Fsas Technologies

P3AG-6202-02Z0

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 ベストプラクティス / インプリメンテーションガイド

© 2021-2025 Fsas Technologies Inc.

目次

1. ONTAP における NAS の進化 13
1.1 フレキシブルボリューム : 実証済みのソリューション 13
1.2 FlexGroup ボリューム : NAS の進化 14

2. 用語 15
2.1 大容量ファイルとは 16

3. ONTAP FlexGroup のメリット 17
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レーテンシー 17
3.2 すべてのクラスタハードウェアの効率的な使用 17
3.3 シンプルで管理しやすいアーキテクチャとバランシング 17
3.4 ビッグデータの高密度化 18

4. 使用例 19
4.1 理想的な使用例 19
4.2 非理想的なケース 19
4.3 FlexGroup ボリュームの使用例 19
4.3.1 FlexGroup の使用例 1: Active IQ インフラストラクチャ 19
4.3.2 FlexGroup の使用例 2: SQL Server のリポジトリをバックアップする 20
4.3.3 結論 21

5. FlexGroup 機能のサポートと最大化 22
5.1 サポートされていない SMB 機能の動作 24
5.2 最大値と最小値 25

6. FlexGroup ボリュームが適しているかどうかの判断 26
6.1 スケールアウトパフォーマンス 26
6.2 機能の互換性の制限 27
6.3 パフォーマンスの合理化 28
6.3.1 AFF A700 テスト 28
6.3.2 ビッグデータのワークロードによる FlexGroup のパフォーマンス 30
6.3.3 ワークロードの自動適応 31
6.3.4 インジェストアルゴリズムの改善 31
6.3.5 パフォーマンス機能 32
6.4 ワークロードと動作 34
6.4.1 最適なワークロード 35
6.4.2 良好なワークロード 36
6.4.3 理想的でないワークロード : 大容量ファイル 36

6.4.4 期待されるパフォーマンス : 読み取り負荷の高いワークロード 40
6.4.5 FlexGroup ボリューム内のデータ不均衡 41

7. FlexGroup の初期設計に関する考慮事項 42
7.1 クラスタに関する考慮事項 42
7.2 ONTAP バージョンに関する考慮事項 42
7.3 障害ドメイン 43
7.4 アグリゲートレイアウトの考慮事項 43
7.4.1 既存の FlexVol ボリュームを持つアグリゲートへの FlexGroup ボリュームの展開 44
7.5 Flash Cache と Flash Pool 46
7.6 高度なディスクパーティション設定 46
7.7 SyncMirror (ミラーリングされたアグリゲート) 46
7.8 MetroCluster 46
7.9 Cloud Volumes ONTAP 47
7.10 容量に関する考慮事項 47
7.10.1 最大値と最小値 48
7.10.2 理論上の最大値または絶対的な最大値 49
7.10.3 FlexVol メンバーボリュームのレイアウトに関する考慮事項 50
7.10.4 クライアントが領域不足エラーになるのはどのような場合ですか？ 54
7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか？ 56
7.10.6 アグリゲートの空き領域に関する考慮事項 60
7.10.7 初期ボリュームサイズに関する考慮事項 61
7.10.8 Snapshot コピーと Snapshot Reserve 62
7.10.9 ボリュームの自動サイズ設定 (自動拡張と自動縮小) 63
7.10.10 エラスティックサイジング 66
7.10.11 プロアクティブなサイズ変更 71
7.11 ネットワークに関する考慮事項 76
7.11.1 LACP の考慮事項 76
7.11.2 DNS ロードバランシングの考慮事項 77
7.11.3 Border Gateway Protocol (BGP) 77
7.12 セキュリティとアクセス制御リストのスタイルに関する考慮事項 78
7.12.1 ボリュームセキュリティスタイルの基本的なガイダンス 78

8. FlexGroup 管理の考慮事項 80
8.1 FlexGroup ボリュームの表示 80
8.1.1 ONTAP System Manager 80
8.1.2 Active IQ Unified Manager 81
8.1.3 コマンドライン 84
8.2 FlexGroup ボリューム容量の表示 84
8.2.1 FlexGroup の総容量 85
8.2.2 FlexGroup ボリュームでの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 85

8.3 FlexGroup ボリュームへの容量の追加 86
8.3.1 容量の追加に関する推奨事項 86
8.3.2 ボリューム拡張 89
8.3.3 ディスク、アグリゲート、およびノードの追加 90
8.3.4 クラスタからのノードの削除または退避 91
8.4 クラスタからのノードの削除または退避 91
8.4.1 無停止でのボリューム移動を使用する場合 92
8.4.2 無停止でのボリューム移動の使用 92
8.5 FlexGroup ボリュームを削除する際の考慮事項 93
8.6 ボリューム名の変更に関する考慮事項 94

9. qtree 95
9.1 qtree とファイル移動 95
9.1.1 qtree ID と名前変更の動作 95
9.1.2 qtree エクスポートのファイルハンドルの効果 96
9.2 FlexGroup によるクォータの管理 96
9.2.1 ユーザー / グループのクォータに関する考慮事項 97
9.2.2 ユーザーまたはグループのクォータの作成 97
9.2.3 コマンドラインからのツリーレポートクォータの作成 99
9.2.4 クォータの適用の例 100
9.2.5 クォータの使用によるパフォーマンスへの影響 102
9.2.6 クォータスキャンの完了時間 103
9.2.7 クォータに関するユーザーマッピングの考慮事項 103
9.2.8 ツリークォータに関する考慮事項 104
9.2.9 クォータが有効なときにクライアントが領域を認識する方法 105

10. NAS とファイル数が多い場合の一般的な考慮事項 106
10.1 ファイル数が多い場合の考慮事項 106
10.1.1 デフォルトおよび最大 inode 数 106
10.1.2 最大ファイル数の増加 : 考慮事項 106
10.1.3 デフォルトおよび最大 inode 数：FlexGroup ボリュームに関する考慮事項 107
10.1.4 ファイル数を多くして容量を少なくするニーズ 108
10.1.5 ONTAP でファイル数が多い場合の計画 109
10.1.6 使用済み inode と合計 inode の表示 109
10.1.7 inode が足りなくなったらどうなりますか？ 109
10.1.8 非同期削除 109
10.2 64 ビットのファイル識別子 110
10.2.1 このオプションを変更するとどうなりますか？ 111
10.2.2 NFSv3 と NFSv4.x: ファイル ID 112
10.2.3 クォータの適用を使用してファイル数を制限する 112
10.2.4 ONTAP System Manager: 9.7 113
10.2.5 ONTAP System Manager: 9.8 以降 113

10.2.6 ファイル ID の競合の影響 ... 114
10.3 ONTAP でのファイルシステム ID の変更の影響 ... 115
10.3.1 ファイル数の多い環境の SVM での FSID の操作 ... 115
10.3.2 FSID による Snapshot コピーの処理方法 ... 115
10.4 ディレクトリサイズに関する考慮事項 :maxdirsize ... 116
10.4.1 maxdirsize はどのようなディレクトリ構造に影響されますか？ ... 116
10.4.2 フラットディレクトリ構造が FlexGroup ボリュームに与える影響 ... 117
10.4.3 使用されている maxdirsize 値の問い合わせ ... 118
10.4.4 デフォルトの maxdirsize で 1 つのディレクトリに収まるファイルの数 ... 119
10.4.5 maxdirsize を超えると送信されるイベント管理システムメッセージ ... 119
10.4.6 maxdirsize 値の増加の効果 ... 120
10.4.7 FlexGroup ボリュームは maxdirsize 制限をバイパスしますか？ ... 120
10.4.8 maxdirsize を超えることの影響 ... 120
10.5 ファイルシステムの分析 ... 120
10.6 特殊文字の考慮事項 ... 122
10.6.1 utf8mb4 ボリューム言語のサポート ... 122
10.6.2 多数のファイルが存在する環境で NFS を使用して低速のディレクトリ一覧を管理 ... 123
10.7 ファイルの削除 /FlexGroup メンバーボリュームのバランシング ... 124
10.7.1 FlexGroup ボリューム内のデータの再バランシング ... 125
10.8 メンバーボリュームがスペース不足の場合のファイルのリスト ... 126
10.9 ファイル名の変更に関する考慮事項 ... 127
10.10 シンボリックリンクの考慮事項 ... 127
10.11 NFS バージョンに関する考慮事項 ... 127
10.11.1 ネットワーク接続の同時実行性 : NFSv3 ... 127
10.11.2 NFS 書き込みの追加 ... 129
10.11.3 nconnect ... 129
10.11.4 NFS 接続クライアントのボリューム名へのマッピング ... 130
10.11.5 FlexGroup ボリュームでの NFSv4.x の有効化と使用 ... 130
10.11.6 FlexGroup ボリュームにおける NAS メタデータの影響 ... 135
10.12 CIFS/SMB に関する考慮事項 ... 135
10.12.1 SMB バージョンに関する考慮事項 ... 135
10.12.2 SMB での変更通知の使用 ... 135
10.12.3 大容量 MTU ... 136
10.12.4 SMB マルチチャネル ... 136
10.12.5 継続的に利用可能な共有 (CA 共有) ... 136
10.12.6 その他の考慮事項 ... 137
10.13 仮想化ワークロードに関する考慮事項 ... 137
10.13.1 ONTAP tools for VMware vSphere(以前の Virtual Storage Console) ... 138
10.13.2 コピーオフロード ... 138
10.13.3 考慮事項 ... 138
10.14 FlexGroup ボリューム上のデータベース ... 139

10.15 FlexCache ボリュームの考慮事項 140
10.16 FlexClone 141
10.16.1 異なるストレージ仮想マシン (SVM) への FlexClone 142
10.16.2 ボリュームのリホスト 142
10.16.3 FlexClone の削除 142

11. 保存中の暗号化 143
11.1 FlexGroup ボリュームのキー更新または既存の FlexGroup ボリュームの暗号化 143
11.2 ドライブレベルの暗号化 (NSE と SED) 143

12. FlexGroup サンプル設計 144
12.1 ボリュームアフィニティと CPU の飽和 144
12.1.1 FlexGroup サンプル設計 1: FlexGroup ボリューム、クラスタ全体 (24 ノード) 145
12.1.2 FlexGroup サンプル設計 2: 複数ノード、アグリゲート、部分クラスタ 146
12.1.3 FlexGroup サンプル設計 3: FlexGroup、単一ノード 147
12.1.4 FlexGroup サンプル設計 4: FlexGroup ボリュームを FlexGroup ボリュームにマウント 152
12.1.5 FlexGroup ボリュームにマウントされている FlexVol ボリューム 152

13. 一般的なトラブルシューティングと修復 153
13.1 障害シナリオ 153
13.1.1 ストレージのフェイルオーバー 153
13.1.2 ネットワーク障害 153
13.1.3 スナップショット障害 153
13.1.4 ハードウェア障害 154
13.1.5 時刻同期 154

14. 容量の監視とアラート 155
14.1 コマンド・ラインを使用した容量の監視とアラート 155
14.1.1 イベント管理システムメッセージ 155
14.2 シンプロビジョニングにおけるクライアント側の容量に関する考慮事項 157
14.2.1 Windows 容量レポート 158
14.3 ONTAP コマンド行からの FlexVol メンバー容量の表示 159
14.4 FlexGroup 容量ビューア 159
14.5 論理スペース計算 161
14.6 FlexGroup パフォーマンスの監視 161
14.6.1 コマンドラインからのパフォーマンスの監視 162
14.6.2 flexgroup show 165
14.6.3 パフォーマンスアーカイバ 166
14.6.4 パフォーマンスの監視 (Active IQ Unified Manager) 166

15. FlexGroup データ保護のベストプラクティス 167

16. ONTAP FlexGroup への移行 168

16.1 NDMP を使用した移行 168

16.2 FlexVol から FlexGroup への変換 169

16.2.1 FlexVol ボリュームを FlexGroup ボリュームに変換する理由 169

16.2.2 FlexVol ボリュームを変換しない場合 170

16.2.3 仕組み 171

16.2.4 その他の考慮事項および注意事項 172

16.3 サード・パーティ製ストレージから FlexGroup への移行 174

16.4 7-Mode で動作する Data ONTAP からの移行 175

16.5 ONTAP での SAN LUN からの移行 175

16.6 XCP Migration Tool 175

16.7 XCP を使用した移行前のファイルのスキャン 176

16.7.1 XCP を使用したディスク使用量 (du) スキャンの実行 178

17. 例 179

17.1 シンプロビジョニングの例 179

17.2 ボリュームの自動サイズ設定の例 180

17.3 Snapshot Spill の例 181

17.4 Active IQ Unified Manager での容量の監視とアラートの例 184

17.4.1 Active IQ Unified Manager でのボリュームしきい値の編集 187

17.4.2 inode 監視 188

17.4.3 Active IQ の「Fix It」 189

17.5 FlexVol から FlexGroup への変換例 189

17.6 既存の SnapMirror 関係における FlexVol の変換例 193

17.7 FlexVol から FlexGroup への変換例：5 億ファイル 197

17.8 イベント管理システムの例 201

17.8.1 inode 関連の EMS の例 201

17.8.2 maxdirsize メッセージの例 202

17.8.3 容量関連のイベント管理システムメッセージの例 203

18. コマンドの例 205

18.1 FlexGroup 容量コマンド 205

18.2 クラスタ全体に対する statistics show-periodic コマンドの例 209

18.3 NFSv3 の読み取り / 書き込み操作に関するリアルタイムの SVM レベルの statistics show-periodic 210

18.4 リアルタイムの FlexGroup のローカルおよびリモート統計情報 210

18.5 FlexGroup ボリュームを作成し、デフォルト値より少ないメンバーボリュームを指定する例 211

18.6 FlexGroup ボリューム作成用のサンプル REST API 212
18.7 FlexGroup ボリュームのサイズを増やす例 215
18.8 FlexGroup ボリュームの拡張例 216
18.9 その他のコマンドライン例 217
18.9.1 flexgroup deploy を使用した FlexGroup ボリュームの作成 217
18.9.2 volume create を使用した複数ノード間での FlexGroup ボリュームの作成 217
18.9.3 FlexGroup Snapshot ポリシーの変更 217
18.9.4 ストレージ QoS の適用 217
18.9.5 ボリュームの自動拡張の適用 217

図目次

図 1.1 100TB 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 ... 13
図 2.1 大容量ファイルとは ... 16
図 4.1 SQL Server バックアップ環境 ... 20
図 4.2 テスト実行中のスループットと合計処理数 ... 20
図 4.3 CPOC スケールアウトスループットの結果 ... 21
図 6.1 FlexGroup ボリューム ... 26
図 6.2 FlexVol ボリュームと FlexGroup ボリューム : ワークロードが増加している場合の最大スループット傾向 ... 29
図 6.3 FlexVol ボリュームと FlexGroup ボリューム : ワークロードが増加している場合の最大スループットの傾向の詳細 ... 29
図 6.4 FlexVol ボリュームと FlexGroup ボリューム：最大平均合計 IOPS ... 30
図 6.5 FlexGroup ボリュームの TeraSort ベンチマーク統計の概要 ... 30
図 6.6 FlexGroup ボリューム上のストレージ QoS：シングルノード接続 ... 32
図 6.7 FlexGroup ボリュームのストレージ QoS：マルチノード接続 ... 33
図 6.8 qtree QoS の使用例 ... 33
図 6.9 容量不均衡と遠隔配置の可能性 ... 35
図 6.10 大きなファイルがいくつかある FlexGroup ボリュームにおいて、使用が最適ではない理由 ... 38
図 6.11 大容量ファイルの最悪のシナリオ：すべてが同じメンバーボリュームに配置 ... 38
図 6.12 容量不均衡の例 ... 39
図 7.1 FlexVol 容量が FlexGroup 負荷分散に与える影響 ... 44
図 7.2 FlexVol と FlexGroup のアーキテクチャの比較 ... 50
図 7.3 ONTAP System Manager FlexGroup ボリュームの作成 ... 52
図 7.4 System Manager で許容される最大値を超えて FlexGroup ボリュームを作成したときのエラー ... 56
図 7.5 容量をメンバーボリューム間で分割する方法 ... 58
図 7.6 FlexGroup メンバーボリューム内のサイズの大きいファイルの影響 ... 58
図 7.7 メンバーボリュームの縮小と拡大 ... 59
図 7.8 FlexGroup ボリューム : メンバーサイズと FlexGroup ボリューム容量 ... 61
図 7.9 ボリュームの自動サイズ設定操作後のメンバーボリュームサイズの割り当て ... 66
図 7.10 エラスティックサイジング前のファイル書き込み動作 ... 68
図 7.11 エラスティックサイジング後のファイル書き込み動作 ... 68
図 7.12 初期 FlexGroup データバランス：プロアクティブなサイズ変更、自動サイズ設定が無効 ... 71
図 7.13 FlexGroup のデータバランス (使用率は最大 68%)：プロアクティブなサイズ変更、自動サイズ設定が無効 ... 71
図 7.14 FlexGroup データバランス、ジョブ完了：プロアクティブなサイズ変更、自動サイズ設定が無効 ... 72
図 7.15 FlexGroup データバランス、新しい大容量ファイル : プロアクティブなサイズ変更、自動サイズ設定が無効 ... 72
図 7.16 FlexGroup データバランス、80GB ファイル：プロアクティブなサイズ変更、自動サイズ設定が無効 ... 72
図 7.17 FlexGroup のデータバランス、領域不足 : プロアクティブなサイズ変更、自動サイズ設定が無効 ... 73
図 7.18 初期 FlexGroup データバランス：プロアクティブなサイズ変更、自動サイズ設定が有効 ... 73
図 7.19 FlexGroup のデータバランス (使用率は最大 68%)：プロアクティブなサイズ変更、自動サイズ設定が有効 ... 74
図 7.20 FlexGroup データバランス、ジョブ完了：プロアクティブなサイズ変更、自動サイズ設定が有効 ... 74
図 7.21 FlexGroup データバランス、2 回目のテスト実行：プロアクティブなサイズ変更、自動サイズ設定が有効 ... 74
図 7.22 FlexGroup データバランス、自動サイズ設定制限：プロアクティブなサイズ変更、自動サイズ設定が有効 ... 75
図 7.23 ONTAP System Manager での FlexGroup ボリュームセキュリティースタイルの変更 ... 79
図 8.1 ONTAP System Manager の FlexGroup ボリュームビュー ... 81
図 8.2 Active IQ Unified Manager、FlexGroup 容量ビュー ... 81
図 8.3 Active IQ Unified Manager での容量の傾向 ... 82
図 8.4 Active IQ Performance Manager の FlexGroup ボリュームビュー ... 82

図 8.5 メンバーボリュームのパフォーマンス表 83
図 8.6 メンバボリュームのグラフ 83
図 8.7 シンプロビジョニングされた FlexGroup ボリュームが、スペース保証された FlexVol ボリュームとともに存在する場合の容量の影響 85
図 8.8 FlexGroup ボリュームでのアグリゲートの追加 90
図 8.9 ノードの追加と FlexGroup ボリュームの拡張 90
図 8.10 FlexGroup メンバーボリュームを含むノードの削除 91
図 9.1 クォータ・レポート：ONTAP System Manager 97
図 9.2 クォータのボリュームステータス：ONTAP System Manager 98
図 9.3 クォータ・ルール：ONTAP System Manager 98
図 9.4 ONTAP 9.5 のパフォーマンス (オペレーション / 秒)：クォータがオンとオフ 102
図 9.5 ONTAP 9.5 のパフォーマンス (MBps)：クォータがオンとオフ 102
図 10.1 ONTAP System Manager 9.8 の 64 ビットファイル ID 113
図 10.2 ファイルシステム分析 : 有効 121
図 10.3 ファイルシステム分析 : ディレクトリおよびファイル情報 121
図 10.4 ファイルシステム分析 : 使用頻度の低いデータと使用頻度の高いデータ 121
図 10.5 サイズの大きいファイルを削除した後の容量の不均衡 124
図 10.6 NFSv3 のパフォーマンスに対する RPC スロットテーブルの影響 128
図 10.7 pNFS 132
図 10.8 pNFS 操作 133
図 10.9 pNFS 操作：FlexGroup ボリューム 134
図 10.10 ONTAP tools for VMware vSphere：FlexGroup データストア 138
図 12.1 FlexGroup ボリューム、クラスタ全体 (24 ノード) 145
図 12.2 複数のノード、部分的なクラスタ 146
図 12.3 Git クローン完了時間の比較 148
図 12.4 平均スループットと最大スループットの比較 148
図 12.5 最大読み取りスループットの比較 149
図 12.6 最大書き込みスループットの比較 149
図 12.7 合計平均 IOPS の比較 150
図 12.8 FlexGroup ボリュームの平均 CPU 使用率、スループット、および IOPS (ETERNUS AX4100 HA ペア、128 スレッド) 150
図 12.9 FlexGroup ボリュームの平均 CPU 使用率、スループット、および IOPS (シングルノード ETERNUS AX4100、128 スレッド) 151
図 12.10 FlexGroup ボリューム、単一ノード 151
図 12.11 FlexGroup ボリュームが FlexGroup ボリュームにマウントされている場合 152
図 14.1 Google シート：FlexGroup の容量ビュー 159
図 14.2 Google シート：FlexGroup の容量ビュー 160
図 14.3 Google シート：平均 inode サイズ 160
図 14.4 論理スペース計算の動作 161
図 14.5 Active IQ Performance Manager のグラフ 166
図 16.1 最大容量に達しそうな FlexVol ボリュームの変換 170
図 16.2 FlexVol ボリュームの FlexGroup への変換とメンバーボリュームの追加 171
図 16.3 XCP レポート作成グラフ 176
図 16.4 XCP レポート 178
図 17.1 FlexGroup の容量の内訳 :Active IQ Unified Manager 179
図 17.2 ボリュームしきい値の編集 187
図 17.3 Active IQ Unified Manager:inode の修正 189
図 17.4 変換プロセスのサンプル統計 198
図 17.5 変換プロセス中の統計の例 : メンバーボリュームの追加 199
図 17.6 変換プロセスの統計の例：2 倍のパフォーマンス 200

表目次

表 5.1 一般的な ONTAP 機能のサポート 22
表 5.2 一般的な NAS プロトコルバージョンのサポート 23
表 5.3 サポートされていない SMB2.x および 3.x の機能 24
表 5.4 サポートされていない SMB 機能が FlexGroup ボリュームでどのように動作するか 24
表 5.5 FlexGroup の最大値 25
表 5.6 FlexGroup の最小値 25
表 6.1 ONTAP ボリュームファミリの比較 27
表 7.1 FlexGroup ボリュームを使用したアグリゲートレイアウトのベストプラクティス 43
表 7.2 FlexGroup の最大値 48
表 7.3 FlexGroup の最小値 48
表 7.4 ONTAP で許可されるボリューム数に基づく FlexGroup の理論上の最大値 49
表 7.5 領域不足エラーが発生する状況 55
表 7.6 容量管理の決定マトリックス 60
表 7.7 最大サイズの自動調整の例 75
表 10.1 FlexVol サイズに応じた inode のデフォルト値と最大値 106
表 10.2 FlexGroup メンバーサイズおよびメンバーボリューム数の結果として生じる inode のデフォルト 108
表 10.3 ファイル数が多い小容量のフットプリントの例 : メンバーボリューム数の増加 108
表 10.4 非同期削除のパフォーマンス 110
表 10.5 nconnect のパフォーマンス結果 130
表 14.1 flexgroup show 出力列の定義 165

はじめに

本書では、ONTAP FlexGroup の概要と、この機能で使用するベストプラクティスと実装のヒントについて説明します。FlexGroup の機能は、拡張型 NAS コンテナを進化させたものであり、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ーテンシーのパフォーマンスを組み合わせます。本書で説明されていない FlexGroup ボリュームについては、当社サポート部門にご連絡ください。必要に応じて、本書に情報を追加します。

第 2 版
2025 年 3 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。

1. 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の時代は終わりました。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。

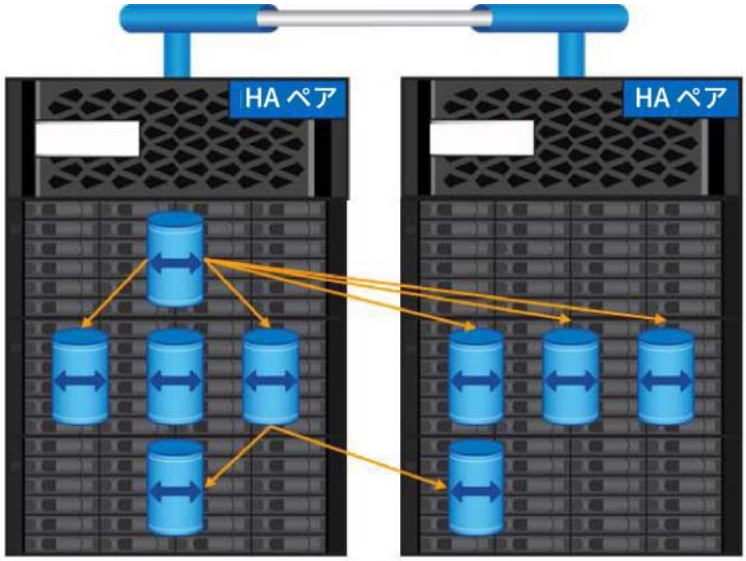
機械学習と人工知能のワークロードには、ペタバイト規模 (何十億ものファイル) まで拡張可能な単一のネームスペースに対するストレージへのニーズが含まれます。これらの技術の台頭と Hadoop のようなビッグデータフレームワークの出現により、NAS ファイルシステムの進化が待ち望まれています。
ONTAP FlexGroup は、このようなタイプのアーキテクチャに最適なソリューションです。

1.1 フレキシブルボリューム : 実証済みのソリューション

FlexVol 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す。

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す。

図 1.1 100TB 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



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 100TB 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するクラスタアーキテクチャを提供します。

FlexGroup 以前では、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、単一のネームスペースとして機能するファイルシステムをクラスタ上に作成していました。

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。

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 ボリューム : NAS の進化

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実質的な制約はありません。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ー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。

