

FUJITSU Storage
ETERNUS AX series オールフラッシュアレイ ,
ETERNUS HX series ハイブリッドアレイ

ONTAP FlexGroup ボリューム技術概要



目次

第 1 章	ONTAP における NAS の進化	8
1.1	ボリューム : 実証済みのソリューション	8
1.2	FlexGroup: NAS の進化	10
第 2 章	用語.....	11
2.1	大容量ファイルとは ?	12
第 3 章	ONTAP FlexGroup のメリット	13
3.1	高メタデータ・ワークロードのための大容量と予測可能な低遅延	13
3.2	すべてのクラスタハードウェアの効率的な使用	13
3.3	シンプルで管理が容易なアーキテクチャとバランシング	13
3.4	優れたビッグデータ密度	14
第 4 章	FlexGroup でサポートされる機能	15
第 5 章	ユースケース	18
5.1	理想的な用途	18
5.2	非理想的なケース	18
第 6 章	パフォーマンス	19
6.1	FlexVol 対 FlexGroup: ソフトウェアのビルド	19
6.2	FlexGroup とスケールアウト NAS の競合システム : 少ないリソースでより多くの成果を達成	21
第 7 章	FlexGroup 技術の概要.....	22
7.1	FlexGroup ボリュームの概要	22
7.2	ファイル作成と自動負荷分散	23
7.2.1	ローカル配置とリモート配置	23
7.2.2	ローカルテストとリモートテスト	29
7.3	エラスティックサイジング	30
7.4	FlexVol から FlexGroup のインプレイス変換	33
7.4.1	FlexVol ボリュームを FlexGroup ボリュームに変換する理由	33

7.5	ボリュームの自動サイズ設定 (自動拡張 / 自動縮小)	33
7.6	64 ビットファイル識別子	34
7.6.1	クォータ強制を使用してファイル数を制限する	35
7.6.2	System Manager での 64 ビットファイル ID オプションのサポート	36
7.6.3	ONTAP におけるファイルシステム ID (FSID) の変更の影響	36
7.6.4	ディレクトリ・サイズの考慮事項	37
第 8 章	FlexGroup の特長	40
8.1	シンプルさ	40
8.1.1	コマンドライン (CLI)	40
8.1.2	ONTAP System Manager	43
8.1.3	Active IQ Performance Manager	46
8.1.4	REST API	48
8.1.5	Cloud Volumes ONTAP	48
8.1.6	単一の透過的な名前空間	48
8.1.7	Qtree	48
8.2	統合されたデータ保護	49
8.2.1	RAID DP および RAID Triple Erasure Coding (RAID-TEC)	49
8.2.2	Snapshot テクノロジー	49
8.2.3	SnapMirror と SnapVault	50
8.2.4	CIFS/SMB または NFS によるテープ・バックアップ	51
8.2.5	MetroCluster	51
8.3	ストレージの効率性	51
8.3.1	FabricPool	53
8.4	保存状態での暗号化	54
8.5	サービス品質 (QoS)	54
8.5.1	FlexGroup ボリュームでのストレージ QoS の最大値の動作	54
8.5.2	サービス品質 (QoS) の最小値	55
8.5.3	Adaptive Quality of Service (QoS)	56
付録 A		57
A.1	コマンドラインの例.....	57
A.2	FlexGroup 統計	58
A.2.1	Qtree 統計	59
A.3	FlexGroup インジェスト分散の表示	59
A.4	FlexGroup ボリューム上にファイルを生成するサンプル Python スクリプト	60

目次

図 1.1	100 TB を超える容量のジャンクションアーキテクチャを使用した FlexVol 設計	9
図 1.2	ONTAP における NAS ファイル・システムの進化.....	10
図 2.1	大容量ファイルとは?.....	12
図 6.1	Git ベンチマーク : FlexGroup と FlexVol での Linux コンパイル.....	19
図 6.2	Git ベンチマーク : FlexGroup と FlexVol での GCC コンパイル.....	20
図 6.3	FlexGroup (2 ノードクラスタ) と競合システム (14 ノードクラスタ): 標準の NAS ワークロード	21
図 7.1	FlexGroup ボリューム.....	22
図 7.2	リモート・ハード・リンクを介したファイルのリモート配置.....	24
図 7.3	FlexGroup ボリューム内のファイルおよびフォルダーの配分 : dd スクリプト.....	26
図 7.4	メンバー・ボリューム間の領域配分 : dd スクリプト	27
図 7.5	エラスティックサイジング前のファイル書き込み動作.....	31
図 7.6	エラスティックサイジング後のファイル書き込み動作.....	32
図 7.7	maxdirsize の警告が表示された ONTAP System Manager のイベント画面	39
図 8.1	ONTAP System Manager で FlexGroup ボリュームへの共有を作成する	44
図 8.2	System Manager での FlexGroup の概要	45
図 8.3	既存の FlexGroup ボリュームの管理	46
図 8.4	Performance Manager の FlexGroup メンバー・ボリューム	46
図 8.5	Performance Manager での FlexGroup メンバー・ボリュームのグラフィック表示	47
図 8.6	Performance Manager での FlexGroup メンバー・ボリュームのグラフィック表示 (拡大表示)	47
図 8.7	FlexGroup Snapshot コピー	50
図 8.8	FlexGroup ボリュームのストレージ QoS: シングルノード接続	55
図 8.9	FlexGroup ボリュームのストレージ QoS: マルチノード接続.....	55
図 A.1	理想的な FlexGroup インジェスト	60

表目次

表 4.1	一般的な ONTAP 機能のサポート	15
表 4.2	一般的な NAS プロトコル・バージョンのサポート	17
表 4.3	サポートされていない SMB 2.x および 3.x の機能	17
表 8.1	FlexGroup で使用するボリュームコマンドオプション	41
表 8.2	ONTAP バージョンでの FlexGroup のストレージ効率に関するガイダンス	52

はじめに

本書では、ONTAP の機能である ONTAP FlexGroup ボリュームの概要について説明します。FlexGroup は、スケールアウト NAS コンテナを進化させたもので、メタデータを多用するワークロードにおいて、ほぼ無限の容量と予測可能な低レイテンシのパフォーマンスを組み合わせます。

Copyright 2020 FUJITSU LIMITED

初版
2020 年 11 月

登録商標

本製品に関連する他社商標については、以下のサイトを参照してください。
<https://www.fujitsu.com/jp/products/computing/storage/trademark/>

本書では、本文中の ™、® などの記号は省略しています。

本書の読み方

対象読者

本書は、ETERNUS AX/HX の設定、運用管理を行うシステム管理者、または保守を行うフィールドエンジニアを対象としています。必要に応じてお読みください。

関連マニュアル

ETERNUS AX/HX に関連する最新の情報は、以下のサイトで公開されています。
<https://www.fujitsu.com/jp/products/computing/storage/manual/>

本書の表記について

■ 本文中の記号

本文中では、以下の記号を使用しています。

注 意

お使いになるときに注意していただきたいことを記述しています。必ずお読みください。

備 考

本文を補足する内容や、参考情報を記述しています。

第 1 章

ONTAP における NAS の進化

ハードディスクのコストが削減され、フラッシュドライブの容量が飛躍的に増加するにつれて、ファイルシステムもそれに追随しています。数十ギガバイト、あるいはテラバイト単位のファイルシステム時代は終わりです。ストレージ管理者は、エンタープライズ・レベルのパフォーマンスを備えた大容量を求めるアプリケーション・オーナーからの要求の高まりに直面しています。

機械学習と人工知能のワークロードには、ペタバイト規模（何十億ものファイル）まで拡張可能な単一のネームスペースに対するストレージニーズが含まれます。これらの技術の台頭と Hadoop のようなビッグデータフレームワークの出現により、NAS ファイルシステムの進化は遅れています。ONTAP FlexGroup は、これらのタイプのアーキテクチャに最適なソリューションです。

1.1 ボリューム：実証済みのソリューション

柔軟なボリュームである FlexVol ソフトウェアは、2005 年に Data ONTAP テクノロジーで Data ONTAP 7.0 (7 モードで動作する Data ONTAP) リリースの一部として導入されました。そのコンセプトは、ストレージ・ファイル・システムをハードウェア構成全体にわたって仮想化し、絶えず変化するデータ・センターで柔軟なストレージ管理を実現することでした。

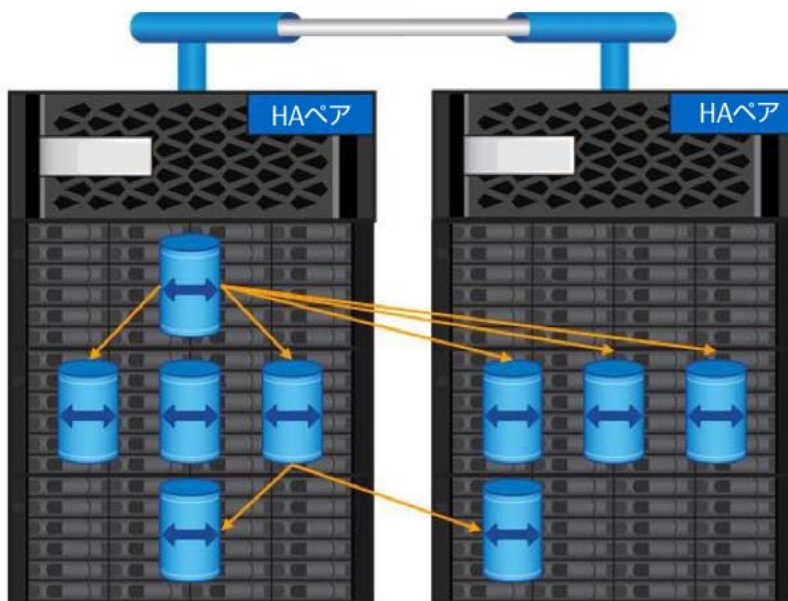
FlexVol ボリュームは無停止で拡張または縮小でき、ストレージシステムのオーバープロビジョニングを可能にするシンプロビジョニングされたコンテナとしてストレージオペレーティングシステムに割り当てることができます。これにより、ストレージ管理者は消費者の要求に応じて容量を自由に割り当てることができました。

しかし、データの増加に伴い、ファイル・システムの拡張が必要になりました。FlexVol は 100 TB の容量でほとんどのストレージニーズに対応でき、Data ONTAP はこれらのボリュームが動作可能なクラスタアーキテクチャを提供しています。しかし、1 つの名前空間で大量のストレージを使用する場合には、ペタバイト規模のストレージが必要でした。

FlexGroup の前に、ONTAP 管理者は、FlexVol ボリュームを相互に接続するジャンクションパスを作成できました。このようにして、単一の名前空間として機能できるファイルシステムをクラスタ上に作成しました。

[図 1.1](#) は、大規模なネームスペースに対する FlexVol ボリューム・ジャンクション設計の例を示しています。

図 1.1 100 TB を超える容量のジャンクションアーキテクチャを使用した FlexVol 設計



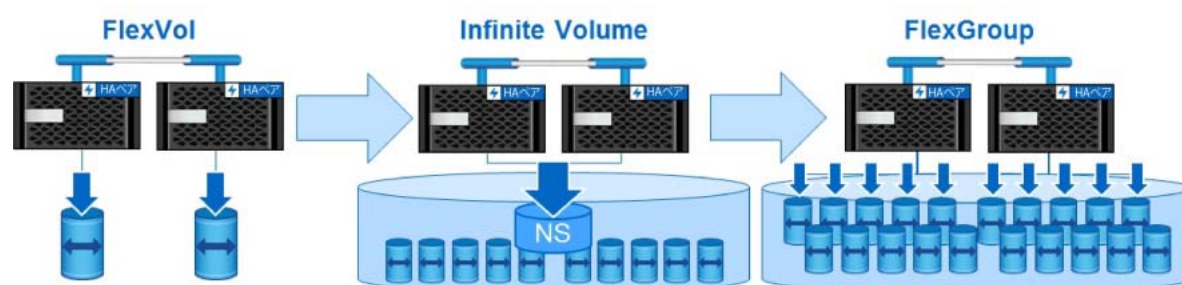
このアーキテクチャは多くの環境で機能しましたが、管理が困難であり、FlexVol ボリュームの容量とファイル数の制約が制約要因となるネームスペースに対して「シングルバケット」アプローチを提供しませんでした。

1.2 FlexGroup: NAS の進化

ONTAP 9.1 は、ONTAP FlexGroup ボリュームというスケールアウト NAS ファイル・システムに革新をもたらしました。

FlexGroup ボリュームを使用すると、ストレージ管理者は 1 つの大規模なネームスペースを数秒で容易にプロビジョニングできます。FlexGroup ボリュームには、ハードウェアの物理的な制限または ONTAP の合計ボリューム制限を超える容量やファイル数の制約は、事実上ありません。制限は、すべてのメンバ間で負荷と領域の割り当てを動的に均等に分散するために共同で動作する構成メンバボリュームの全体数によって決まります。FlexGroup ボリュームに必要なメンテナンスや管理のオーバーヘッドはありません。FlexGroup ボリュームを作成し、NAS クライアントと共有するだけです。残りは ONTAP が行います ([図 1.2](#))。

図 1.2 ONTAP における NAS ファイル・システムの進化



第 2 章

用語

ONTAP FlexGroup 固有の用語については、次のリストで説明します。

- **構成ボリューム / メンバー・ボリューム**

FlexGroup のコンテキストでは、「構成ボリューム」と「メンバー・ボリューム」は交換可能な用語です。これらは、FlexGroup ボリュームを構成する基盤となる FlexVol ボリュームを指し、FlexGroup ボリュームでのみ実現される容量とパフォーマンスの向上を提供します。

- **FlexGroup ボリューム**

FlexGroup ボリュームは、複数の構成ボリューム / メンバーボリュームで構成される単一のネームスペースです。ストレージ管理者によって管理され、FlexVol ボリュームのように機能します。FlexGroup ボリューム内のファイルは、個々のメンバーボリュームに割り当てられ、ボリュームまたはノード間でストライプされません。

- **アフィニティ**

アフィニティは、特定の操作を 1 つのスレッドに結合することを表します。

- **自動インクリメンタル・リカバリ (AIR)**

自動インクリメンタル・リカバリ (AIR) は、FlexGroup の不整合を動的に修復する ONTAP サブシステムです。システム停止や管理者の介入は必要ありません。

- **インジェスト**

インジェストとは、ファイルまたはフォルダの作成によってデータが消費されることです。

- **ジャンクションパス**

FlexGroup をシンプル化してスケールアウトする前に、FlexVol ボリュームの 100 TB の制限を超える容量を提供するために、ジャンクションパスが使用されました。ジャンクションパスは、複数の FlexVol ボリュームを結合してクラスタ全体にスケールアウトし、複数のボリュームアフィニティを提供します。ONTAP でのジャンクションパスの使用は、ONTAP 名前空間内のボリュームの「マウント」として知られています。

- **大容量ファイル**

[\[2.1 大容量ファイルとは?\] \(P.12\)](#) を参照してください。

- **オーバープロビジョニングとシン・プロビジョニング**

オーバープロビジョニング (またはシン・プロビジョニング) ストレージは、ボリュームのスペース保証 (guarantee = none) を無効にする方法です。この方法では、FlexVol ボリュームの仮想空間の割り当てが、そのボリュームが存在するアグリゲートの物理的な制限を超えることができます。たとえば、オーバープロビジョニングの場合、物理サイズがわずか 10 TB のアグリゲートで FlexVol ボリュームが 100 TB になることがあります。オーバープロビジョニングにより、ストレージ管理者は後でボリュームを大きくする必要がないようにボリュームを大きくすることができますが、使用可能な領域を注意深く監視する必要があるため、管理オーバーヘッドが生じます。オーバー・プロビジョニングされたボリュームがある場合、使用可能なスペースは、アグリゲート内の実際の物理使用可能スペースを反映します。したがって、使用率と使用可能な容量の値が少しずれているように見ることがあります。ただし、FlexVol ボリュームで使用可能な仮想スペースと比較して、実際に使用可能なスペースの計算が反映されるだけです。オーバープロビジョニングを使用する場合のスペース割り当てをより正確に表現するには、aggregate show-space コマンドを使用します。

- **リモートアクセスレイヤー (RAL)**

リモートアクセスレイヤー (RAL) は、WAFL システムの機能の 1 つで、FlexGroup ボリュームが複数の FlexGroup 構成要素またはメンバー間でインジェスト作業負荷を分散できるようにします。

- ・ リモートハードリンク

リモートハードリンクは FlexGroup の構成要素です。これらのリンクは通常のハードリンクとして機能しますが、ONTAP に固有です。リンクにより、FlexGroup ボリュームは複数のリモートメンバーまたは構成要素間でワークロードを分散できます。この場合、「リモート」は単に「親ボリュームにないこと」を意味します。リモート・ハード・リンクは、同じアグリゲートまたはノード上の別の FlexVol メンバーになることができます。

