—単一トランザクションなど）を実行したり、長時間のクエリを実行したりすると、ランダムリードが発生する可能性があります。したがって、この場合、元に戻すセグメントをSSDに保存する方が適しています。短いトランザクションまたはクエリのみを実行する場合、ランダムリードは問題ではありません。



- NetAppでは、InnoDBのI/O特性を考慮して、以下のストレージ設計レイアウトを推奨しています。
- MySQLのランダムI/OおよびシーケンシャルI/O指向ファイルを1つのボリュームに格納
- 純粋にシーケンシャルなI/O指向のMySQLファイルを格納するための別のボリューム

このレイアウトは、データ保護のポリシーと戦略の設計にも役立ちます。

## MySQLセツテイハラメエタ

NetAppでは、最適なパフォーマンスを得るために、いくつかの重要なMySQL構成パラメータを推奨しています。

パラメータ	値
<code>innodb_log_file_size</code>	2億5、600万
<code>innodb_flush_log_at_trx_commit</code>	2.
<code>innodb_doublewrite</code>	0
<code>innodb_flush_method</code>	<code>fsync</code>
<code>innodb_buffer_pool_size</code>	11g
<code>innodb_io_capacity</code>	8、192です
<code>innodb_buffer_pool_instances</code>	8
<code>innodb_lru_scan_depth</code>	8、192です
<code>open_file_limit</code>	65535

このセクションで説明するパラメータを設定するには、MySQL構成ファイル（my.cnf）でパラメータを変更する必要があります。NetAppのベストプラクティスは、社内でも実施したテストの結果です。

## innodb\_log\_file\_size

InnoDBログファイルのサイズに適したサイズを選択することは、書き込み処理およびサーバクラッシュ後の適切なリカバリ時間を確保するために重要です。

ファイルに記録されるトランザクションの数が非常に多いため、書き込み処理ではログファイルのサイズが重要になります。レコードが変更されても、変更はすぐにはテーブルスペースに書き込まれません。代わりに、変更はログファイルの最後に記録され、ページはダーティとしてマークされます。InnoDBはログを使用してランダムI/OをシーケンシャルI/Oに変換します。

ログがいっぱいになると、ダーティページが順番にテーブルスペースに書き込まれ、ログファイルのスペースが解放されます。たとえば、トランザクションの途中でサーバがクラッシュし、書き込み処理のみがログファイルに記録されるとします。サーバを再び稼働させるには、ログファイルに記録された変更が再生されるリカバリフェーズを実行する必要があります。ログファイル内のエントリ数が多いほど、サーバのリカバリにかかる時間が長くなります。

この例では、ログファイルのサイズがリカバリ時間と書き込みパフォーマンスの両方に影響します。ログファイルのサイズに適切な数を選択する場合は、リカバリ時間と書き込みパフォーマンスのバランスを取ります。通常、128Mから512Mの間の値は良い値です。

## innodb\_flush\_log\_at\_trx\_commit

データに変更があっても、変更はすぐにはストレージに書き込まれません。

代わりに、データはログバッファに記録されます。これは、InnoDBがログファイルに記録されるバッファ変更割り当てメモリの一部です。InnoDBは、トランザクションがコミットされたとき、バッファがいっぱいになったとき、または1秒に1回のイベントが発生したときに、バッファをログファイルにフラッシュします。このプロセスを制御する構成変数は、`InnoDB_flush_log_at_trx_commit`です。値オプションには次のものがあります。

- 設定時 ``innodb_flush_log_trx_at_commit=0`` では、（InnoDBバッファプール内の）変更されたデータがログファイル（`ib_logfile`）に書き込まれ、ログファイルが1秒ごとにフラッシュされます（ストレージへの書き込み）。ただし、トランザクションがコミットされても何も実行されません。停電またはシステムクラッシュが発生した場合、フラッシュされていないデータはログファイルまたはドライブに書き込まれないため、リカバリできません。
- 設定時 ``innodb_flush_log_trx_commit=1`` InnoDBはログバッファをトランザクションログに書き込み、トランザクションごとに永続的なストレージにフラッシュします。たとえば、すべてのトランザクションコミットについて、InnoDBはログに書き込み、その後ストレージに書き込みます。ストレージの速度が遅いと、パフォーマンスが低下します。たとえば、1秒あたりのInnoDBトランザクション数が減少します。
- 設定時 ``innodb_flush_log_trx_commit=2`` InnoDBはコミットのたびにログバッファをログファイルに書き込みますが、ストレージにデータを書き込むことはありません。InnoDBは、1秒に1回データをフラッシュします。停電やシステムクラッシュが発生した場合でも、オプション2のデータはログファイルに保存され、リカバリ可能です。