2. 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 / 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」は読み替え可能な用語です。これらは、FlexGroup ボリュームを構成し、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する、基盤となる FlexVol ボリュームを指します。
- FlexGroup ボリューム
  FlexGroup ボリュームは、複数の構成 / 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、ジャンクションパスを使用して、FlexVol ボリュームの 100TB の制限を超えた容量が提供されました。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 でジャンクションパスを使用することを、ONTAP ネームスペース内のボリュームの「マウント」と呼びます。
- 大容量ファイル
  「2.1 大容量ファイルとは」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を無効 (`guarantee = none`) 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 10TB のアグリゲートで FlexVol ボリュームが 100TB 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、使用可能な領域を厳密に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では、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えることがあります。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た場合、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「親ボリュームに存在しない」という意味です。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あります。

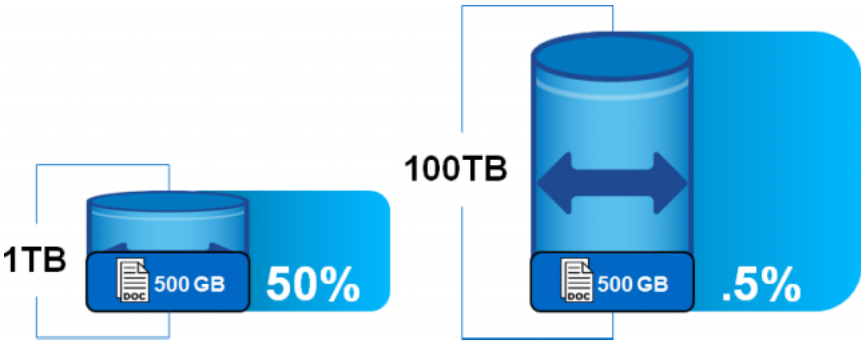
2.1 大容量ファイルとは

このドキュメントでは、「大容量ファイル」という用語を多用します。したがって、FlexGroup のコンテキストで「大容量ファイル」とは何か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、(「4. 使用例」で定義されているとおり) その他のさまざまなワークロードでも正常に機能します。ただし、これまで FlexGroup ボリュームで問題が発生していたワークロードのタイプの 1 つに、データベースファイルなどのサイズの大きいファイルやファイルが時間の経過とともに増大するワークロードがあります。

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積で表されます。したがって、一部の FlexGroup 構成、たとえば、メンバーボリュームサイズが 1TB しかない構成では、「大容量ファイル」は 500GB(メンバーボリュームサイズの 50%) になります。たとえば、メンバーボリュームのサイズが 100TB である他の構成では、同じ 500GB のファイルサイズはボリューム容量の 0.5% にすぎません。

図 2.1 大容量ファイルとは



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。ONTAP は、ファイルが作成されるときのサイズを認識できないため、サイズの大きいファイルの優先順位付けの方法を認識しません。代わりに、受動的に動作して、空きスペースの不一致がある他のメンバーボリュームとデータのバランスをとります。

ONTAP の各リリースでは、FlexGroup ボリューム内の大容量ファイルに対するアプローチが常に改善されています。

- 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、サイズの異なる大容量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。
- ONTAP 9.8 は、プロアクティブなサイズ変更によって容量管理を簡略化しています。

3. ONTAP FlexGroup のメリット

ONTAP FlexGroup は、ワークロードごとにさまざまなメリットを提供します。このメリットについては、次のセクションで説明します。

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レーテンシー

FlexGroup ボリュームを使用すると、ストレージ管理者は大容量のプロビジョニングを簡単に行うことができ、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実現し、ミッションクリティカルなワークロードには低レーテンシーを提供しながら、スループットとトータルオペレーションを向上します。

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と見なしていますが、ONTAP はバックグラウンドで作業し、入ってくるファイル作成を FlexGroup ボリューム全体に均等に分散して、CPU とディスクの効率的な使用率を実現しています。

3.3 シンプルで管理しやすい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 ボリュームなどの 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 ビッグデータの高密度化

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。10U のラックスペースでペタバイトの未フォーマット時容量を得ることができ、冷却コスト、消費電力、ラックのレンタルスペースのコストを削減し、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、パフォーマンスをバランスさせる 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。

4. 使用例

ONTAP FlexGroup の設計は、特定の用途 (電子設計・自動化、ソフトウェア開発など) に最も適しています。

4.1 理想的な使用例

FlexGroup ボリュームは、インジェスト負荷が高く (高レベルの新規データ作成)、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 (EDA)
- AI と機械学習のログファイルのリポジトリ
- ソフトウェア構築 / 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 (ビデオ監視など)
- 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)
- 仮想ワークロード (ONTAP 9.8 以降)

4.2 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。

- ファイルの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方法で制御する必要があるワークロード
- 大量のファイル名変更を伴うワークロード
- 1 つのディレクトリに数百万個のファイルがあり、すべてのファイルを頻繁にスキャンする必要があるワークロード
- 数千のシンボリックリンクを持つワークロード
- 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、遠慮なく当社サポート部門にご連絡ください。

4.3 FlexGroup ボリュームの使用例

次のセクションでは、実際の使用例を 2 つ説明します。

4.3.1 FlexGroup の使用例 1: Active IQ インフラストラクチャ

ソリューションの詳細については、当社サポート部門にお問い合わせください。

4.3.2 FlexGroup の使用例 2: SQL Server のリポジトリをバックアップする

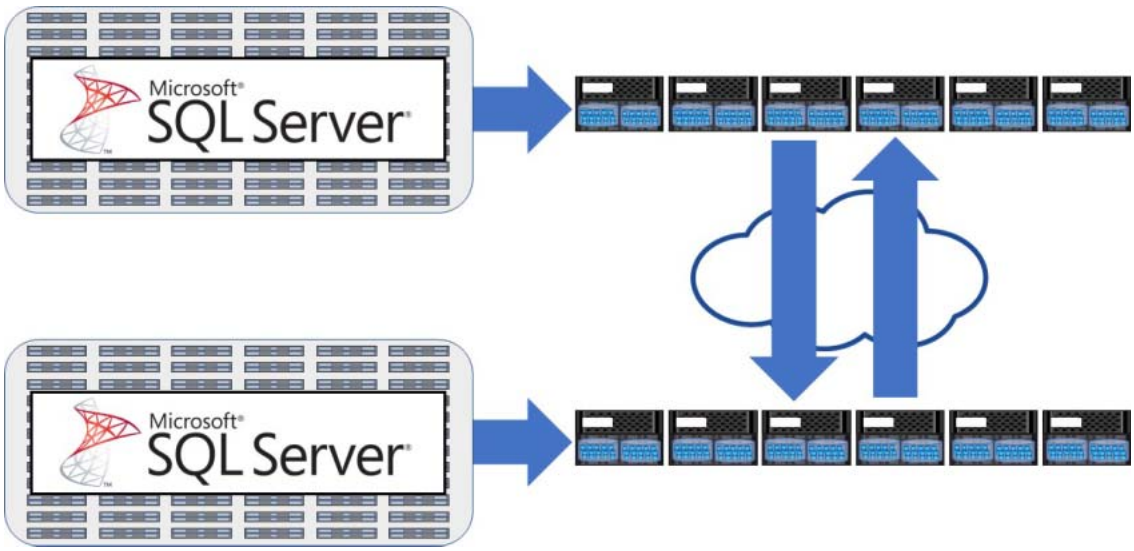
この環境では、お客様は SMB 上で 5,000 台の Microsoft SQL Server の圧縮バックアップを実行することを希望していました。このテストは、ソリューションを検証するために約 200 台のサーバを使用して実施され、数か月の間にゆっくりと開始されました。

しかし、このデータベースは単なるバックアップターゲットではなく、データ保護を強化するために SnapMirror を使用して災害復旧サイトにもレプリケートされます。

各サイトには、6TB のニアライン SAS (NL-SAS) 暗号化ドライブを使用して ONTAP を実行する 6 ノードの ETERNUS HX series クラスタがあります。各クラスタの有効容量は 3PB です。クラスタは 30 の FlexGroup ボリュームとボリューム内の qtree をデータ編成に使用します。

FlexGroup ボリュームはそれぞれ 64TB で、メンバーボリュームはそれぞれ 2.6TB で、6 つのノードにわたって各ノードに 4 つのメンバー (FlexGroup ボリュームあたり合計 24 メンバー) があります。

図 4.1 SQL Server バックアップ環境



結果

このお客様は、12 時間で最大 150TB の MSSQL バックアップデータを収集できる単一のネームスペースを必要としていました。1 時間あたり最大 12TB、1 秒あたり最大 3.5GB です。

テストでは、サイト A で 222 台のサーバ、サイト B で 171 台のサーバを使用しました。テスト中、各クラスタの CPU 使用率は 95% で、バックアップジョブ (シーケンシャル書き込み) では、ジョブに必要なスループットの 2.4 倍に相当する 8.4GB/ 秒を達成できました。この場合、バックアップは 12 時間ではなく約 5 時間で完了します。また、この SMB ワークロードは約 120,000 IOPS でした。このワークロードにさらにクライアントが追加されると、スループットは最大で約 9GBps になると予想されます。

図 4.2 テスト実行中のスループットと合計処理数

cpu avg	cpu busy	total ops	nfs-ops	cifs-ops	fcache ops	spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	disk read	disk write	pkts recv	pkts sent
56%	81%	54530	0	54530	0	54420	6.16GB	2.65GB	44%	3.34GB	28.3MB	22%	2.82GB	2.42GB	128MB	3.31GB	968237	898917
65%	78%	70482	0	70482	0	70407	8.03GB	3.44GB	47%	4.33GB	30.9MB	24%	3.70GB	3.41GB	114MB	4.79GB	1178768	1102912
74%	87%	88725	0	88725	0	88105	10.2GB	4.30GB	49%	5.44GB	37.1MB	36%	4.78GB	4.26GB	157MB	5.56GB	1389743	1324559
86%	92%	111577	0	111577	0	110569	12.8GB	5.88GB	53%	6.84GB	41.9MB	31%	6.00GB	5.84GB	153MB	6.77GB	1726469	1679506
88%	92%	115036	0	115036	0	113509	13.2GB	6.44GB	51%	7.06GB	45.9MB	49%	6.14GB	6.40GB	142MB	7.65GB	1845760	1814549
92%	95%	118148	0	118148	0	117104	13.6GB	6.11GB	45%	7.26GB	49.9MB	42%	6.34GB	6.07GB	149MB	8.11GB	1802929	1769902
95%	98%	122953	0	122953	0	122123	14.3GB	7.10GB	47%	7.54GB	45.9MB	43%	6.75GB	7.04GB	135MB	8.29GB	1978205	1952416
96%	99%	126241	0	126241	0	125104	14.6GB	6.43GB	53%	7.75GB	54.3MB	44%	6.80GB	6.37GB	135MB	8.28GB	1865375	1849777
95%	97%	121948	0	121948	0	120719	13.9GB	7.25GB	44%	7.47GB	47.3MB	40%	6.41GB	7.20GB	108MB	8.30GB	1995908	1967271
95%	98%	123079	0	123079	0	121113	13.9GB	5.71GB	41%	7.56GB	49.0MB	38%	6.37GB	5.66GB	129MB	8.40GB	1761097	1712061
95%	97%	120567	0	120567	0	120493	13.7GB	7.01GB	42%	7.41GB	47.6MB	36%	6.34GB	6.96GB	116MB	8.48GB	1888934	1882711
95%	98%	119573	0	119573	0	119458	13.6GB	5.74GB	37%	7.33GB	44.4MB	35%	6.28GB	5.69GB	111MB	8.19GB	1702969	1671363
95%	97%	119538	0	119538	0	119829	13.5GB	6.98GB	41%	7.34GB	46.2MB	35%	6.17GB	6.93GB	120MB	8.44GB	1880298	1873821
95%	98%	118119	0	118119	0	118373	13.4GB	5.56GB	37%	7.25GB	45.4MB	37%	6.17GB	5.52GB	118MB	8.42GB	1666066	1630785
95%	98%	118862	0	118862	0	118327	13.6GB	6.29GB	39%	7.29GB	47.1MB	33%	6.30GB	6.24GB	114MB	8.31GB	1784134	1759266
96%	99%	121039	0	121039	0	121136	13.7GB	6.67GB	38%	7.44GB	44.5MB	34%	6.21GB	6.63GB	120MB	8.35GB	1832520	1827158
96%	99%	120852	0	120852	0	120920	13.7GB	5.77GB	39%	7.42GB	47.8MB	33%	6.24GB	5.72GB	111MB	8.51GB	1706939	1678778
94%	97%	119819	0	119819	0	120129	13.7GB	7.05GB	41%	7.36GB	42.6MB	35%	6.29GB	7.01GB	118MB	8.49GB	1882656	1877381

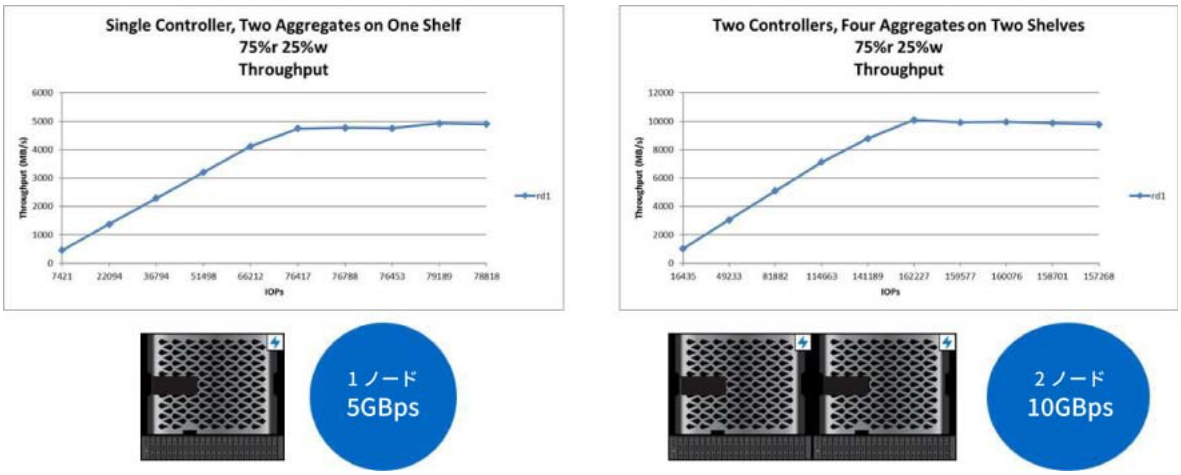
データ保護

本番ワークロードの FlexGroup ボリュームで見られるパフォーマンスに加えて、このお客様は、サイト間の SnapMirror 関係で高い転送レート (SnapMirror 転送で 8.4GB/ 秒) を達成できました。このレートでは、150TB のデータセットのレプリケーション時間は、最初の転送で約 5.5 時間になります。その後、デルタは必要な転送時間内に完全に完了できるようになり、MSSQL バックアップのための堅牢な災害復旧プランが提供されます。

スケールアウトパフォーマンス

この 6 ノードクラスタは、毎秒 8.4GB 以上を FlexGroup ボリュームにプッシュできました。Customer Proof of Concept (CPOC) ラボでは、クラスタにノードを追加することでほぼ直線的にパフォーマンスが向上しています。次のグラフは、シングルノードの AFF A700 オールフラッシュストレージシステムと 2 ノードの AFF A700 のスループット結果を示しています。

図 4.3　CPOC スケールアウトスループットの結果



備考

バックアップワークロードのパフォーマンスを向上させたい場合は、ノードを追加します。

4.3.3 結論

FlexGroup ボリュームは、EDA やソフトウェアビルドなど、ファイル数の少ないワークロードや多いワークロードに適しているだけでなく、大容量のストリーミングファイルの高スループット要件にも対応できます。また、複数のノード間でストレージをスケールアウトすることでバックアップウィンドウを短縮し、すべてのクラスタリソースを適用し、回転中のドライブでもパフォーマンスを維持します。

5. FlexGroup 機能のサポートと最大化

この章では、FlexGroup ボリュームで使用でき ONTAP 機能について説明します。機能サポートが追加された ONTAP バージョンも記載されています。機能がこのセクションにリストされていない場合、情報については、当社サポート部門に連絡してください。データ類に固有の機能については、「ETERNUS AX/HX series ONTAP FlexGroup ボリュームのデータ保護とバックアップ」を参照してください。

表 5.1　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る ONTAP バージョン
Snapshot コピーテクノロジー	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 テクノロジー	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	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
AFF	ONTAP 9.1
クォータのレポート	ONTAP 9.1
SnapMirror テクノロジー	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
SMB 共有の Access Based Enumeration (ABE)	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
Volume Encryption (VE)	ONTAP 9.2
SnapVault テクノロジー	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 と統合レプリケーション	ONTAP 9.3
SMB のアンチウイルススキャン	ONTAP 9.3
自動ボリュームの自動拡張	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
インジェストヒューリスティックスの改善	ONTAP 9.3
SMB の変更 / 通知	ONTAP 9.3
ファイル監査	ONTAP 9.4
FPolicy	ONTAP 9.4
適応 QoS	ONTAP 9.4
QoS の最小値 (AFF のみ)	ONTAP 9.4
緩和された SnapMirror の制限	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	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

サポートされている機能	最初にサポートされる ONTAP バージョン
SMB コピーオフロード (オフロードデータ転送)	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
ボリューム名の変更	ONTAP 9.6
ボリュームの縮小	ONTAP 9.6
MetroCluster	ONTAP 9.6
エラスティックサイジング	ONTAP 9.6
継続的に使用可能な共有 (SMB)* * SQL Server および Hyper-V ワークロードのみ	ONTAP 9.6
Aggregate Encryption (AE)	ONTAP 9.6
Cloud Volumes ONTAP	ONTAP 9.6
FlexClone	ONTAP 9.7
FlexVol から FlexGroup へのインプレース変換 (「7.4.1 既存の FlexVol ボリュームを持つアグリゲートへの FlexGroup ボリュームの展開」を参照してください。)	ONTAP 9.7
アレイ統合用の vStorage API (VAAI)	ONTAP 9.7
NDMP	ONTAP 9.7
NFSv4.0 および NFSv4.1(パラレル NFS または pNFS を含む)	ONTAP 9.7
FexCache のオリジンボリュームとしての FlexGroup ボリューム	ONTAP 9.7
ファイルのクローン作成	ONTAP 9.8
プロアクティブなサイズ変更	ONTAP 9.8
NFS v4.2(基本プロトコルのサポート)	ONTAP 9.8
NDMP の機能強化 : EXCLUDE、RBE (再起動可能なバックアップ拡張機能)、MULTI_SUBTREE_NAMES	ONTAP 9.8
1,023 スナップショット	ONTAP 9.8
Qtree QoS	ONTAP 9.8

表 5.2　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る ONTAP バージョン
NFSv3	ONTAP 9.0
SMB2.1、SMB3.x	ONTAP 9.1 RC2
NFSv4.0、NFSv4.1、pNFS	ONTAP 9.7
NFSv4.2(基本プロトコルのサポート)	ONTAP 9.8

表 5.3 サポートされていない SMB2.x および 3.x の機能

サポートされていない SMB2.x の機能	サポートされていない SMB3.x の機能
SMB リモートボリュームシャドウコピーサービス (VSS)	• SMB ファイル共有の VSS • SMB ディレクトリリース • 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) **備考** SMB3.0 暗号化がサポートされています。

備考

リモート VSS は、[SMB Previous Versions] タブとは異なります。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、[SMB Previous Versions] タブが導入されて以来サポートされています。

5.1 サポートされていない SMB 機能の動作

通常、SMB 機能が ONTAP でサポートされていない場合は、機能しません。ONTAP FlexGroup では、サポートされていない SMB の機能に関する考慮事項があります。

表 5.4 サポートされていない SMB 機能が FlexGroup ボリュームでどのように動作するか

特徴	FlexGroup ボリュームでの動作
SMBv1.0	SMB 1.0 にアクセスするすべての共有で、アクセスが失敗するか、アクセスが拒否される。これは、Windows XP、Windows 2003、および SMB を使用して NAS に接続しようとするスキャナやコピー機などのオフィス機器に影響する可能性があります。
変更通知 /SMB ウォッチ	変更通知の詳細については、「10.12.2 SMB での変更通知の使用」を参照してください。 **備考** ONTAP 9.7 以降では、変更通知操作の並列処理により、全体的な操作性が向上しています。変更通知を使用する場合は、最適な結果を得るために ONTAP 9.7 以降を使用してください。
オフロードデータ転送 (ODX)	ODX の詳細については、この TechNet 記事を参照してください。
リモートボリュームシャドウコピーサービス (VSS)	警告はありません。リモート VSS が機能しないだけです。Remote VSS の主な用途は Hyper-V であり、FlexGroup ボリュームのワークロードとしては推奨されないため、影響は小さいはずです。 リモート VSS の詳細については、この TechNet 記事を参照してください。

5.2 最大値と最小値

このセクションでは、ONTAP FlexGroup ボリュームに固有の最大値と最小値について説明します。表 5.5 は、最大値をリストし、最大値がハードコード / 強制値であるか、推奨値またはテスト済みの値であるかを示しています。

表 5.5 FlexGroup の最大値

	値	値のタイプ
FlexGroup ボリュームサイズ	20PB	テスト済み / 推奨 *
FlexGroup 合計ファイル数	4000 億	テスト済み / 推奨 *
クラスタノード数	24 (12 組の HA ペア)	ハードコード / 強制
FlexVol メンバーボリュームのサイズ	100TB	ハードコード / 強制
FlexVol メンバのボリュームファイル数	20 億	ハードコード / 強制
SnapMirror ボリューム数 (FlexGroup ごとのメンバー)	200	ハードコード / 強制
SnapMirror ボリューム数 (クラスタごとの FlexGroup 合計)	6,000	ハードコード / 強制
ファイルサイズ	16TB	ハードコード / 強制
FlexVol メンバー構成数	200	テスト済み / 推奨 *
アグリゲートサイズ / 数	ONTAP の制限と同じ	ハードコード / 強制

表 5.6 FlexGroup の最小値

	値	値のタイプ
FlexVol メンバーのサイズ	100GB	テスト済み / 推奨 *
データ集計カウント	1	ハードコード / 強制
SnapMirror スケジュール	30 分	テスト済み / 推奨 *
スナップショットコピースケジュール	30 分	テスト済み / 推奨 *

備考

* テスト済み / 推奨として記載されている制限は、10 ノードクラスタに基づくテスト済みの制限です。プラットフォームで許可されている場合、実際の制限はハードコードされず、最大 24 ノードまでこの制限を超えることができます。詳細は、「7.10.2 理論上の最大値または絶対的な最大値」を参照してください。ただし、メンバーボリュームの公式サポート数は 200 です。この制限を超える必要がある場合は、当社サポート部門に連絡して、追加メンバーボリュームの認定プロセスを開始してください。

6. FlexGroup ボリュームが適しているかどうかの判断

ONTAP FlexGroup ボリュームは、多くの用途、特に「4.1 理想的な使用例」に記載されている用途に最適です。ただし、すべての使用例が FlexGroup ボリュームに適しているわけではありません。本章では、FlexGroup ボリュームがワークロードに適しているかどうかを判断するための情報を提供します。

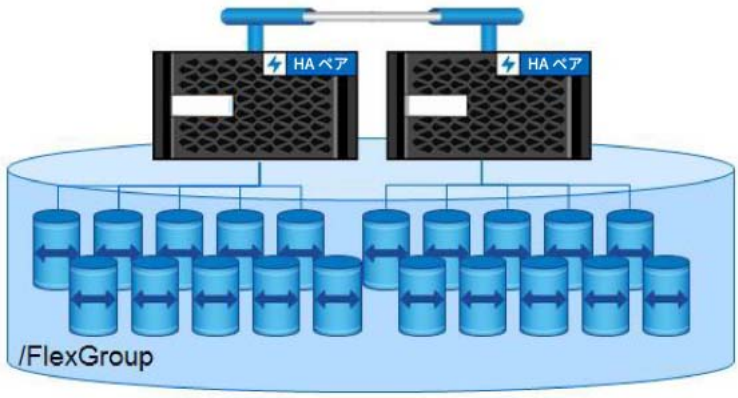
6.1 スケールアウトパフォーマンス

FlexGroup ボリュームは、集合的な FlexGroup ボリュームを構成する複数の構成要素間でデータと負荷を分散します。このモデルでは、FlexGroup ボリュームはワークロードに対応するために、各ノード内でより多くのリソース (CPU、ネットワークアダプタ、ディスクなど) を使用でき、クラスタ内でより多くのノードを使用できます。

さらに、この概念は ONTAP クラスタアーキテクチャと緊密に連携しているため、ノードとディスクを無停止で追加しても、アプリケーションに悪影響を及ぼすことなくパフォーマンスを向上させることができます。FlexGroup ボリュームでは、FlexGroup を拡張してメンバーを追加したり、無停止ボリューム移動テクノロジーを使用して新しいノード間でメンバーボリュームを再配分したりできます。

単一の FlexGroup ボリュームは、内部的に複数の個別の FlexVol ボリュームで構成されます。これらのボリュームは、任意のアグリゲートに保存でき、クラスタ内の複数のノードにまたがることができます。

図 6.1　FlexGroup ボリューム



さらに、この概念は ONTAP クラスタアーキテクチャと緊密に連携しているため、ノードとディスクを無停止で追加しても、アプリケーションに悪影響を及ぼすことなくパフォーマンスを向上させることができます。FlexGroup ボリュームでは、FlexGroup を拡張してメンバーを追加したり、無停止ボリューム移動テクノロジーを使用して新しいノード間でメンバーボリュームを再配分したりできます。

単一の FlexGroup ボリュームは、内部的に複数の個別の FlexVol ボリュームで構成されます。これらのボリュームは、任意のアグリゲートに保存でき、クラスタ内の複数のノードにまたがることができます。

クライアントが FlexGroup ボリュームにファイルとサブディレクトリを追加すると、ONTAP は新しいファイルとサブディレクトリの保存に最適な FlexVol メンバを自動的に決定します。FlexGroup ボリュームは、最高のパフォーマンスとデータおよび負荷分散の両方のために、データの編成を試みます。