2.1 大容量ファイルとは？

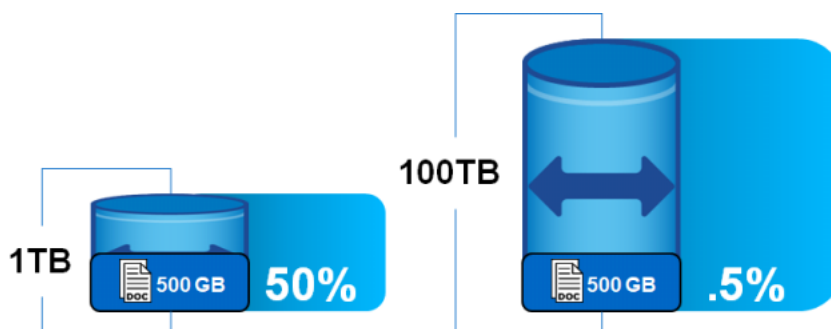
このドキュメントでは、「大容量ファイル」という用語をよく使用します。したがって、FlexGroup のコンテキストでは、大容量ファイルとはどのようなものを正確に定義することが重要です。

FlexGroup ボリュームは、ワークロードが多数の小さなファイルを取り込んでいるときに最適に動作します。これは、FlexGroup ボリュームがシステムリソースを最大化して、FlexVol ボリューム内のシリアル処理によってボトルネックとなる可能性のある特定のワークロードに対処するためです。FlexGroup ボリュームは、他のさまざまなワークロードで適切に機能します (セクション「ユースケース」で定義されています)。ただし、問題を引き起こす可能性のあるワークロードのタイプの 1 つは、大容量ファイルまたは時間の経過とともに増大するファイル (データベース・ファイルなど) を含むワークロードです。

FlexGroup ボリュームでは、大容量ファイルは、特定のファイルサイズではなく、割り当てられた容量のパーセンテージの積で表されます。したがって、一部の FlexGroup 構成では、たとえばメンバーボリュームサイズが 1 TB だけの場合、「大容量ファイル」は 500 GB になります (メンバーボリュームサイズの 50%)。メンバーボリュームサイズが 100 TB の場合など、他の構成では、同じ 500 GB のファイル・サイズでもボリューム容量の 0.5% しか使用されません。このタイプのファイルは、FlexGroup ボリューム内のインジェストヒューリスティックスを発生させるのに十分な大きさである可能性があります。または、後でメンバーボリュームがいっぱいに近づいたときに問題が発生する可能性があります ([図 2.1](#))。

ONTAP 9.6 以降では、エラスティックサイジングにより、大容量ファイルに関する懸念を緩和できます。ONTAP は、他のメンバーボリュームから領域を借用して、大きなファイルの書き込みを完了できるようにします。また、ONTAP 9.7 ではインジェストアルゴリズムの変更が導入され、サイズの異なる大規模ファイルやデータセットのバランスを取るのに役立ちます。これらの機能により、FlexGroup ボリュームはほとんどのワークロードにとって現実的なランディングスペースとなります。([\[7.3 エラスティックサイジング\]](#) (P.30) を参照してください。)

図 2.1 大容量ファイルとは？



第 3 章

ONTAP FlexGroup のメリット

3.1 高メタデータ・ワークロードのための大容量と予測可能な低遅延

ONTAP FlexGroup は、ストレージ管理者が大容量のプロビジョニングを容易に行うための方法を提供し、無停止で容量を拡張できます。FlexGroup はまた、高メタデータワークロードの並列パフォーマンスを可能にし、ミッションクリティカルなワークロードに低レイテンシを提供しながら、スループットと総処理能力を向上させます。

3.2 すべてのクラスタハードウェアの効率的な使用

FlexGroup ボリュームを使用すると、ストレージ管理者は、複数の物理的なアグリゲートやノードをメンバーの FlexVol ボリュームで簡単に拡張できます。また、アプリケーションやユーザーがデータをダンプできるように、真の意味で単一のネームスペースを維持できます。クライアントとユーザはこの領域をモノリシックなものに見なしていますが、ONTAP はバックグラウンドで作業し、FlexGroup ボリューム全体に着信ファイル作成を均等に分散して、CPU とディスクの効率的な使用率を実現しています。

3.3 シンプルで管理が容易なアーキテクチャとバランシング

大容量の導入を容易にするために、FlexVol ボリュームのように FlexGroup ボリュームを管理できます。ONTAP は、基盤となるメンバー・ボリュームの作成とクラスタ・ノード間のバランス調整を処理し、NAS 共有に対する単一のアクセス・ポイントを提供します。

3.4 優れたビッグデータ密度

FlexGroup ボリュームを使用すると、以下のような ONTAP の優れたストレージ効率機能を利用して、大量のデータを小規模なデータセンターに集約できます。

- シン・プロビジョニング
- データコンパクション
- データ圧縮
- 重複排除

さらに、ONTAP は大容量 SSD をサポートしているため、24 ドライブのシェルフ・エンクロージャ 1 台で大量の未フォーマット時容量を実現できます。10 U のラックスペースでペタバイト規模の容量 (未フォーマット時) を実現し、冷却コスト、消費電力、ラックのレンタルスペースを削減し、ストレージ環境で優れた密度を提供します。これらの機能と、クラスタ全体で容量を効率的に使用し、パフォーマンスのバランスをとる FlexGroup ボリュームの機能を組み合わせることで、ビッグデータ向けに作成されたソリューションが実現します。

第 4 章

FlexGroup でサポートされる機能

表 4.1 に、FlexGroup でサポートされている ONTAP 機能の現在のリストを示します。サポートされていない機能に関するご質問は、富士通サポートにお問い合わせください。

表 4.1 一般的な ONTAP 機能のサポート

サポートされている機能	最初にサポートされたONTAPのバージョン
Snapshotテクノロジー	ONTAP 9.0
SnapRestoreソフトウェア(FlexGroupレベル)	ONTAP 9.0
ハイブリッドアグリゲート	ONTAP 9.0
構成要素またはメンバーのボリューム移動	ONTAP 9.0
後処理重複排除	ONTAP 9.0
RAID-TEC技術	ONTAP 9.0
アグリゲート単位の整合性ポイント	ONTAP 9.0
FlexGroupとFlexVolを同じSVMで共有	ONTAP 9.0
Active IQ Unified Managerのサポート	ONTAP 9.1
インライン適応圧縮	ONTAP 9.1
インライン重複排除	ONTAP 9.1
インラインデータコンパクション	ONTAP 9.1
シン・プロビジョニング	ONTAP 9.1
ETERNUS AX	ONTAP 9.1
クォータのレポート	ONTAP 9.1
SnapMirror技術	ONTAP 9.1
ユーザーおよびグループ・クォータのレポート作成 (強制なし)	ONTAP 9.1
アグリゲートインライン重複排除 (ボリューム間の重複排除)	ONTAP 9.2
ボリューム暗号化(VE)	ONTAP 9.2
SnapVault技術	ONTAP 9.3
Qtree	ONTAP 9.3
重複排除スケジュールの自動化	ONTAP 9.3
バージョンに依存しないSnapMirror/ 統合レプリケーション	ONTAP 9.3
SMBのウイルス対策スキャン	ONTAP 9.3
ボリュームに対する自動調整(自動拡張/自動縮小)	ONTAP 9.3
QoSの最大値/上限値	ONTAP 9.3
SnapMirror再ベースラインを使用しないFlexGroup拡張	ONTAP 9.3
取り込みに考慮されるinode数	ONTAP 9.3
SMBの変更/通知	ONTAP 9.3
ファイル監査	ONTAP 9.4

サポートされている機能	最初にサポートされたONTAPのバージョン
FPolicy	ONTAP 9.4
アダプティブQoS	ONTAP 9.4
QoS最小値(ETERNUS AXのみ)	ONTAP 9.4
Relaxed SnapMirror制限	ONTAP 9.4
SMB 3.xマルチチャネル	ONTAP 9.4
FabricPool	ONTAP 9.5
クォータの適用の例	ONTAP 9.5
Qtree統計	ONTAP 9.5
継承されたSMBウォッチと変更通知	ONTAP 9.5
SMBコピーオフロード(オフロードデータ転送(ODX))	ONTAP 9.5
ストレージ・レベルのアクセス・ガード	ONTAP 9.5
FlexCache (キャッシュのみ；ONTAP 9.7でサポートされる起点としてのFlexGroup)	ONTAP 9.5
エラスティックサイジング	ONTAP 9.6
SMB常時使用可能な共有(SQL/Hyper-Vのみ)	ONTAP 9.6
MetroCluster	ONTAP 9.6
ボリューム名の変更	ONTAP 9.6
ボリュームの縮小	ONTAP 9.6
Aggregate Encryption(AE)	ONTAP 9.6
Cloud Volumes ONTAP	ONTAP 9.6
FlexCache	ONTAP 9.7
NDMP	ONTAP 9.7
vStorage APIs for Array Integration (VAAI)	ONTAP 9.7
NFS V 4.0およびNFS V 4.1 (パラレルNFSまたはpNFSを含む)	ONTAP 9.7
FlexVolからFlexGroupのインプレース変換	ONTAP 9.7
FlexCacheの元ボリュームとしてのFlexGroupボリューム	ONTAP 9.7

表 4.2 一般的な NAS プロトコル・バージョンのサポート

サポートされているNASプロトコルのバージョン	最初にサポートされたONTAPのバージョン
NFSv3	ONTAP 9.0
SMB 2.1, SMB 3.x	ONTAP 9.1 RC2
NFSv4.x	ONTAP 9.7

表 4.3 サポートされていない SMB 2.x および 3.x の機能

サポートされていないSMB 2.xの機能	サポートされていないSMB 3.x
SMBリモートボリュームシャドウコピーサービス (VSS)	• SMB 透過フェイルオーバー • SMB スケールアウト • SMB リモート VSS • SMB ディレクトリリリース • SMB ダイレクトまたはリモートダイレクトメモリアクセス (RDMA) 注 意 FlexGroup ボリュームでは、SMB 3.0 暗号化がサポートされています。

注 意

リモート VSS は、[SMB Previous Versions] タブと同じではありません。リモート VSS はアプリケーション対応のスナップショット機能であり、Hyper-V ワークロードで最もよく使用されます。FlexGroup ボリュームでは、導入されて以来、[SMB Previous Versions] タブがサポートされています。

第 5 章

ユースケース

ONTAP FlexGroup の設計は、理想的と考えられる特定の用途に最も適しています。FlexGroup ボリュームのその他の使用例も可能ですが、通常は機能のサポートに依存します。ほとんどの場合、このユースケースはサポートされている機能セットに限定されます。たとえば、仮想化ワークロードは FlexGroup ボリュームで動作しますが、現時点では SIS クローン作成をサポートしておらず、Virtual Storage Console とは統合されていません。

5.1 理想的な用途

FlexGroup ボリュームは、取り込み負荷が高く (高度な新規データ作成)、並行負荷が高く、サブディレクトリ間に均等に分散されているワークロードに最適です。

- 電子設計自動化
- 人工知能と機械学習のログファイルのリポジトリ
- ソフトウェア・ビルドおよびテスト環境 (Git など)
- 地震、石油、ガス
- メディア資産または HIPAA アーカイブ
- ファイルストリーミングワークフロー
- 非構造化 NAS データ (ホームディレクトリなど)
- ビッグデータとデータレイク (NFS コネクターを持つ Hadoop)

5.2 非理想的なケース

FlexGroup ボリュームでは、一部のワークロードは現在推奨されていません。これらのワークロードには、次のものがあります。

- 仮想化されたワークロード
- ストライピングが必要なワークロード (複数のノードまたはボリュームにまたがる大きなファイル)
- FlexVol ボリュームに対するデータの関係のレイアウトを特定の 방법으로制御する必要があるワークロード
- FlexGroup ボリュームでは現在利用できない特定の機能が必要なワークロード

ご質問がございましたら、富士通サポートにお問い合わせください。

第 6 章

パフォーマンス

ONTAP FlexGroup テクノロジーはキャパシティ・プレイとして位置づけられていますが、パフォーマンス・プレイでもあります。FlexGroup ボリュームでは、トレードオフはありません。同じストレージコンテナで、大容量、予測可能な低レイテンシ、高スループットのパフォーマンスを実現できます。FlexGroup ボリュームは、ワークロードに並行性を追加し、クライアントやストレージ管理者が何も管理しなくても 1 つのストレージコンテナに複数のボリュームのアフィニティを提供することで、この目標を達成できます。ファイル数が多く、メタデータを多用するワークロードでは、複数のボリュームとクラスタノードをすばやく簡単に提供できるため、ONTAP は複数のハードウェア資産と CPU コアを使用して、より高いパフォーマンスしきい値で実行できます。ETERNUS AX/HX も同じ性能傾向となります。

6.1 FlexVol 対 FlexGroup: ソフトウェアのビルド

ソフトウェアビルドツール (Git) を使用した単純なワークロードベンチマークでは、同じハードウェア上の単一の FlexVol に対して、FlexGroup 内の SAS ドライブの 2 つのアグリゲートと 8 つの FlexVol メンバ構成要素を持つ ONTAP 9.1 を実行している単一の FAS8080 ノード上で Linux カーネルがコンパイルされました。測定対象のメトリックは、単純な、完了までの時間テストです。このベンチマークでは、複数の Git 操作で、FlexGroup ボリュームのパフォーマンスが、FlexVol ボリュームの 2 ～ 6 倍優れていました。さらに、同じ Git テストが ETERNUS AX ([図 6.1](#) および [図 6.2](#)) で gcc コンパイルで実行されました。

注意

GCC コンパイルでは多数のファイルを処理するため、完了時間が異なります。

図 6.1 Git ベンチマーク : FlexGroup と FlexVol での Linux コンパイル

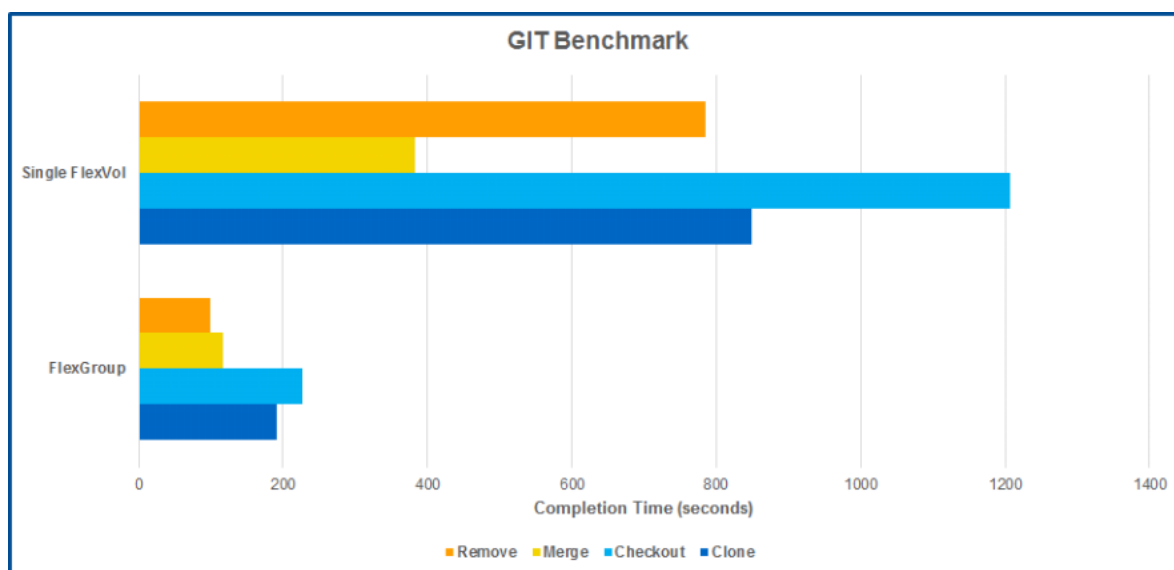
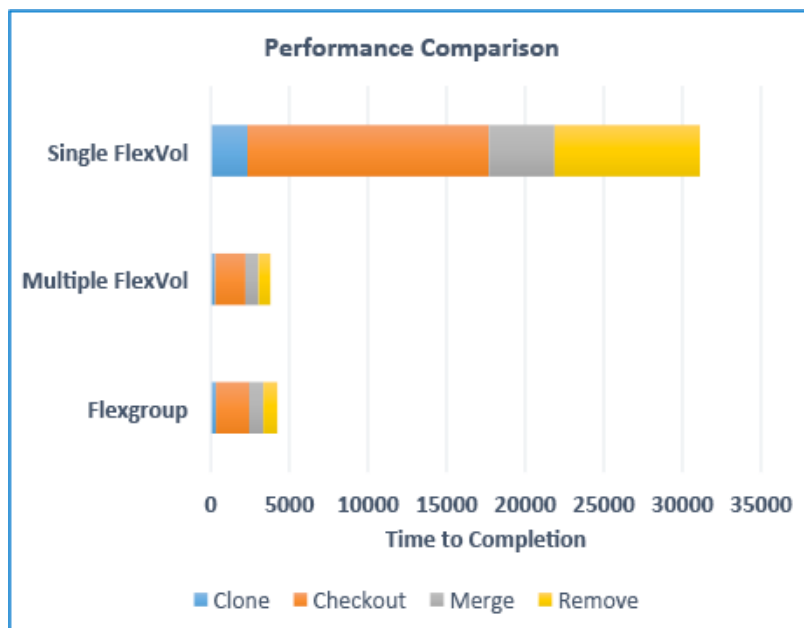