パフォーマンスが主な目標である場合は、値を2に設定します。InnoDBは、トランザクションコミットごとではなく、1秒に1回ドライブに書き込みを行うため、パフォーマンスが大幅に向上します。停電やクラッシュが発生した場合は、トランザクションログからデータをリカバリできます。

データの安全性が主な目的である場合は、トランザクションコミットごとにInnoDBがドライブにフラッシュされるように、値を1に設定します。ただし、パフォーマンスに影響する可能性があります。



\* NetAppでは、パフォーマンスを向上させるために、`innodb_flush_log_trx_commit`値を2に設定することを推奨しています。

## innodb\_doublewrite

いつ `innodb_doublewrite` イネーブル(デフォルト)では、InnoDBはすべてのデータを2回格納します。最初にダブルライトバッファに格納し、次に実際のデータファイルに格納します。

このパラメータは、次のコマンドで無効にできます。 `--skip-innodb_doublewrite` ベンチマークの場合や、データの整合性や障害の可能性よりも最高のパフォーマンスに関心がある場合。InnoDBでは、ダブルライトと呼ばれるファイルフラッシュ技術が使用されています。InnoDBは、ページをデータファイルに書き込む前に、ダブル書き込みバッファと呼ばれる連続領域にページを書き込みます。書き込みとダブル書き込みバッファへのフラッシュが完了すると、InnoDBはページをデータファイル内の適切な位置に書き込みます。ページの書き込み中にオペレーティングシステムまたはmysqldプロセスがクラッシュした場合、InnoDBは後でクラッシュリカバリ中にダブルライトバッファからページの適切なコピーを見つけることができます。



\* NetAppでは、ダブル書き込みバッファを無効にすることを推奨しています。ONTAP NVRAMは同じ機能を果たします。ダブルバッファリングは不必要にパフォーマンスを低下させます。

## innodb\_buffer\_pool\_size

InnoDBバッファプールは、チューニングアクティビティの中で最も重要な部分です。

InnoDBは、インデックスのキャッシュとデータのローディング、アダプティブハッシュインデックス、挿入バッファ、および内部で使用される他の多くのデータ構造をバッファプールに大きく依存しています。また、データへの変更もバッファに格納されるため、書き込み処理をストレージに対してすぐに実行する必要はありません。これにより、パフォーマンスが向上します。バッファプールはInnoDBの不可欠な部分であり、それに応じてサイズを調整する必要があります。バッファプールサイズを設定するときは、次の点を考慮してください。

- 専用のInnoDB専用マシンの場合は、バッファプールサイズを使用可能なRAMの80%以上に設定します。
- MySQL専用サーバーでない場合は、サイズをRAMの50%に設定します。

## innodb\_flush\_method

`innodb_flush_method`パラメータは、InnoDBがログファイルとデータファイルを開いてフラッシュする方法を指定します。

### 最適化

InnoDB最適化では、このパラメータを設定すると、必要に応じてデータベースのパフォーマンスが調整されます。

次のオプションは、InnoDBを使用してファイルをフラッシュするためのものです。

- `fsync`。InnoDBでは、`fsync()` データファイルとログファイルの両方をフラッシュするシステムコール。このオプションはデフォルト設定です。
- `O_DSYNC`。InnoDBでは、`O_DSYNC` ログファイルを開いてフラッシュし、データファイルをフラッシュするには`fsync()`を使用します。InnoDBが使用しない `O_DSYNC` 直接、なぜなら、多くの種類のUNIXで問題があったからです。
- `O_DIRECT`。InnoDBでは、`O_DIRECT` オプション（または `directio()`（Solarisの場合）データファイルを開くには、次のコマンドを使用します。`fsync()` データファイルとログファイルの両方をフラッシュします。このオプションは、一部のGNU/Linuxバージョン、FreeBSD、Solarisで利用できます。
- `O_DIRECT_NO_FSYNC`。InnoDBでは、`O_DIRECT` オプションはI/Oのフラッシュ時に使用されますが、`fsync()` 後でシステムコール。このオプションは、一部のタイプのファイルシステム（XFSなど）には適していません。ファイルシステムに `fsync()` システムコール（たとえば、すべてのファイルメタデータを保持する場合は、`O_DIRECT` 代わりにオプション。