このようにワークロードが分散されているため、FlexGroup ボリュームは FlexVol ボリュームよりも多くのメタデータトラフィックを処理できます。したがって、FlexGroup ボリュームは、メタデータ集約型のワークロードや大量のスループットを必要とするさまざまなワークロードに便利です。

6.2 機能の互換性の制限

FlexGroup ボリュームは、NFSv3、NFSv4.0、v4.1 および v4.2、SMB2.x、SMB3 などの一般的な NAS プロトコルをサポートします。これらのプロトコルのサポートおよびサポートされている ONTAP リリースの詳細については、「5. FlexGroup 機能のサポートと最大化」の表 5.1、表 5.2、表 5.3 を参照してください。

さらに、FlexGroup ボリュームは FlexCache ボリュームと ONTAP S 3 オブジェクトバケットを構成します。ただし、NAS プロトコルの相互作用を利用できるのは FlexCache ボリュームだけです。S3 バケットは、S3 プロトコルを介してのみアクセスできます。ONTAP S3 は ONTAP 9.8 で一般的に使用できます。

FlexGroup ボリュームでは、SMB1.0 はサポートされていません。FlexGroup ボリュームは、ブロックプロトコル /SAN アクセス (iSCSI、FCP、NVMe) をサポートしません。

表 6.1 に、ONTAP で現在使用可能なコンテナタイプを比較して、FlexGroup ボリュームが環境に適しているかどうかを判断するための情報を示します。

表 6.1 ONTAP ボリュームファミリの比較

	FlexVol ボリューム	FlexGroup ボリューム
クライアントアクセスプロトコル (現在のサポート)	SAN (FCP、iSCSI、NVMe) NAS • SMB1.0、2.1、3.x • NFSv3 • NFSv4.0、NFSv,4.1、NFSv4.2	S3 NAS • SMB2.x、3.x • NFSv3 • NFSv4.0、NFSv,4.1、NFSv4.2
容量の拡張	• 単一の FlexVol ボリューム • ネームスペース内の FlexVol ボリュームまたは FlexGroup ボリュームにマウント可能 • 100TB、20 億ファイルの制限	• ネームスペース内の FlexGroup ボリュームまたは FlexVol ボリュームにマウント可能 • 20PB* • 4000 億ファイル * • 無停止での容量増加 *10 ノードクラスタで現在テストされている制限。これらの値を超えて拡張することも可能。
メタデータのスケーリング	FlexVol ボリュームは、メタデータ処理とメタデータのシリアル処理のために 1 つのノードに制限されているため、ノードの CPU スレッドを最大限に活用しません。	FlexGroup ボリュームは、複数のノード (およびそのリソース) と複数のアグリゲートを使用できます。さらに、FlexGroup は複数のボリュームアフィニティを使用して CPU スレッドの使用率を最大にします。
ONTAP 機能の互換性	すべての ONTAP 機能との互換性	ほとんどの ONTAP 機能をサポートします。 • 詳細は、表 5.1 を参照してください。
スループットのスケーリング	制限対象 : • 1 つのノード (CPU、RAM、ネットワークポート、接続制限などのセット) • 1 つのアグリゲート	FlexGroup ボリュームは、I/O を処理するクラスタ全体のリソースを使用でき、単一の FlexVol ボリュームよりもはるかに高いスループットを提供します。ノードを FlexGroup に追加すると、パフォーマンスがリニアに向上します。
クラウドのサポート	次のようなすべてのクラウド統合 : • Cloud Volumes ONTAP • Cloud Volumes Services • SnapMirror Cloud	Cloud Volumes ONTAP (CVO): CLI のみ • CVO に適用される容量制限
ONTAP のアップグレードと復帰	ONTAP バージョンの変更中は、どのボリュームファミリに格納されているデータも安全に保持されます。	
GUI 互換性	• ONTAP System Manager • Active IQ Performance Manager • Active IQ Unified Manager • Cloud Insights	

6.3 パフォーマンスの合理化

単一の FlexGroup ボリュームは、複数の FlexVol メンバーボリュームで構成できます。これらのボリュームは、クラスタ内の任意のアグリゲートおよび任意のノードに配置できます。クライアントが FlexGroup ボリュームに対してトラフィックを誘導すると、ONTAP はそのトラフィックを、構成要素の異なる FlexVol ボリュームが実行するタスクに自動的に分割します。このアプローチでは、単一の FlexVol ボリュームでは処理できない操作を並行して実行できます。

このスケールアウト動作の利点は、ONTAP クラスタにノードを追加すると処理能力が劇的に向上することです。1 つの FlexGroup ボリュームは、1 つの FlexVol ボリュームよりもはるかに負荷の高いワークロードを、予測可能なレーテンシーで処理できます。

6.3.1 AFF A700 テスト

ソフトウェア構築ツール (Git) を用いた簡単な作業負荷ベンチマークにおいて、Linux カーネルを 2 ノード A700 クラスタ上でコンパイルした。

次の構成が使用されました。

- 2 ノード AFF A700 クラスタ
- ノードあたり 800GB SSD で構成された単一のアグリゲート
- FlexVol ボリューム : シングルノード、100% ローカル
- FlexGroup ボリューム : 高可用性 (HA) ペアにまたがる、ノードあたり 8 メンバー (合計 16 メンバー)

ワークロードは以下のとおりです。

- GCC ライブラリのコンパイル
- 複製操作のみ (これらの操作は、FlexVol と FlexGroup の両方で最高のスループットを示しました。)
- 4 台の物理サーバ
- 4 から 224 の範囲のクライアント上のユーザワークロード / スレッド

図 6.2 では、Git クローン操作で達成される最大スループット (読み取り + 書き込み) について、単一の FlexVol ボリュームと、2 つのノードにまたがる単一の FlexGroup ボリュームを比較しています。

備考

FlexGroup ボリュームの最大スループットは、FlexVol ボリュームの量の約 5 倍に達します。ワークロードが 64 スレッドに達する場合の FlexVol ボリュームと同じようにパフォーマンスが低下することはありません。

図 6.2　FlexVol ボリュームと FlexGroup ボリューム : ワークロードが増加している場合の最大スループット傾向

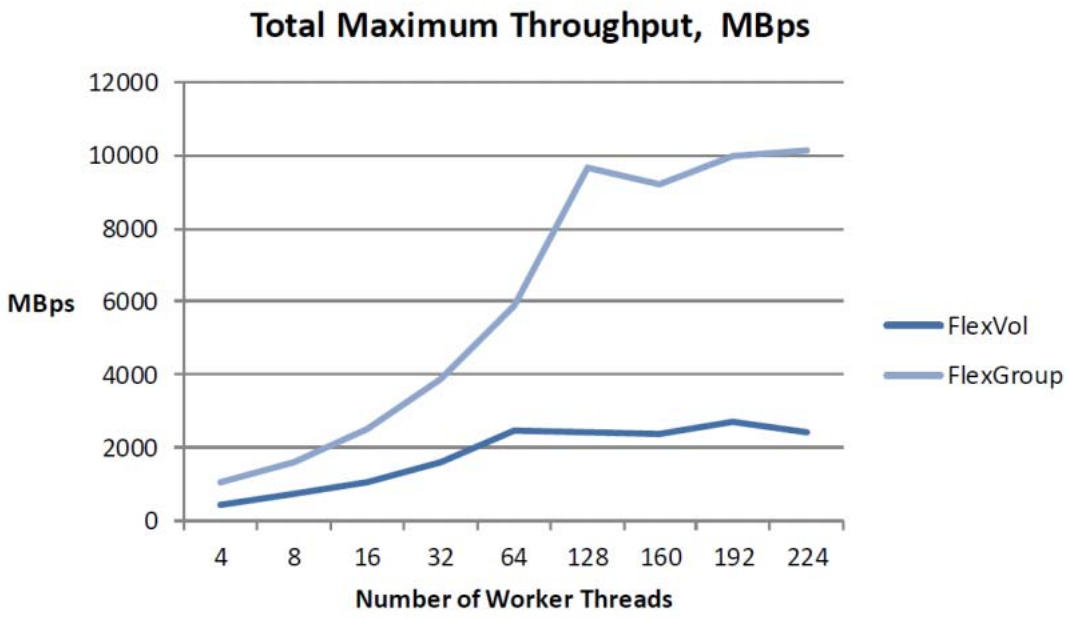


図 6.3 では、同じ構成内の FlexVol ボリュームと FlexGroup ボリュームを比較しています。今回は、最大読み取りスループットと最大書き込みスループットを個別に分析し、FlexVol ボリュームと FlexGroup ボリュームの平均スループットと比較します。

図 6.3　FlexVol ボリュームと FlexGroup ボリューム : ワークロードが増加している場合の最大スループットの傾向の詳細

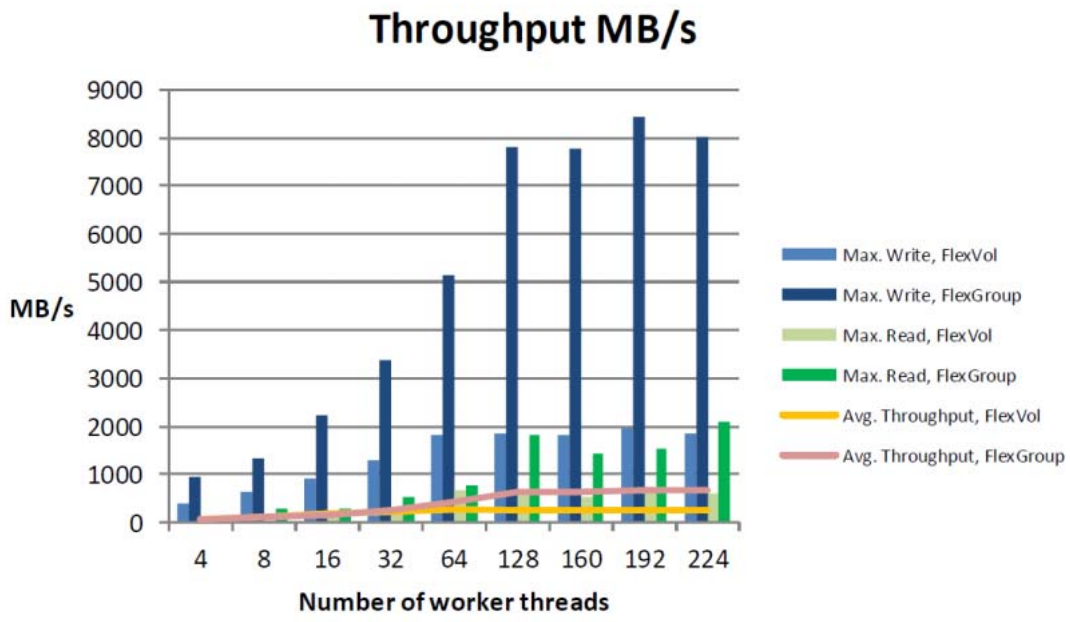
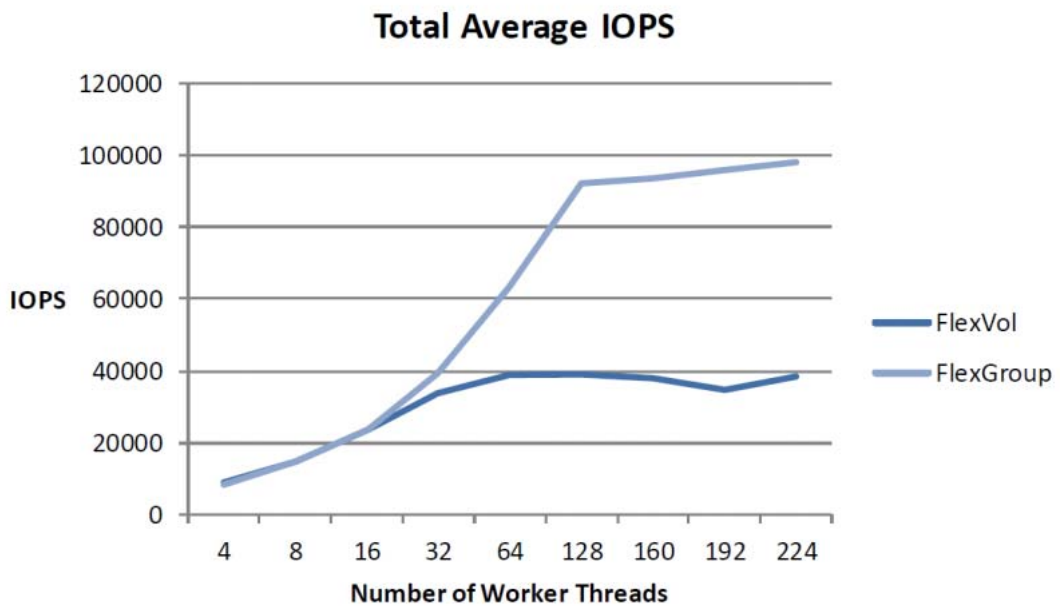


図 6.4 は、AFF A700 の FlexGroup ボリュームと FlexVol ボリュームの最大合計平均 IOPS を示しています。ここでも、FlexGroup ボリュームの IOPS が大幅に増加していることと、FlexVol ボリュームの IOPS が 64 スレッドで低下していることに注目してください。

図 6.4　FlexVol ボリュームと FlexGroup ボリューム：最大平均合計 IOPS



6.3.2 ビッグデータのワークロードによる FlexGroup のパフォーマンス

FlexGroup ボリュームは、クラスタ内の複数のコンピュートノード間で単一のネームスペースを拡張できる容量と機能を備えているため、Apache Hadoop、Splunk、Apache Spark などのビッグデータワークロードに興味深いユースケースを提供します。これらのアプリケーションでは、通常、1 つまたは 2 つのディレクトリだけが大量のデータと多数のファイルをダンプし、低レーテンシーで高スループットを必要とします。FlexVol ボリュームはこのパフォーマンスを達成できましたが、複数のボリュームを認識するようにアプリケーションを調整する必要がありました。

また、Customer Proof-of-Concept (CPOC) ラボでは、TeraSort ベンチマークを使用したパフォーマンステストを実施しました。このベンチマークは、Apache のビッグデータワークロードのテストに使用されます。このテストでは、ONTAP 9.2 を実行している 2 ノード AFF A700 クラスタを使用して、両方のノードの平均 CPU 使用率を約 55% に維持しながら、平均読み取りレーテンシーを約 3ms から 5ms、平均書き込みレーテンシーを約 4ms から 8ms にして、最大 8GBps をクラスタに入出力しました。ビッグデータのワークロードに FlexGroup ボリュームを使用すると、使用可能なすべてのハードウェアを使用でき、必要に応じてワークロードにノードを追加することで、容量とパフォーマンスを無停止で拡張できます。

図 6.5　FlexGroup ボリュームの TeraSort ベンチマーク統計の概要

ノード 1

Sysstat System Utilization

Min	Avg	Max	σ	Cv	Counter
13	42	75	12.4	0.3	cpu
3,447	25,991	54,412	9,913.2	0.4	nfs
752,703	2,179,080	3,572,454	597,657.6	0.3	net_kb_in
828,342	2,417,590	4,107,351	710,786.3	0.3	net_kb_out

ノード 2

Sysstat System Utilization

Min	Avg	Max	σ	Cv	Counter
24	56	84	13.3	0.2	cpu
14,403	38,608	68,636	10,024.7	0.3	nfs
254,777	2,216,472	4,283,503	766,327.7	0.4	net_kb_in
340,437	2,392,841	4,439,274	784,810.8	0.3	net_kb_out

さらに、ONTAP FlexGroup ボリュームで実行されているビッグデータワークロードでは、インラインアグリゲート重複排除、インラインデータコンパクション、インライン圧縮などのストレージ効率化機能により、約 50% の領域節約が実現されています。

6.3.3 ワークロードの自動適応

FlexGroup ボリュームは、クラスタ内の現在の状態に継続的に適応し、使用量を均等に維持し、動的な負荷を均等に分散するように常に動作を変更します。トレードオフはこの継続的な均衡のための行動に暗黙的に含まれます。この自動バランシングの代償は、FlexGroup ボリュームが、完全にバランスの取れた手動で編成された FlexVol ボリュームのコレクションが達成できるパフォーマンスと同じ理論上の最大パフォーマンスを達成できないことです。ただし、FlexGroup ボリュームはその最大値に非常に近い値になる可能性があり、その作業を完了するためにワークロードの事前知識は必要ありません。さらに、FlexGroup ボリュームは、複数の FlexVol アーキテクチャでは実現できないシンプルさを大規模なデータレイアウトにもたらします。

FlexGroup ボリュームは、さまざまなワークロードや高いデータ作成率に直面した場合に、より優れたパフォーマンス (負荷と使用率のよりスムーズなバランス) を発揮します。したがって、多くの異なる役割を実行する単一の FlexGroup ボリュームは、異なるワークロードに異なる FlexGroup ボリュームを使用する場合よりも、クラスタのリソースをより効果的に使用できます。ただし、データの制御性と柔軟性を高める必要がある場合は、複数の FlexVol ボリュームと FlexGroup ボリュームを同じ ONTAP SVM に結合できます。

ワークロードによって多数の小さなファイルが作成される場合、FlexGroup ボリュームは、パフォーマンスを向上させるためにフォルダのローカリティを重視しながら、それらのファイルをボリューム間で均等に分散するように配置します。ワークロードのサイズの大きいファイルの数が少ない場合、ONTAP はその違いを認識します。ONTAP では、ローカルフォルダの配置 (複数の大容量ファイルが同じメンバーボリュームに配置され、人為的にデータの不均衡が生じる場合もある) を優先するのではなく、ファイルをラウンドロビン方式で配置して、領域の均等な割り当てを保証します。これにより、FlexGroup ボリュームでさまざまなワークロードを最適に実行できるようになり、スペースの不均衡のシナリオを防ぎ、管理者の介入の必要性を減らすことができます。

ONTAP 9.8 では、プロアクティブなサイズ変更と呼ばれる、容量の管理方法の変更が導入されています。この変更により、容量のしきい値に達したときに、すべてのメンバーボリュームにわたって空きスペースバッファが効果的に維持され、メンバーボリュームがいっぱいになりすぎないように保護したり、異なる空きスペースを持つボリュームを保護したりできます。また、9.8 より前のバージョンでは、FlexGroup のメンバーボリュームの容量が 90% のしきい値に達した場合、ONTAP が新しいファイル用にリモートハードリンクを作成し始めるため、新しいファイル作成のパフォーマンスが低下していました。ONTAP 9.8 では 90% のしきい値が削除されています。これは、プロアクティブなサイズ変更では、トラフィックをリダイレクトする必要がないように十分な空き領域が維持されるためです。

6.3.4 インジェストアルゴリズムの改善

すべての ONTAP リリースでは、FlexGroup ボリュームのインジェストアルゴリズムがさらに改善されており、ONTAP が FlexGroup ボリュームに新しいデータを配置する方法をより適切に決定できるようになります。このアルゴリズムにより、メンバーボリュームが「ほぼいっぱい」の状態に近づいたときの FlexGroup ボリュームの対応方法も改善されます。

ベストプラクティス 1: 常に最新バージョンの ONTAP を実行する
最善のインジェスト結果を得るために FlexGroup ボリュームを使用する場合は、パッチ適用済みの最新バージョンの ONTAP を実行することを強く推奨します。最新リリースは、FTI (Fsas Technologies Inc.) のダウンロードサイトで入手できます。

さまざまなリリースで行われたインジェストの変更には、次のようなものがあります。

- ONTAP 9.7 での混合ワークロードタイプの処理の改善
- ONTAP 9.8 でのプロアクティブなサイズ変更とリモート配置トリガーの調整

6.3.5 パフォーマンス機能

ONTAP には、パフォーマンスを制御および監視するためのさまざまな機能が用意されています。

6.3.5.1 サービス品質 (QoS)

最大ストレージ QoS ポリシーを適用して、FlexGroup ボリュームのワークロードが他のボリュームワークロードを超過するのを防ぐことができます。ONTAP 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。

ONTAP は FlexGroup ボリュームの QoS 最小値 (「保証」または「フロア」とも呼ばれる) をサポートします。これにより、指定され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。

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、ワークロードに保証されたパフォーマンスを提供し、ワークロードのパフォーマンスを監視します。ワークロードを反応的に制限してパフォーマンスの問題に対処したり、ワークロードのパフォーマンスをプロアクティブに管理して問題を回避したりできます。

FlexGroup でのストレージ QoS ポリシーの動作

FlexGroup では、ストレージ QoS ポリシーはメンバーボリュームレベルではなく、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ができるため、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 6.6 と図 6.7 は、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 6.6 FlexGroup ボリューム上のストレージ QoS：シングルノード接続

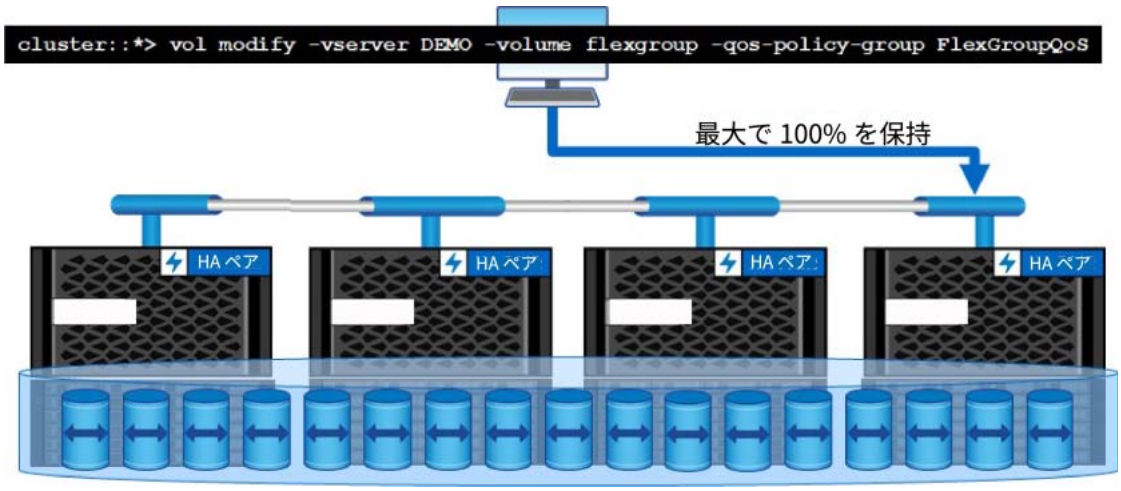
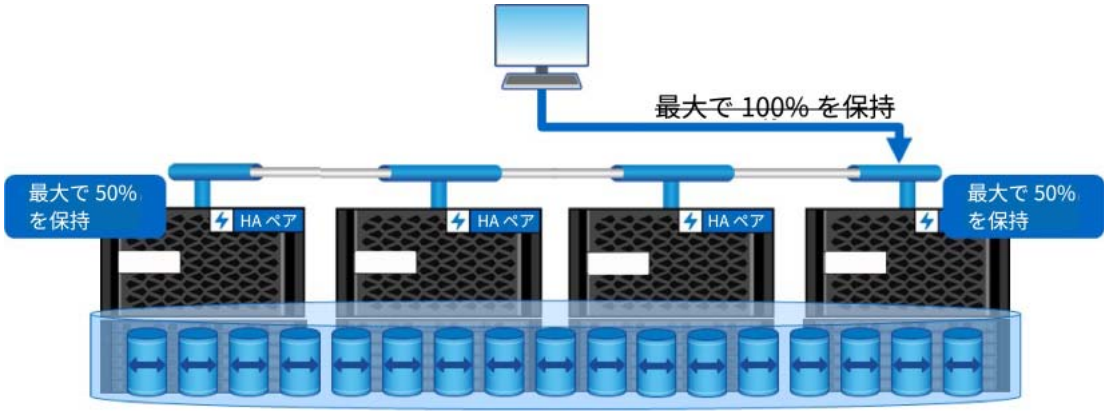


図 6.7　FlexGroup ボリュームのストレージ QoS：マルチノード接続



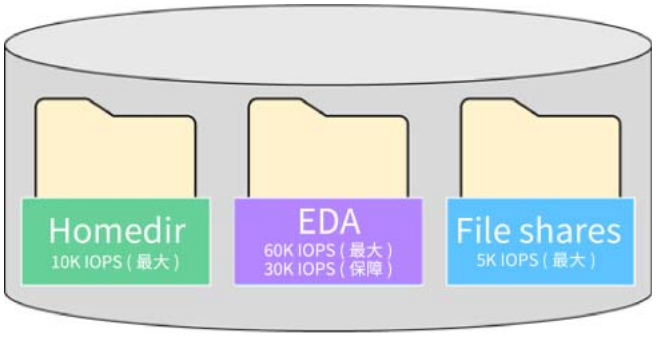
備考

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。

6.3.5.2　qtree QoS

ONTAP 9.8 では、qtree レベルで QoS ポリシーを適用する機能が導入されています。

図 6.8　qtree QoS の使用例



つまり、複数の FlexVol ボリュームまたは FlexGroup ボリュームを作成してワークロードを分割するのではなく、qtree を使用して FlexGroup ボリュームをプロビジョニングし、そのボリュームのパフォーマンスを管理できます。

qtree QoS では、以前の qtree 統計よりも詳細なレベルの統計情報も提供されます。

ONTAP 9.8 の qtree QoS は FlexGroup ボリュームおよび FlexVol ボリュームで使用できますが、次の制限があります。

- NFS のみ
- CLI/REST API のみ。現在の GUI サポートなし
- アダプティブ QoS のサポートなし

qtree QoS では、パフォーマンス監視のための統計情報も拡張されています。これは、特定のワークロードの理解に役立ちます。

```
Policy Group             IOPS      Throughput    Latency
-------------------- -------- --------------- ----------
qtree                     113      113.00MB/s     2.82ms
```

6.3.5.3 アダプティブ QoS

ONTAP は、FlexGroup ボリュームに対するアダプティブ QoS を導入しました。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 値と TB 値を調整できるようになりました。

備考

qtree は拡張または縮小できるオブジェクトではないため、アダプティブ QoS は qtree QoS でサポートされていません。

6.3.5.4 qtree 統計

FlexGroup ボリュームで qtree 統計が利用可能です。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