図 6.2 Git ベンチマーク : FlexGroup と FlexVol での GCC コンパイル

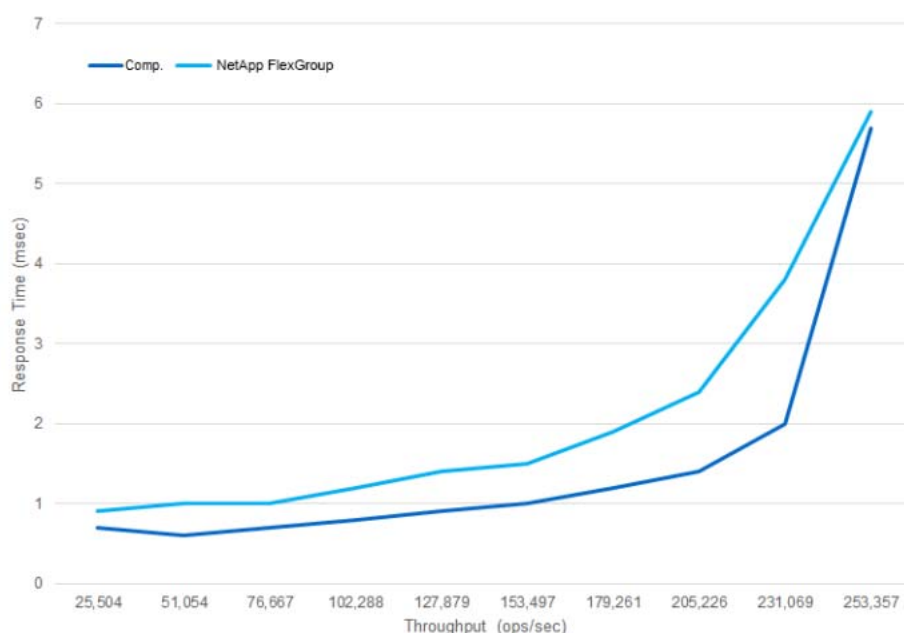


6.2 FlexGroup とスケールアウト NAS の競合システム : 少ないリソースでより多くの成果を達成

別のベンチマークでは、SAS ドライブを使用して ONTAP 9.1 を実行する 2 ノード FAS8080 クラスタ上の FlexGroup ボリュームを、14 ノードを使用する競合システムと比較しました。競合システムでも、メタデータのキャッシュにソリッドステート・ドライブ (SSD) を使用していました。このテストでは、標準的な NAS ワークロード生成ツールを使用してワークロードをシミュレートしました。

テストでは、8 つのメンバー構成要素を持つ単一の FlexGroup ボリュームが、競合他社の 14 ノードクラスタ (図 6.3) と基本的に同じレイテンシ曲線で、ほぼ同じ 1 秒あたりの操作数を取り込むことができたことがわかりました。

図 6.3 FlexGroup (2 ノードクラスタ) と競合システム (14 ノードクラスタ) : 標準の NAS ワークロード



第 7 章

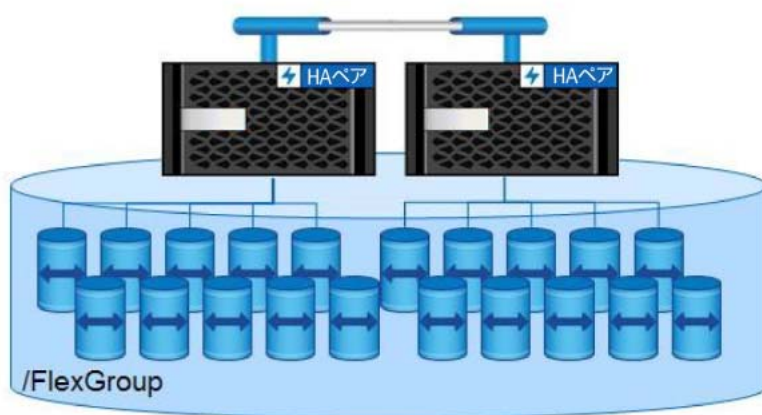
FlexGroup 技術の概要

ONTAP FlexGroup 技術では、ONTAP 内の FlexVol ボリュームの概念を取り入れ、Remote Access Layer (RAL) と呼ばれる WAFL サブシステムを適用しました。RAL は、新しいファイルやフォルダの取り込みを管理し、既存のファイルやフォルダを追跡して、読み取り時のリダイレクトを高速化します。この機能により、参加メンバー・ボリューム間での着信書き込みの自動ロード・バランシングが可能になります。

7.1 FlexGroup ボリュームの概要

高いレベルでは、FlexGroup ボリュームは、単一のエンティティとして機能する FlexVol ボリュームの集合に過ぎません。NAS クライアントは、通常の FlexVol ボリュームと同様に、エクスポートまたは CIFS/SMB 共有 ([図 7.1](#)) から FlexGroup ボリュームにアクセスします。ONTAP は、クライアントに対して透過的な ONTAP 内のシンボリックリンクを使用して、着信 NAS 要求をリダイレクトします。

図 7.1 FlexGroup ボリューム



FlexGroup ボリュームの基本構成は 1 つの FlexVol ボリュームですが、FlexGroup ボリュームには通常の FlexVol ボリュームにはない利点がいくつかあります。詳細は、[「第 3 章 ONTAP FlexGroup のメリット」 \(P.13\)](#) を参照してください。

FlexGroup ボリュームでは、FlexVol ごとにファイルが作成されます。ファイルのストライピングは行われません。FlexGroup ボリュームのスループットとパフォーマンスの向上は、複数の FlexVol ボリューム、アグリゲート、およびノードにわたるオペレーションの同時実行によって確認できます。FlexGroup ボリュームが存在するすべてのハードウェアで、一連の操作を並行して実行できます。FlexGroup ボリュームは、クラスタ化された ONTAP スケールアウトアーキテクチャを補完する理想的なボリュームです。

7.2 ファイル作成と自動負荷分散

ファイルが FlexGroup ボリューム内に作成されると、そのファイルは FlexGroup ボリューム内の「最も可用性が高い」FlexVol メンバーに送られます。この場合の「最も可用性が高い」とは、「使用可能な空き領域が最大」、「利用可能なフリー inode 数が最大」および FlexVol メンバーの最近の負荷を意味します。ONTAP は、管理者の介入なしにこれらの決定を行います。FlexGroup ボリュームの目標は、メンバーボリュームの容量をできるだけ均等に割り当てることです。また、リモートのハードリンクの数をできるだけ少なくして、FlexGroup ボリュームの取り込みワークロードをメンバー間でできるだけ均等に分散することも目標です。

このアプローチは「自動負荷分散」と呼ばれ、クライアントおよびストレージ管理者に対して透過的です。この概念により、FlexGroup のストーリーの全体的なシンプルさが高まります。ストレージ管理者は数秒でストレージをプロビジョニングできるため、通常は設計やレイアウトを考慮する必要はありません。

ONTAP の各種バージョンでは、次の機能に注意してください。

- ONTAP 9.7 以降のバージョンでは、ボリューム自動拡張がサポートされているため、「スペース不足」(ENOSPC)の問題を軽減できます。
- ONTAP 9.6 以降のバージョンでは、メンバーがいっぱいになったときに大きなファイルへの書き込みが失敗する場合に備えて、エラスティックサイジングを使用して保護レイヤーを追加します。
- ONTAP 9.7 では、ストリーミング I/O ワークロードの配置をより適切に処理するために、取り込みの変更が導入されています。

7.2.1 ローカル配置とリモート配置

FlexGroup ボリュームには複数の構成ボリュームがあり、FlexGroup はすべての構成ボリュームにデータを均等に配置するように設計されているため、ファイルの「リモート配置」という概念があります。ONTAP は、NAS のみの構成で最大 24 ノードのクラスターで動作できるため、「リモートトラフィック」という概念もあります。

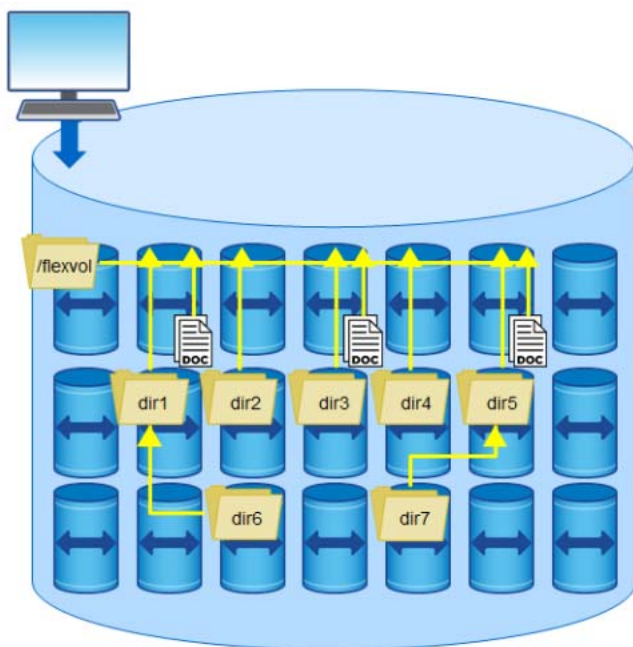
しかし、この 2 つの概念は同義ではありません。リモートトラフィックは、クラスターインターコネクトネットワーク経由のトラフィックであり、FlexGroup ボリュームだけでなく、すべての ONTAP 構成に存在します。リモート配置は、FlexGroup ボリュームに固有で、親ディレクトリを持たない FlexGroup メンバー上にファイルまたはフォルダが作成されたときに発生します。リモート配置は、RAL によって割り当てられたリモートハードリンクを経由して、親ボリュームを所有するノード上でも実行できます。

ローカル配置でオブジェクトを作成する場合、オブジェクトが親ディレクトリと同じ構成要素に格納されるようにします。ローカルに配置すると、今後の操作でそのファイルまたはサブディレクトリに対して最高のメタデータ・パフォーマンスが得られます。対照的に、リモートに配置すると、そのファイルまたはサブディレクトリは、わずかなメタデータアクセスペナルティを受けます。ただし、FlexGroup ボリュームでは、新しいコンテンツを親ディレクトリとは別の構成要素に配置することで、集合的なデータセットとワークロードの分散を改善できます。ファイルやフォルダをリモートに配置することで生じるペナルティは、ワークロードに対して複数のボリュームアフィニティ ([図 7.2](#)) を持つことで得られるパフォーマンスとスループットの向上によって十分に相殺されます。

ワークロードの均等なバランスを実現するために、ONTAP はステータスとメンバのボリューム属性を毎秒監視して、データの最適な配置を実現します。このプロセスでは、複数の目標のバランスを達成しようとします。

- スペースの使用率のバランスが保たれます。
- データ作成パターンを分析することにより、メンバー構成要素の過剰利用が回避されます。
- ファイルはローカルに配置されるため、レイテンシーは可能な限り低くなります。

図 7.2 リモート・ハード・リンクを介したファイルのリモート配置



■ ファイルとフォルダの取り込みの例

GitHub では、クライアント上で dd 操作を並列に行うため、dd スクリプトを追加しました。このスクリプトは、シングルスレッドの dd コマンドよりも優れたスループットテストを提供します。このスクリプトを使用して、取り込み時に、ONTAP が FlexGroup 全体でファイルをどのように配置するかを説明しています。次に、診断レベルのコマンドを使用して inodes に基づいてそのファイルの場所を特定し、次のスクリプトでファイルの場所の一部をマッピングしました。

構成は次のとおりです。

- ONTAP 9.7
- 8 つのメンバーを持つ単一ノードの FlexGroup
- dd スクリプトを実行する単一のクライアント

スクリプトを実行する前の FlexGroup ボリュームサイズの出力は次のとおりです。

```
cluster::*> vol show -vserver DEMO -volume flexgroup_local* -fields used,percent-used,size
vserver volume size used percent-used
-----
DEMO flexgroup_local__0003 2.50TB 57.28MB 5%
DEMO flexgroup_local__0004 2.50TB 57.28MB 5%
DEMO flexgroup_local__0005 2.50TB 57.28MB 5%
DEMO flexgroup_local__0007 2.50TB 57.28MB 5%
DEMO flexgroup_local__0002 2.50TB 57.28MB 5%
DEMO flexgroup_local__0001 2.50TB 57.29MB 5%
DEMO flexgroup_local__0006 2.50TB 57.29MB 5%
DEMO flexgroup_local__0008 2.50TB 57.29MB 5%
DEMO flexgroup_local 20TB 458.2MB 61%
9 entries were displayed.
```

このスクリプトは、スクリプトで指定された最上位レベルにフォルダを作成し、指定された数のファイルを特定のサイズで各フォルダに作成します。このテストでは、FlexGroup メンバーのボリューム数をミラーリングするために、8 つのフォルダと 8 つのファイルを選択しました。ファイルサイズを 1 GB にしました。

スクリプトを実行すると、FlexGroup メンバーは次のようになります (重複排除と圧縮により、ファイル・サイズが 1 GB 未満に保たれる)。

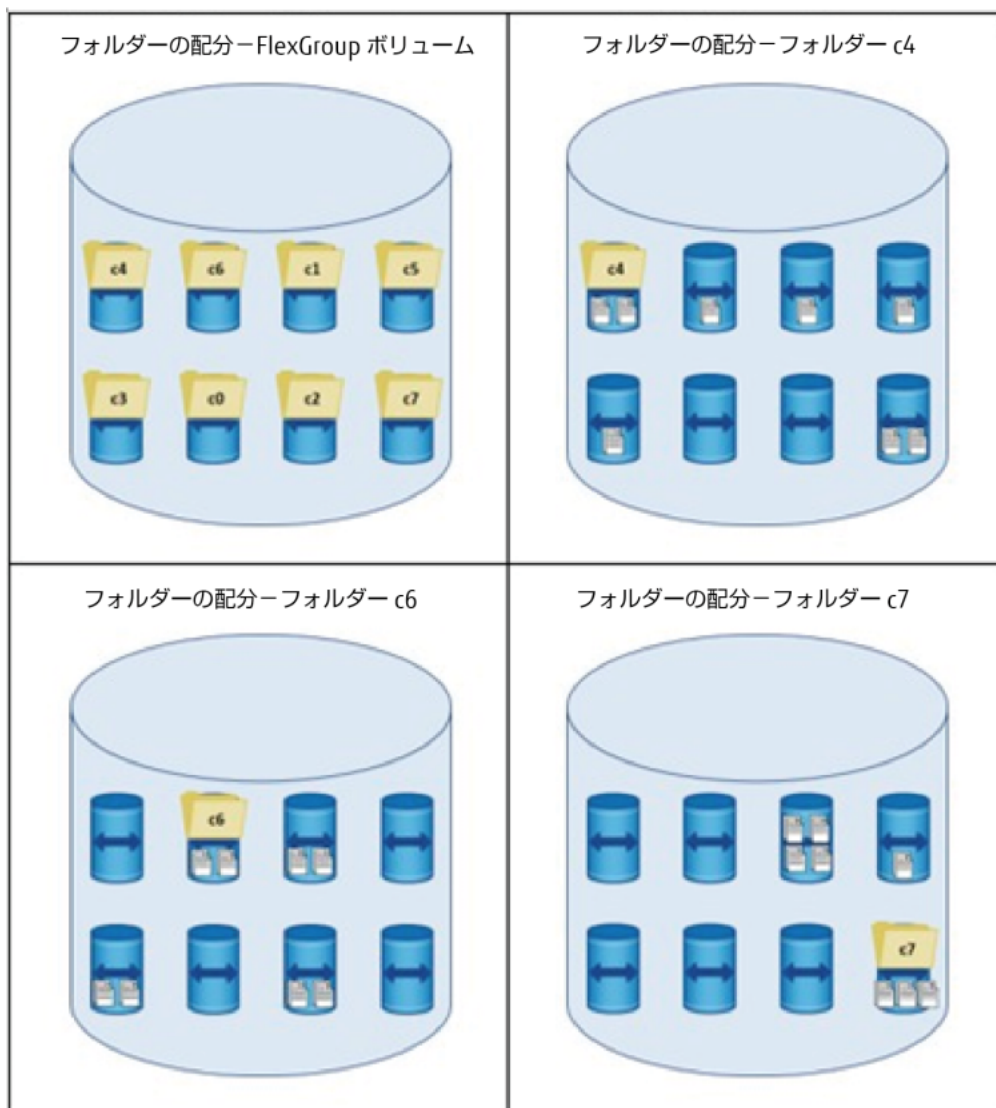
```
cluster::*> vol show -vserver DEMO -volume flexgroup_local* -fields used,percent-  
used,size,files,files-used
```

vserver	volume	size	used	percent-used	files	files-used
DEMO	flexgroup_local__0004	2.50TB	86.88MB	5%	21251126	106
DEMO	flexgroup_local__0007	2.50TB	90.85MB	5%	21251126	106
DEMO	flexgroup_local__0005	2.50TB	91.20MB	5%	21251126	107
DEMO	flexgroup_local__0006	2.50TB	91.23MB	5%	21251126	108
DEMO	flexgroup_local__0008	2.50TB	91.58MB	5%	21251126	107
DEMO	flexgroup_local__0002	2.50TB	92.48MB	5%	21251126	107
DEMO	flexgroup_local__0001	2.50TB	92.93MB	5%	21251126	114
DEMO	flexgroup_local__0003	2.50TB	96.44MB	5%	21251126	108
DEMO	flexgroup_local	20TB	733.6MB	61%	170009008	863