## 観察

NetAppラボテストでは、`fsync` NFSとSANではdefaultオプションが使用されており、`O_DIRECT`。flushメソッドを使用している場合 `O_DIRECT` ONTAPでは、クライアントが4096ブロックの境界で大量のシングルバイト書き込みをシリアル方式で書き込みました。この書き込みにより、ネットワーク経由のレイテンシが増加し、パフォーマンスが低下します。

## innodb\_io\_capacity

InnoDBプラグインでは、MySQL 5.7からInnoDB\_io\_capacityという新しいパラメータが追加されました。

InnoDBが実行する最大IOPS（ダーティページのフラッシュレートと挿入バッファ[ibuf]バッチサイズを含む）を制御します。innodb\_io\_capacityパラメータは、バッファプールからのページのフラッシュや変更バッファからのデータのマージなど、InnoDBバックグラウンドタスクによるIOPSの上限を設定します。

innodb\_io\_capacityパラメータに、システムが1秒あたりに実行できるI/O処理のおおよその数を設定します。理想的には、設定をできるだけ低くしてください、バックグラウンドアクティビティが遅くなるように低くしないでください。この設定が高すぎると、データがバッファプールから削除され、キャッシュするにはバッファの挿入が早すぎて大きなメリットが得られません。



\* NetAppでは\*この設定をNFSで使用する場合は、IOPSのテスト結果（SysBench / Fio）を分析し、それに応じてパラメータを設定することを推奨します。InnoDBバッファプールで必要以上に変更されたページやダーティなページが表示されない場合は、フラッシュとページにできるだけ小さい値を使用してください。



ワークロードに対して低い値では不十分であることを証明した場合を除き、20、000以上などの極端な値は使用しないでください。

innodb\_io\_capacityパラメータは、フラッシュ速度と関連するI/Oを調整します。



このパラメータまたはinnodb\_io\_capacity\_maxパラメータを設定すると、パフォーマンスに深刻な影響を与える可能性があります。このパラメータが高すぎて、早期フラッシュでI/O処理が無駄になります。



## innodb\_lru\_scan\_depth

。 `innodb_lru_scan_depth` パラメータは、InnoDBバッファプールのフラッシュ操作のアルゴリズムとヒューリスティックに影響を与えます。

このパラメータは、I/O負荷の高いワークロードを調整するパフォーマンスのエキスパートが主に興味を持っています。このパラメータは、バッファプールインスタンスごとに、Least Recently Used (LRU) ページリスト内でページクリーナースレッドがスキャンを続行し、フラッシュするダーティページを探す距離を指定します。このバックグラウンド処理は1秒に1回実行されます。

値を上下に調整して、空きページ数を最小限に抑えることができます。必要以上の値を設定しないでください。スキャンを実行すると、パフォーマンスが大幅に低下する可能性があります。また、バッファプールインスタンスの数を変更する場合は、このパラメータを調整することを検討してください。

`innodb_lru_scan_depth * innodb_buffer_pool_instances` ページクリーナースレッドが毎秒実行する作業量を定義します。

デフォルトよりも小さい設定は、ほとんどのワークロードに適しています。一般的なワークロードでスペアのI/O容量がある場合にのみ、値を増やすことを検討してください。逆に、大量の書き込みが発生するワークロードでI/O容量が飽和状態になった場合は、特にバッファプールが大きい場合は値を小さくしてください。

## open\_file\_limits

。 `open_file_limits` パラメータは、オペレーティングシステムがmysqldを開くことを許可するファイルの数を決定します。

実行時のこのパラメータの値は、システムで許可されている実際の値であり、サーバの起動時に指定した値とは異なる場合があります。MySQLが開いているファイルの数を変更できないシステムでは、値は0です。効果的な `open_files_limit` 値は、システムの起動時に指定された値（存在する場合）と `max_connections` および `table_open_cache` 次の式を使用します。

- 10以上 `max_connections` [ $+$ ] (`table_open_cache`  $\times$  2)
- `max_connections`  $\times$  5
- オペレーティングシステムの制限（正の場合）
- オペレーティングシステムの制限が無限大の場合： `open_files_limit` 起動時に値が指定されます。指定されていない場合は5,000

サーバは、これら4つの値の最大値を使用してファイル記述子の数を取得しようとし、これだけ多くのディスクリプタを取得できない場合、サーバはシステムが許可する数だけのディスクリプタを取得しようとし、す。

## 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。