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

6.4 ワークロードと動作

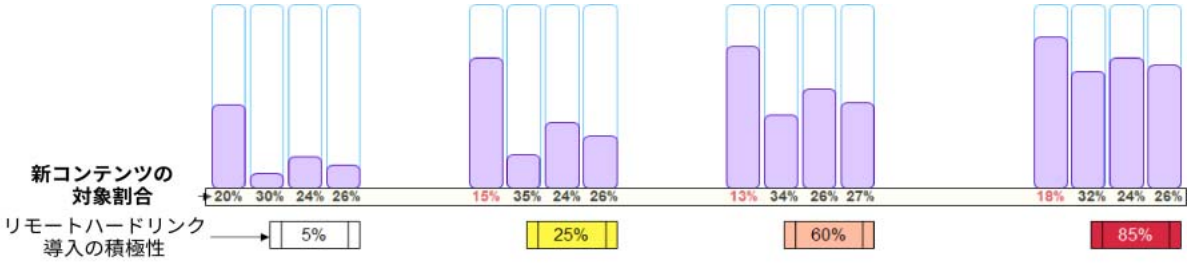
最適にバランシングされた FlexGroup ボリュームでは、すべての構成要素がほぼ同じ量のデータと負荷を持ち、ボリュームは最高のパフォーマンスを得るために高い頻度のローカル配置を使用しながらその状態を維持できます。フォルダと同様のサイズのファイルのバランスが良いワークロードでは、ローカルの親フォルダの配置を維持しながら、容量のバランスを比較的均等に保つことができます。

最適でない FlexGroup ボリュームには、ピアよりも多くまたは少なくデータを保持する構成要素や、より多くまたはより少ないトラフィックを受信する構成要素が含まれる場合があります。1 つのフォルダに多数のファイルが格納されている少数のフォルダしかないワークロードや、ファイルサイズが大きく異なるワークロードでは、FlexGroup 内でデータ使用量の不均衡が発生する可能性があります。

ただし、容量バランスは FlexGroup ボリュームの最も重要な機能ではありません。FlexGroup ボリュームは、パフォーマンスを向上させるためのローカル配置と、容量と inode 数のバランスが混在している場合に最適に機能します。メンバーボリューム間で容量のバランスを完全に保つためにパフォーマンスを犠牲にすることはしません。

次の図に、FlexGroup ボリューム内の容量不均衡の例と、FlexGroup ボリューム全体の使用量に基づいたリモート配置の程度の変化を示します。たとえば、データの不均衡がある比較的使用量が少ないの FlexGroup は、使用量が高い FlexGroup よりもリモートハードリンクを使用する可能性が低くなります。

図 6.9　　容量不均衡と遠隔配置の可能性



ONTAP は、メンバーボリュームの進行状況を常に監視し、FlexGroup ボリュームの現在の状態に基づいて配置の決定を調整します。あるメンバーボリュームのバランスが他のメンバーボリュームと少しずれている場合、インジェストの調整は行われない可能性があります。しかし、そのメンバーボリュームの容量が 90% に近づき始めた場合、または容量の差が 512GB を超える場合、ONTAP はこの差を修正するために新しいデータの配置をより積極的に選択します。この調整により、他のメンバーボリュームへのファイルの配置がよりリモートになり、FlexGroup ボリュームのパフォーマンスに悪影響 (目立たない可能性がある — 約 5% から 10%) を及ぼす可能性があります。

場合によっては、FlexGroup ボリュームの使用と負荷が完全に均衡しているように見えることがありますが、その状態を維持するためには、頻繁にリモートに配置する必要がありました。この状況は、FlexGroup メンバーボリュームの使用率が 100% に近づいたときに発生します。

ベストプラクティス 2: 容量の不均衡を心配しないでください。
メンバーボリューム間の容量の不均衡がある FlexGroup は、それ自体は問題ではなく、そのように扱うべきではありません。代わりに、FlexGroup が期待どおりに動作していない場合や、容量の不均衡が極端でクラスタの容量が不足している場合は、容量の不均衡を潜在的な原因として考えてください。FlexGroup のボリューム容量の不均衡がパフォーマンスの問題の原因であると考えられる場合は、必ず当社のサポートを受けてください。

ワークロードによって、FlexGroup ボリュームの最適な動作の度合いが決まります。ほとんどのワークロードは ONTAP 9.8 の FlexGroup ボリュームで使用できますが、一部のワークロード (EDA/ ソフトウェア開発など) は他のワークロードよりも最適に動作します。

6.4.1　最適なワークロード

FlexGroup ボリュームは、インジェスト負荷が高いとき、つまりファイルやディレクトリの作成率が高いときに最適に動作します。ONTAP は、新しいファイルとディレクトリが作成されるたびに配置を決定するため、このアクションが頻繁に発生するほど、負荷と使用率の不均衡を修正する機会が多くなります。ワークロードが既存のファイルに対して大量の読み取りまたは書き込みの追加を行う場合、FlexGroup の配置はそれほど影響しません。ファイルがいったん配置されると、それらは配置された場所に残ります。「6.3.4 インジェストアルゴリズムの改善」で説明したように、それぞれの新しい ONTAP リリースでは、より多様なワークロードに対応できるように FlexGroup ボリュームに改良と調整が加えられています。

一般的に、最も最適な FlexGroup ワークロードの属性を次に示します。

- FlexGroup ボリュームは、多数の小さなサブディレクトリで最適に動作
  これは、ディレクトリごとに数十から数百のファイルが存在することを意味します。FlexGroup ボリュームでは、新しい子サブディレクトリをリモートで配置しながら、個々のファイルを親ディレクトリのローカルに保持して最高のパフォーマンスを得ることができるからです。より多くのファイルを含むディレクトリでは、容量とファイル数のバランスをとるために、他のメンバーボリュームへの配置がよりリモートになります。

- FlexGroup ボリュームは、同時に発生する大量のトラフィックに適切に対応
  FlexGroup ボリュームに対して同時に負荷をかけるために、より多くのワークロード (特に、同時に異なる処理を行う複数のクライアントからのトラフィック) を導入すると、全体的なパフォーマンスを向上させることができます。つまり、FlexGroup ボリュームを限界までプッシュして、このドキュメントで説明されているパフォーマンスの可能性を 1 から数台のクライアントで実現することは期待しないでください。
- 空き領域が十分にある場合は、FlexGroup ボリュームが最適に機能
  構成要素が満杯になり始めると、FlexGroup ボリュームはリモート配置をより頻繁に採用し始め、ピアが満杯になる前に 1 つの構成要素が満杯にならないようにします。このようにリモート配置の使用が増えると、メタデータのパフォーマンスが低下します。
- FlexGroup ボリュームは、メタデータの書き込み処理のレートが高い場合に最適
  ONTAP FlexVol ボリュームは、すでに読み取り / 書き込み I/O とメタデータ読み取り操作 (GETATTR など) を並行して処理しています。ただし、ONTAP は書き込みメタデータ (SETATTR や CREATE) を連続的に処理するため、通常の FlexVol ボリュームにボトルネックが発生する可能性があります。FlexGroup ボリュームは、これらのタイプのワークロードに並列処理オプションを提供します。その結果、これらのタイプのワークロードのパフォーマンスが 2 ～ 6 倍向上します。

6.4.1.1 パフォーマンスと容量に関する考慮事項

最高のパフォーマンスを得るには、負荷が高いときに FlexGroup ボリュームに十分な空き容量 (10% 以上の空きスペース) を確保してください。

FlexGroup のボリューム容量を管理しやすくするには、ONTAP 9.8 以降を使用します。このリリースには容量を管理するための機能を多く備えています。これには、ボリュームの自動拡張、エラスティックサイジング、特に ONTAP 9.8 では、プロアクティブなサイズ変更などがあります。

FlexVol メンバーまたは構成要素ボリュームの空きスペースは、admin 権限レベルで次のコマンドを使用して監視できます。

```
cluster::> vol show -vserver SVM -volume [flexgroupname__]* -is-constituent true -fields
available,percent-used
```

備考

また、Active IQ Unified Manager などの GUI ユーティリティを使用するか、ONTAP を設定してアラートを生成することで、空き領域を監視することもできます。詳細は、「14. 容量の監視とアラート」を参照してください。

6.4.2 良好なワークロード

ワークロードが前述のパラメータに適合しない場合でも、FlexGroup ボリュームで容易に対応できる可能性があります。「6.4.1 最適なワークロード」では、FlexGroup ボリュームの最適な動作に役立つ状況について説明していますが、ほとんどのユースケースでは、最適でない状況でも優れたスループット、拡張性、および負荷分散が得られます。

6.4.3 理想的でないワークロード：大容量ファイル

いくつかのアクティビティは、FlexGroup ボリュームが構成要素間の負荷と使用のバランスを維持するのをより困難にします。これらのアクティビティのほとんどは、いろいろな点で大きなファイルに関連しています。これらのワークロードは FlexGroup ボリュームを使用できますが、実装する前にワークロードの平均ファイルサイズと最大ファイルサイズを理解しておく必要があります。「理想的な」または「良い」ワークロード定義に適合しないワークロードをデプロイする場合は、次の点を考慮する必要があります。

6.4.3.1 考慮事項 1: ONTAP では、ファイルの将来のサイズを予測できません。

大容量ファイルワークロードの主な課題の 1 つは、ストレージシステムが時間の経過とともにファイルのサイズがどの程度大きくなるかを認識していないことです。クライアントもこの情報を持っていないことがよくあります。ストレージシステムでは、ファイルは小さな inode として開始され、ファイルの作成 / 書き込みが完了するまでデータが書き込まれます。これは、パフォーマンスを考慮して、FlexGroup ボリュームがファイルの配置を親フォルダに対してローカルに維持する傾向があるため、さらに悪化します。ファイルの書き込み速度と、ファイルの作成に関与するクライアントの数によっては、500MB のファイルを 100 個含むフォルダがすべて同じメンバーに配置されることがあります。同様に、4K サイズのファイルを 100 個含むフォルダが同じメンバーに配置されることがあります。その結果、このシナリオでは、1 つのメンバーボリュームが 50GB の使用スペースになり、もう 1 つのメンバーボリュームが 400KB しか使用されていない可能性があります。

前述したように、これは本質的には問題ではありませんが、ストレージ管理者にとっては明らかな違いであり、FlexGroup ボリュームのサイズが適切に設定されていないと問題が発生する可能性があります。たとえば、メンバーボリュームのサイズがすべて 100GB の場合はどうなりますか。この例では、1 つのメンバーの使用率が 50% で、残りのメンバーの使用率が 0% になります。

一般に、このようなワークロードは時間の経過とともにバランスがとれ、FlexGroup ボリュームは均等な分散と良好なパフォーマンスを維持します。これらのワークロードに関するベンチマークは、「私のメンバーボリュームの容量は不均等です。」ではなく、「FlexGroup ボリュームのパフォーマンスが期待どおりではありません。」である必要があります。

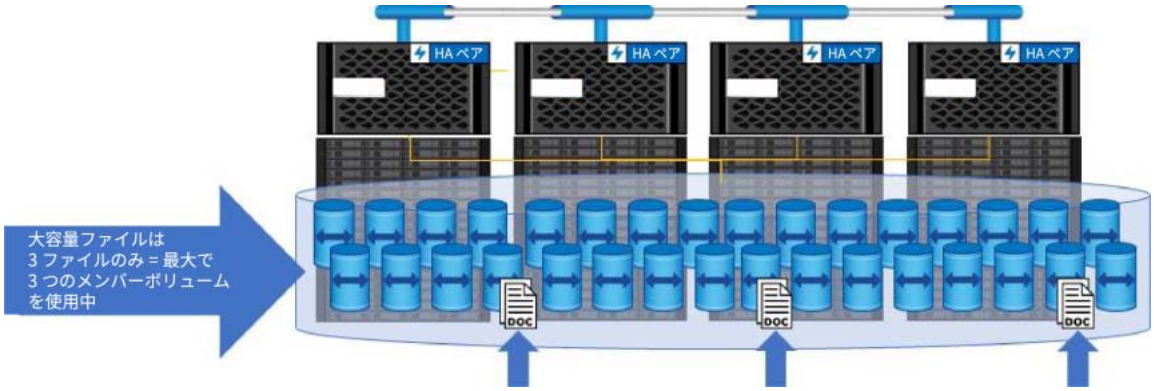
6.4.3.2 考慮事項 2: 大容量ファイルのワークロードは一般的にファイル数の少ないワークロード

大きなファイルは、小さなファイルよりも FlexGroup ボリュームの処理がわずかに難しくなります。これは主に、大きなファイルを使用すると、全体的に使用するファイルの数が少なくなるためです。前述したように、FlexGroup ボリュームは、新しいファイルやディレクトリが頻繁に作成される場合に最適なパフォーマンスを発揮します。ワーキングセットが、ほぼ同じサイズの多数の大きなファイルで構成されている場合、FlexGroup ボリュームでは、構成要素間の使用率と負荷分散の維持に問題はありません。大容量ファイルのワークロードの場合のパフォーマンスは、FlexVol のパフォーマンスに似ています。これは、一度に多数のファイルをインジェストするワークロードではないワークロードでは、並列インジェストのメリットが生かされないためです。

6.4.3.3 考慮事項 3: 大容量ファイルのワークロードの均等な分散が保証されない

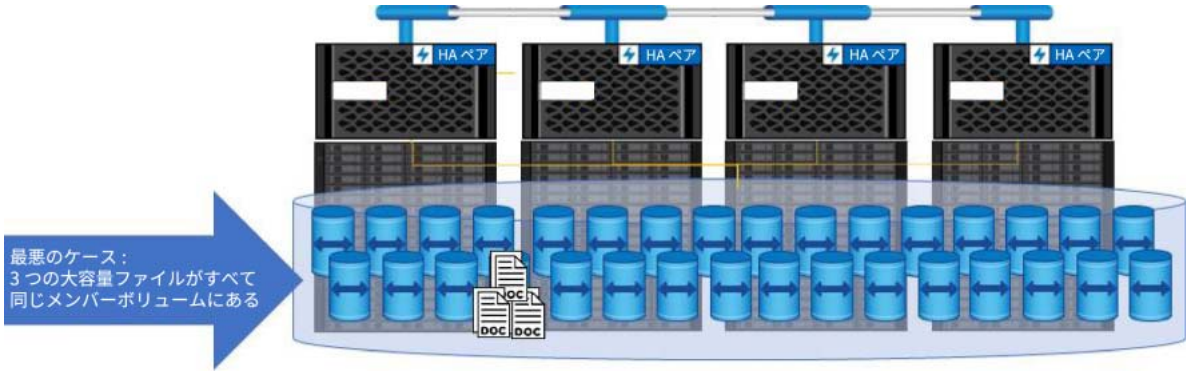
大きなファイルには、大量の情報を保持する特性もあります。それだけの情報を読み書きするには、長い時間がかかります。ワークロードがこれらのサイズの大きいファイルの一部 (大規模な単一ファイルデータベースの読取りまたは書込みなど) に集中している場合は、これらのファイルをホストする構成要素によってのみ、すべてのトラフィックが処理されます。その時点では他の構成要素がワークロードに参加していないため、この状況では FlexGroup ボリュームの使用が最適ではなくなります。一般的に、FlexGroup ボリューム内の大容量ファイルまたはストリーミングワークロードのパフォーマンスは、FlexVol ボリュームの場合とほぼ同じです。

図 6.10 大きなファイルがいくつかある FlexGroup ボリュームにおいて、使用が最適ではない理由



また、大きなファイルがすべて FlexGroup ボリューム全体に均等に分散されるという保証はありません。上の図では、3 つの大きなファイルが 3 つの異なるメンバーボリュームに配置されています。このシナリオは、これらのファイルのいずれかを一度に書き込み、次のファイルが書き込まれる前にファイル全体が書き込まれる場合に発生する可能性があります。ただし、3 つのファイルすべてが同時に書き込まれる場合は、3 つのファイルすべてが同じメンバーボリュームに格納される可能性があります。

図 6.11 大容量ファイルの最悪のシナリオ：すべてが同じメンバーボリュームに配置



ここでも、FlexGroup ボリュームに十分な容量がある場合は、必ずしも問題ではありませんが、大容量ファイルワークロードを処理する際には、いくつかの要因を考慮する必要があります。

- 合計 FlexGroup サイズ
- メンバーボリューム数
- メンバーボリュームサイズ (最大ファイルサイズとの比較)

メンバーボリュームを大きくすることで、大容量ファイルによって発生する可能性のある問題を回避できます。また、プロアクティブなサイズ変更やエラスティックサイジングなどの機能により、データの可用性に影響を及ぼす可能性のある容量の問題を緩和できます。

6.4.3.4 考慮事項 4: 大きなファイルでは、パフォーマンスに影響を与える可能性のある不均衡が生じる

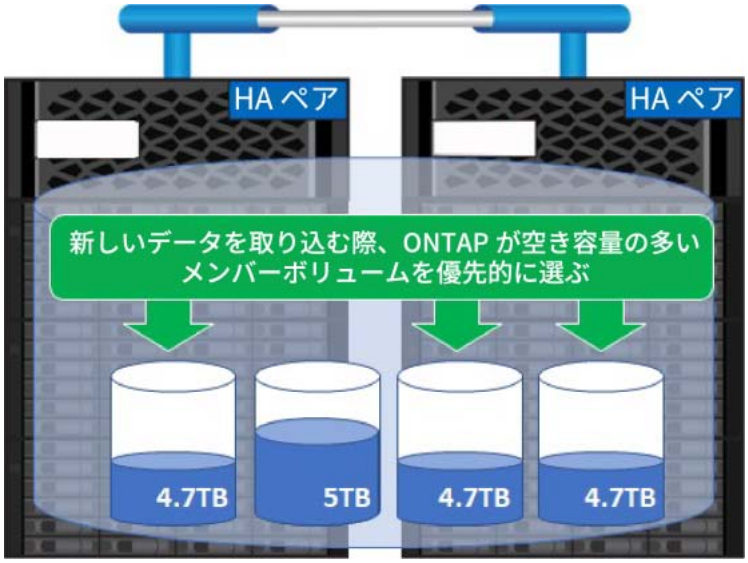
大きなファイルに関するもう 1 つの懸念は、1 つのファイルが構成要素上の十分なスペースを消費し、構成要素間の使用バランスに大きな影響を与える可能性があることです。場合によっては、いくつかのファイルが平均ファイルサイズよりも桁違いに大きくなることがあります。その結果、一部の構成要素 (異常に大きなファイルを保持しているもの) のデータ量は、他の構成要素のデータ量よりもはるかに多くなります。これに応じて、FlexGroup ボリュームは、他の新しいコンテンツ作成を、十分に活用されていない構成要素に転用し始めます。その結果、構成要素のサブセットがほとんどのトラフィックにサービスを提供することになります。通常、この問題は深刻ではありません。単に最適ではない行動を表しています。ONTAP 9.7 以降のバージョンでは、これらのタイプのファイルとワークロードの配置の処理が大幅に向上しているため、メンバーボリューム間のバランスが向上しています。

ベストプラクティス 3: 大容量ファイルのサイズに関する考慮事項
FlexGroup ボリュームのサイズを決定する前に、分析を実行してワークロード内で可能な最大ファイルサイズを決定します。その場合、メンバーボリュームのサイズには大きなファイルサイズが反映されるため、大きなファイルが FlexGroup メンバーボリュームの 1% ～ 5% を超える量を消費することはありません。このベストプラクティスに従うことは、「スペースがない」懸念を避けるのに役立ちます。また、ONTAP を実行すると、エラスティックサイジング機能により「スペースがない」の問題を回避できます。FlexGroup ボリュームでは、パッチが適用された最新バージョンの ONTAP を実行することが常に最善です。

もう 1 つの問題は、FlexGroup ボリュームが常にフルに近い状態で実行されることです。FlexGroup ボリュームがいっぱいになると、ONTAP は、空き容量のある構成要素に新しいコンテンツを積極的に配置します。ワーキングセットが主に小さなファイルで構成されている場合、この動作は、全体的な FlexGroup ボリュームが実際にいっぱいになるまで、クライアントが `Volume Full` エラーを受信しないようにするのに適しています。

エラスティックサイジング機能によって、ONTAP が空きメンバーボリュームから領域を借用し、メンバーボリュームへのファイル書き込みを完了できます。ONTAP 9.8 では、プロアクティブなサイズ変更も導入されており、フルメンバーボリュームの容量管理がさらに向上しています。

図 6.12 容量不均衡の例



6.4.3.5 FlexGroup ボリュームで大容量ファイルを使用する場合のベストプラクティス

FlexGroup ボリュームは、多数の小さなファイルを処理する場合に最適です。ただし、FlexGroup ボリュームが事前にそのワークロードを考慮するように設定されている場合は、大きなファイルを保存するときにも効果的です。大容量ファイル用に FlexGroup ボリュームのサイズを設定する場合、大容量ファイルとは何か、ワークロード内の最大および平均ファイルサイズは何かを考慮することが重要です。

FlexGroup ボリュームを設計するときは、ファイルサイズを考慮して、メンバーボリュームのサイズを適切に設定する必要があります。ONTAP 9.7 およびそれ以降のバージョンでは、エラスティックサイジングが追加されているため、このプロセスは不要です。また、ONTAP 9.8 では、プロアクティブなサイズ変更により、容量管理がさらにシンプルになります。一般に、大容量ファイルのワークロードには、次のベストプラクティスを適用できます。

- ファイルサイズが大きい場合は、FlexGroup ボリュームあたりのメンバー数が少ない、より大きなメンバーボリュームの導入を検討してください。詳細は、初期ボリュームサイズに関する考慮事項を参照してください。
- ONTAP 9.7 を実行している場合、エラスティックサイジングは既定で有効になっています。ボリュームの自動拡張を使用すると、これらのリリースで FlexGroup ボリュームのエラスティックサイジングが無効になるため、容量の管理方法を決定します。
- プロアクティブなサイズ変更のメリットを得るには、ONTAP 9.8 以降を使用することをお勧めします。ONTAP 9.8 以降では、エラスティックサイジングとボリュームの自動拡張を共存させることができます。
- クォータの適用を使用して、qtree またはユーザーごとに容量を制限および監視します。
- FlexGroup ボリュームを導入する前に、XCP を使用してファイルシステムをスキャンし、ファイルサイズを分析して平均ファイルサイズ、最大ファイルサイズなどを把握します。
- メンバーボリュームが不均衡になる可能性が低くなるように、FlexGroup ボリュームのサイズを設定する必要があります。最大ファイルサイズは、メンバーボリュームの容量の 1% ～ 5% を超えないようにする必要があります。ただし、メンバーボリュームサイズが小さい場合、ボリューム容量の 1% ～ 5% に対して、最大ファイルは比較的小さくする必要があることに注意してください。小さいファイルほど、ストレージシステムへの書き込みが速く完了し、インジェストアルゴリズムを大幅に変更するほどの不一致は生じません。ファイルサイズが 50GB 以下の場合は、可能であれば、メンバーボリュームサイズを 1TB 未満にしないでください (メンバーボリュームの最小サイズは 100GB)。「大容量ファイル」の概念はメンバーボリューム領域の 1 ～ 5% であるため、「大容量ファイル」のサイズ値は、100GB のメンバーボリューム (1GB ～ 5GB) の方が、1TB のメンバーボリューム (10GB ～ 50GB) よりもはるかに小さくなります。

6.4.4 期待されるパフォーマンス : 読み取り負荷の高いワークロード

FlexGroup ボリュームのパフォーマンスは、多くのファイルを取り込む書き込みメタデータの多いワークロード (高レベルの CREATE および SETATTR 呼び出し) では、FlexVol ボリュームや競合システムのパフォーマンスを大幅に上回ることがあります。ただし、ファイルストリーム、ファイルアペンド、読み取り負荷の高いワークロードなどの他のワークロードでは、取り込み負荷の高いワークロードで FlexVol ボリュームを大きく超える性能は出せません。これは、FlexGroup ボリュームが、書き込みメタデータのワークロードに対するボリュームアフィニティを高めることによって、これらのワークロードのシリアル処理のボトルネックを克服するように設計されているためです。基本的な読み取りと書き込みは、このシリアル処理のボトルネックには直面しません。

場合によっては (特にすべてのローカルトラフィックで)、ランダムおよびシーケンシャルな読み取り / 書き込みワークロードでは、複数の FlexVol ボリュームのセットの方が FlexGroup ボリュームよりもパフォーマンスがわずかに優れていることがあります。ただし、1 つの FlexGroup ボリュームではなく、複数の FlexVol ボリュームの作成と管理に伴う複雑さが、わずかなパフォーマンス向上を上回る場合があります。

読み取り負荷の高いワークロードの場合、FlexGroup ボリュームを使用すると、単一の FlexVol ボリュームを使用するよりも次のような利点があります。

- 複数の CPU およびノード間での拡張により、複数のファイルへの読み取りをロードバランシング
- 大容量バケット用の単一ネームスペース

FlexGroup ボリュームを使用するかどうかを決定する場合は、特定の機能のサポートを検討してください。FlexGroup ボリュームで現在サポートされているものとサポートされていないものについては、前のセクションを参照してください。

また、読み取り負荷の高いワークロードを使用する場合は、FlexGroup のオリジンボリュームに接続された FlexCache ボリュームを導入して、クラスタ内のより多くのボリュームアフィニティ間、または同じクラスタ内の他の SVM 間、または複数のサイトやクラウド内の他のクラスタ間でワークロードを分散することも検討してください。

6.4.5 FlexGroup ボリューム内のデータ不均衡

まれに、FlexGroup ワークロードのメンバーボリュームの容量が不均衡になることがあります。これだけでは、問題を示すものではありません。これが問題になるのは、パフォーマンスが著しく低下している場合や、容量の不均衡によってクラスタの空き領域が不足している場合だけです。ほとんどの場合、容量の不均衡に対処する必要はありません。新しいファイルの作成に一貫性がある場合、ONTAP は作業負荷を分散するために作業を実行します。これは、プロアクティブなサイズ変更のため、ONTAP 9.8 以降で特に当てはまります。データが静的である場合、不均衡は解消されません。ただし、パフォーマンスは、データの不均衡がない場合とほぼ同じです。