9 entries were displayed

出力に示すように、ファイルはすべてのメンバーボリュームに比較的均等に配置されています。レイアウトを詳しく見ると、フォルダがメンバー・ボリューム全体に均等に割り当てられていることがわかります。ただし、フォルダ単位では、ファイルが均等に割り当てられていません。これは、ファイルのサイズが 1 GB であり、サイズが大きいために親フォルダから離れていく可能性が高いためです ([図 7.3](#))。

図 7.3 FlexGroup ボリューム内のファイルおよびフォルダーの配分 : dd スクリプト



ファイルはフォルダごとに均等に配分されていませんでしたが、その結果、すべてのフォルダ全体では、ファイルは時間の経過とともに均等に割り当てられました。この結果は、領域配分出力 ([図 7.4](#)) に表示されます。

```
cluster::*> vol show -vserver DEMO -volume flexgroup_local* -fields used,percent-used,size
vserver volume size used percent-used
-----
DEMO flexgroup_local__0002 2.50TB 85.77MB 5%
DEMO flexgroup_local__0008 2.50TB 85.92MB 5%
DEMO flexgroup_local__0003 2.50TB 86.04MB 5%
DEMO flexgroup_local__0006 2.50TB 89.72MB 5%
DEMO flexgroup_local__0004 2.50TB 90.34MB 5%
DEMO flexgroup_local__0007 2.50TB 94.10MB 5%
DEMO flexgroup_local__0005 2.50TB 94.18MB 5%
DEMO flexgroup_local__0001 2.50TB 94.23MB 5%
DEMO flexgroup_local 20TB 720.3MB 61%
9 entries were displayed.
```

図 7.4 メンバー・ボリウム間の領域配分 : dd スクリプト



■ FlexGroup ロードバランシングの概念

次の用語と概念は、FlexGroup ボリュームがデータを取り込み、メンバーボリューム間で負荷を分散する方法の中心となります。

- **リモートアクセスレイヤー**
RAL は、FlexGroup 機能を備えた ONTAP の新しいメカニズムです。RAL を使用すると、1 つのメンバー・ボリュームに対して実行される 1 つの WAFL メッセージで、そのメンバー・ボリュームおよび他のメンバー・ボリューム内の inode を操作できます。
- **インジェストヒューリスティクス**
インジェストヒューリスティクスを使用すると、ONTAP は一連の決定ポイントに基づいてファイルインジェストに関する知的意思決定を行うことができます。FlexGroup ボリュームに参加しているメンバーボリュームは、毎秒更新され、ボリュームの現在の状態の最新のビューを提供します。各 ONTAP リリースにはインジェストヒューリスティクスの改善が施されているため、FlexGroup ボリュームを使用している場合は、必ず最新のパッチ適用済みリリースを実行してください。
- **リモート割り当て**
FlexGroup ボリュームは、次の項目に基づいてワークロードを割り当てます。
 - メンバーの構成要素が保持するデータ量 (使用率)
 - メンバー内の使用可能な空きスペースの量
 - 最後の N 秒分の新しいコンテンツ割り当て要求
 - 最後の N 秒分の inode 割り当てとその割り当て元
 - メンバー・ボリューム内の空き inode/ ファイルの数 (ONTAP 9.3 以降)

通常、FlexGroup ボリュームは可能であればローカル配置を優先しますが、場合によってはデータのリモート割り当ての方が優先されることもあります。これらのシナリオには、次のものがあります。

- FlexGroup ボリュームの最上位ジャンクションの近くにサブディレクトリを作成する
- すでに多くのファイルまたはフォルダが存在するディレクトリにファイルまたはフォルダを作成する
- メンバー構成要素のスペース割り当てが、容量使用率と大きな差異を持つ状況
- メンバー構成ボリュームの容量または inode 使用率が 90% に近づく状況
- 負荷不平衡状態 (1 人のメンバーが他のメンバーよりも多くのトラフィックを得ている)

メンバーが親ディレクトリとは異なる FlexVol ボリュームである場合もありますが、FlexGroup のインジェストヒューリスティックは、リモートトラフィックよりもファイルのローカルトラフィックを優先する傾向があります。ディレクトリの場合、FlexGroup ボリュームは、FlexVol メンバーのローカル作成よりもリモート作成を優先する傾向があります。FlexGroup ボリューム上でトラフィックが正規化され、ボリュームがファイルの割り当てを開始すると、ローカルトラフィックの方が優先され、リモートトラフィックは 5% から 10% 程度にとどまります。

ONTAP 9.7 以降では、インジェストアルゴリズムによってワークロードタイプが検出され、ワークロードタイプを支配する操作の種類に基づいて、よりインテリジェントな配置決定が行われます。この変更により、ボリュームは、ストリーミングや大容量ファイルのワークロードの配置をよりインテリジェントに決定できるようになり、データの取り込み時のメンバーボリューム間のワークロードのバランスが向上します。

- **緊急度**

FlexGroup ボリュームの緊急度は、データの取り込みで使用される可能性がどの程度高いかではなく、メンバーがどの程度満たされている (または、ONTAP 9.3 の時点で、ボリュームが最大ファイル数にどれくらい近い) かを示します。各ノードは、新しいデータが FlexGroup ボリューム内のメンバーボリュームに到達する可能性を判断するために使用される 2 つのグローバル変数を保持します。

- **空き容量警告**

この変数は、デフォルトで 50 GB に設定されています。この数値は、FlexGroup メンバーで使用可能な空き容量を表し、リモートメンバーにコンテンツを配置することがどれだけ緊急であるかの確率を計算するために使用されます。

- **空き容量危険**

この変数は、デフォルトで 10GB に設定されています。この数値は、メンバー・ボリュームの緊急度が 100% に設定されるしきい値を表します。すべての取り込みトラフィックは、メンバーに十分な空き領域が追加されるか、データが削除されるまで、メンバーボリュームを回避します。

- **許容値**

FlexGroup ボリュームの許容値は、メンバー間の使用率の不均衡、または FlexGroup ボリュームがリモート割り当ての決定を生成する前に許容できるメンバー間の使用スペースの不均衡の割合を示します。

- **最大許容値**

この値はデフォルトの 10% に設定されています。つまり、メンバボリュームは、リモートトラフィックを、その時点のより高いパーセンテージで送信する前に、使用済み領域のワーキングセット値 (100 G バイト) の 10% まで許容できます。たとえば、1 つのメンバーが別のメンバーより 10% 多く埋まっている場合、トラフィックは他の場所に転送されます。空のメンバボリュームでは、常に最大許容値が使用されます。

- **ミニトレランス**

この値はデフォルトの 0% に設定されています。メンバー・ボリュームがいっぱいになると、最小許容値が適用され、スペース配分を均等にするために 100% の時間でトラフィックがリモートに送信されます。

- **ワーキングセット**

この変数は、最大許容値と最小許容値のパーセンテージが計算に使用する空き領域レベルを定義します。この値のデフォルトは 100 GB です。

■ 注意事項

メンバー・ボリュームが他のメンバー・ボリュームよりも「いっぱい」になり始める場合、ワークロードにより多くのリモート・ハード・リンクが作成されるため、FlexGroup ボリュームのパフォーマンスが低下する可能性があります。

FlexGroup ボリューム内のメンバーボリュームが完全にいっぱいになった場合、FlexGroup ボリューム全体が ENOSPC エラーをクライアントに報告します。問題を修正するには、修正手順を実行する必要があります (FlexGroup メンバーの拡張やデータの削除など)。これは、inode が不足しているメンバー・ボリュームにも適用されます。ONTAP 9.3 では、ファイルを割り当てるときにメンバーボリュームの inode 数を考慮するインジェスト計算が改善されました。

ONTAP 9.6 以降のバージョンには、このドキュメントの後半で説明するエラスティックサイジング機能が含まれています。これにより、メンバーボリュームの空き容量が不足すると書き込み操作が失敗するというシナリオを回避できます。

7.2.2 ローカルテストとリモートテスト

ファイルおよびディレクトリのローカル配置とリモート配置を測定するために、単純なファイルおよびディレクトリの作成テストを実施した。次の設定が使用されました。

- 2 つの FAS8040 ノード
- SSD アグリゲート (非 ETERNUS AX) × 2
- アグリゲートごとに 4 つの FlexVol メンバー構成要素、合計 8 メンバー
- 10 万ディレクトリ
- 100 億ファイル (ディレクトリあたり 10 万 × 10 万ディレクトリ)
- Red Hat 7.x クライアント
- ループに対する単純な mkdir および truncate コマンド:

```
for x in `seq 1 100000`; do mkdir dir$x; done  
for x in dir{1..100000}; do truncate -s 1k /mnt/$x/file{1..100000}; done
```

上記のシナリオでは、ディレクトリのリモート割り当ては 90% でした。

remote_dirs	90
-------------	----

ファイルのリモート割り当てはわずか 10% でした。

remote_files	10
--------------	----

上記の統計情報は、高度な権限レベルで statistics show-object flexgroup コマンドから取得したものです。FlexGroup 統計の収集および表示方法については、付録を参照してください。

7.3 エラスティックサイジング

FlexGroup ボリュームに書き込まれるファイルは、個々のメンバーボリューム内に存在します。メンバーボリューム間ではストライプが行われなため、ファイルが書き込まれて時間の経過とともに増大する場合や、大きなファイルが FlexGroup ボリュームに書き込まれる場合は、メンバーボリューム内の領域不足が原因で書き込みが失敗することがあります。

FlexGroup ボリューム内のメンバーボリュームがいっぱいになる理由はいくつかあります。

- メンバー・ボリュームの使用可能なスペースを超える単一のファイルを書き込む場合。たとえば、10 GB のファイルが 9 GB の使用可能なメンバーボリュームに書き込まれます。
- ファイルが時間の経過とともに追加されると、最終的にメンバー・ボリュームがいっぱいになることがあります (たとえば、データベースがメンバー・ボリュームに存在する場合)。
- Snapshot コピーは、使用可能なアクティブなファイル・システム領域を消費します。

FlexGroup ボリュームは、メンバー・ボリューム間でスペースを適切に割り当てることができますが、ワークロードの異常が発生した場合には、悪影響を及ぼす可能性があります。(たとえば、ボリュームが 4 K ファイルで構成されているが、一部を圧縮して巨大な単一ファイルを作成する場合)。

1 つのソリューションは、ボリュームの拡張またはデータの削除です。しかし、管理者は、手遅れになって「スペースがない」エラーが発生するまで問題を認識しないことがよくあります。

たとえば、FlexGroup ボリュームのサイズは数百テラバイトになりますが、基礎となるメンバーボリュームとその空き容量によって、個々のファイルに使用できる容量が決まります。200 TB の FlexGroup ボリュームに 20 TB (体積の 10%) が残っている場合、1 つのファイルで書き込み可能な容量は 20 TB ではありません。20 TB/[FlexGroup ボリューム内のメンバー・ボリューム数] になります。

2 ノードクラスタでは、両方のノードにまたがる FlexGroup には、通常 16 個のメンバーボリュームがあります。つまり、FlexGroup ボリュームで 20 TB が使用可能な場合、メンバー・ボリュームでは 1.25 TB が使用可能です。ONTAP 9.6 より前のバージョンでは、ボリューム自動拡張を有効にしないと、サイズが 1.25 TB を超えるファイルは FlexGroup ボリュームに書き込めませんでした。

ONTAP 9.6 以降では、エラスティックサイジング機能により、このようなシナリオでの「スペースがない」エラーを回避できます。この機能はデフォルトで有効になっており、管理者による設定や操作は必要ありません。

■ エラスティックサイジング：データ用エアバッグ

FlexGroup のボリューム開発者の一人は、エラスティックサイジングを「エアバッグ」と呼んでいます。これは、事故に遭うのを防ぐためのものではありませんが、事故が起きたときに着陸を軟らかくするのに役立ちます。つまり、大きなファイルの書き込みや容量不足を防ぐことはできませんが、これらの書き込みが完了する方法を提供します。

その仕組みは次のとおりです。

- 1 ファイルが ONTAP に書き込まれると、システムはそのファイルのサイズがどれだけ大きくなるかを認識できません。クライアントにはわかりません。通常、アプリケーションはそれを認識しません。分かっているのは「ファイルに書き込みたい」ということだけです。

- 2 FlexGroup ボリュームが書き込み要求を受け取ると、空き容量、inode 数、ファイルが最後に作成されてからの時間、メンバーのボリューム・パフォーマンス (ONTAP 9.6 の新機能) などのさまざまな要因に基づいて、最も利用可能なメンバーに割り当てられます。
- 3 ファイルが配置されるとき、ONTAP はファイルのサイズを認識しないため、ファイルが使用可能な領域よりも大きくなるかどうか分かりません。したがって、書き込みを許可するスペースがある限り、書き込みは許可されます。
- 4 メンバー・ボリュームの領域が不足した場合、ONTAP が「領域不足」エラーをクライアントに送信する直前に、FlexGroup ボリューム内の他のメンバボリュームに対して照会が実行され、借用可能な領域があるかどうかを確認されます。ある場合、ONTAP が、そのボリュームの総容量の 1%(10 MB ~ 10 GB の範囲) を、フルになっているボリュームに追加し (その一方で同じ FlexGroup ボリューム内の別のメンバー・ボリュームから同じ量を取り出し)、ファイルの書き込みは続行されます。
- 5 ONTAP が借用する領域を探している間、そのファイルの書き込みは一時停止されます。これは、パフォーマンスの問題としてクライアントに現れます。しかし、全体的な目標は、書き込みを迅速に完了させることなく、書き込みを完了させることです。通常、メンバーボリュームは 10 GB の増分 (1 TB の 1% は 10 GB) を提供するのに十分な大きさになり、多くの場合、ファイルの作成を完了するために十分な大きさよりも大きくなります。メンバー・ボリュームが小さいほど、システムはより頻繁にスペースを借りるために照会する必要があるため、パフォーマンスへの影響が大きくなる可能性があります。
- 6 容量の借用により、FlexGroup の全体的なサイズが維持されます。たとえば、FlexGroup ボリュームのサイズが 40 TB の場合、40 TB のままです。

図 7.5 エラスティックサイジング前のファイル書き込み動作

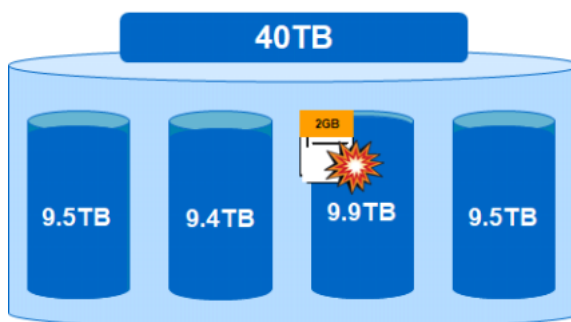
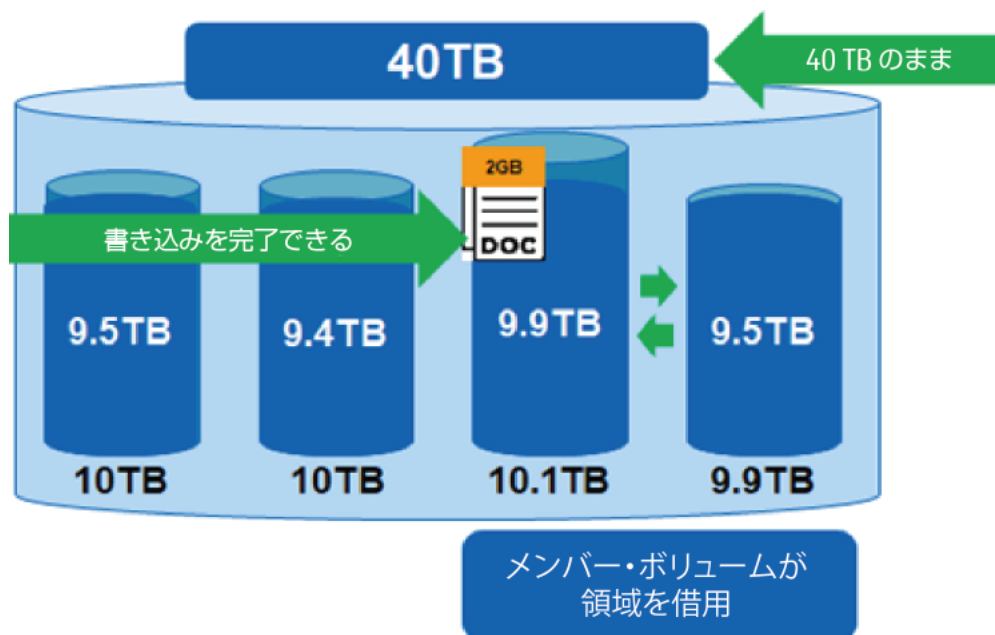


図 7.6 エラスティックサイジング後のファイル書き込み動作



ファイルが削除されるか、ボリュームが拡張され、そのメンバー・ボリューム内の領域が再び使用可能になると、ONTAP は、領域の均一性を維持するために、メンバー・ボリュームを元のサイズに再調整します。

最終的には、エラスティックサイジングによって、管理者のスペース管理のオーバーヘッドを排除し、FlexGroup ボリュームの初期サイジングと導入に関する心配を軽減できます。必要なメンバー・ボリュームの数、ボリュームのサイズなどについて考える時間を短縮できます。