データの不均衡は、次のシナリオで発生する可能性があります。

- サイズの大きいファイルと小さいファイルが混在して FlexGroup ボリュームに書き込まれる場合。
- ファイルは FlexGroup に書き込まれ、後で追加されるため、使用される容量が増加します。
- 複数の大きいファイルが同じフォルダに一度に書き込まれます。ONTAP では、これらのファイルがどれくらい大きくなるかわからないため、パフォーマンスを考慮してローカルメンバーボリュームに配置されます。
- 大規模なデータセットが削除され、1 つのメンバーボリュームで他のメンバーボリュームよりも多くのファイルが削除される場合。
- ユーザーが非常に大きなファイル (多数の既存ファイルの zip ファイルなど) を作成する場合。

各 ONTAP リリースでは、インジェストアルゴリズムを調整して、より広範なワークロードシナリオに対応できるようにしているため、利用可能な最新の ONTAP リリースを使用してください。問題が発生している場合は、テクニカルサポートケースを開いて問題を特定し、修正します。

6.4.5.1 配置後リバランス

現在、ONTAP には、すでに取り込んだファイルを再配分する方法はありません。効果的にリバランスする唯一の方法は、FlexGroup ボリュームから新しい空の FlexGroup ボリュームにデータを一度にコピーすることです。データの移動後、クライアントとアプリケーションは新しい FlexGroup ボリュームを参照する必要があるため、このプロセスは中断されます。また、この処理はファイルレベルで行われるため、かなりの時間がかかる場合があります。ファイルのリバランスは、ボリュームの不均衡によって本番環境に影響を与える問題が発生した場合にのみ考慮する必要があります。前述したように、容量の不均衡は通常、クライアントのアクティビティには認識されません。ほとんどのお客様は、容量の閾値に関するアラートを受け取るまで、不均衡には気づきません。データの再配分が必要な場合は、XCP Migration Tool を使用してファイルのコピープロセスを高速化することを検討してください。

7. FlexGroup の初期設計に関する考慮事項

この章では、ONTAP FlexGroup のボリュームの初期設計に関する考慮事項について説明します。この情報を表示する際、クラスタ上に以前の FlexGroup ボリュームが作成されていないことを前提としています。また、ユーザーが CLI および GUI を使用した ONTAP の管理に関する経験と知識があり、ストレージシステムへの管理者レベルのアクセス権を持っていることを前提としています。

7.1 クラスタに関する考慮事項

NAS 機能 (CIFS/SMB および NFS) のみを使用する ONTAP クラスタは、最大 24 ノード (12 組の HA ペア) まで拡張できます。各 HA ペアは同種のシステム (つまり、2 つの ETERNUS AX　series ノードなど) ですが、クラスタ自体に混在したシステムタイプを含めることができます。たとえば、10 ノードのクラスタでは、4 つの ETERNUS AX　series ノード、4 つの ETERNUS HX series ノード、ストレージ階層化機能用の 2 つのハイブリッドノードを混在させることができます。

FlexGroup ボリュームは、24 ノードのクラスタ全体に及ぶ場合があります。ただし、次の点に注意してください。

- FlexGroup ボリュームは、理想的には同一のハードウェアシステムだけを対象とします。
  ハードウェアシステムは、CPU、RAM、および全体的なパフォーマンス機能の点で大きく異なる可能性があるため、同種のシステムのみを使用すると、FlexGroup ボリューム全体で予測可能なパフォーマンスが向上します。FlexGroup ボリュームにメンバーボリュームが導入されていれば、どこでもデータのバランスをとることが可能ですが、ストレージ管理者はこの配置を制御しません。
- FlexGroup ボリュームは、同一のディスクタイプにのみスパンします。
  ハードウェアシステムと同様に、ディスクタイプのパフォーマンスも大きく異なります。FlexGroup ボリュームはクラスタ内の複数のノードにまたがることができ、ストレージ管理者はデータが配置される場所を制御できないため、使用されているアグリゲートがすべて SSD、すべてスピン、またはすべてハイブリッドであることを確認してください。ディスクタイプを混在させると、パフォーマンスが予測できなくなる可能性があります。
- ディスクサイズはそれほど重要ではありません。
  FlexGroup がスパンする可能性のあるアグリゲートでは、同様のディスクタイプを使用することが重要ですが、ディスクサイズはそれほど重要ではありません。たとえば、アグリゲートに 3TB のディスクがあるが、新しい 16TB のディスクセットを購入した場合、メディアタイプが同じであれば、それらのディスク全体に FlexGroup を導入しても構いません。ここでの主な注意点は、展開するメンバーボリュームのサイズが、他のメンバーボリュームのサイズと同じである必要があることです。
- FlexGroup ボリュームは、クラスタの一部にまたがることができます。
  FlexGroup ボリュームは、クラスタ内の任意のノードの組み合わせ、つまり 1 つのノードから HA ペア、さらに 24 のすべてのノードにまたがるように構成できます。クラスタ全体にまたがるように FlexGroup ボリュームを構成する必要はありません。ただし、そのように構成すると、使用可能なすべてのハードウェアリソースを利用することができます。

7.2 ONTAP バージョンに関する考慮事項

ONTAP の各リリースには、FlexGroup ボリュームに関する新機能と改良が含まれています。パッチが適用された最新の ONTAP リリースを使用することを推奨していますが、多くのストレージ管理者はそれを実行できないか、実行したくないと考えています。

古いバージョンの ONTAP を実行する必要がある場合は、「5. FlexGroup 機能のサポートと最大化」を参照して、そのリリースの機能の違いをよく理解し、可能であれば、本番環境に展開する前に FlexGroup ボリュームでワークロードをテストしてください。

7.3 障害ドメイン

障害ドメインは、障害が発生した場合にワークロードに悪影響を及ぼす可能性のあるエンティティです。たとえば、ONTAP クラスタでは、HA ペアの両方のノードに障害が発生した場合 (まれな出来事)、それらのノード上のボリュームはフェイルオーバー先がないため使用できなくなります。その結果、HA ペアはクラスタ内の障害ドメインと見なされます。ただし、パートナーが引き継ぐことができるため、HA ペア内の 1 つのノードが故障しても使用できなくなることはありません。この状況では、単一のノードは障害ドメインとは見なされません。

障害ドメイン内のエラー (RAID エラー、ディスクの消失、マルチパス構成エラー、メタデータの不整合など) は ONTAP で処理され、FlexGroup ボリュームに悪影響を与えません。

FlexGroup ボリュームは、複数のノードおよび HA ペアにまたがることができるため、複数の障害ドメインにまたがることができます。ただし、FlexGroup ボリュームが 10 ノードクラスタ全体にまたがる場合でも、障害ドメインは HA ペアのままです。FlexGroup ボリューム内のメンバーにアクセスできなくなった場合 (HA ペアのまれな障害の場合など)、それらのメンバーがすべて修復されて FlexGroup ボリュームに再導入されるまで書き込みアクセスは無効になります。FlexGroup ボリュームがまたがる HA ペアの数が多いほど、障害が発生する可能性が高くなります。これは、またがる障害ドメインが増えているためです。使用される HA ペアが少ないほど、障害が発生する可能性は低くなります。ただし、ワークロードに使用できるハードウェアリソースが少ないため、FlexGroup 全体のパフォーマンスは低下します。

そのため、導入を計画するときは、FlexGroup ボリューム内にまたがるノードの数と許容できる SLA を検討し、必要な容量と必要なパフォーマンスを比較検討します。

7.4 アグリゲートレイアウトの考慮事項

アグリゲートとは、RAID グループに配置される物理ディスクの集合であり、FlexVol ボリュームや FlexGroup ボリュームなどの仮想エンティティのバックエンドストレージリポジトリを提供します。各アグリゲートは特定のノードによって所有され、ストレージのフェイルオーバーイベント中に再割り当てされます。

アグリゲートは整合性ポイントのために専用の NVRAM パーティションを持ち、アグリゲートが遅くなったり劣化したりしてノード全体に問題が発生するシナリオを回避します。これらの整合性ポイントは、アグリゲート単位の整合性ポイントとも呼ばれ、同じノード上でディスクシェルフタイプを混在させて、ストレージシステムの設計の柔軟性を高めることができます。

ベストプラクティス 4: FlexGroup によるアグリゲートの使用
FlexGroup ボリュームまたは複数の FlexVol ボリュームを使用している場合に一貫したパフォーマンスを得るには、FlexGroup ボリュームまたは FlexVol ボリュームの設計が、アクティブなワークロードに対して同じディスクタイプおよび RAID グループ構成を持つアグリゲートのみに及ぶことを確認します。コールドデータの階層化では、予測可能なパフォーマンスはそれほど重要ではないため、ディスクタイプやアグリゲートを混在させても大きな影響はありません。

表 7.1 は、FlexGroup ボリュームを使用する場合のアグリゲートレイアウトに推奨するベストプラクティスを示しています。これらのプラクティスはハード要件では**ない**ことに注意してください。ETERNUS AX series システムではノードごとに 1 つのアグリゲートを推奨しています。これは、RAID Triple Erasure Coding (RAID-TEC) に関連するディスクコストに起因しています。これは、複数の RAID グループに必要な追加のパリティドライブ用に高価な SSD 容量を使い果たしたくない場合があるためです。ただし、ADP では、パーティションはデータディスク全体に分散されるため、このような場合には、ETERNUS AX series ではノードごとに 2 つのアグリゲートの方が適しています。これは、2 つのアグリゲートが存在するノードごとの使用可能なボリュームアフィニティが多いためです。

表 7.1 FlexGroup ボリュームを使用したアグリゲートレイアウトのベストプラクティス

スピニングディスクまたはハイブリッドアグリゲート	ETERNUS AX series
ノードごとに 2 つのアグリゲート	ノードごとに 2 つのアグリゲート (ADP なし) ノードごとに 2 つのアグリゲート (ADP あり)

備考

一貫したパフォーマンスを得るには、アグリゲートでは、FlexGroup ボリューム全体でドライブと RAID グループの数が同じである必要があります。

既存の FlexVol ボリュームを処理する際のアグリゲートレイアウトの詳細については、「7.3 障害ドメイン」を参照してください。

7.4.1 既存の FlexVol ボリュームを持つアグリゲートへの FlexGroup ボリュームの展開

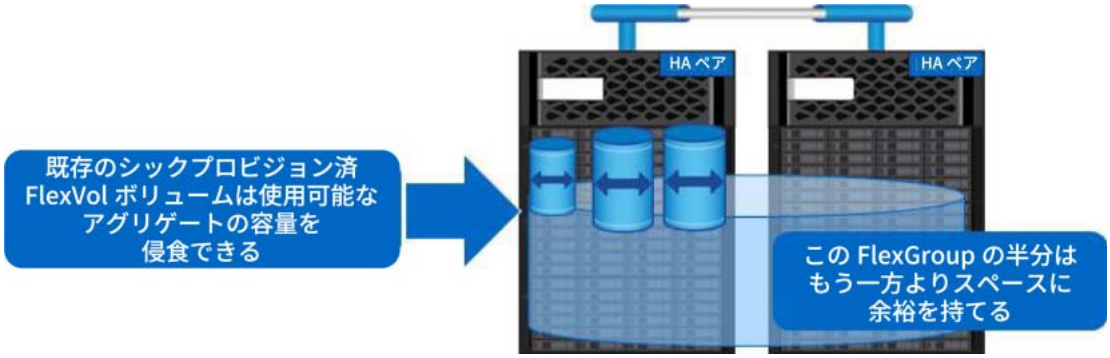
FlexGroup ボリュームはクラスタ内の複数のアグリゲートにまたがることができ、通常の FlexVol ボリュームと同じ SVM に共存できるため、FlexGroup ボリュームが既存の FlexVol ボリュームとアグリゲートを共有しなければならない場合があります。したがって、FlexGroup ボリュームを展開する際には、このセクションで説明する要素を考慮する必要があります。

7.4.1.1 既存の FlexVol ボリュームの容量フットプリントを考慮する

FlexGroup ボリュームは複数のアグリゲートにまたがることができますが、これらの各アグリゲートに同じ数の FlexVol ボリュームがない場合があります。したがって、既存の FlexVol ボリューム容量が FlexGroup ボリュームの容量を食いつぶしてしまう可能性があるため、容量保証が無効になっている FlexGroup ボリュームのインジェスト分散に影響を及ぼす可能性がある、異なる空き容量がアグリゲートに存在する可能性があります。

たとえば、node1 上の aggr1 にそれぞれ 1TB の FlexVol ボリュームが 4 つあり、node2 上の aggr2 にそれぞれ 1TB の FlexVol ボリュームが 2 つある場合、node1 のアグリゲートの容量は node2 よりも 2TB 少なくなります。両方のノードにまたがる FlexGroup ボリュームを展開し、両方のアグリゲートを満たすためにオーバープロビジョニングされている場合、node1 のメンバーボリュームの容量レポートにはすでに「使用スペース」が含まれています。これにより、使用されている容量がすべてのメンバーボリュームに均等に渡るまで、node2 のメンバーはデータのインジェストの大部分を吸収することになります。

図 7.1 FlexVol 容量が FlexGroup 負荷分散に与える影響



備考

これは、FlexGroup ボリュームがシンプロビジョニングされている場合にのみ問題になります。スペースが保証された FlexGroup ボリュームでは、他のボリュームがスペースのフットプリントを圧迫することはありません。ただし、システム内の他のボリュームがスペースの割り当てを妨げている場合は、スペースが保証されている FlexGroup ボリュームを必要なサイズに作成できないことがあります。

7.4.1.2 既存の FlexVol ボリュームのパフォーマンスへの影響を考慮する

FlexGroup ボリュームを導入する際には、既存の FlexVol ボリュームの作業量を考慮することも重要です。1 つのノード上の一連の FlexVol ボリュームが特定の時間に大量にヒットする場合、既存の FlexVol ボリュームと同じノードにまたがって集約される FlexGroup ボリュームのパフォーマンスに悪影響を及ぼす可能性があります。これは、FlexVol ボリュームで見られる影響と似ていますが、FlexGroup ボリュームは複数のノードにまたがることができるため、データ I/O が発生しているノードによっては、パフォーマンスの効果がクライアント側からは断続的に見える場合があります。

この影響を緩和する 1 つの方法は、ストレージ QoS ポリシーを使用して、これらのボリュームの IOPS とスループットを制限するか、FlexGroup ボリュームの QoS を最小限に抑えてパフォーマンスを保証することです。また、無停止でのボリューム移動を使用して、ノード間でボリュームを再配分し、パフォーマンスへの影響を分散させることもできます。

7.4.1.3 ボリューム数制限を考慮する

ONTAP には、使用中のノードのタイプ (ETERNUS AX series または ETERNUS HX series) とノードのパーソナリティに依存する、ノードあたりのボリューム数の制限があります。データ保護パーソナリティを備えたシステムは、備えていないシステムよりも多くのボリュームを使用できます。
また、使用しているノードタイプに関係なく、クラスタ全体のボリューム制限は 12,000 です。これは、ONTAP クラスタごとにプロビジョニングできる FlexVol ボリュームの数に影響します。

通常、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るため、これらのメンバーボリュームはこの合計制限にカウントされます。さらに、多くの ONTAP 機能では、アーキテクチャに FlexGroup ボリュームを活用しています。たとえば、1 つの FlexGroup ボリュームが 16 個のメンバー FlexVol ボリュームを使用するとします。そのボリュームに FlexClone を使用する場合は、32 個の FlexVol ボリュームを使用しています。そのクラスタ内に FlexCache ボリュームを作成する場合 (FlexGroup でもあります)、そのキャッシュボリュームに対して ONTAP が選択した数を使用します。

ほとんどの場合、複数の小さい FlexGroup ボリュームを作成する必要はありません。代わりに、より大きな FlexGroup ボリュームをプロビジョニングし、qtree を使用してワークロードを分離します。

ベストプラクティス 5: 既存の FlexVol ボリュームを使用した FlexGroup ボリュームの導入
FlexGroup ボリュームを導入する前に、以下の手順に従ってください。 • 既存の FlexVol ボリュームがある場合は、複数の FlexGroup ボリュームと、それに対応する機能をミックスに追加しても、ボリューム数の制限を超えないことを確認してください。 • Active IQ Unified Manager および ONTAP System Manager のパフォーマンスヘッドルーム機能を使用して、使用率の高いノードを確認してください。 • 不均衡がある場合は、無停止でのボリューム移動を使用して、使用率の低い他のノードに「ホット」ボリュームを移行し、ノード間で可能な限りワークロードのバランスをとります。 • FlexGroup ボリュームで使用するアグリゲートの空き領域を評価し、使用可能な領域がほぼ同等であることを確認してください。 • ボリューム数制限の影響が潜在的な要因である場合は、新しいボリュームを追加する余裕のあるノード間で FlexGroup ボリュームを作成するか、無停止ボリューム移動を使用してボリュームを再配置し、ボリューム数のバランスをとります。 • または、ボリューム数の制限が懸念される場合は、より少ないメンバーボリュームで FlexGroup ボリュームを作成します。

7.5 Flash Cache と Flash Pool

Flash Cache カードと Flash Pool のアグリゲートは FlexGroup ボリュームでサポートされていますが、これらを使用する場合は、一貫したパフォーマンス結果を得るために、FlexGroup ボリュームに参加しているすべてのノードに 1 つずつ配置してください。Flash Cache カードは、FlexGroup ボリュームに対して、FlexVol ボリュームに対する場合と同じパフォーマンス上のメリットを提供することが期待されています。

7.6 高度なディスクパーティション設定

FlexGroup ボリュームは、ADP の使用には関係ありません。特別な配慮は必要ありません。

7.7 SyncMirror (ミラーリングされたアグリゲート)

FlexGroup ボリュームは、SyncMirror 構成に参加するアグリゲート上に配置できます。これは、追加のデータ保護機能のためにアグリゲートを内部でレプリケートする方法です。FlexGroup は、SyncMirror アグリゲート上に完全に存在する必要があります (たとえば、すべてのメンバーボリュームが SyncMirror アグリゲート上に存在するか、すべてが存在しません)。それ以外の場合、SyncMirror は使用できません。

SyncMirror の詳細については、当社サポート部門にお問い合わせください。

備考

SyncMirror は StrictSync (SnapMirror Synchronous) と同じ機能を提供しません。FlexGroup ボリュームは現在、StrictSync または SnapMirror Synchronous をサポートしていません。SyncMirror は MetroCluster に近いものです。FlexGroup ボリュームに関する最新のデータ保護情報については、「ETERNUS AX /HX series ONTAP FlexGroup ボリュームのデータ保護とバックアップ」を参照してください。

7.8 MetroCluster

ONTAP は、MetroCluster デプロイメント (IP) での FlexGroup ボリュームをサポートしています。

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て、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。

備考

Volume Encryption (VE) または Aggregate Encryption (AE) を MetroCluster で使用する場合は、VE/AE を有効にする前に必ず MetroCluster 構成を完了してください。

7.9 Cloud Volumes ONTAP

ONTAP では、クラウド上で動作する ONTAP ソリューションである Cloud Volumes ONTAP を公式にサポートしています。Cloud Volumes ONTAP (CVO) を使って FlexGroup ボリュームをデプロイできます。

Cloud Volumes ONTAP で動作する FlexGroup ボリュームは、Cloud Volumes ONTAP インスタンスに導入された ONTAP バージョンで利用可能なものと同じ機能セットを使用できます。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) インスタンスで使用する EDA リポジトリ
- オンプレミス SnapMirror で使用するデータのバックアップとアーカイブ

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 368TB で、FlexGroup ボリュームは複数のクラスタインスタンスにまたがることはできません。CVO 内の FlexGroup ボリュームは、CLI または ONTAP System Manager を使用してのみ作成できます。現時点では、Cloud Manager を使用して FlexGroup ボリュームを作成することはできません。

7.10 容量に関する考慮事項

FlexGroup は大容量とファイル数の可能性を実現しますが、FlexGroup ボリューム自体は、基盤となるハードウェアの物理的な最大数に制限されています。現在指定されている最大値 (20PB、4000 億ファイル) は、テストされた最大値にすぎません。理論上の最大値はもう少し大きくなるかもしれませんが、FlexGroup ボリューム内で公式にサポートされているメンバーボリューム数は現在 200 です。FlexGroup ボリューム内に 200 を超えるメンバーボリュームが必要な場合は、担当者に問い合わせるか、当社サポート部門にご連絡ください。

また、ノード固有のアグリゲートサイズの制限により、100TB の FlexVol ボリュームのセット数のみが許可されます。詳細については、ハードウェアの物理容量の制限を確認してください。

たとえば、ETERNUS HX6100 は 400TB のアグリゲートが許可されています。これは、アグリゲートあたり最大 4 つの 100TB ボリュームが許可されることを意味します。
ただし、メンバーボリュームの 100TB の制限に達しないことを推奨しています。これは、メンバーボリュームの容量が不足した場合に、将来メンバーボリュームをそれ以上拡張することができなくなるためです (その場合は、容量を増やすために新しい 100TB のメンバーボリュームを追加する必要があります)。その代わりに、FlexVol メンバの最大合計領域の 10% から 20% 以上の保存領域を残すことを目標とし、自動拡張、エラスティックサイジング、プロアクティブなサイズ変更などの緊急スペース割り当て機能を有効にします。

これらの数値は、Snapshot リザーブ、WAFL リザーブ、ストレージ効率、FabricPool クラウド階層化などの機能が考慮される前の未フォーマット時容量です。FlexGroup ソリューションのサイズを正しく設定するには、当社サポート部門に連絡して支援を受けてください。

7.10.1 最大値と最小値

このセクションでは、ONTAP FlexGroup ボリュームに固有の最大値と最小値について説明します。表 7.2 は、最大値をリストし、最大値がハードコード / 強制値であるか、推奨値 / テスト済みの値であるかを示しています。

表 7.2 FlexGroup の最大値

	値	値のタイプ
FlexGroup ボリュームサイズ	20PB	テスト済み / 推奨 *
ファイル数	4000 億	テスト済み / 推奨 *
クラスタノード数	24 (12 組の HA ペア)	ハードコード / 強制
FlexVol メンバーボリュームのサイズ	100TB	ハードコード / 強制
FlexVol メンバのボリュームファイル数	20 億	ハードコード / 強制
SnapMirror ボリューム数 (FlexGroup ごとのメンバー)	200	ハードコード / 強制
SnapMirror ボリューム数 (クラスタごとの FlexGroup 合計)	6,000	ハードコード / 強制
ファイルサイズ	16TB	ハードコード / 強制
FlexVol メンバー構成数	200	テスト済み / 推奨 *
アグリゲートサイズ / 数	ONTAP の制限と同じ	ハードコード / 強制

表 7.3 FlexGroup の最小値

	値	値のタイプ
FlexVol メンバーのサイズ	100GB	テスト済み / 推奨 *
データアグリゲート数	1	ハードコード / 強制
SnapMirror スケジュール	30 分	テスト済み / 推奨 *
スナップショットコピースケジュール	30 分	テスト済み / 推奨 *

備考

* テスト済み / 推奨として記載されている制限は、10 ノードクラスタに基づくテスト済みの制限です。プラットフォームで許可されている場合、実際の制限はハードコードされず、最大 24 ノードまでこの制限を超えることができます。詳細は、「7.10.2 理論上の最大値または絶対的な最大値」を参照してください。ただし、メンバーボリュームの公式サポート数は 200 です。この制限を超える必要がある場合は、担当営業に連絡して、追加メンバーボリュームの認定プロセスを開始してください。

7.10.1.1 クラスタ内の FlexGroup ボリュームの最大数

FlexGroup ボリュームは、単一の FlexVol メンバーボリュームまたは数百の FlexVol メンバーボリュームで構成できます。FlexVol メンバーボリュームの最大数は、クラスタ内の合計ボリューム数によってのみ物理的に制約されます。その結果、FlexGroup ボリュームは理論上、24 ノードのクラスタ内に最大 24,000 のメンバーボリュームを持つことができます。

FlexGroup ボリュームの合計数も、クラスタ内の合計ボリューム数によって同様に制限されます。各 FlexGroup ボリュームのメンバーボリュームはボリュームカウントの一部であるため、クラスタで許可される FlexGroup ボリュームの数は、メンバーボリュームの数によって異なります。

たとえば、2 ノードのクラスタでは、2,000 個のボリュームを使用できます。その結果、(他にも可能性はあるが) 次のいずれかの構成にすることができます。これらはすべて、最大 2,000 のボリュームになります。

- 10 個の FlexGroup ボリューム中に 200 個のメンバーボリューム
- 20 個の FlexGroup ボリューム中に 100 個のメンバーボリューム
- 40 個の FlexGroup ボリューム中に 50 個のメンバーボリューム
- 200 個の FlexGroup ボリューム中に 10 個のメンバーボリューム

クラスター内の他の FlexVol ボリューム (SVM ルートボリュームを含む) が、使用可能なメンバーボリュームの合計数に影響することに注意してください。FlexVol メンバーボリューム制限は、SnapMirror 関係に参加している FlexGroup ボリュームによっても制限されます。これらの制限の最新の詳細については、「ETERNUS AX/HX series ONTAP FlexGroup ボリュームのデータ保護とバックアップ」を参照してください。

ほとんどの場合、ONTAP はデフォルトで複数のメンバーボリュームを作成します。FlexGroup 作成方法の詳細は、「7.10.3 FlexVol メンバーボリュームのレイアウトに関する考慮事項」を参照してください。複数の FlexGroup ボリュームを作成すると、クラスタ内のボリューム数が意図せずに使い果たされてしまうことがあります。一般に、CLI コマンド `volume create -auto-provision-as` を使用して新しい FlexGroup ボリュームを作成すると、メンバーボリューム数の詳細に行き詰まることはありません。全体的に見て、作成する FlexGroup ボリュームの数を減らして容量を増やし、qtree を使用してワークロードを分割することをお勧めします。ONTAP 9.8 の qtree では、クォータの適用、詳細な統計情報、QoS ポリシーが提供されています (現在、qtree QoS は NFS でのみサポートされています)。

7.10.2 理論上の最大値または絶対的な最大値

FlexGroup ボリュームで公式にサポートされる制限は、構成ボリューム 200 個、20PB、ファイル 4000 億個です。ただし、これらは単に 10 ノードのクラスタでテストされた制限です。クラスタ内のノードごとに許可される最大ボリュームを考慮すると、制限が大幅に拡大する可能性があります。