ONTAP 9.6 の伸縮性のあるサイジング機能を自動拡大 / 縮小などの機能と組み合わせると、ONTAP で容量を管理し、緊急時のスペース問題を回避できます。

7.4 FlexVol から FlexGroup のインプレース変換

ONTAP 9.7 では、単一の FlexVol ボリュームを、単一のメンバボリュームを持つ FlexGroup ボリュームに変換できるようになりました。ボリューム内のデータ容量やファイル数に関係なく、変換の中断時間は 40 秒未満です。クライアントを再マウントしたり、データをコピーしたり、メンテナンスウィンドウを作成する可能性のあるその他の変更を行う必要はありません。FlexVol ボリュームを FlexGroup ボリュームに変換したら、新しいメンバボリュームを追加して容量を拡張できます。

7.4.1 FlexVol ボリュームを FlexGroup ボリュームに変換する理由

FlexGroup ボリュームには、FlexVol ボリュームよりも次のような利点があります。

- 1 つのボリュームで 100 TB および 20 億ファイルを超える拡張が可能
- 無停止で容量またはパフォーマンスを拡張可能
- 取り込みの多いワークロード向けのマルチスレッド・パフォーマンス
- ボリューム管理と導入の合理化

たとえば、ワークロードが急速に増大しており、データを移行する必要はないが、より多くの容量を提供する必要があるとします。あるいは、FlexVol ボリュームではワークロードのパフォーマンスがよくないため、FlexGroup を使用してパフォーマンス処理を向上させたい場合もあります。ここでは変換が役立ちます。

FlexVol から FlexGroup への変換の詳細については、この[ブログ](#)、この[ポッドキャスト](#)を参照してください。

7.5 ボリュームの自動サイズ設定 (自動拡張 / 自動縮小)

ONTAP 9.3 で、FlexGroup ボリュームに対するボリューム自動拡張のサポートが追加されました。このサポートにより、ストレージ管理者は FlexGroup ボリュームの自動拡張ポリシーを設定して、ボリュームが最大容量に近づいたときに ONTAP が FlexVol サイズを事前に定義されたしきい値まで増やすことができます。この機能は、FlexGroup ボリュームで特に役立ちます。これは、ボリューム自動拡張によって、FlexGroup ボリューム全体でメンバボリュームが途中でいっぱいになったり、早すぎる「スペースがない」シナリオが発生したりするのを防ぐことができるためです。FlexGroup ボリュームへのボリューム自動拡張の適用は、FlexVol ボリュームの場合と同じ方法で行います。ボリューム自動拡張の適用方法の例については、付録を参照してください。

■ Elastic Sizing との自動サイズ調整の相互作用

ONTAP 9.6 以降、Elastic Sizing では、他のメンバー・ボリュームから領域を借用することによって、ほぼ満杯のメンバー・ボリュームへのファイル書き込みを完了する方法を実現しています。これは、FlexGroup ボリュームの合計サイズを増やすことなく行われます。充填されたメンバー・ボリューム内のスペースが解放されると、エラスティックサイジングによってメンバー・ボリュームのサイズが元の容量への平常化が開始されます。

一方、ボリュームの自動サイズ設定では、メンバー・ボリュームがスペースしきい値に達したときに自動的に拡張されるため、FlexGroup ボリュームの合計サイズにスペースが追加されます。

ボリュームの自動サイズ調整が有効になっている場合、そのボリュームの弾性サイズ調整は無効になります。ボリュームに弾性サイズを使用する場合は、自動サイズ調整を無効にします。

7.6 64 ビットファイル識別子

デフォルトでは、ONTAP の NFS は 32 ビットのファイル ID を使用します。32 ビットのファイル ID は、最大 2,147,483,647 の符号なし整数に制限されます。FlexVol の 20 億 inode 制限では、この値はアーキテクチャにうまく適合します。

しかし、FlexGroup ボリュームは公式には 1 つのコンテナで最大 4000 億個のファイルをサポートできるため (理論的には、それ以上)、64 ビットのファイル ID を実装する必要がありました。64 ビットのファイル ID は、最大 9,223,372,036,854,775,807 個の符号なし整数をサポートします。

既定では、64 ビットファイル識別子オプションは「オフ / 無効」に設定されています。これは、管理者が環境を適切に評価する前に、32 ビットのファイル識別子を必要とするレガシーアプリケーションやオペレーティングシステムがONTAPの変更によって予期せぬ影響を受けないようにするための設計でした。64 ビットファイル ID を有効にする前に、使用しているアプリケーションと OS のベンダーに確認してください。

または、テスト用の SVM を作成して、有効にし、アプリケーションとクライアントが 64 ビットのファイル ID に対してどのように反応するかを確認します。最近のアプリケーションやオペレーティングシステムのほとんどは、64 ビットのファイル ID を問題なく処理できます。

このオプションは、現在、コマンドラインで高度な権限レベルを指定した場合にのみ有効にできます。

```
cluster::> set advanced
cluster::*> nfs server modify -vserver SVM -v3-64bit-identifiers enabled
```

このオプションを有効または無効にした後、すべてのクライアントを再マウントする必要があります。そうしないと、ファイルシステム ID が変更されるため、クライアントが NFS 操作を試みたときに古いファイルハンドルメッセージを受け取ることがあります。ファイル数の多い環境で FSID 変更オプションを有効または無効にすると SVM にどのような影響があるかについての詳細は、[\[■ ファイル数の多い環境での FSID と SVM の動作\] \(P.37\)](#) を参照してください。

FlexGroup ボリュームのファイル数が 20 億を超えない場合は、この値を変更しないでください。ただし、ファイル ID の競合を防ぐために、FlexGroup ボリュームの inode の最大数は、2,147,483,647 以下まで増やす必要があります。

```
cluster::*> vol show -vserver SVM -volume flexgroup -fields files
```

注意

このオプションは SMB 操作には影響しないため、SMB のみを使用するボリュームでは不要です。

■ NFSv3 と NFSv4.x の比較

NFSv3 と NFSv4.x では、異なるファイル ID セマンティクスが使用されます。FlexGroup ボリュームが NFSv4.x をサポートするようになったため、ONTAP 9.7 には 64 ビットファイル ID を有効 / 無効にする 2 つの異なるオプションが用意されています。

64 ビット ID オプションを両方のプロトコルに適用する SVM で、NFSv3 と NFSv4.x の両方を使用する場合は、両方のオプションを設定する必要があります。

1 つのオプションしか設定されておらず、両方のプロトコルがボリュームにアクセスする場合、プロトコル間で望ましくない動作が発生する可能性があります。たとえば、NFSv3 では 20 億を超えるファイルを作成および表示できますが、NFSv4.x ではエラーがスローされます。

オプションは次のとおりです。

```
-v3-64bit-identifiers [enabled/disabled]  
-v4-64bit-identifiers [enabled/disabled]
```

7.6.1 クォータ強制を使用してファイル数を制限する

ONTAP 9.5 以降では、32 ビットのファイルハンドルがクォータ適用のために使用されている場合に、FlexGroup ボリュームが 20 億ファイルを超えないようにクォータポリシーを設定できます。

クォータ・ポリシーは親ボリュームの下位に作成されたファイルには適用されないため、FlexGroup ボリューム内に qtree を作成します。次に、ユーザーが 32 ビットファイル ID の制限を超過するリスクを減らすために、制限として 20 億ファイルの既定のクォータの規則を作成します。

```
cluster::*> qtree create -vserver DEMO -volume FG4 -qtree twobillionfiles -security-style  
unix -oplock-mode enable -unix-permissions 777  
cluster::*> quota policy rule create -vserver DEMO -policy-name files -volume FG4 -type tree  
-target "" -file-limit 2000000000  
cluster::*> quota on -vserver DEMO -volume FG4  
[Job 15906] Job is queued: "quota on" performed for quota policy "tree" on volume "FG4" in  
Vserver "DEMO".  
cluster::*> quota resize -vserver DEMO -volume FG4  
[Job 15907] Job is queued: "quota resize" performed for quota policy "tree" on volume "FG4"  
in Vserver "DEMO".  
cluster::*> quota report -vserver DEMO -volume FG4  
Vserver: DEMO
```

Volume	Tree	Type	ID	----Disk----	Used	Limit	----Files-----	Used	Limit	Quota
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
FG4	twobillionfiles	tree	1		0B	-		1		
								2000000000		twobillionfiles
FG4		tree	*		0B	-		0		
								2000000000		*

2 entries were displayed.

その後、ファイルのアクセス許可を使用してアクセスを制限し、ユーザーがボリュームレベルでファイルを作成できないようにします。SMB 共有をボリュームではなく qtree に適用し、マウントは qtree レベルで実行する必要があります。

その後、qtree でファイルが作成されると、この制限に対してカウントされます。

```
[root@centos7 home]# cd /FG4/twobillionfiles/
[root@centos7 twobillionfiles]# ls
[root@centos7 twobillionfiles]# touch new1
[root@centos7 twobillionfiles]# touch new2
[root@centos7 twobillionfiles]# touch new3
[root@centos7 twobillionfiles]# ls
new1 new2 new3
cluster:.*> quota report -vserver DEMO -volume FG4
Vserver: DEMO
```

Volume	Tree	Type	ID	----Disk----		----Files-----		Quota Specifier
				Used	Limit	Used	Limit	
FG4	twobillionfiles	tree	1	0B	-	4	2000000000	twobillionfiles
		tree	*	0B	-	0	2000000000	*

7.6.2 System Manager での 64 ビットファイル ID オプションのサポート

ONTAP 9.7 では、REST API 機能に基づく新しい System Manager インタフェースが導入されました。64 ビットファイル ID オプションは現在 REST API に存在しないため、System Manager で NFS サーバ・オプションを有効または無効にする際には CLI を使用します。

■ ファイル ID の衝突の影響

64 ビットのファイル ID が有効になっていない場合、ファイル ID が競合するリスクが高くなります。ファイル ID の競合が発生した場合の影響は、クライアント上の「古いファイルハンドル」エラーから、ディレクトリやファイルの一覧表示の失敗、アプリケーション全体の失敗まで、さまざまです。通常、FlexGroup ボリュームを使用する場合は、64 ビットファイル ID オプションを有効にする必要があります。

7.6.3 ONTAP におけるファイルシステム ID (FSID) の変更の影響

NFS は、クライアントとサーバー間の通信にファイルシステム ID (FSID) を使用します。この FSID によって、NFS クライアントは、NFS サーバのファイル・システム内のデータの場所を認識できます。ONTAP はジャンクションパスを使用して複数のノードにわたる複数のファイル・システムにまたがることのできるため、この FSID はデータがどこに存在するかによって変わる可能性があります。一部の古い Linux クライアントでは、これらの FSID の変更を区別する際に問題が発生する可能性があり、その結果、chown や chmod などの基本属性の操作中に障害が発生します。

NFSv3 で FSID の変更を無効にする場合は、ONTAP 9 で 64 ビットファイル識別子を有効にしてください。ただし、このオプションは、32 ビットのファイル ID を必要とする古いレガシーアプリケーションに影響する可能性があることに注意してください。

■ ファイル数の多い環境での FSID と SVM の動作

NFSv3 および NFSv4.x の FSID 変更オプションでは、FlexVol ボリュームと FlexGroup ボリュームに独自のファイルシステムが提供されます。つまり、SVM で許可されるファイルの数はボリュームの数によって決まります。ただし、FSID 変更オプションを無効にすると、32 ビットまたは 64 ビットのファイル識別子が SVM 自体に適用されます。つまり、32 ビットでの 20 億ファイル制限がすべてのボリュームに適用されます。

したがって、SVM のファイル数は、FlexVol や FlexGroup のボリュームではなく、20 億に制限されます。FSID 変更オプションを有効にしておくと、ボリュームは専用のファイル・カウントを持つ独立したファイル・システムとして動作できます。

ファイル ID の競合を防ぐために、FlexGroup ボリュームでは FSID 変更オプションを有効のままにしておくことを推奨します。

■ FSID と Snapshot コピーの動作

ボリュームの Snapshot コピーが作成されると、ファイルの inode のコピーがファイル・システムに保存され、後でアクセスできるようになります。このファイルは、理論的には 2 つの場所に存在します。

NFSv3 では、基本的に同じファイルのコピーが 2 つあっても、それらのファイルの FSID は同じではありません。ファイルの FSID は、WAFL の inode 番号、ボリューム ID、スナップショット ID を組み合わせで作成されます。すべての Snapshot コピーには異なる ID が割り当てられるため、NFSv3 では、オプション `-v3-fsid-change` の設定に関係なく、ファイルの Snapshot コピーごとに異なる FSID が割り当てられます。NFS RFC 仕様では、ファイルの FSID がファイルバージョン間で同一である必要はありません。

注意

NFSv4 は現在 FlexGroup ボリュームでサポートされていないため、`-v4-fsid-change` オプションは FlexGroup ボリュームには適用されません。

7.6.4 ディレクトリ・サイズの考慮事項

ONTAP では、ディスク上のディレクトリの最大サイズに制限があります。この制限は `maxdirsize` と呼ばれます。ボリュームの `maxdirsize` 値の上限は、プラットフォームに関係なく 320 MB です。つまり、ディレクトリー・サイズのメモリー割り当てが最大 320 MB に達すると、ディレクトリーはそれ以上大きくなりません。

■ デフォルトの最大サイズで 1 つのディレクトリに収まるファイル数

デフォルトの `maxdirsize` 設定で 1 つのディレクトリに格納できるファイル数を調べるには、次の式を使用します。

- メモリ (KB) *53*25%

`maxdirsize` の値は、大規模なシステムではデフォルトで 320 MB に設定されているため、SMB と NFS の場合、1 つのディレクトリ内の最大ファイル数は 4,341,760 になります。`maxdirsize` の値をできる限り小さくし、320 MB の制限値の 80% (256 MB (約 340 万ファイル)) を超えないようにすることを強く推奨します。

■ maxdirsize を超えた場合に送信されるイベント管理システム・メッセージ

次のイベント管理システムメッセージは、maxdirsize 値を超えたとき、または超えそうになったときにトリガーされます。警告は maxdirsize 値の 90% で送信され、event log show コマンドまたは ONTAP System Manager event セクションで表示できます。