最終的に、FlexGroup ボリュームのアーキテクチャ上の制限は、基盤となるハードウェア容量と、1 つのクラスタ内で許容されるボリュームの総数です。

表 7.4 ONTAP で許可されるボリューム数に基づく FlexGroup の理論上の最大値

クラスタの最大サイズ	クラスタごとの現在のアーキテクチャの最大メンバーボリューム (ONTAP 9.8)	FlexGroup ボリュームあたりの理論上の最大容量	FlexGroup ボリュームあたりの理論上の最大 inode 数
24 ノード	12,000	最大 1200PB (メンバーボリュームあたり 100TB×11,999 個の FlexGroup メンバーボリュームに基づく)	最大 24 兆 inode (20 億個の inode×11,999 個の FlexGroup メンバーボリュームに基づく)

備考

1 つのクラスタで使用できる 100TB のメンバーボリューム数の主な制限要因は、基盤となる物理ハードウェアの制限であり、これはプラットフォームによって異なります。

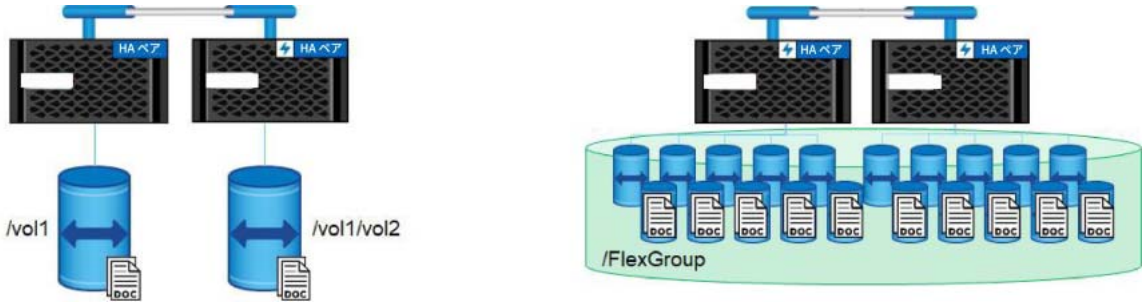
表示されている 20PB、4000 億ファイル、200 メンバーボリュームの制限を超える場合は、担当営業に連絡して、認定プロセスを開始してください。

7.10.3 FlexVol メンバーボリュームのレイアウトに関する考慮事項

FlexVol ボリュームは、FlexGroup ボリュームの構成要素です。各 FlexGroup ボリュームには、複数のメンバー FlexVol ボリュームが含まれています。これにより、パフォーマンスを同時に発揮し、単一の FlexVol ボリュームの通常の上限である 100TB を超えてボリュームの容量を拡張できます。

標準 FlexVol ボリュームは、アグリゲートの利用可能なストレージからプロビジョニングされます。FlexVol ボリュームは柔軟性があり、環境に影響を与えたり、障害を与えたりすることなく動的に増加または減少させることができます。FlexVol ボリュームは、アグリゲート内の特定のディスクセットには関連付けられておらず、アグリゲート内のすべてのディスクに存在します。ただし、個々のファイル自体はストライプされません。個々の FlexVol メンバーボリュームに割り当てられます。

図 7.2 FlexVol と FlexGroup のアーキテクチャの比較



このアーキテクチャと、FlexGroup の操作に影響を与える可能性のある大容量ファイルのために、FlexGroup ボリュームをプロビジョニングする際に留意すべき考慮事項がいくつかあります。

FlexGroup ボリュームを設計するときは、基礎となる FlexVol メンバーボリュームについて、次の点を考慮してください。

- `volume create -auto- provision-as flexgroup` や ONTAP System Manager などの自動 FlexGroup 作成方法を使用する場合、FlexGroup ボリューム内のメンバー FlexVol ボリュームのデフォルト数は、このセクションで説明するいくつかの要因によって異なります。

> **備考**
>
> ほとんどすべての使用例で、より大きな FlexGroup ボリューム (10TB 以上) を作成する場合は、ノードごとのメンバーボリューム数を ONTAP で決定することを推奨します。FlexGroup ボリュームのサイズが小さい場合は、ワークロード内のファイルサイズと、それらが使用する可能性のあるメンバーボリュームあたりの容量の割合に一層の注意が必要です。

FlexGroup ボリュームの作成方法に関係なく、FlexVol メンバーボリュームは均等な容量で導入されます。たとえば、8 つのメンバーで構成される 800TB の FlexGroup ボリュームを作成した場合、各メンバーは 100TB で導入されます。導入時に必要な FlexVol メンバーボリュームの量が多いか少ない場合は、`-aggr-list` および `-aggr-list-multiplier` オプションを指定した `volume create` コマンドを使用して、アグリゲートごとに導入されるメンバーボリュームの数をカスタマイズします。「7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか ?」を参照してください。

7.10.3.1 導入方法 1: コマンドライン

FlexGroup ボリュームを導入するには、ONTAP コマンドラインを使用することをお勧めしています。ただし、CLI からこれを行うには、手動と自動の 2 つの方法があります。
どちらのコマンドでも、`volume create` コマンドセットを使用します。

FlexGroup の自動作成 (auto-provision-as): CLI

これは、使いやすさと、作成中の予測可能なデプロイメントロジックおよび警告を組み合わせて、誤って設定された FlexGroup ボリュームを防止できるため、FlexGroup の作成に推奨される方法です。自動 CLI 方式を使用するには、`volume create -auto-provision-as flexgroup` コマンドを実行します。デフォルトでは、このコマンドは FlexGroup ボリュームを次のパラメータでプロビジョニングします。

- N 個のメンバーボリューム (各 100GB、クラスタ内のアグリゲートごとに 4 つ、8 つのメンバーボリュームごとに最大 2 つのアグリゲート)
- 合計サイズ = メンバーあたり 100GB × メンバーボリュームの数 (16 個のメンバーボリューム =1.6TB)
- ノードまたはハードウェアのタイプに関係なく、使用されるクラスタ内のすべてのノードとデータのアグリゲート

コマンドを実行すると、設定を知らせる警告が表示されます。Y と入力する前に、メンバーボリュームのサイズが適切であることを必ず確認してください。リストされているアグリゲートは正しく、同じメディアタイプを使用していますか？

ほとんどの場合、オプションなしで指定したデフォルト値では不十分です。デプロイメントで使用するアグリゲートまたはノードを指定したい場合があります。

このため、FlexGroup の自動プロビジョニング用にいくつかの追加設定オプションフラグがあります。

```
-support-tiering                    -use-mirrored-aggregates
-encryption-type                    -nodes
-size                               -state
-policy                             -user
-group                              -security-style
-unix-permissions                   -junction-path
-comment                            -max-autosize
-min-autosize                       -autosize-grow-threshold-percent
-autosize-shrink-threshold-percent  -autosize-mode
-space-guarantee                    -type
-percent-snapshot-space             -snapshot-policy
-language                           -foreground
-nvfail                             -qos-policy-group
-qos-adaptive-policy-group          -caching-policy
-encrypt                            -is-space-reporting-logical
-is-space-enforcement-logical       -tiering-policy
-tiering-object-tags                -analytics-state
```

これらのフラグを使用して、FlexGroup のサイズ、階層化ポリシー、スペース保証、ノード、アグリゲートなどをカスタマイズできます。

自動化 CLI コマンドの一般的な動作を次に示します。

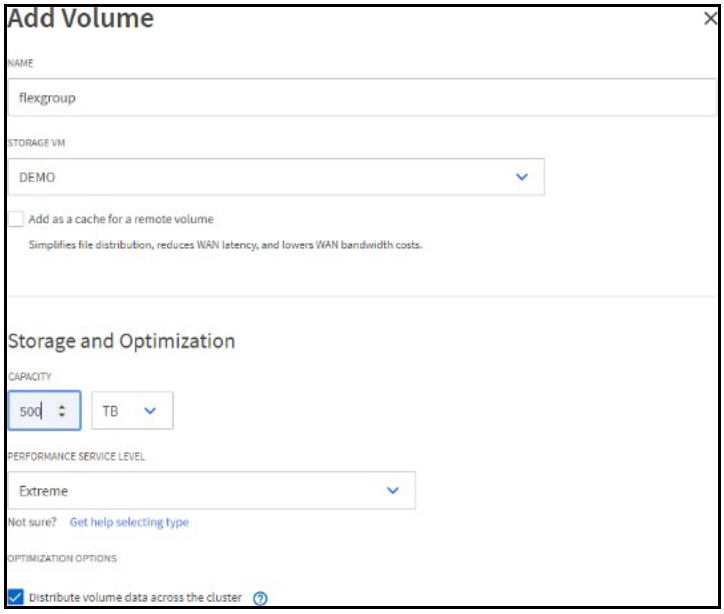
- 可能であれば、ノードごとに 2 つのアグリゲートを使用します。そうでない場合は、ノードごとに 1 つのアグリゲートを使用します。
- 各ノードで同じ数のアグリゲートを使用します。
- 自動化されたコマンドは、空きスペースが最も多いアグリゲートを選択します。
- ノードが 8 つ以下の場合、自動化コマンドはノードごとに 8 つの構成要素を作成します。
- 8 つより多くのノードを持つクラスタでは、ノードごとに 4 つのメンバーボリュームにスケールバックします。
- 実行可能な最速のアグリゲートを使用します。まず SSD、ハイブリッド、それからスピニングディスクを試行します。
- CPU 使用率、ノード・パフォーマンス、アグリゲート容量などは現在考慮されていません。

■ 手動による FlexGroup の作成 : CLI

CLI では、FlexGroup ボリュームを手動で作成する方法も提供されます。ほとんどのユースケースをカバーしているので、ほとんどの場合は自動化コマンドを使用します。ただし、1 つのアグリゲートあたりのメンバーボリュームの数をカスタマイズする必要がある場合、使用するアグリゲートを指定する必要がある場合、または「7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか ?」で説明されているその他の理由がある場合は、`volume create` コマンドを使用します。ただし、`-auto-provision-as` オプションの代わりに `-aggr-list` を一緒に指定する必要があります。`-aggregate` を指定すると、通常の FlexVol ボリュームが作成されますが、`-aggr-list` を指定することはできません。アグリゲートごとのメンバーボリュームの数を制御するには、`-aggr-list-multipler` を使用します。メンバーボリューム数は、指定したアグリゲートの数に `-aggr-list-multiplier` を乗算した数です。

7.10.3.2 導入方法 2: ONTAP System Manager

図 7.3 ONTAP System Manager FlexGroup ボリュームの作成



ONTAP System Manager には、ボリューム作成用の使いやすい GUI があります。ただし、System Manager GUI を使用して導入する場合は、FlexGroup ボリュームのプロビジョニングに CLI を使用する方が適しているという注意事項があります。

ONTAP System Manager で FlexGroup ボリュームを作成するために、ボリュームが FlexGroup であり、FlexVol でないことを確認するには、[More Options] をクリックし、[Distribute Volume Data Across the Cluster] ボックスを選択します。

複数のノードにまたがる FlexGroup ボリュームを作成するよう System Manager に指示します。アグリゲートまたはノードの指定は必要ありません。

System Manager は、次のルールに従って FlexGroup ボリュームをデプロイします。

- メンバーボリュームが 100GB を下回ることはない
- 最小許容 FlexGroup は 100GB (100GB メンバーボリューム ×1)
- 小さい FlexGroup ボリュームでは、100GB ルールに準拠する必要がある場合に、デプロイされるメンバーボリュームが少なくなります。たとえば、200GB の FlexGroup は、2 つの 100GB メンバーボリュームをデプロイします。
- FlexGroup のアグリゲートとノード選択は自動的に実行されます。ノードまたはアグリゲートを指定するには、CLI または REST API を使用します。
- FlexGroup ボリュームが十分な大きさであれば、メンバーのボリューム数は、ノードごとに使用可能なボリュームアフィニティ数に制限されます。

- System Manager は、空き領域が保証されているかのように、初期 FlexGroup サイズを使用可能な合計領域に制限します。System Manager に戻り、編集機能を使用してボリュームを大きくすることができます。
- System Manager は、FlexGroup に対して同様のアグリゲートのみを使用します。言い換えれば、SSD と HDD のアグリゲートは混在しません。

7.10.3.3 導入方法 3: REST API

REST API を使用すると、FlexGroup ボリュームを作成、監視、および管理できます。REST API を使用して FlexGroup ボリュームをプロビジョニングするには、「7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか ?」で説明したものと同じガイダンスを使用します。

導入方法 1: コマンドラインたとえば、ONTAP に構成を選択させるか、オプションを手動で指定するかを決定します。

REST API のドキュメントは `https://[ クラスター IP またはクラスター名 ]/docs/api` にあります。このサイトでは、サンプルと、独自の REST API を生成するための対話型の Try It Out 機能を提供します。

たとえば、FlexGroup ボリュームを作成するには、`/storage/volumes` の下にある `POST REST API` を使用します。このコールで FlexGroup を FlexGroup (FlexVol ではない) に指定するのは、次の値の 1 つまたは組み合わせです。

- **Aggregates**
  複数を指定した場合、REST API は FlexGroup ボリュームを作成します。これは、CLI の `-aggr-list` と同じ動作です。
- **consolubients_per_aggregate**
  FlexGroup ボリュームが作成または拡張されるときに、`aggregates.name` または `aggregates.uuid` でリストされたアグリゲートを順に処理する回数を指定します。ボリュームが単一のアグリゲート上に作成されている場合、`conboxives_per_aggregate` フィールドが指定されていないと、システムはフレキシブルボリュームを作成します。このフィールドを指定すると、FlexGroup ボリュームが作成されます。ボリュームが複数のアグリゲート上に作成されている場合、システムは常に FlexGroup ボリュームを作成します。これは、CLI の `-aggr-list-multiplier` と同じ動作です。
- **Style**
  `flexgroup` として `style` を指定し、`constibilities_per_aggregate` 値または複数のアグリゲートを設定しない場合、ONTAP はアグリゲートごとに 4 つのメンバーからなる FlexGroup ボリュームを自動的にプロビジョニングします。これは、CLI の `-auto-provision-as` と同じ動作です。

REST API のマニュアルでは、Try It Out 機能を使用して正しい REST API 文字列を作成する方法を説明しています。ミスをすると、インターフェースにエラーメッセージとエラーコードの一覧が表示されます。また、REST API コマンドが正しいが、別の理由 (たとえば、メンバーのサイズが小さすぎる FlexGroup ボリュームの作成) でジョブが失敗する場合は、ジョブ文字列 URL が指定されます。次の URL を使用して、ブラウザからジョブ文字列にアクセスできます。

```
https://[クラスターIPまたはクラスター名]/api/cluster/jobs/job_uuid]
```

失敗メッセージは次のようになります。

```
{
   "uuid": "b5b04f0b-82ea-11e9-b3aa-00a098696eda",
   "description": "POST /api/storage/volumes/b5b02a66-82ea-11e9-b3aa-00a098696eda",
   "state": "failure",
   "message": "Unable to set parameter \"-min-autosize\" to specified value because it is
   too small. It must be at least 160MB (167772160B).",
   "code": 13107359,
   "start_time": "2019-05-30T10:53:39-04:00",
   "end_time": "2019-05-30T10:53:39-04:00",
   "_links": {
      "self": {
         "href": "/api/cluster/jobs/b5b04f0b-82ea-11e9-b3aa-00a098696eda"
      }
   }
}
```

成功したジョブは次のようになります。

```
{
   "uuid": "ac2155d1-82ec-11e9-b3aa-00a098696eda",
   "description": "POST /api/storage/volumes/ac2131c5-82ec-11e9-b3aa-00a098696eda",
   "state": "success",
   "message": "success",
   "code": 0,
   "start_time": "2019-05-30T11:07:42-04:00",
   "end_time": "2019-05-30T11:07:46-04:00",
   "_links": {
      "self": {
         "href": "/api/cluster/jobs/ac2155d1-82ec-11e9-b3aa-00a098696eda"
      }
   }
}
```

FlexGroup ボリュームを作成する REST API 文字列の例については、「18. コマンドの例」を参照してください。

7.10.4 クライアントが領域不足エラーになるのはどのような場合ですか？

一般的に、NAS クライアントで out of space（領域不足）エラーが発生した場合は、ボリュームの領域が実際に不足していることを意味し、df および volume show コマンドで確認できます。

ただし、NAS ストレージシステムから out of space（領域不足）エラーを取得することは必ずしも容易ではありません。これは、使用可能なリソースがもうないことをクライアントに通知するサーバからの一般的なエラーであるためです。NFS や SMB には、out of inodes(inode 不足) エラーや reached maximum directory size(ディレクトリ最大サイズ到達) という概念がないため、ONTAP は標準的な out of space(領域不足) を使用して、それ以上データを書き込めないことをクライアントに通知します。

FlexGroup ボリュームを使用すると、メンバーボリュームが 100% に達したときに、クライアントは out of space(領域不足) エラーを受け取ることもあります。しかし、ONTAP 9.7 以降では、このシナリオは事実上存在しません。

次の表に、out of space(領域不足) エラーが発生する状況、その原因、および対処方法を示します。

表 7.5 領域不足エラーが発生する状況

状況	識別および解決方法
Volume or aggregate has no available space to honor writes.	• クラスタ CLI からの df または volume show-space 出力 • ONTAP System Manager からの容量の表示 • Active IQ Unified Manager アラート • EMS メッセージ 解決方法 : ボリュームに容量を追加します。アグリゲートにディスクを追加します。FlexGroup ボリュームを使用して、ノード間で拡張します。
Quota limit reached	• クラスタ CLI からの df または volume show-space 出力 • ONTAP System Manager からの容量の表示 • クラスタ CLI からの quota report 出力 • ONTAP System Manager からのクォータレポート • Active IQ Unified Manager アラート • EMS メッセージ 解決方法 : クォータの制限値を上げるか、またはクォータの制限値を維持するためにデータを削除する必要があることをクライアントに通知します。
Out of inodes (inode 不足)	• クラスタ CLI からの df または volume show-space 出力 • ONTAP System Manager からの容量の表示 • クラスタ CLI からの df -i または volume show -fields files、files-used コマンド。 • Active IQ Unified Manager アラート • EMS メッセージ 解決方法 : ボリューム内の合計ファイルの値を増やします。詳細は、「10.1 ファイル数が多い場合の考慮事項」を参照してください。
Maxdirsize exceeded	• クラスタ CLI からの df または volume show-space 出力 • ONTAP System Manager からの容量の表示 • クラスタ CLI からの df -i または volume show -fields files、files-used コマンド。 • Active IQ Unified Manager アラート • EMS メッセージ • (「10.4.3 使用されている maxdirsize 値の問い合わせ」に示すように) ディレクトリ・サイズを表示するクライアント側コマンド 解決方法 : 影響を受けるファイルパスを見つけるには、クラスタ CLI から volume file show-inode コマンドを使用します。問題のあるディレクトリのファイル数を減らすか、サポートに連絡して、maxdirsize の値を増やしても安全かどうかを確認してください。maxdirsize についての詳細は、「10.4 ディレクトリサイズに関する考慮事項 :maxdirsize」を参照してください。
Member volume at 100%	• クラスタ CLI からの df または volume show-space 出力 • Active IQ Unified Manager アラート • EMS メッセージ • ONTAP バージョン情報 解決方法 : ONTAP では、メンバーボリュームがいっぱいになったときのファイル書き込みエラーに対する保護手段として、エラスティックサイジングが導入されています。ONTAP 9.8 では、プロアクティブなサイズ変更によってメンバーボリュームのサイズを事前に変更し、メンバーボリューム内の空き領域のバランスを均等に維持します。最大の容量使用率を得るには、ONTAP 9.8 以降にアップグレードしてください。

7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか？

ほとんどの場合、FlexGroup ボリュームを作成するときは、ONTAP にメンバーボリュームを選択させることをお勧めします。言い換えれば、FlexGroup 作成のための CPU 数とボリュームアフィニティのベストプラクティスについて心配する必要はありません。ONTAP に任せてください。ボリューム数を操作しようとすると、混乱を招き、後で FlexGroup に影響を及ぼしかねない問題が発生するおそれがあります。

ただし、場合によっては、手動で作成する必要があります。次の各項では、FlexGroup ボリュームを手動で作成する必要がある場合のシナリオについて説明します。

ボリューム数の過剰なプロビジョニングに関する懸念

ONTAP では、各ノードとクラスタには有限数の FlexVol ボリュームを使用できます。制限はプラットフォームと ONTAP のバージョンに依存しますが、FlexGroup ボリュームは複数の FlexVol ボリュームで構成されるため、各 FlexVol メンバーのボリューム数が合計ボリューム数の制限に対してカウントされるため、これらの制限は FlexGroup ボリュームにも適用されます。多数のメンバーボリュームを持つ FlexGroup ボリュームがある場合や、クラスタ内に多数の FlexGroup ボリュームを作成する場合は、ノードごとの全体的なボリューム制限を考慮する必要があります。また、FlexGroup ボリュームを手動で作成してデフォルトのボリューム数を変更したり、総 FlexVol ボリューム数をノード制限を下回るようにアグリゲート配置を変更したりする必要がある場合もあります。

容量に制限のある大容量ファイル

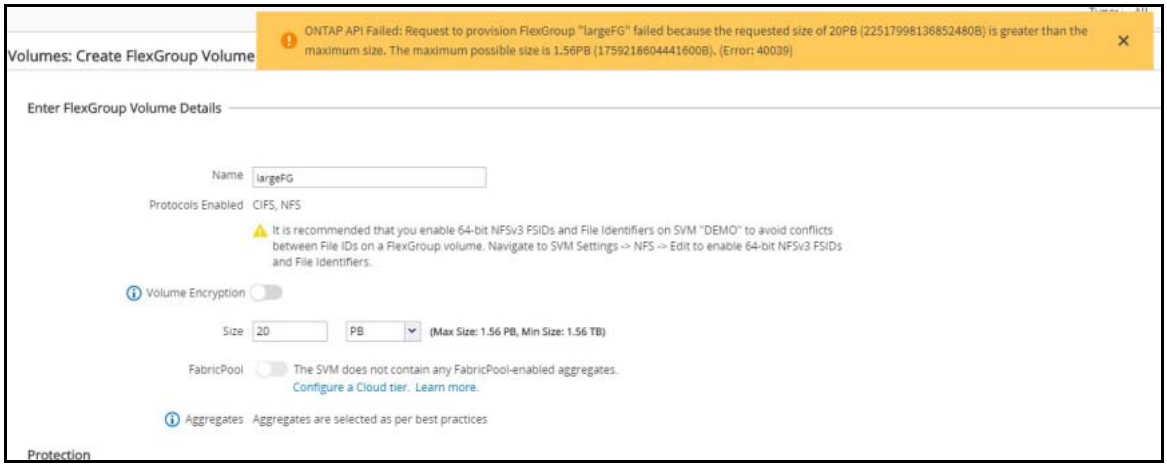
大容量のファイルを使用するワークロードがあり、数十 TB または数百 TB のボリュームをプロビジョニングできない場合に、大容量ファイルのベスト・プラクティスに準拠するには、メンバーボリューム数を調整して、大容量の個別容量で作成するメンバーの数を減らす必要があります。

たとえば、4 つのノードにまたがって 16TB の FlexGroup ボリュームを作成する場合、ONTAP のデフォルトの方法を使用すると、メンバーボリューム数を最小限に抑えるため、500GB のメンバーボリュームが 32 個作成されます。250GB の場合、500GB のメンバーボリュームではデータを効果的に分散するには十分ではありません。このような場合は、FlexGroup ボリュームを手動で作成してメンバーボリュームの数を調整する方が効率的です。

大容量または多数のファイルを必要とする

FlexVol ボリュームのサイズは 100TB に制限されており、最大 20 億個のファイルを格納できます。2 ノードのクラスターがあり、ONTAP に FlexGroup ボリュームを作成させると、場合によっては、1 つの FlexGroup ボリュームに最大 16 個のメンバーボリュームが作成されます。これは、1 ノードごとに 8 個というコード制限があるためです。次の例では、2 ノードのクラスターで作成できる FlexGroup ボリュームの最大容量は 1.56PB (1 ノードあたり 8 メンバー、合計 16 メンバー、メンバーボリュームあたり 100TB) です。

図 7.4 System Manager で許容される最大値を超えて FlexGroup ボリュームを作成したときのエラー



auto-provision-as オプションを使用すると、同じエラーが表示されます。

```
cluster::*> vol create -vserver DEMO -volume largeFG -auto-provision-as flexgroup -size 2PB

Error: command failed: Request to provision FlexGroup volume "largeFGT" failed because the
requested size of 2PB (2251799813685248B) is greater than the maximum size. The maximum
possible size is 1.56PB (1759218604441600B).
```

自動化ツールで許可されているよりも大きな FlexGroup ボリュームが必要な場合は、-aggr-list-multiplier オプションを使用して手動で FlexGroup を作成し、より多くのメンバーボリュームを許可する必要があります。20PB の FlexGroup ボリュームの場合、少なくとも 200 個のメンバーボリュームが必要です。理想的には、FlexGroup メンバーボリュームのサイズを 100TB 未満に設定し、後でそれらのボリュームを拡張するためのスペースが必要になった場合に備えます。80 ～ 90TB が最大メンバーボリュームの対象値になります。

ファイル数が許可されている最大ファイル数を超える必要がある場合も、同様の考慮が必要です。メンバー数が 16 の FlexGroup の例では、最大 320 億個のファイルを使用できます。さらにファイルが必要な場合は、まず maxfiles の値を増やします。それが不可能な場合は (たとえば、maxfiles の値が上限に達している場合)、メンバーボリュームをさらに追加します。

CLI から FlexGroup ボリュームを作成し、メンバー数を指定する方法の例については、このドキュメントの後の「18. コマンドの例」を参照してください。

クラスターネットワークを回避する

あまり一般的ではありませんが、1 つのノードに FlexGroup ボリュームを作成することでクラスタネットワークを回避したり、障害ドメインの影響を軽減したい場合があります。この使用例では、ONTAP CLI を使用して -aggr-list および -aggr-list-multiplier オプションを volume create コマンドに使用して、指定するアグリゲートを管理します。