Active IQ Unified Manager を使用すると、maxdirsize の監視、アラームのトリガー、および 90% 近い値の前の通知が可能です (図 7.7 を参照してください)。

```
Message Name: wafl.dir.size.max
Severity: ERROR
Corrective Action: Use the "volume file show-inode" command
with the file ID and volume name information to find the file path. Reduce the number of files
in the directory. If not possible, use the (privilege:advanced) option "volume modify -volume
vol_name -maxdir-size new_value" to increase the maximum number of files per directory. How-
ever, doing so could impact system performance. If you need to increase the maximum directory
size, work with technical support.
Description: This message occurs after a directory has
reached its maximum directory size (maxdirsize) limit.
Supports SNMP trap: true
Destinations: -
Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0

Message Name: wafl.dir.size.max.warning
Severity: ERROR
Corrective Action: Use the "volume file show-inode" command
with the file ID and volume name information to find the file path. Reduce the number of files
in the directory. If not possible, use the (privilege:advanced) option "volume modify -volume
vol_name -maxdir-size new_value" to increase the maximum number of files per directory. How-
ever, doing so could impact system performance. If you need to increase the maximum directory
size, work with technical support.
Description: This message occurs when a directory has
reached or surpassed 90% of its current maximum directory size (maxdirsize) limit, and the
current maxdirsize is less than the default maxdirsize, which is 1% of total system memory.
Supports SNMP trap: true
Destinations: -
Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0

Message Name: wafl.dir.size.warning
Severity: ERROR
Corrective Action: Use the "volume file show-inode" command
with the file ID and volume name information to find the file path. Reduce the number of files
in the directory. If not possible, use the (privilege:advanced) option "volume modify -volume
vol_name -maxdir-size new_value" to increase the maximum number of files per directory. How-
ever, doing so could impact system performance. If you need to increase the maximum directory
size, work with technical support.
Description: This message occurs when a directory sur-
passes 90% of its current maximum directory size (maxdirsize) limit.
Supports SNMP trap: true
Destinations: -
Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

図 7.7 maxdirsize の警告が表示された ONTAP System Manager のイベント画面

[illegible]

■ maxdirsize 値を増加させる影響

1つのディレクトリに多数のファイルが格納されている場合、検索（「見つける」操作など）はCPUとメモリを大量に消費します。ONTAP 9.2から、「ディレクトリ索引」は、2MBを超えるディレクトリ・サイズ用のインデックス・ファイルを作成して、多くの検索を実行してキャッシュ・ミス回避の必要性を軽減します。通常、これは大きなディレクトリのパフォーマンスを向上させます。ただし、ワイルドカード検索とreaddir操作では、インデックスはあまり使用されません。

■ FlexGroup ボリュームは maxdirsize の制限を回避するか？

FlexGroup ボリュームでは、各メンバボリュームに同じ maxdirsize 設定があります。ディレクトリが複数の FlexVol メンバボリュームおよびノードにまたがる可能性がある場合でも、maxdirsize のパフォーマンスへの影響が発生する可能性があります。これは、ディレクトリサイズが重要なコンポーネントであり、個々の FlexVol ボリュームではないためです。したがって、ディレクトリのサイズは依然として問題になります。したがって、FlexGroup ボリュームでは、maxdirsize の制限に直面している環境にとって軽減策にはできません。新しいプラットフォームではより多くのメモリと CPU が提供され、ETERNUS AX システムではパフォーマンスが向上しますが、多数のファイルがあるディレクトリでパフォーマンスへの影響を軽減する最善の方法は、ファイルシステム内のより多くのディレクトリにファイルを分散することです。

■ maxdirsize 値を超えた場合の影響

ONTAP で maxdirsize 値を超えると、「スペースがない」エラー (ENOSPC) がクライアントに発行され、イベント管理システム・メッセージがトリガされます。この問題を解決するには、ストレージ管理者が maxdirsize の設定値を増やすか、ディレクトリからファイルを移動する必要があります。

第 8 章

FlexGroup の特長

ONTAP FlexGroup の機能は、次のカテゴリに分類されます。

- シンプルさ
- データ保護
- ストレージの効率性

さらに、FlexGroup の機能とメリットは、それ自体が機能であり、[「第 3 章 ONTAP FlexGroup のメリット」\(P.13\)](#) で確認できます。

8.1 シンプルさ

FlexGroup ボリュームの主なメリットの 1 つは、通常の FlexVol ボリュームと同じように容易に優れたパフォーマンスを提供する大容量コンテナを作成して容量を確保できることです。FlexGroup は、ONTAP System Manager、Active IQ Performance Manager、および volume create -auto-provision-as flexgroup などの CLI の自動化コマンドをサポートしており、FlexGroup ボリュームの迅速かつ簡単な導入が可能になっています。

8.1.1 コマンドライン (CLI)

ほとんどの人は単純さを考えるときに GUI を思い浮かべますが、コマンドラインもタスクをより簡単に行える場所です。FlexGroup CLI は、全体的なユーザビリティを向上させるいくつかの方法を提供します。

■ ボリューム作成 (-auto-provision-as flexgroup を使用)

ONTAP 9.2 では、管理者権限レベルで FlexGroup ボリュームをプロビジョニングするためのボリュームオプションが導入されました。このオプションを指定すると、デフォルトでノードごとに 8 つのメンバー FlexVol ボリュームが作成されます。サイズが指定されていない場合は、それぞれ 200 MB のメンバー FlexVol ボリュームが作成されるため、コマンドでサイズを指定することが重要です。式は (8 * 指定されたアグリゲート数 / 指定された合計サイズ) であることに注意してください。

メンバーボリュームは 100 GB 以上でなければならず、100 TB を超えることはできないため、これは重要です。

アグリゲートが指定されていない場合、ONTAP は指定された SVM で使用可能なすべてのアグリゲートを選択しようとします。したがって、単純化されたコマンドを実行することは可能ですが、FlexGroup の設定については、できるだけ規定に従うことをお勧めします。

少なくとも、次のように指定します。

- FlexGroup としての自動プロビジョニング (-auto-provision-as)
- ボリューム名 (-volume)
- SVM (-vserver)
- ボリューム・サイズ (-size)
- エクスポート / マウントポイント (-junction-path)

オプションで次を指定します。

- ボリュームセキュリティスタイル (-security-style {unix|ntfs|mixed})
- UNIX 権限 (-unix-permissions (セキュリティ・スタイルが UNIX の場合))
- シン・プロビジョニング (-space-guarantee none)

■ ボリューム作成 (詳細設定)

ベストプラクティス以外のカスタマイズが必要な場合 (必要なメンバー・ボリュームが少ない / 多い場合など) は、FlexGroup ボリュームの作成に使用するコマンドが、volume create -auto-provision-as flexgroup ではない可能性があります。クラスタに 4 つ以上のノードがある場合、または FlexGroup 構成メンバーの設計と配置をより細かく制御する場合は、-auto-provision-as オプションを指定せずに volume create コマンドを使用します。FlexGroup の作成に固有のオプションがいくつか追加されました。

表 8.1 FlexGroup で使用するボリュームコマンドオプション

ボリュームオプション	機能
-aggr-list	FlexGroup 構成要素に使用するアグリゲート名の配列を指定します。リスト内の各エントリは、指定されたアグリゲートに構成要素を作成します。アグリゲートは、複数の構成要素を作成するために複数回指定できます。このパラメータは、FlexGroup にのみ適用されます。
-aggr-list-multiplier	FlexGroup ボリュームの作成時に、-aggr-list パラメータでリストされたアグリゲートを繰り返す回数を指定します。アグリゲートリストは指定した回数だけ繰り返されます。
-max-constituent-size	オプションで、FlexGroup 構成要素の最大サイズを指定します。デフォルト値は、FlexGroup ボリュームによって使用されるすべてのノードの最大 FlexVol サイズ設定をチェックすることによって決定されます。見つかった最小値が、FlexGroup 構成要素の最大構成要素サイズのデフォルトとして選択されます。このパラメーターは、FlexGroup ボリュームにのみ適用されます。

■ ボリューム修正

FlexGroup ボリュームの作成後、CLI コマンドの volume modify または ONTAP System Manager GUI を使用して、ボリュームオプションまたはサイズを変更する必要があります。

■ ボリューム拡張

FlexGroup ボリュームを管理するもう 1 つのコマンドは、`volume expand` です。このコマンドを使用すると、ストレージ管理者は `-aggr-list` および `-aggr-list-multiplier` オプションを使用して、既存の FlexGroup ボリュームにメンバーボリュームを追加できます。メンバーを追加するアグリゲートと、アグリゲートごとに必要なメンバーの数を指定するだけです。残りは ONTAP が行います。

SnapMirror 関係におけるボリュームのボリューム拡張

ONTAP 9.3 より前の SnapMirror 関係に参加している FlexGroup ボリュームでは、SnapMirror 関係の再ベースラインが必要なため、`volume expand` コマンドはネイティブでは機能しません。ONTAP 9.3 では、再ベースラインを行わなくても、SnapMirror 関係に参加している FlexGroup ボリュームのボリューム拡張を可能にする機能拡張が導入されました。ONTAP 9.3 の時点では、ONTAP は次の SnapMirror 更新時に FlexGroup メンバーのボリューム数を調整します。

注意

富士通では、FlexGroup ボリュームで SnapMirror を使用する場合、ONTAP 9.3 以降にアップグレードすることを推奨しています。

ONTAP 9.3 より前の SnapMirror 関係における FlexGroup ボリュームの拡張

ONTAP 9.3 より前の SnapMirror 関係でボリュームを拡張 (メンバーを追加) するには、次の手順に従います。

手順 ▶▶▶

- 1 デスティネーション上の既存の関係に対して `snapmirror delete` を実行します。
- 2 ソースで `snapmirror release` を実行します。
- 3 デスティネーション FlexGroup データ保護ボリュームで `volume delete` を実行します。
- 4 ソース FlexGroup ボリュームで `volume expand` を実行します。
- 5 `volume create` を使用して、ソース FlexGroup ボリュームと同じサイズと構成要素数で新しいデスティネーション FlexGroup データ保護ボリュームを作成します。
- 6 新しいリレーションシップに対して `snapmirror initialize` を実行します (再ベースライン)。

関係を再ベースライン化することなくメンバボリュームを拡張することは、SnapMirror および FlexGroup でサポートされています。

8.1.2 ONTAP System Manager

ONTAP 9.1 は、ONTAP System Manager で FlexGroup をすぐにサポートしました。GUI の「Volumes」ページに「FlexGroup」タブが追加されました。このページでは、ストレージ管理者は、2 回クリックするだけで、既存の FlexGroup ボリュームを管理したり、FlexGroup ボリュームを作成したりできます。ONTAP 9.4 は、FlexGroup ボリュームをサポートするためのより堅牢な GUI を備え、FlexVol ボリュームと同じタスクをほぼすべて実行できる機能を備えています。

■ FlexGroup ボリュームの作成

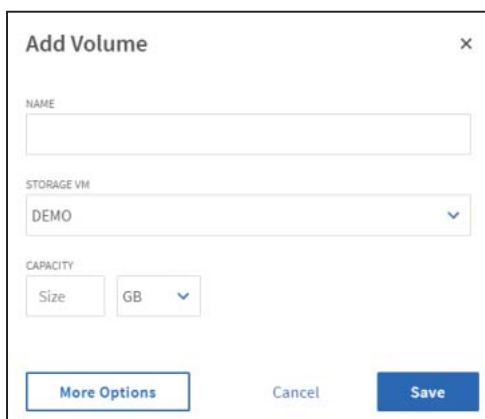
ONTAP 9.7 では、ONTAP System Manager の再設計が導入され、設定操作が簡素化されました。CLI または REST API を使用して高度な設定を行います。

■ FlexGroup ボリュームの作成 (ONTAP 9.7 以降)

ONTAP 9.7 では、ONTAP System Manager の再設計が導入され、設定操作が簡素化されました。この節では、ONTAP System Manager 9.7 以降で FlexGroup ボリュームを作成する方法について説明します。

手順 ▶▶▶

- 1 ボリュームは、ONTAP System Manager の複数の場所から作成できます。以前のバージョンでは、ウィザードの起動時に FlexVol または FlexGroup を指定していました。現在は、両方のボリューム・タイプに対して 1 つのウィザードしかありません。基本的なウィザードは name/SVM/size のように単純ですが、デフォルトでは FlexVol ボリュームが作成されます。FlexGroup ボリュームを作成するには、[More Options] ボタンをクリックします。



- 2 [More Options] を選択したら、サイズ、共有、エクスポートポリシー、データ保護などの使用可能なオプションを使用して、ボリュームを構成します。ボリュームを FlexGroup ボリュームにするには、[Optimization Options] の [Distribute Volume Data Across the Cluster] チェックボックスを使用します。

3 [Save] をクリックします。

■ FlexGroup ボリュームの管理

FlexGroup ボリュームの作成が完了すると、CIFS 共有を作成するか、[Done] をクリックして処理を終了するかを選択できるダイアログ・ボックスが開きます (図 8.1)。

図 8.1 ONTAP System Manager で FlexGroup ボリュームへの共有を作成する

ここから、管理者は「FlexGroup」タブから FlexGroup ボリュームを管理できます。ONTAP System Manager には、FlexGroup ボリュームの概要が表示されます (図 8.2)。

- ボリュームの概要
- スペースの割り当て
- データ保護ステータス (SnapMirror)
- 現在のパフォーマンス統計

図 8.2 System Manager での FlexGroup の概要

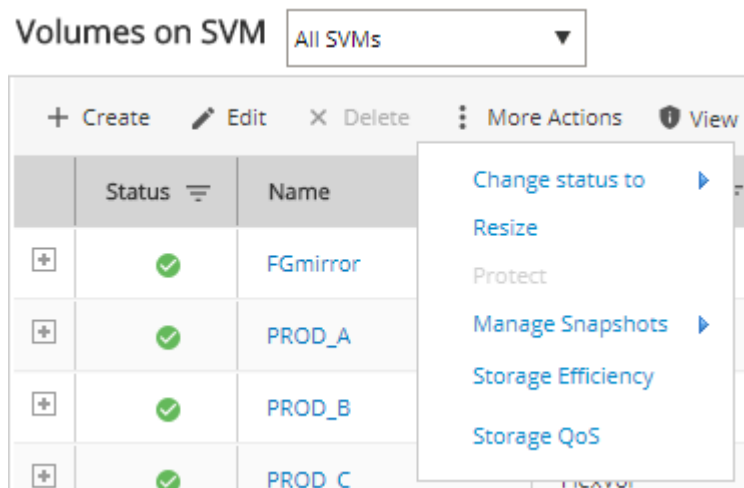