多数のメンバーボリュームが必要ですか。

通常、FlexGroup ボリュームのベストプラクティスのボリューム数を超える必要はありません。ただし、容量やファイル数を増やす必要がある場合は、初期展開時にメンバーボリュームの数を増やすか、後で volume expand コマンドを使用して増やすことができます。一般的には、データが既存のメンバーボリュームに格納される前に、メンバーボリューム数を増やすことをお勧めします。後でメンバーボリュームを追加すると不均衡が生じますが、ONTAP で調整する必要があり、ワークロードに影響する可能性があります。メンバのボリューム数に関する ONTAP のベストプラクティスから外れる必要がある場合の詳細については、上記のセクションを参照してください。

大きいファイルと小さいファイルのメンバー数に関する考慮事項

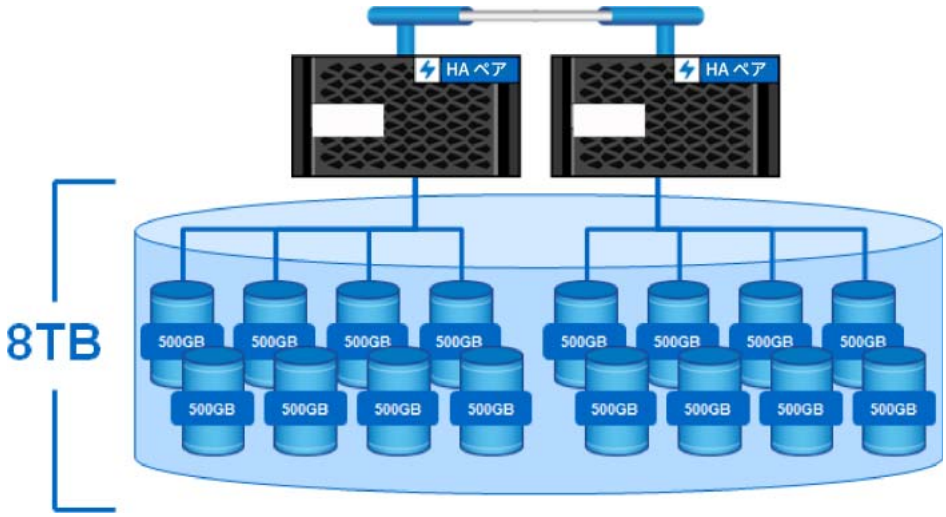
FlexGroup ボリュームは、多数の小さなファイルで構成されるファイル数の多い環境で最適に機能します。ただし、サイズの大きいファイルにも適しています。「2.1 大容量ファイルとは」で説明したように、サイズの大きいファイルは、メンバーボリュームに割り当てられた合計領域の割合で考慮する必要があります。

ワークロードに大きなファイルが存在する場合は、FlexGroup の初期導入サイズを考慮する必要があります。デフォルトでは、FlexGroup はノードごとに 8 つのメンバーボリュームを導入するため、FlexGroup レベルで定義された容量のフットプリントは、実質的に [合計スペース / メンバーボリューム数 (n)] に分割されます。

たとえば、クラスタ内の 2 つのノードに 8TB の FlexGroup を導入し、メンバー数が 16 の場合、各メンバーボリュームのサイズは約 500GB になります。

多くのワークロードでは、図 7.5 に示すような分散がうまく機能します。ただし、ワークロード内のサイズの大きいファイルによって、メンバーボリュームで使用されている容量の大きなチャンクが埋まる可能性がある場合は、パフォーマンスやアクセス可能性に影響が及ぶ可能性があります。ONTAP 9.8 では、「プロアクティブなサイズ変更」がこれらの「member volume full (メンバーボリュームを使い切った)」シナリオの頻度をさらに減らすのに役立ちます。

図 7.5　容量をメンバーボリューム間で分割する方法

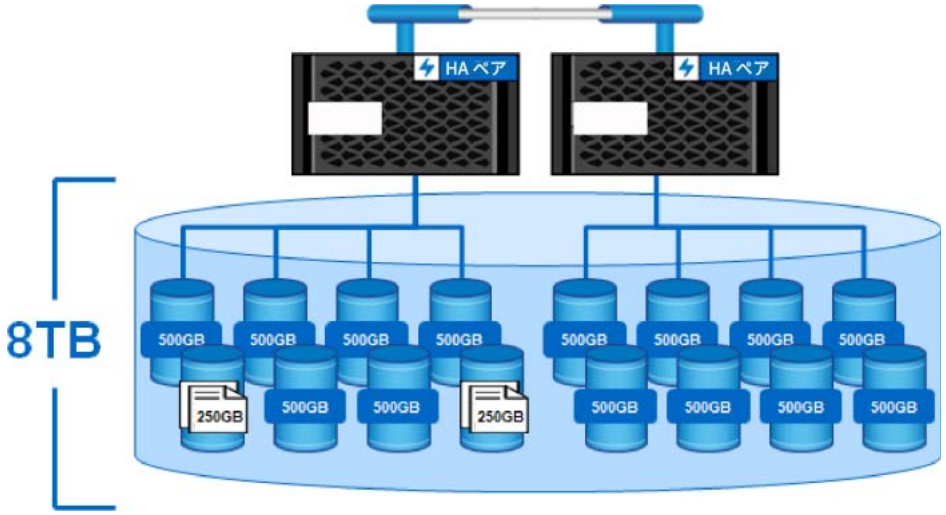


ベストプラクティス 6: 大容量ファイルワークロードの実行に最適な ONTAP バージョン
一般的に、使用可能な最新の ONTAP リリースは FlexGroup ボリュームを使用する場合に実行するのに最適なバージョンですが、すべてのユーザーが最新の ONTAP リリースにアップグレードできるわけではありません。ONTAP 9.8 の最新のパッチ・リリースの使用を推奨しています。そのバージョンが使用できない場合は、ONTAP 9.7 の最新のパッチリリースを使用してください。

たとえば、ワークロード内の一部のファイルが 250GB の場合、500GB のメンバーを持つ FlexGroup ボリュームにファイルが書き込まれるたびに、メンバーボリュームの合計容量の 50% が使用されます。

2 番目の 250GB のファイルがその 500GB のメンバーボリュームに書き込もうとすると、ファイルの書き込みが完了する前にメンバーボリュームの空き容量が不足します。

図 7.6　FlexGroup メンバーボリューム内のサイズの大きいファイルの影響



FlexGroup ボリューム内のファイルはストライプされないことに注意してください。常に単一の FlexVol メンバーボリュームに書き込みます。したがって、書き込みを実行するには、1 つのメンバーボリュームに十分なスペースが必要です。

エラスティックサイジングでは、クライアントに out of space(領域不足) エラーを送信する前に一時停止して、可能であれば同じ FlexGroup 内の他のメンバーボリュームから空き領域を借用することで、いくらか軽減されます。ただし、ボリュームの自動サイズ変更が有効になっている場合、そのボリュームのエラスティックサイジングは無効になります。エラスティックサイジングは、容量管理を回避するための方法ではなく、容量の問題の影響を軽減するための事後対応的な方法です。最適な結果を得るためには、メンバーボリュームの容量を 80~90% 未満に抑えることが依然として必要です。

ONTAP 9.8 でのプロアクティブなサイズ変更は、エラスティックサイジングとボリュームの自動拡張の利点を兼ね備えています。ONTAP は、ファイル作成で空き領域が不足するのを待つのではなく、空き領域しきい値で事前にメンバーボリュームサイズを増やして、容量不均衡の影響を減らし、個々のメンバーボリュームから容量を管理する必要性を減らします。さらに、ボリュームの自動拡張をプロアクティブなサイズ変更と組み合わせて使用できるため、FlexGroup の総容量がしきい値に達した場合、ONTAP は自動的にサイズを指定された値まで増加させます。

最終的には、物理的なスペースが不足したときにストレージを追加する必要があります。また、FlexGroup ボリュームへのノードまたはディスクの追加は、無停止で容易かつ迅速に実行できます。

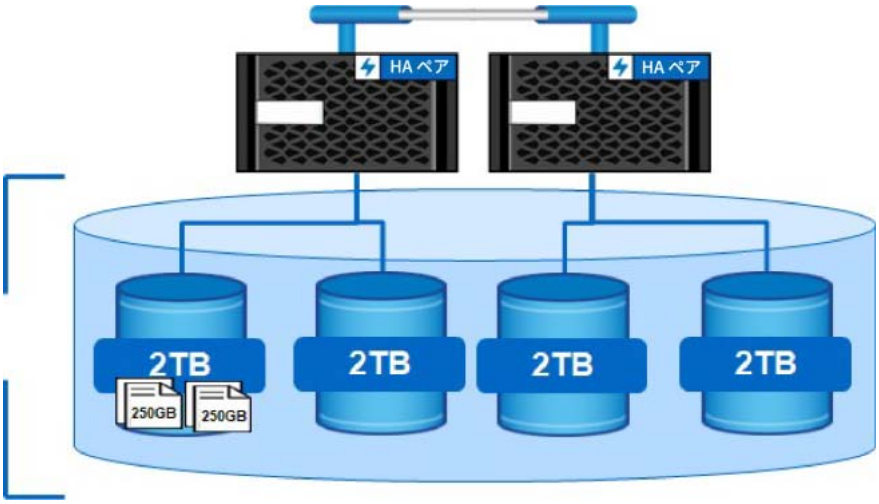
FlexGroup ボリュームのサイズを決定するより良い方法は、新しい FlexGroup ボリュームを導入する前、または新しいワークロードが既存の FlexGroup ボリュームにアクセスできるようにする前に、ワークロードと平均ファイルサイズを分析することです。ファイルを迅速に分析してサイズに関するレポートを作成できる XCP Migration Tool が提供されています。XCP についての詳細は、「16. ONTAP FlexGroup への移行」を参照してください。

FlexGroup ボリュームに格納されるファイルのサイズを把握したら、初期導入時にボリュームのサイズをどのように設定するかを設計上で決定できます。

次のようなオプションがあります。

- **メンバーのボリューム数はデフォルトのままにして、FlexGroup ボリュームを拡張**
  ワークロードを処理できるメンバーボリュームサイズに対応できる大きさになるように、FlexGroup ボリュームの合計サイズを設定します。この例では、FlexGroup ボリュームは 80TB であり、ボリュームあたり 5TB で 16 個のメンバーボリュームを提供します。ただし、このアプローチでは、より多くの物理容量が必要です (シンプロビジョニングを使用しない場合)。

図 7.7　　メンバーボリュームの縮小と拡大



- **メンバーボリューム数を手動で減らして、FlexGroup 容量は変更しない**
  自動化されたコマンドからデフォルト値を受け入れる代わりに、CLI を使用して、総容量は同じであるが、含まれるメンバーボリュームの数が少ない (あるいは多い)FlexGroup ボリュームを作成できます。この例では、8TB FlexGroup のノードあたりのメンバーボリューム数を 2 に減らすと、それぞれ 2TB のメンバーボリュームサイズになります。これにより、使用可能なボリュームアフィニティの数は減ります (ファイル取り込みのための FlexGroup ボリュームの全体的なパフォーマンスが低下する可能性があります) が、より大きなファイルを配置できます。

大きなファイルをメンバーボリュームに配置した後のパフォーマンスは、より多くのメンバーボリュームを持つ FlexVol ボリュームまたは FlexGroup ボリュームの場合と同様です。

容量管理機能

次の表は、ONTAP リリースに応じて使用できる容量管理機能を示しています。

表 7.6 容量管理の決定マトリックス

ONTAP バージョン	容量管理機能
ONTAP 9.7	• シンプロビジョニング • 容量アラート • ストレージの効率性 • ボリュームの自動サイズ設定 (自動拡張 / 自動縮小) • qtree と監視クォータ • クォータの適用 • FabricPool 自動階層化 • エラスティックサイジング
ONTAP 9.8 以降	• シンプロビジョニング • 容量アラート • ストレージの効率性 • ボリュームの自動サイズ設定 (自動拡張 / 自動縮小) • qtree と監視クォータ • クォータの適用 • FabricPool 自動階層化 • エラスティックサイジング • プロアクティブなサイズ変更

7.10.6 アグリゲートの空き領域に関する考慮事項

FlexGroup ボリュームを作成する場合、FlexGroup が導入されている 1 つまたは複数のアグリゲートには次の特性があることが理想的です。

- 複数のアグリゲートにわたるほぼ均等な空き容量がある (シンプロビジョニングを使用する場合には特に重要)
- 10GB 以上または 0.6% 以上の空き容量がある (どちらか少ない方)

ONTAP 9.7 以降のバージョンでは、重複排除メタデータのチェックが行われなくなりました。

メンバーボリュームの容量が重要な理由

FlexGroup ボリュームの目的は、基礎となるメンバーボリュームにあまり注意を払う必要がなく、FlexGroup レベルから管理することです。ほとんどの場合、FlexGroup ボリュームはこのように動作します。
ただし、ONTAP 9.8 より前のリリースでは、FlexGroup ボリュームを作成および管理する際に、メンバーボリュームの容量をより頻繁に考慮する必要がありました。

メンバーボリューム内の使用可能な空き領域は、FlexGroup ボリューム内で新しいファイルをローカルまたはリモートで取り込む頻度に影響します。これは、新しいファイルを作成するための FlexGroup ボリューム内のパフォーマンスと容量の配分にも影響します。

初期 FlexGroup ボリュームを設計する際には、ワークロード内の平均ファイルサイズと最大ファイルサイズを考慮することが重要です。これは、大容量ファイルの方が個々のメンバーボリュームをより速く満たすことができるため、新しいファイル作成のリモート割り当てが多くなり、他のメンバーボリュームよりも先にメンバーボリュームの領域が不足する場合があるためです。ただし、すでに配置されているファイルでは、バランスの取れていない FlexGroup に対しても、新しいファイル作成時と同じようなパフォーマンス上の影響はありません。

容量の不均衡はそれ自体が問題ではありませんが、パフォーマンスの問題や容量使用率の問題の根本原因となる可能性があります。パフォーマンスの問題があり、容量の不均衡が疑われる場合は、当社サポート部門にパフォーマンスデータの分析支援を依頼してください。場合によっては、容量の不均衡が問題のように見えることがありますが、実際の問題には別の根本原因があります。

ベストプラクティス 7: メンバーボリュームサイズの推奨事項
最大ファイルがメンバーボリュームの容量の 1% ～ 5% を超えないように、メンバーボリュームのサイズを設定することを推奨しています。ノードあたりのメンバー数が 2 未満の FlexGroup ボリュームは作成しないでください。

7.10.7 初期ボリュームサイズに関する考慮事項

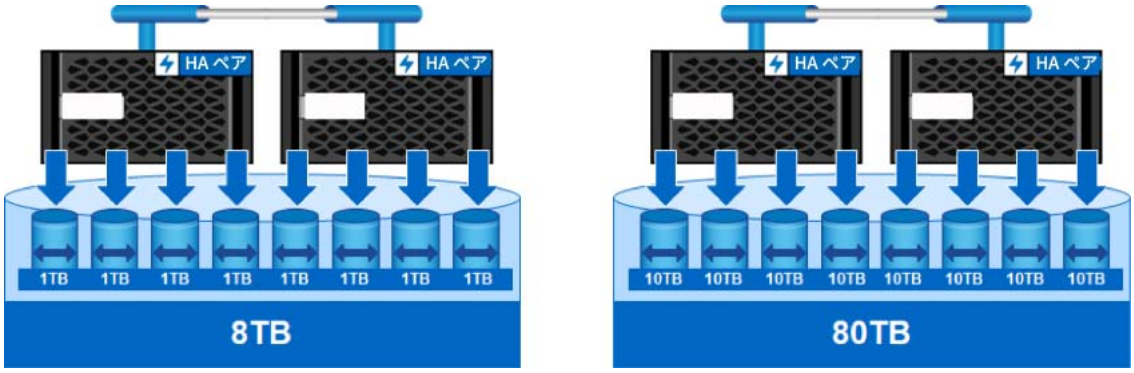
一般的な導入の問題は、FlexGroup ボリュームのメンバーボリューム容量の設定が小さすぎることです。これは、ストレージ管理者には知られずに、よく行われます。関心があるのは総容量であり、通常は、基礎となるメンバーボリュームについてよく考慮されることはありません。80TB は 80TB です。しかし、FlexGroup では、80TB は実質的に、メンバーボリュームの総数で割った値が 80TB になります。

FlexGroup ボリュームはほとんどすべての容量で作成できますが、FlexGroup ボリュームの合計サイズは複数の FlexVol メンバーボリュームで構成されることに注意してください。デフォルトでは、自動化された FlexGroup コマンドは、使用する展開方法に応じて、デフォルト数のメンバーボリュームを作成します (詳細は 「7.10.3 FlexVol メンバーボリュームのレイアウトに関する考慮事項」を参照してください)。

ベストプラクティス 8: FlexGroup 導入の簡素化
メンバーボリュームのことをあまり気にする必要がない場合は、ONTAP 9.8 にアップグレードして、プロアクティブなサイズ変更のメリットを活用してください。これにより、メンバーボリュームのサイズをあまり気にする必要がなくなります。

たとえば、8 つのメンバーボリュームを持つ 80TB の FlexGroup ボリュームでは、各メンバーボリュームのサイズは 10TB です。10TB というサイズは処理を行うにはかなりの大きさがあるため、これらのメンバーボリュームのサイズがほとんどのワークロードケースに影響しないと考えられていますが、容量計画をたてるにはワークロードのファイルサイズを把握することが重要です。たとえば、ワークロードに 500GB のファイルがある場合、10TB のメンバーボリュームでは問題ありませんが、1TB のメンバーボリュームでは問題があります。

図 7.8 FlexGroup ボリューム : メンバーサイズと FlexGroup ボリューム容量



新しい FlexGroup ボリュームの初期消費容量

各 FlexGroup メンバーボリュームは、内部使用のために少量のスペース (約 50MB) を確保します。メンバーボリュームのサイズが 100GB 上に設定されている場合、使用済み領域は約 0.05% になります。これは ONTAP では無視されます。ただし、空の FlexGroup ボリュームの出力にはまだ使用済み領域が表示されるため、FlexGroup ボリュームを導入する際には、このことは問題にならないことを念頭に置いておいてください。
例 :

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -fields used
vserver volume           used
------- ---------------- -------
DEMO    fgautogrow__0006 57.48MB
DEMO    fgautogrow__0008 57.48MB
DEMO    fgautogrow__0001 57.50MB
DEMO    fgautogrow__0004 57.50MB
DEMO    fgautogrow__0005 57.52MB
DEMO    fgautogrow__0007 57.52MB
DEMO    fgautogrow__0002 57.57MB
DEMO    fgautogrow__0003 57.57MB
DEMO    fgautogrow         460MB
```

FlexGroup ボリュームの縮小

ボリュームの縮小がサポートされている場合でも、最初の作成時にボリュームのサイズを大きくしすぎないようにしてください。サイズを大きくしすぎると、後で容量を増やす必要があるときに管理オプションが制限されることがあります。同じメンバーボリュームサイズで追加する必要があるため、新規メンバーボリュームを追加する必要があります。

7.10.8 Snapshot コピーと Snapshot Reserve

ONTAP Snapshot コピーは、データが上書きされるまで領域を使用せずにファイルシステムのポイント・イン・タイム・コピーを作成するように設計されています。データが変更または削除されると、そのスペースはアクティブなファイルシステム (AFS) から削除済みとしてマークされ、Snapshot コピーへのポインタが ONTAP によってリダイレクトされます。この処理が完了すると、Snapshot コピーに使用済み領域が表示されます。

7.10.8.1 Snapshot Reserve

デフォルトでは、ボリュームには 5% の Snapshot コピー予約領域が割り当てられます。つまり、100TB のボリュームをプロビジョニングすると、そのボリュームの 5% (5TB) が Snapshot コピーに割り当てられます。その結果、ストレージシステム上のボリュームサイズ出力 (`volume show` や `df` など) には、このシナリオで 95TB の使用可能な領域が示されます。これは、クライアントが使用可能な領域と見なしている容量でもあります。FlexGroup ボリュームでは、Snapshot リザーブは FlexGroup レベルで設定されますが、各メンバーボリュームに適用されます。100TB の FlexGroup ボリュームの Snapshot リザーブは 5TB になる場合がありますが、個々のメンバーボリュームはその 5TB を均等に共有します。メンバーボリュームが 8 つある場合、それぞれのボリュームには、Snapshot コピー用に予約された 640GB の容量があります。

7.10.8.2 Snapshot Spill

Snapshot コピーの使用量が Snapshot リザーブのサイズを超えた場合、その領域は AFS の容量を使用し始めます。ONTAP は、Snapshot コピーによって使用されている領域が 100% を超えていると報告し、ボリューム内に物理データが存在しない場合でも、ボリューム内の使用済み領域の合計が増加します。

例については、「17.3 Snapshot Spill の例」を参照してください。

Snapshot Spill とスナップショットスキャナ

ONTAP は、ボリュームを定期的にスキャンして、WAFL で使用されたブロックと変更されたブロックをチェックします。これは、ボリューム内の使用済みスペースを正しく計算するために使用されます。

これらのスキャナーは優先順位の低いジョブであり、実動ワークロードに従うため、スキャナーの完了速度はシステムの負荷に依存します。これらのスキャナは、次のように表示できます。

```
Mon Dec 21 19:24:30 -0600 [CLUSTER: scan_ownblocks_calc_wkr: wafl.scan.ownblocks.done:info]:
Completed block ownership calculation on volume vol__0003@vserver:ebf4370e-208a-11eb-921e-
d039ea2020c8. The scanner took 202 ms.
```

ボリュームが小さい (例 :100GB) 場合は、よりはやく容量が使い切られます。このような場合、特に高速な ETERNUS AX series システムでは、入力データ書き込みがスキャナより高速に実行される可能性があります。その結果、容量レポートは、ストレージ管理者が容量の追加やスナップショットの削除によって容量のニーズに対応できるようには対応しません。このような場合、ボリュームの自動拡張が使用可能になっていないと、他のメンバーボリュームから借用するスペースがないため、ボリュームが `out of space`(領域不足) を報告することがあります。

Snapshot Spill の修復に関するヒント

Snapshot Spill は、スナップショットリザーブが超過したときに Snapshot コピーが ONTAP でどのように動作するかを示す通常の機能です。Snapshot のオーバーフローの影響と、Snapshot の増加速度は、ボリュームの合計サイズによって異なります。ボリュームが小さいほど、Snapshot リザーブ領域の合計が少なくなり、Snapshot のオーバーフローが発生しやすくなります。Snapshot のオーバーフローの影響を最小限に抑えるには、次の 1 つ以上の操作を行います。

- FlexGroup ボリュームの合計サイズを増やします。これにより、使用可能な Snapshot リザーブの合計が増え、Snapshot のオーバーフローが発生しにくくなります。
- Snapshot コピーを使用している場合は、小さい FlexGroup ボリュームを作成しないでください。
- データ量が多い WAFL 場合は、Snapshot リザーブの割合を大きくします。クライアントはアクティブファイルシステム内の使用可能な領域のみを認識し、予約済みの Snapshot 領域は認識しません。
- 可能な場合は、より大きな Snapshot を削除します。FlexGroup ボリュームは現在 Snapshot の自動削除をサポートしていないため、Snapshot コピーを手動またはスクリプトで削除する必要があります。
- Snapshot リザーブを 0 に設定します。これにより、Snapshot の使用済みスペースが AFS にただちに反映されます。

7.10.9 ボリュームの自動サイズ設定 (自動拡張と自動縮小)

FlexGroup ボリュームに対するボリュームの自動拡張をサポートしています。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。FlexGroup ボリュームへのボリュームの自動拡張の適用は、FlexVol ボリュームの場合と同じ方法で行います。しきい値を指定し、さまざまなオプションを構成します。詳細については、FTI のマニュアルサイトを参照してください。

構成オプションは FlexVol と同じで、次のものがあります。

```
[-max-autosize {<integer>[KB|MB|GB|TB|PB]}] - Maximum Autosize
This parameter allows the user to specify the maximum size to which a volume can grow. The
default for volumes is 120% of the volume size. If the value of this parameter is
invalidated by manually resizing the volume, the maximum size is reset to 120% of the volume
size. The value for -max-autosize cannot be set larger than the platform-dependent maximum
volume size. If you specify a larger value, the value of -max-autosize is automatically
reset to the supported maximum without returning an error.

[-min-autosize {<integer>[KB|MB|GB|TB|PB]}] - Minimum Autosize
This parameter specifies the minimum size to which the volume can automatically shrink. If
the volume was created with the grow_shrink autosize mode enabled, then the default minimum
size is equal to the initial volume size. If the value of the -min-autosize parameter is
invalidated by a manual volume resize, the minimum size is reset to the volume size.

[-autosize-grow-threshold-percent <percent>] - Autosize Grow Threshold Percentage
This parameter specifies the used space threshold for the automatic growth of the volume.
When the volume's used space becomes greater than this threshold, the volume will
automatically grow unless it has reached the maximum autosize.

[-autosize-shrink-threshold-percent <percent>] - Autosize Shrink Threshold Percentage
This parameter specifies the used space threshold for the automatic shrinking of the
volume. When the amount of used space in the volume drops below this threshold, the volume
will shrink unless it has reached the specified minimum size.

[-autosize-mode {off|grow|grow_shrink}] - Autosize Mode
This parameter specifies the autosize mode for the volume. The supported autosize modes
are:

o off - The volume will not grow or shrink in size in response to the amount of used space.
o grow - The volume will automatically grow when used space in the volume is above the grow
threshold.
o grow_shrink - The volume will grow or shrink in size in response to the amount of used
space.

By default, -autosize-mode is off for new volumes, except for DP mirrors, for which the
default value is grow_shrink. The grow and grow_shrink modes work together with Snapshot
autodelete to automatically reclaim space when a volume is about to become full. The volume
parameter -space-mgmt-try-first controls the order in which these two space reclamation
policies are attempted.