詳細については、ハイパーリンクが設定されたボリューム名または [Show More Details] をクリックしてください。このビューには、次のような情報が表示されます。

- ボリュームの概要
- Snapshot コピー
- データ保護の詳細
- ストレージ効率の詳細
- パフォーマンスの詳細 (リアルタイムのみ)

また、FlexGroup ボリュームは、System Manager GUI の [Edit] および [More Actions] ボタンを使用して管理できます (図 8.3)。

図 8.3 既存の FlexGroup ボリュームの管理



8.1.3 Active IQ Performance Manager

ONTAP System Manager のサポートに加えて、Active IQ Performance Manager を使用して FlexGroup ボリュームとそのメンバーを詳細レベルで監視できます。

Performance Manager では、SVM 内の他のボリュームとともに FlexGroup ボリュームを見つけることができます。目的のオブジェクトをクリックすると、メンバボリュームの概要を示す画面が表示されます。これらのメンバは、指定した時間範囲でグラフィカルビューに追加することもできます (図 8.4)。

図 8.4 Performance Manager の FlexGroup メンバー・ボリューム

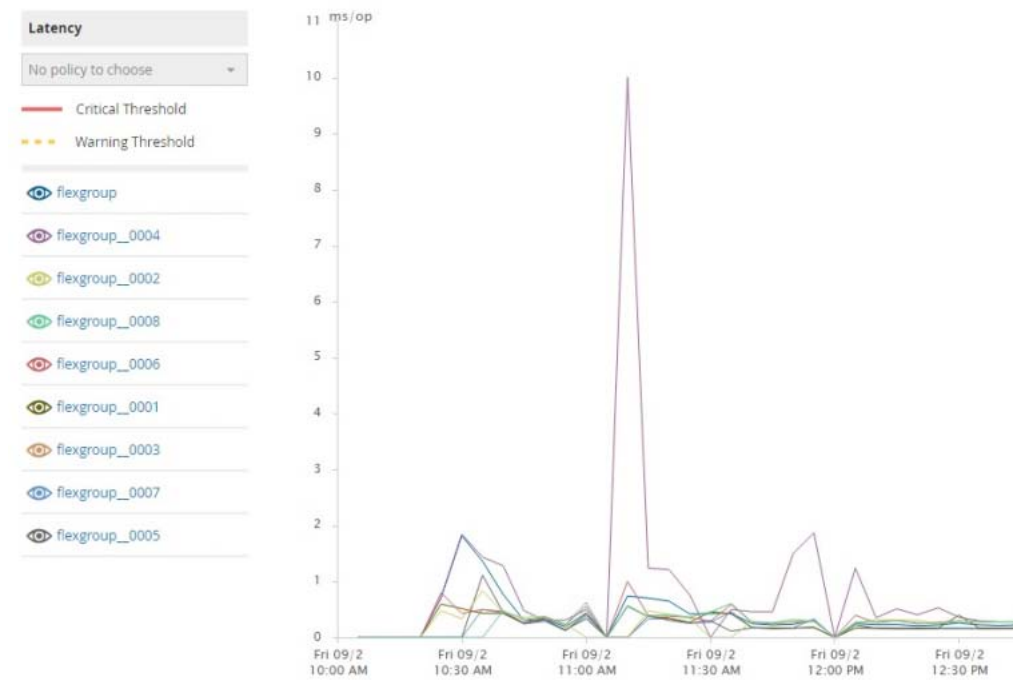
View and compare		Constituents of this FlexGroup		Filtering	
Volume	Latency	IOPS	MBps		
flexgroup_0004	0.311 ms/op	56.9 IOPS	< 1 MBps	Add →	
flexgroup_0002	0.273 ms/op	49.4 IOPS	< 1 MBps	Add →	
flexgroup_0008	0.271 ms/op	107 IOPS	< 1 MBps	Add →	
flexgroup_0006	0.27 ms/op	103 IOPS	< 1 MBps	Add →	
flexgroup_0001	0.164 ms/op	9.12 IOPS	< 1 MBps	Add →	
flexgroup_0003	0.161 ms/op	64.9 IOPS	< 1 MBps	Add →	
flexgroup_0007	0.159 ms/op	82.8 IOPS	< 1 MBps	Add →	
flexgroup_0005	0.159 ms/op	47.5 IOPS	< 1 MBps	Add →	

Performance Manager のグラフィカル・ビューに追加されたメンバー・ボリュームには、グラフ内で区別するために異なるグラフの線の色が割り当てられます。このようにして、期待されるパフォーマンス出力から逸脱しているすべてのメンバを調査し、修正することができます (図 8.5 および 図 8.6)。

図 8.5 Performance Manager での FlexGroup メンバー・ボリュームのグラフィック表示



図 8.6 Performance Manager での FlexGroup メンバー・ボリュームのグラフィック表示 (拡大表示)



8.1.4 REST API

REST API のサポートは ONTAP 9.6 で導入されました。独自のインターフェース (Manageability SDK など) をナビゲートするのではなく、REST API を使用することで、クラスターへのアクセスと対話に汎用標準を使用することができます。

REST API のドキュメントは [http://\[your_cluster_IP_or_name\]/docs/api](http://[your_cluster_IP_or_name]/docs/api) にあり、独自の REST API を生成できる例と対話式「試してみる」機能を提供しています。

8.1.5 Cloud Volumes ONTAP

ONTAP 9.6 は、Cloud Volumes ONTAP (クラウドで動作する ONTAP ソリューション) の公式サポートを導入しました。これは、FlexGroup ボリュームを Cloud Volumes ONTAP にデプロイできることを意味します。

Cloud Volumes ONTAP で動作する FlexGroup ボリュームは、Cloud Volumes ONTAP インスタンスにデプロイされた ONTAP バージョンと同じ機能セットを使用することができます。Cloud Volumes ONTAP と FlexGroup に見られる一般的なユースケースには、次のようなものがあります。

- 分析のためのデータレイク
- Amazon Elastic Compute Cloud (Amazon EC2) 計算インスタンスで使用する EDA リポジトリ
- オンプレミス SnapMirror で使用するデータバックアップ / アーカイブ

FlexGroup ボリュームは、オンプレミスでの導入のために 1 つのネームスペースで複数ペタバイトをサポートできますが、Cloud Volumes ONTAP インスタンスの最大容量はインスタンスあたり 368 TB で、FlexGroup ボリュームは複数のインスタンスにまたがることはできません。また、FlexGroup を作成するには、現在 System Manager または CLI を使用する必要があります。現時点では、Cloud Manager で FlexGroup ボリュームを作成する方法はありません。

8.1.6 単一の透過的な名前空間

FlexGroup は、複雑なアーキテクチャを実装することなく、通常の FlexVol ボリュームを超える大容量のメリットを提供します。ボリューム全体を 1 つのエクスポートまたは共有としてマウントできるため、ストレージを追加しても、アプリケーション側で変更を加える必要がありません。このメリットにより、クライアント側から多数のコンテナと多数のマウント・ポイントまたは共有を管理する際の管理の複雑さが軽減されます。

8.1.7 Qtree

ONTAP 9.3 では、qtree と呼ばれる論理ディレクトリに対する FlexGroup ボリュームのサポートが導入されました。Qtree を使用すると、ストレージ管理者は ONTAP GUI または CLI からフォルダを作成して、大きなバケット内でデータを論理的に分離できます。Qtree は、データにアクセスするユーザーのユーザー名を反映するようにフォルダに名前を付けることができ、ユーザー名に基づいてアクセスを提供するように動的共有を作成できるため、ホームディレクトリのワークロードに役立ちます。Qtree は、通常のフォルダとほぼ同じ方法で FlexGroup ボリューム全体に分散されます。クォータの

監視は qtree レベルで適用できます。ONTAP 9.5 以降では、クォータ強制ポリシーを適用できます。Qtree は、FlexVol qtree が管理されるのと同じ方法で作成および管理されます。FlexGroup ボリュームごとに最大 4,995 の qtree がサポートされます。

■ クォータの適用

クォータの適用が qtree またはユーザーに対して有効になっている場合、ONTAP は、クォータを超えた後の新しいファイルの作成や書き込みを禁止します。さらに、EMS メッセージは DEBUG 重大度レベルでログに記録され、ストレージ管理者にクォータ違反を通知します。これらの EMS メッセージについては、システムが SNMP トラップまたは syslog メッセージとして転送するように設定できます。

8.2 統合されたデータ保護

FlexGroup は、RAID DP ソフトウェア、RAID Triple Erasure Coding (RAID-TEC テクノロジー)、スナップショット・テクノロジー、SnapMirror レプリケーション・テクノロジー、NFS または CIFS でマウントされたテープ・バックアップなど、いくつかのデータ保護方法をサポートしています。

8.2.1 RAID DP および RAID Triple Erasure Coding (RAID-TEC)

RAID DP は「二重パリティ」RAID と呼ばれ、RAID グループごとに同時に 2 つのディスク障害に耐えることができます。つまり、ドライブに障害が発生しても、データは別のパリティドライブで保護されます。

RAID Triple Erasure Coding (RAID-TEC) は、ONTAP 9.0 で新しく導入されたもので、大容量ドライブを使用する RAID グループに追加のパリティドライブを提供します。この機能は、大容量ドライブの再構築時間が長い場合に、ドライブ障害からの保護に役立ちます。また、RAID-TEC は、ドライブ番号の点からも、より大きな RAID グループを提供します。

FlexGroup では、すべての RAID 保護機能がサポートされています。

8.2.2 Snapshot テクノロジー

Snapshot コピーは、自動的にスケジュール設定されるポイント・イン・タイム・コピーであり、作成時に領域を消費せず、パフォーマンスのオーバーヘッドも生じません。Snapshot コピーは、アクティブなファイル・システムに対する変更のみが書き込まれるため、時間の経過とともに最小限のストレージ領域を消費します。個々のファイルおよびディレクトリを任意の Snapshot コピーから容易にリカバリでき、ボリューム全体を数秒で任意のスナップショット状態にリストアできます。

Snapshot コピーは、FlexGroup での使用がサポートされています。各 Snapshot コピーは、FlexVol メンバーのコンシステンシ・グループとして作成されます。このグループでは、FlexGroup ボリューム内のすべてのメンバーの整合性のあるポイント・イン・タイム・コピーを確保するために、すべてのメンバーが静止し、Snapshot コピーの準備が行われます。

FlexGroup ボリューム内のいずれかのメンバーがスナップショット管理を実行できない場合 (容量不足、オフライン、ビジーのため完了できません)、FlexGroup Snapshot コピー全体が部分的であると

見なされ、障害が発生しています。ONTAP は、試行された Snapshot コピーの残りをクリーンアップし、EMS イベント (図 8.7) を発行します。

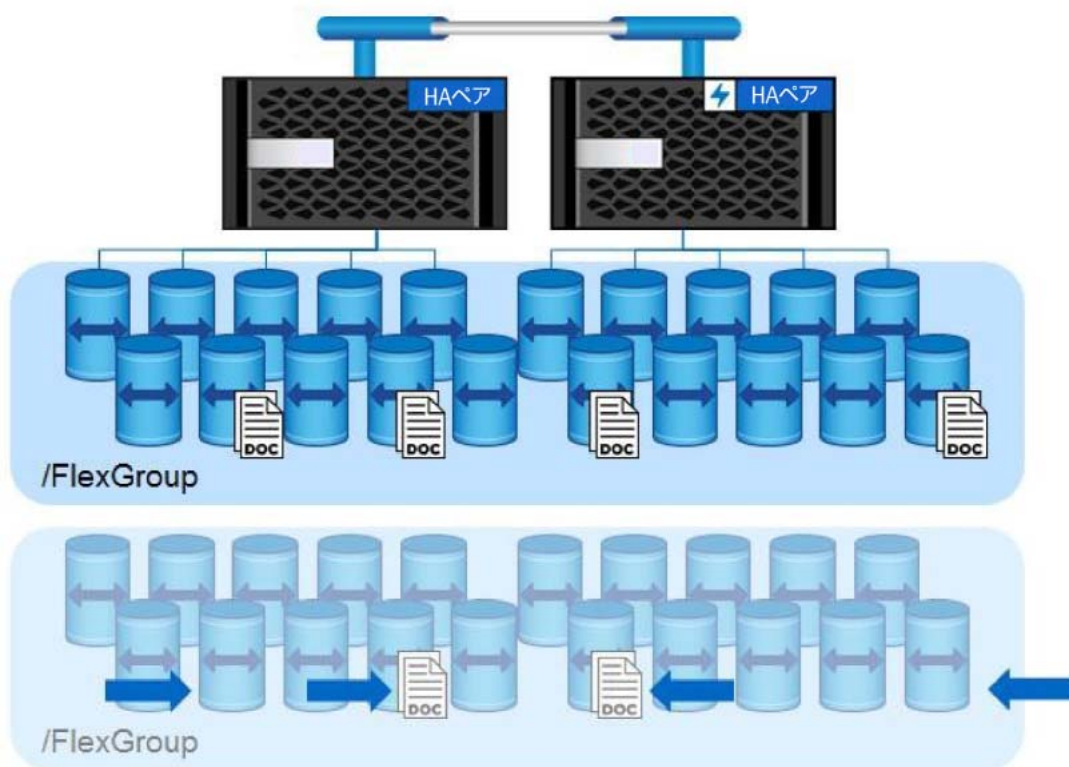
Snapshot コピーの作成後、ストレージ管理者はリストア用に次のオペレーションを実行できます。

- SnapRestore テクノロジーを使用して、FlexGroup ボリューム全体をリストアします。
- .snapshot(NFS) または ~snapshot(CIFS/SMB) フォルダを使用して、スナップショットディレクトリを移動して、個々のファイルとフォルダをリストアすることができます。

注意

単一ファイル SnapRestore はサポートされておらず、単一の FlexGroup メンバーボリュームをリストアすることもできません。FlexGroup ボリュームに対して SnapRestore を使用すると、単一のメンバーボリュームだけでなく、ボリューム全体がリストアされます。現在、フルスナップショットリストアを使用できるのは、diag 権限がある場合のみです。

図 8.7 FlexGroup Snapshot コピー



8.2.3 SnapMirror と SnapVault

SnapMirror は、データ保護および災害復旧のためのプロトコルとは無関係に、ボリュームの非同期レプリケーションを提供します。SnapVault は、データ保護およびバックアップ / アーカイブの用途に非同期 Snapshot コピー保存機能を提供します。

FlexGroup による SnapMirror のサポートは、ONTAP 9.1 以降で提供されています。FlexGroup ボリュームの SnapVault サポートは、ONTAP 9.3 で追加されました。

SnapMirror と SnapVault は、Snapshot コピーと同様に機能します。すべてのメンバー・ボリュームに正常な Snapshot コピーを作成し、すべてのメンバー・ボリュームを DR サイトに同時にレプリケートする必要があります。このオペレーションのいずれかのコンポーネントに障害が発生すると、SnapMirror によるミラーリングも失敗します。

FlexGroup ボリュームによる SnapMirror および SnapVault のベストプラクティスと制限の詳細については、「FUJITSU Storage ETERNUS AX/HX series FlexGroup ボリュームデータ保護のベストプラクティス」を参照してください。

8.2.4 CIFS/SMB または NFS によるテープ・バックアップ

FlexGroup ボリュームのテープバックアップは、Commvault Simpana や Symantec NetBackup などの外部バックアップアプリケーションを使用して、CIFS または NFS マウント経由で実行できます。ONTAP 9.7 以降のバージョンでは、NDMP もサポートされています。FlexGroup ボリューム上の NDMP のパフォーマンスは、FlexVol ボリュームのパフォーマンスと類似していると予想されます。バックアップ・サービス・レベル目標 (SLO) を維持するためにバックアップされるファイルの総数に注意してください。必要に応じて、フォルダレベルでバックアップすることにより、バックアップジョブをより小さなチャンクに分割します。

FlexGroup ボリュームでの NDMP サポートの詳細については、「FUJITSU Storage ETERNUS AX/HX series FlexGroup ボリュームデータ保護ベストプラクティス」を参照してください。

8.2.5 MetroCluster

ONTAP 9.6 では、MetroCluster デプロイメントでの FlexGroup のサポートが導入されています (ファイバ・チャンネルと IP)。

MetroCluster ソフトウェアは、アレイ・ベースのクラスタリングと同期レプリケーションを組み合わせ、継続的な可用性とデータ消失ゼロを最小のコストで実現するソリューションです。MetroCluster を使用する FlexGroup ボリュームには、記載されている制限や注意事項はありません。

MetroCluster の詳細は、「FUJITSU Storage ETERNUS AX/HX series MetroCluster ソリューションの設計とアーキテクチャ」を参照してください。

8.3 ストレージの効率性

FlexGroup は、次のストレージ効率化テクノロジーもサポートしています。

- ・ **インラインおよびポスト・プロセスの重複排除**では、プライマリ・ストレージとセカンダリ・ストレージの重複データ・ブロックが削除され、一意のブロックのみが保存されます。これにより、ストレージ容量とコストを節約できます。重複排除はカスタマイズ可能なスケジュールで実行されます。

- **インラインアグリゲート重複排除 (またはボリューム間の重複排除)** は、アグリゲートレベルでのインライン・ストレージ効率を提供します。これにより、重複ブロックが同じアグリゲート内の複数の FlexVol ボリュームに存在する場合に、重複ブロックを削減できます。FlexGroup メンバボリュームは、ONTAP 9.2 で導入されたこの機能の優れたユースケースです。ONTAP 9.3 自動スケジューリングとスケジュールされたバックグラウンドのインラインアグリゲート重複排除。
- ONTAP 8.3 .1 では、データベースやデスクトップの仮想化などのプライマリワークロード向けに**インライン適応型圧縮**が導入されました。インライン圧縮は、8.3 .1 以降の ETERNUS AX 製品ファミリではデフォルトでオンになっています。
- **インラインデータコンパクション**は ONTAP 9.0 で導入され、データの格納に必要な物理的な使用スペースをさらに削減します。データコンパクションは、当社のストレージ効率ポートフォリオの重要な追加機能であり、重複排除および圧縮テクノロジーを補完します。データコンパクションでは、通常は物理ストレージ上の 4 K ブロックを消費する I/O を取得し、そのような複数の I/O を、1 つの物理 4 K ブロックにパックします。
- **シン・プロビジョニング**は、何年も前から存在しており、ストレージ管理者は物理ストレージ (アグリゲート) 上の仮想コンテナ (FlexVol ボリューム) を過剰にプロビジョニングできます。FlexGroup を使用すると、シンプロビジョニングは、FlexGroup ボリュームの初期導入の処理方法において重要な役割を果たすことができます。シンプロビジョニングでは、メンバー構成要素を物理的なアグリゲート対応物よりもはるかに大きくすることができ、コンテナの設計に柔軟性を持たせることができます。

これらの機能は、ONTAP によってメンバボリュームレベルで個別に適用されますが、ストレージ管理者によって FlexGroup レベルで設定されるため、管理が容易です。ONTAP の以前のリリースでは、これらの機能は FlexVol メンバレベルで適用されていました。そのため、[表 8.2](#) に、ONTAP のどのバージョンが FlexGroup レベルでこれらの機能の管理をサポートしているか、および ONTAP のどのバージョンがメンバボリュームあたりの効率をより細かく管理する必要があるかについてのガイダンスを示します。

注意

富士通では、FlexGroup ボリュームで最大のストレージ効率を得るために、ONTAP 9.3 以降の使用を強く推奨しています。

表 8.2 ONTAP バージョンでの FlexGroup のストレージ効率に関するガイダンス

	9.1RC1	9.1 RC 2以降	9.1 RC 2以降
シン・プロビジョニング	FlexGroupレベル	FlexGroupレベル	FlexGroupレベル
インライン重複排除	FlexVolメンバー	FlexGroupレベル	FlexGroupレベル
後処理重複排除	FlexVolメンバー	FlexGroupレベル	FlexGroupレベル
インラインデータコンパクション	FlexVolメンバー	FlexGroupレベル	FlexGroupレベル
インラインデータ圧縮	FlexVolメンバー	FlexGroupレベル	FlexGroupレベル
後処理データ圧縮	FlexVolメンバー	FlexGroupレベル	FlexGroupレベル
アグリゲートインライン重複排除	該当なし	該当なし	FlexGroupレベル

■ FlexGroup メンバーボリュームごとのストレージ効率の適用

FlexGroup ボリュームが FlexGroup レベルでストレージ効率を有効にする機能を現在サポートしていない場合は、次のコマンドを使用してすべての FlexVol メンバーでストレージ効率を有効にします。これは ONTAP 9.1 RC 1 でのみ必要です。

```
cluster::*> volume efficiency on -vserver SVM -volume flexgroup4*
Efficiency for volume "flexgroup4TB__0001" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0002" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0003" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0004" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0005" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0006" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0007" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0008" of Vserver "SVM" is enabled.
```

変更するには：

```
cluster::*> volume efficiency modify -vserver SVM -volume flexgroup4* -compression true -
data-compaction true -inline-compression true -inline-dedupe true

cluster::*> volume efficiency show -vserver SVM -volume flexgroup4* -fields data-compac-
tion,compression,inline-compression,inline-dedupe
vserver volume                compression inline-compression inline-dedupe data-compaction
-----
SVM      flexgroup4TB__0001 true          true          true          true
SVM      flexgroup4TB__0002 true          true          true          true
SVM      flexgroup4TB__0003 true          true          true          true
SVM      flexgroup4TB__0004 true          true          true          true
SVM      flexgroup4TB__0005 true          true          true          true
SVM      flexgroup4TB__0006 true          true          true          true
SVM      flexgroup4TB__0007 true          true          true          true
SVM      flexgroup4TB__0008 true          true          true          true
```

8.3.1 FabricPool

ONTAP 9.2 では、SSD アグリゲート上のコールドデータブロックをクラウドあるいはオンプレミスの Amazon Simple Storage Service (Amazon S3) オブジェクトストレージに自動的に階層化する機能が、FlexVol ボリューム用に追加されました。この機能により、ストレージ管理者は、よりコストのかかる SSD をアクティブなワークロード用に保持できるようになりました。一方、コールドデータまたは未使用のデータは、よりコスト効率の高いクラウド階層に移動されました。この機能は、FabricPool と呼ばれます。この機能の詳細については、「FabricPool のベストプラクティス」を参照してください。

ONTAP 9.5 では、FlexGroup ボリュームの FabricPool がサポートされるようになりました。FlexGroup ボリュームについて考慮すべき特別な事項はありません。同じ FlexVol の考慮事項が適用されます。

8.4 保存状態での暗号化

ONTAP 9.2 では、FlexGroup ボリューム用に Volume Encryption (VE) がサポートされています。この機能を FlexGroup ボリュームに実装する場合は、FlexVol ボリュームに記載されているものと同じ推奨事項とベストプラクティスに従います。ONTAP 9.5 以降では、既存の FlexGroup ボリュームの「鍵を変更」できます。ONTAP 9.5 より前のバージョンで既存の FlexGroup ボリュームを暗号化するには、暗号化を有効にしたボリュームを作成し、そのボリュームにファイルレベルでデータをコピーする必要があります。これを行うには、XCP Migration Tool などのユーティリティを使用します。

一般的に、VE には次のものがが必要です。

- 有効な VE ライセンス
- 鍵管理サーバー (ONTAP 9.3 時点のオンボックスまたはオフボックス)
- クラスタ全体のパスフレーズ (32 から 256 文字)
- AES-NI オフロードをサポートする ETERNUS AX/HX ハードウェア
- ONTAP 9.5 以降 (既存の FlexGroup ボリュームのキー変更用)

FlexGroup および FlexVol ボリュームを使用した VE の実装と管理については、富士通マニュアルサイトにある、「FUJITSU Storage ETERNUS AX/HX series Encryption パワー ガイド」および「FUJITSU Storage ETERNUS AX/HX series FlexGroup ボリュームを使用した 拡張性とパフォーマンス パワー ガイド」を参照してください。

ONTAP 9.6 では Aggregate Encryption (AE) が追加され、アグリゲートレベルでの暗号化が可能になりました。メンバー・ボリュームを含むすべてのアグリゲートが暗号化されていれば、FlexGroup ボリュームは AE を使用できます。

8.5 サービス品質 (QoS)

ONTAP 9.3 以降では、最大ストレージ QoS ポリシーを FlexGroup ボリュームに適用して、FlexGroup ボリュームが ONTAP で過度のワークロードとして機能しないようにすることができます。ストレージ QoS は、パフォーマンス目標の達成に関するリスクの管理に役立ちます。ストレージ QoS を使用して、スループットをワークロードに制限し、ワークロードのパフォーマンスを監視します。ワークロードを事前に制限してパフォーマンスの問題に対処したり、ワークロードを事前に制限してパフォーマンスの問題を回避したりできます。

8.5.1 FlexGroup ボリュームでのストレージ QoS の最大値の動作

FlexGroup ボリュームでは、ストレージ QoS ポリシーが FlexGroup ボリューム全体に適用されます。FlexGroup ボリュームには複数の FlexVol メンバーボリュームが含まれ、複数のノードにまたがることのできるため、QoS ポリシーは、クライアントがストレージ・システムに接続するときにノード間で均等に共有されます。[図 8.8](#) と [図 8.9](#) は、クラスタ内の複数のノードにまたがる FlexGroup ボリュームにストレージ QoS がどのように適用されるかを示しています。

図 8.8 FlexGroup ボリュームのストレージ QoS: シングルノード接続

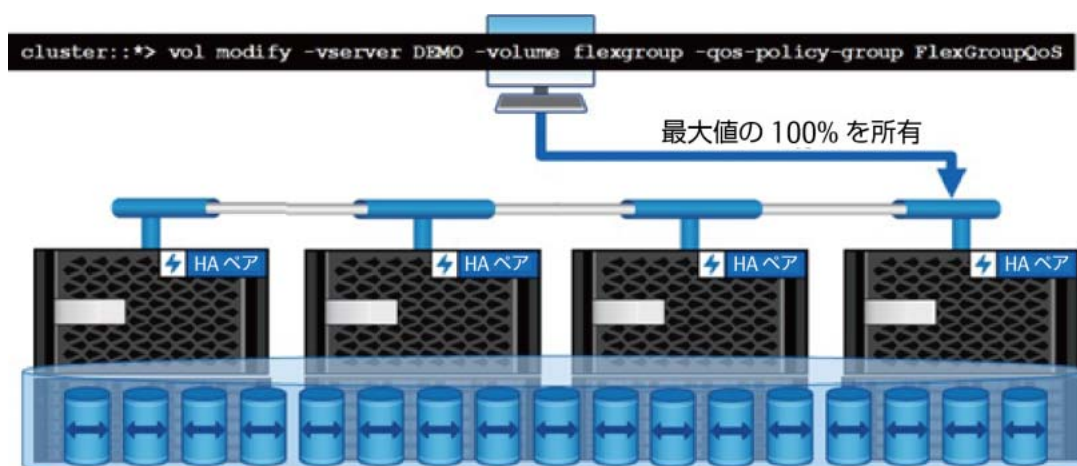
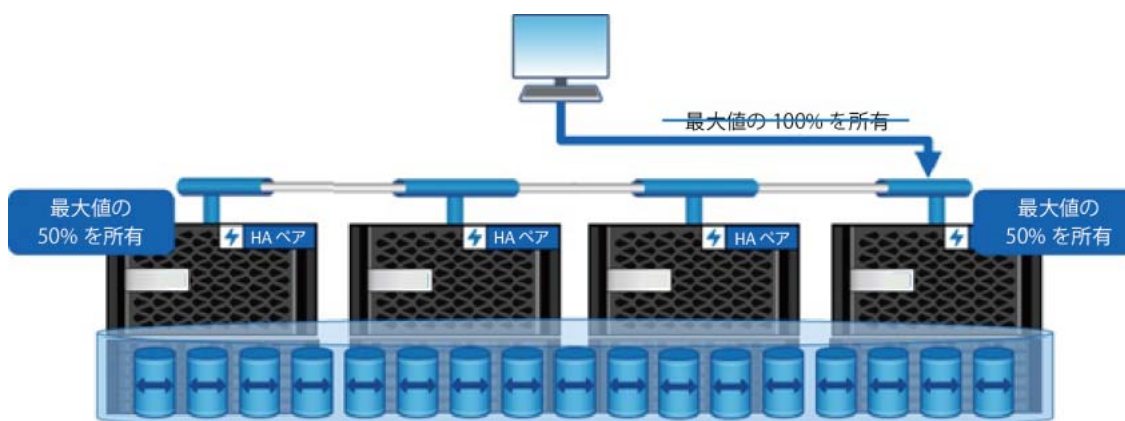


図 8.9 FlexGroup ボリュームのストレージ QoS: マルチノード接続



■ FlexGroup ボリュームでのストレージ QoS に関する考慮事項

現在、ストレージ QoS は FlexGroup ボリュームレベルでのみ適用され、QoS の最大値のみをサポートしています。QoS の最小値、適応型 QoS、ファイルレベル QoS、およびネストされたポリシーは、FlexGroup ボリュームでは現在サポートされていません。ポリシーは現在、コマンドラインでのみ適用されます。FlexGroup ボリューム QoS の GUI サポートは、将来の ONTAP リリースに含まれる予定です。

8.5.2 サービス品質 (QoS) の最小値

ONTAP 9.4 では、FlexGroup ボリュームの QoS 最小化 (「保証」または「フロア」とも呼ばれる) のサポートが追加されました。これにより、指定したオブジェクトに割り当てられるパフォーマンスのしきい値が設定されます。この機能は、ETERNUS AX システムでのみサポートされます。

8.5.3 Adaptive Quality of Service (QoS)

ONTAP 9.4 では、FlexGroup ボリュームに対する適応型 QoS サポートも導入されました。これにより、ボリューム容量の調整に合わせて、ONTAP が QoS ポリシーの IOPS/TB 値を調整できるようになりました。

付録 A

次の項では、このドキュメントのこれまで項で説明しなかった FlexGroup 情報について説明します。

- ONTAP FlexGroup の作成と管理のコマンドライン例
- FlexGroup 統計の収集
- CLI を使用した FlexGroup のインジェスト使用量分布の表示
- クライアントから多数のファイルを生成するためのサンプル Python スクリプト

A.1 コマンドラインの例

このセクションでは、FlexGroup に関連するさまざまな基本操作のコマンドライン例を示します。

ONTAP 9.2auto-provision-as オプションの使用：

```
cluster::> vol create -auto-provision-as flexgroup -vserver SVM -volume flexgroup92  
-junction-path /flexgroup92 -size 100t -space-guarantee none -security-style unix
```

volume create を使用した、複数のノード間での FlexGroup ボリュームの作成：

```
cluster ::> volume create -vserver SVM -volume flexgroup -aggr-list aggr1_node1,aggr1_node2 -  
policy default -security-style unix -size 20PB -space-guarantee none -junction-path  
/flexgroup
```

FlexGroup スナップショットポリシーの変更：

```
cluster::> volume modify -vserver SVM -volume flexgroup -snapshot-policy [policynone|none]
```

FlexGroup ボリュームのサイズ変更：

```
cluster::> volume size -vserver SVM -volume flexgroup -new-size 20PB
```

FlexGroup ボリュームへのメンバーの追加：

```
cluster::> volume expand -vserver SVM -volume flexgroup -aggr-list aggr1_node1,aggr1_node2 -  
aggr-list-multiplier 2
```

ストレージ QoS の適用：

```
cluster::> volume modify -vserver DEMO -volume flexgroup -qos-policy-group FlexGroupQoS
```

ボリューム自動拡張の適用：

```
cluster::> volume autosize -vserver DEMO -volume Tech_ONTAP -mode grow -maximum-size 20t -  
grow-threshold-percent 80  
  
cluster::> volume autosize -vserver DEMO -volume Tech_ONTAP  
Volume autosize is currently ON for volume "DEMO:Tech_ONTAP".  
The volume is set to grow to a maximum of 20t when the volume-used space is above 80%.  
Volume autosize for volume 'DEMO:Tech_ONTAP' is currently in mode grow.
```

A.2 FlexGroup 統計

ONTAP 9 では、flexgroup と呼ばれる統計オブジェクトが追加されました。オブジェクトは diag 権限でのみ使用できます。このオブジェクトは、次のカウンタを収集します。

cat1_tld_local	cat1_tld_remote
cat2_hld_local	cat2_hld_remote
cat3_dir_local	cat3_dir_remote
cat4_fil_local	cat4_fil_remote
dsidlist_factory_enomem	groupstate_analyze
groupstate_create	groupstate_delete
groupstate_enomem	groupstate_insert
groupstate_preupdate_fail	groupstate_update
indextable_factory_enomem	indextableload_factory_enomem
indextablesave_factory_enomem	instance_name
instance_uuid	memberstate_create
memberstate_delete	memberstate_enomem
memberstate_expired	memberstate_factory_enomem
memberstate_unhealthy	monitor_receive
monitor_respond	node_name
node_uuid	process_name
refresh_enomem	refreshclient_create
refreshclient_delete	refreshserver_create
refreshserver_delete	remote_dirs
remote_files	snapclient_create
snapclient_delete	snapcoord_create
snapcoord_delete	snapserver_create
snapserver_delete	snapserver_fail_fence_down
snapserver_fail_fence_raise	snapserver_fail_snapid
snapshot_create	snapshot_enomem
snapshot_restore	tally_enomem
vldb_enomem	vldb_enorecord
vldbclient_create	vldbclient_delete
vldbclient_factory_enomem	

カウンタは FlexGroup ボリュームに固有のもので、リモート割り当ての割合、ローカルとリモートのファイルとディレクトリの数、リフレッシュカウンタ、およびその他のさまざまなオブジェクトを測定します。

FlexGroup 統計は、他の統計と同じ方法でキャプチャできます。統計収集は、statistics start で開始する必要があります。これにより、sample_id ファイルが作成されます。この後、統計情報を表示するには、statistics show を使用します。

複数のオブジェクトまたはカウンタを指定する場合は、パイプ記号 (|) を使用します。

FlexGroup および NFSv3 統計の statistics start の例：

```
cluster::> set diag
cluster::*> statistics start -object nfsv3|flexgroup
Statistics collection is being started for sample-id: sample_2144
```

FlexGroup カウンタの statistics show の例：

```
cluster::*> statistics show .object flexgroup

Object: flexgroup
Instance: 0
Start-time: 8/9/2016 13:00:22
End-time: 8/9/2016 15:22:29
Elapsed-time: 8527s
Scope: node1
  Counter                                     Value
-----
cat4_fil_local                               3623435
cat4_fil_remote                              600298
groupstate_analyze                           293448
groupstate_update                           59906297
instance_name                                0
node_name                                    node1
process_name                                  -
refreshclient_create                         146724
refreshclient_delete                         146724
refreshserver_create                         146724
refreshserver_delete                         146724
remote_files                                  10
```

統計コマンドの詳細は、CLI で man statistics start コマンドを使用してください。

A.2.1 Qtree 統計

ONTAP 9.5 以降、FlexGroup ボリュームで qtree 統計が利用可能になりました。これらの統計は、FlexGroup ボリュームとその qtree に関する詳細なパフォーマンス情報を提供します。次の例は、大規模な NFS ワークロードを実行している FlexGroup ボリュームの統計情報の取得を示しています。

```
cluster::> statistics qtree show -interval 5 -iterations 1 -max 25 -vserver DEMO -volume
flexgroup_local

cluster : 11/7/2018 15:19:15
```

Qtree	Vserver	Volume	NFS Ops	CIFS Ops	Internal Ops	*Total Ops
DEMO:flexgroup_local/	DEMO	flexgroup_local	22396	0	0	22396
DEMO:flexgroup_local/qtree						
	DEMO	flexgroup_local	0	0	0	0

A.3 FlexGroup インジェスト分散の表示

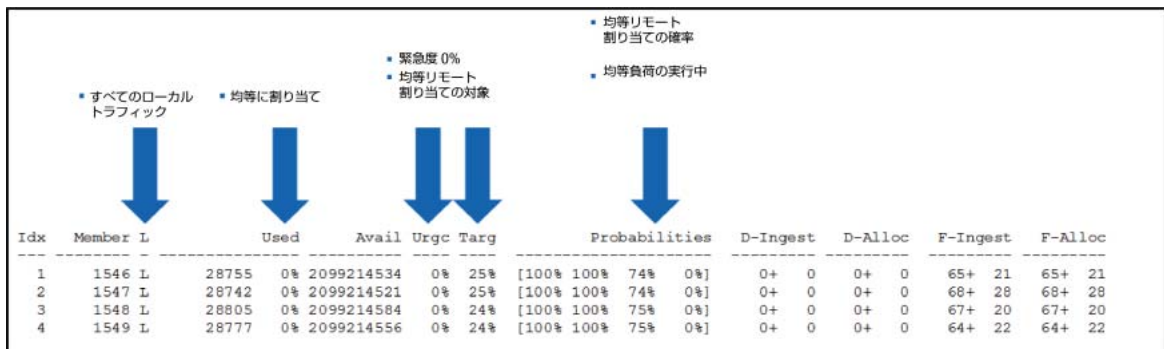
コマンドラインを使用して、ワークロード中の FlexGroup データ取り込みのリアルタイムビューを取得し、diag 特権ノードレベルコマンド flexgroup show を使用して、メンバーボリュームがどの程度均等に割り当てられているかを確認できます。さらに、このコマンドを使用すると、緊急度と許容度の割合、およびファイルとフォルダのリモート配置とローカル配置の計算された確率を表示できます。詳細は、「FlexGroup ロードバランシングの概念」を参照してください。

このコマンドは、クラスタシェル CLI で複数のノードに対して実行できます。

```
cluster:::> set diag
cluster:::> node run * flexgroup show
```

次の図は、トラフィックが均等に分散されている「理想的な」flexgroup show の出力を示しています (図 A.1)。

図 A.1 理想的な FlexGroup インジェスト



A.4 FlexGroup ボリューム上にファイルを生成するサンプル Python スクリプト

FlexGroup ボリュームをテストする場合は、通常の負荷生成ユーティリティを使用できます。私たちのラボテストで使用したベンチマークの 1 つは、Linux ソースコードのコンパイルを使用した [基本的な Git ベンチマーク](#) でした。このタイプのテストは誰でも利用できますが、ほとんどのストレージ管理者が行うよりも複雑な場合があります。

逆に、dd などの一般的なファイル作成ユーティリティを使用したり、bash スクリプトを作成してファイルやフォルダーを作成したりするのは理想的ではありません。これらはシングル・スレッド・テストであり、クライアントまたはストレージの CPU とスループット機能のメリットを十分に活用していません。

多数のファイルを作成し、FlexGroup ボリュームに十分な負荷をかけてその利点を確認する簡単な方法の 1 つは、Python スクリプトを使用することです。

このスクリプトはマルチプロセッサ呼び出しを使用して 1,000 個のディレクトリを作成し、それぞれに 1,000 個のサブディレクトリがあります。その下には小さなテキストファイルが 5 つ、合計 500 万ファイル書き込まれます。このスクリプトを変更して、作成するファイルとフォルダの数を変更できます。

このスクリプトは GitHub から入手できますが、富士通サポートでは正式にサポートされていません。このスクリプトは、負荷生成の測定や FlexGroup ボリュームのパフォーマンスの最大化を目的としたものではありません。

FUJITSU Storage
ETERNUS AX series オールフラッシュアレイ ,
ETERNUS HX series ハイブリッドアレイ
ONTAP FlexGroup ボリューム技術概要

P3AG-5652-01Z0

発行年月 2020 年 11 月
発行責任 富士通株式会社

- 本書の内容は、改善のため事前連絡なしに変更することがあります。
- 本書の内容は、細心の注意を払って制作致しましたが、本書中の誤字、情報の抜け、本書情報の使用に起因する運用結果に関しましては、責任を負いかねますので予めご了承ください。
- 本書に記載されたデータの使用に起因する第三者の特許権およびその他の権利の侵害については、当社はその責を負いません。
- 無断転載を禁じます。

FUJITSU