[-autosize-reset [true]] } - Autosize Reset
This allows the user to reset the values of autosize, max-autosize, min-autosize, autosize-
grow-threshold-percent, autosize-shrink-threshold-percent and autosize-mode to their
default values. For example, the max-autosize value will be set to 120% of the current size
of the volume.
```

7.10.9.1 FlexGroup ボリュームでのボリュームの自動サイズ設定の仕組み

ONTAP はクライアントにエラーを送信する代わりに、他のメンバーボリュームで使用可能な空き領域を検索している間、操作を一時停止します。使用可能な空き領域がある場合、メンバーボリュームは拡大し、別のメンバーボリュームは同じ量縮小します。これにより、FlexGroup ボリュームの合計サイズが維持されます。これはエラスティックサイジングと呼ばれ、後で詳しく説明します。

ボリュームの自動サイズ設定が有効になっている場合は、ボリュームの自動拡張によって、一時停止して、同じ総容量を維持したまま他のメンバーボリュームの領域を使用する (これにより、ワークロードの遅延が増加する) のではなく、容量のしきい値に達したときに、構成済みの容量だけメンバーボリュームを拡張します。これにより、メンバーボリュームの増加分だけ、FlexGroup ボリュームの合計サイズが増加します。

たとえば、10TB の FlexGroup ボリュームがあり、メンバーボリュームが自動的に 1TB 増加した場合、ボリュームの自動拡張によって 11TB の FlexGroup ボリュームができます。

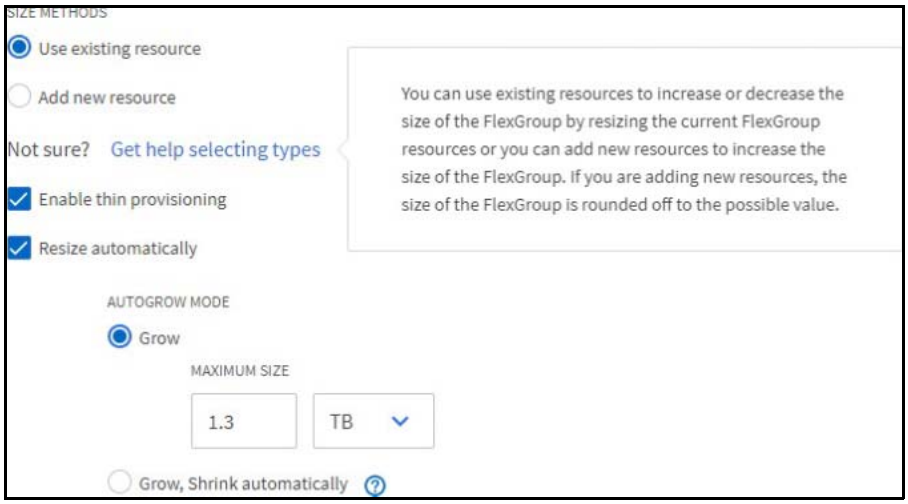
ボリュームの自動縮小

ボリュームの自動サイズ設定機能には、自動拡張だけでなく、自動縮小機能もあります。これは、`-autosize-mode` オプションを使用して有効または無効にできます。自動縮小を使用すると、ONTAP は、使用されている容量が構成済みの `-autosize-shrink-threshold-percent` 値に達した場合に、メンバーボリュームを通常のサイズに縮小できます。その 11TB FlexGroup ボリュームで、拡張したメンバーボリューム内に 1TB の空き容量がある場合、構成した分だけ縮小されますが、デフォルトでは元のボリュームサイズより小さくなることはありません。

ボリュームの自動サイズ設定を有効にする方法

ボリュームの自動サイズ設定を有効にするには、いくつかの方法があります。

1. ボリュームの作成中に ONTAP System Manager を使用するか、[Edit] を使用します。



2. CLI から `volume autosize` コマンドを使用できます。

```
cluster::> volume autosize -vserver DEMO -volume fgautogrow -maximum-size 100g
-grow-threshold-percent 80 -autosize-mode grow
```

- 次のコマンドを使用して、有効になっていることを確認できます。

```
cluster::> vol autosize -vserver DEMO -volume fgautogrow
Volume autosize is currently ON for volume "DEMO:fgautogrow".
The volume is set to grow to a maximum of 100g when the volume-used space is above
80%. Volume autosize for volume 'DEMO:fgautogrow' is currently in mode grow.
```

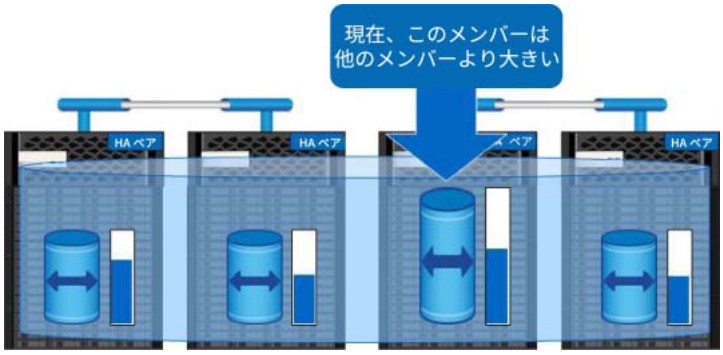
3. CLI から `volume modify` コマンドを使用できます。

```
cluster::> volume modify -vserver DEMO -volume fgautogrow -autosize-mode grow_shrink
-autosize-grow-threshold-percent 95% -autosize-shrink-threshold-percent 50% -max-
autosize 1.20PB -min-autosize 1PB
```

自動拡張によってメンバーボリュームが拡張されると、使用可能なメンバーボリュームのサイズ / 割り当てが不均衡になります。これは仕様です。

さらに、メンバーボリュームのサイズが大きくなったため、FlexGroup の合計サイズが大きくなりました。FlexGroup ボリュームの合計サイズを大きくしたくない場合は、ボリュームの自動拡張を無効のままにしておき、ONTAP のもう 1 つの容量管理機能、エラスティックサイジングやプロアクティブなサイズ変更 (ONTAP 9.8 以降) などを使用できます。

図 7.9　ボリュームの自動サイズ設定操作後のメンバーボリュームサイズの割り当て



■ ボリュームの自動サイズ調整とエラスティックサイジングの相互作用

エラスティックサイジングにより、他のメンバーボリュームから領域を借用することで、ほぼ一杯になった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が提供されています。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。いっぱいになっ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に正規化されます。

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。

FlexGroup ボリュームのエラスティックサイジングは、デフォルトで有効になっています。ONTAP 9.7 でボリュームの自動サイズ設定を有効にした場合、そのボリュームのエラスティックサイジングは有効になりません。ONTAP 9.8 以降では、エラスティックサイジングによるボリュームの自動サイズ設定の使用が可能です。

■ ボリュームの自動サイズ調整とプロアクティブなサイズ変更の相互作用

プロアクティブなサイズ変更は ONTAP 9.8 以降で使用可能で、「7.10.11 プロアクティブなサイズ変更」で説明されています。

ボリュームの自動サイズ設定は、プロアクティブなサイズ変更と連動して機能します。プロアクティブなサイズ変更では、メンバーボリュームの容量が調整され、自動サイズ設定の容量しきい値に達すると、ONTAP はボリュームの自動サイズ設定を適用します。ボリュームの自動サイズ設定が無効になっている場合、プロアクティブなサイズ変更は自動的に機能します。自動サイズ設定を有効にした場合のプロアクティブなサイズ変更の詳細については、「■ プロアクティブなサイズ変更動作 : ボリュームの自動サイズ設定が有効」を参照してください。

7.10.10　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。ファイルはメンバーボリューム間でストライピングされません。

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領域を超える単一のファイルを書き込もうとしています。たとえば、10GB のファイルが 9GB の使用可能なメンバーボリュームに書き込まれます。
- 時間の経過とともにファイルが追加されると、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。例えば、ボリュームは 4,000 個のファイルで構成されていますが、ユーザーがいくつかを圧縮して巨大な単一の tarball ファイルを作成したとします。

1 つの解決策は、手動で、またはボリュームの自動拡張を使用して、ボリュームを拡張することです。もう 1 つの解決策は、データを削除することです。しかし、管理者は、手遅れになるまで容量の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。200TB の FlexGroup ボリュームに 20TB(ボリュームの 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 20TB ではありません。FlexGroup ボリューム内のすべてのメンバーボリュームに容量が均等に分散していれば、容量は 20TB/ [FlexGroup 内のメンバーボリューム数] に近くなります。

2 ノードクラスタでは、両方のノードにまたがる FlexGroup ボリュームには、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 20TB が使用可能な場合、メンバーボリュームでは 1.25TB が使用可能です。

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「out of space (領域不足)」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は万能ではありません。ファイルの書き込みが失敗しないように、問題が起きてから対応することを目的としています。エラスティックサイジングが有効になっている場合でも、十分なメンバーボリュームスペースを使用可能な状態に維持するための容量管理が必要です。

エラスティックサイジング : データ用のエアバッグ

FlexGroup の開発者の 1 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。事故に遭わないように設計されているわけではありませんが、事故が起きたときに被害を軽減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、これらの書き込みを完了する方法を提供します。実際、場合によっては、エラスティックサイジングにより、FlexGroup 全体が容量不足になるまで、またはパフォーマンスの問題が発生するまで、容量の問題を無視することができます。

この機能の概要は次のとおりです。

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を認識できません。クライアントにはそんなことはわかりません。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「ファイルを書き込みたい」という指示だけです。
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、最後にファイルを作成してからの経過時間など、さまざまな要因に基づいて、最も利用可能なメンバーに配置されます。
3. ファイルを配置すると、ONTAP はファイルのサイズを認識しないため、ファイルのサイズが使用可能な領域よりも大きくなるかどうかも認識しません。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が領域不足になると、ONTAP が `out of space` (領域不足) エラーをクライアントに送信する直前に、ONTAP は FlexGroup ボリューム内のもう一方のメンバーボリュームを照会し、借用できるスペースがあるかどうかを確認します。スペースがある場合、ONTAP は、フルになっているボリュームに、(10MB ～ 10GB の範囲で) ボリュームの総容量の 1% を追加し (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す場合)、ファイルの書き込みを続行します。
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、クライアントから見るとパフォーマンスの問題であり、通常はレーテンシーの問題です。しかし、ここでの全体的な目標は、書き込みを迅速に完了することではなく、書き込みを完了できるようにすることです。通常、メンバーボリュームは 10GB の増分 (1TB の 1% は 10GB) を提供するのに十分な大きさであり、多くの場合、ファイルの作成を完了するのに十分な大きさです。小さいメンバーボリュームでは、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。これは、増分が小さいため、システムがより頻繁にスペースを借りるためにクエリーを実行する必要があり、ボリュームを満たすためにファイルのサイズを大きくする必要がないためです。
6. 容量の借用により、FlexGroup ボリューム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 40TB の場合、40TB のままです。

7 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った後、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。ただし、メンバーボリュームの容量が FlexGroup 内の他のメンバーボリュームの平均空き領域の 75% 以内にある場合に限ります。

図 7.10 エラスティックサイジング前のファイル書き込み動作

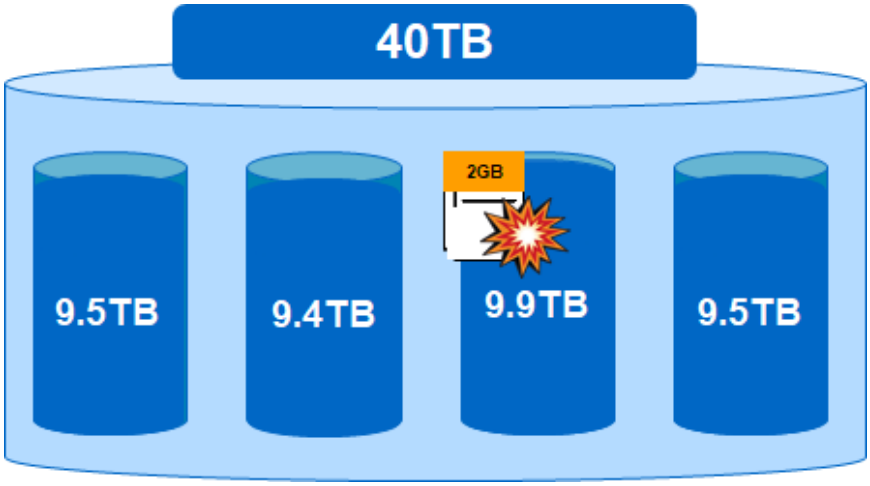
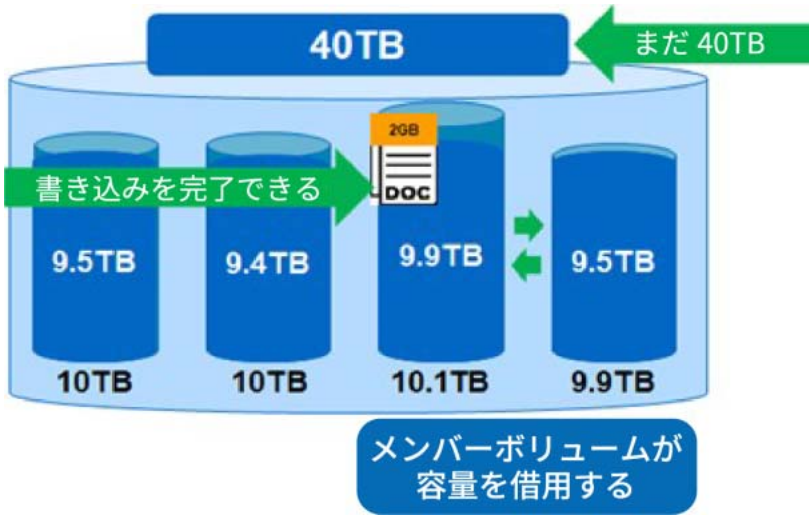


図 7.11 エラスティックサイジング後のファイル書き込み動作



最終的には、いっぱいになったメンバーが緊急のイベントではなくなったため、エラスティックサイジングによって、いっぱいになったメンバーボリュームでのファイル書き込みの失敗を軽減し、容量管理の管理オーバーヘッドの一部を排除できます。
ONTAP は、他のメンバーボリュームに使用可能な空き領域がある限り、ファイルを書き込むことができます。
ただし、事後対応型であるため、プロアクティブなサイズ変更を利用するには ONTAP 9.8 にアップグレードすることをお勧めします。

ボリュームの自動拡張とエラスティックサイジングのどちらを使用するか

FlexGroup ボリュームでボリュームの自動拡張が有効になっている場合、ONTAP 9.7 ではそのボリュームのエラスティックサイジングは無効になります。この 2 つの機能は基本的に冗長であるためです。

ただし、動作方法や、どちらか一方を優先して使用するタイミングには違いがあります。

- 新しいデータを書き込むために FlexGroup ボリュームの総容量を拡張できる場合は、ボリュームの自動拡張を使用する必要があります。
- エラスティックサイジングはデフォルトで有効になっているため、FlexGroup ボリュームの合計サイズが指定された容量を超えて大きくならないようにする場合に使用します。

エラスティックサイジングのパフォーマンスの効果

ONTAP が FlexGroup ボリュームの容量を増やすためにファイルの書き込みを一時停止するたびに、クライアントのレーテンシーが発生します。ファイルへの書き込み操作で発生するレーテンシーの量は、より多くの領域を見つけるために書き込みを停止する必要がある回数によって異なります。たとえば、メンバーボリュームに 10GB しか使用できないが、100GB のファイルが書き込まれている場合、エラスティックサイジングにより、書き込みが何度も一時停止され、書き込みが完了できるようになります。この数は、メンバーボリュームの合計サイズによって決まります (10MB~10GB)。

次の例は、ファイルが FlexGroup ボリュームにコピーされたテストを示しています。最初のテストでは、FlexGroup 構成要素はファイルを保持するのに十分な大きさではなかったため、エラスティックサイジングを使用しました。6.7GB のファイルのコピーには約 2 分かかりました。

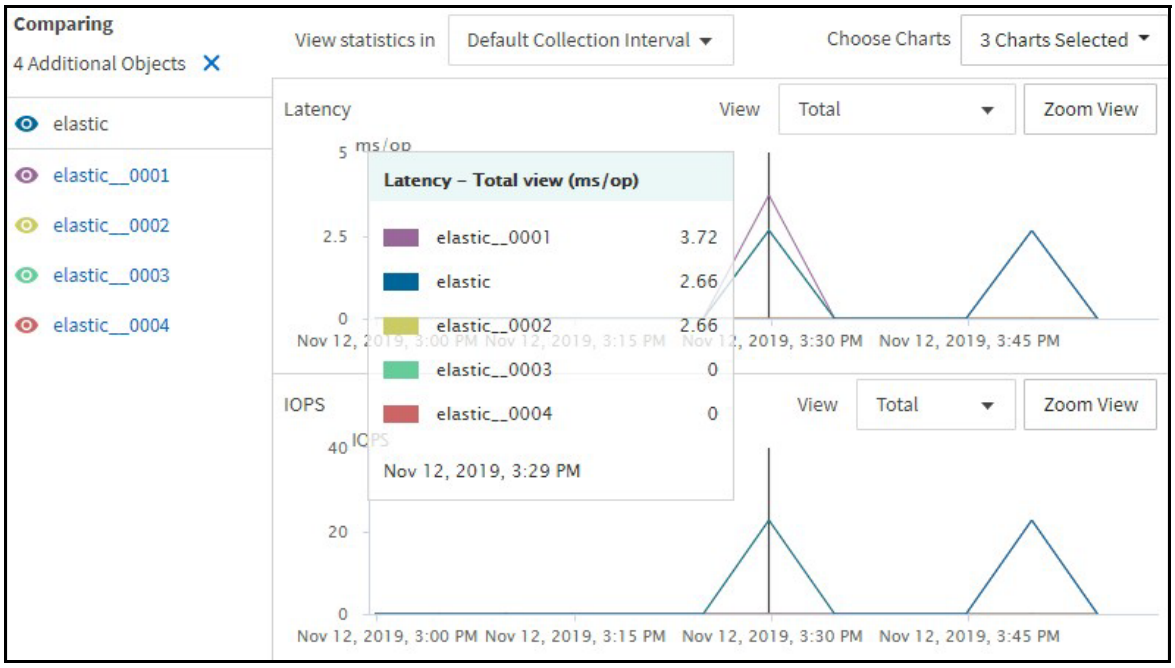
```
[root@centos7 /]# time cp Windows.iso /elastic/
real    1m52.950s
user    0m0.028s
sys     1m8.652s
```

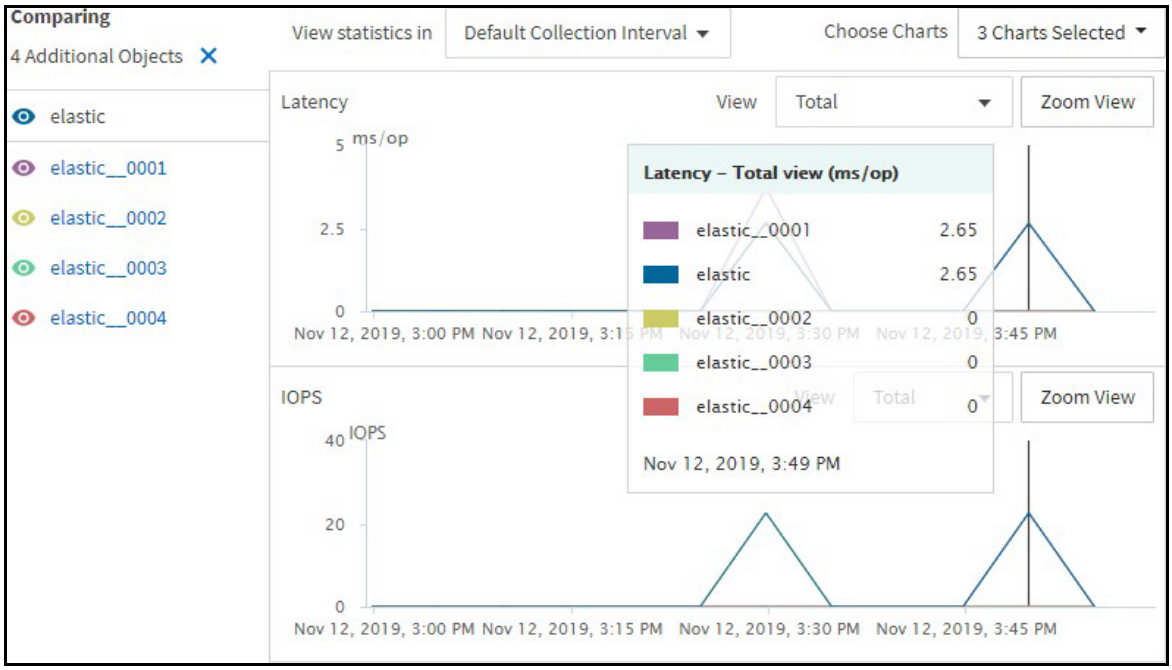
FlexGroup 構成ボリュームがエラスティックサイジングを回避するのに十分な大きさである場合、同じコピーにかかる時間は 15 秒短縮されました。

```
[root@centos7 /]# time cp Windows.iso /elastic/
real    1m37.233s
user    0m0.052s
sys     0m54.443s
```

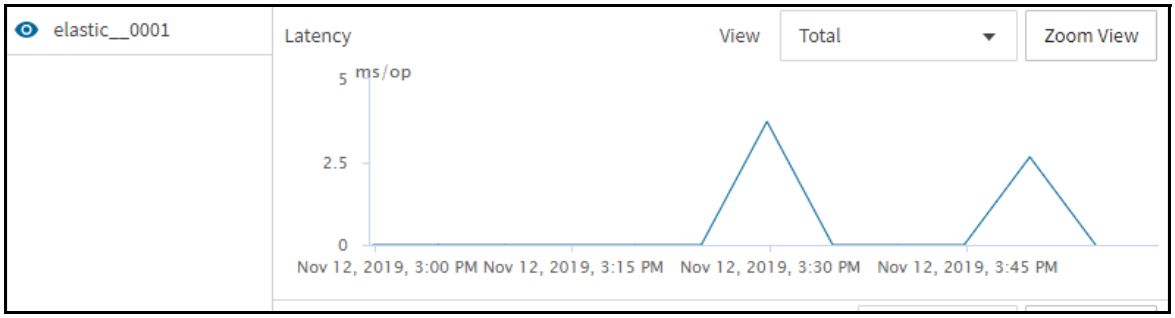
これは、エラスティックサイジングには実際に遅延効果があることを示しています。

次のグラフは、構成ボリュームの遅延ヒットを示しています。





エラスティックサイジングが有効な場合、構成ボリューム 0001 ではレーテンシーが約 0.5 ミリ秒増加します。



エラスティックサイジングがパフォーマンスの問題の原因になっていると考えられる場合は、次のいずれかの操作を行います。

- サポートケースを開き、現象とログを確認します。
- FlexGroup ボリュームを大きくして、方程式からエラスティックサイジングを削除するのに十分なスペースがあることを確認します。

ONTAP 9.8 以降では、FlexGroup メンバーボリュームでエラスティックサイジングが行われたことを通知する新しい EMS イベント (`fg.member.elastic.sizing`) が導入されています。

ただし、プロアクティブなサイズ変更では、すべてのメンバーボリュームのサイズ変更イベントはエラスティックサイジングと見なされ、パフォーマンスの問題を示すものではありませんが、FlexGroup に容量を追加して対処する必要のある容量の問題が FlexGroup に存在します。

エラスティックサイジング EMS の例については、「17.8 イベント管理システムの例」を参照してください。

FlexGroup メンバーボリュームに 20% を超える空き容量を確保することは、エラスティックサイジングの必要性を回避するための理想的な方法です。このためには、ONTAP 9.7 で容量を厳密に管理する必要があります。ただし、ONTAP 9.8 では、プロアクティブなサイズ変更によって ONTAP がユーザーのために空き領域を維持します。

7.10.11 プロアクティブなサイズ変更

ONTAP 9.8 には容量管理のための新機能が導入されており、ストレージ管理者から容量管理タスクを除外し、代わりに ONTAP に FlexGroup メンバーボリュームの容量を管理させることを目的としています。

プロアクティブなサイズ変更に関しては、次の点を考慮する必要があります。

- メンバーボリュームの容量が 60% 未満の場合、容量の差が大きくてもメンバーボリュームのサイズは変わりません。
- プロアクティブなサイズ変更では、メンバーボリュームのサイズを使用済み容量の 60% ～ 80% の範囲で少しずつ調整し、使用可能なスペースの比較的均等なバランスを維持するようにします。
- 容量が 80% 使用された後は、メンバボリュームの合計サイズを上下に調整して、容量使用率を均等に維持することが目標です。
- サイズ変更が行われる場合、サイズは大きくなりません。範囲は 10MB ～ 10GB です。しかし、空き領域をチェックするために一時停止する必要がないため、エラスティックサイジングのようにパフォーマンスに影響を与えません。容量の問題が発生する前にサイズ変更が行われます。
- ボリュームの自動サイズ設定を有効にしている場合、メンバーボリュームが自動拡張しきい値に達すると、ボリュームの自動サイズ設定が実装されます。

この空き容量バッファは、ボリューム全体でファイルの均等な取り込みを維持し、容量の不均衡を軽減し、ONTAP の FlexGroup ボリュームの容量管理を向上させます。

プロアクティブなサイズ変更動作 : ボリュームの自動サイズ設定が無効

次の例では、4 つの 100GB メンバーボリュームを持つ 400GB FlexGroup ボリュームが使用されています。クライアントは、4 つのフォルダにわたって 32 個の 10GB ファイルを FlexGroup ボリュームに作成します。FlexGroup ボリュームのボリュームの自動サイズ設定が無効になっています。つまり、容量が 100% に達しても、指定した FlexGroup ボリュームサイズはそのまま維持されます。

最初のファイルが書き込まれた後のジョブの開始時には、容量バランスは次のように表示されます。

図 7.12 初期 FlexGroup データバランス：プロアクティブなサイズ変更、自動サイズ設定が無効

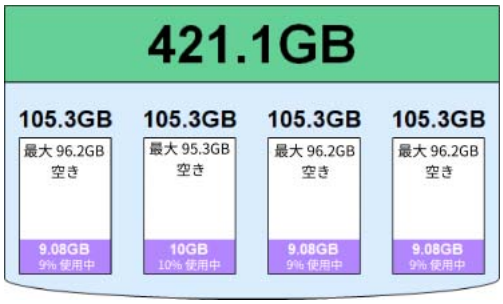
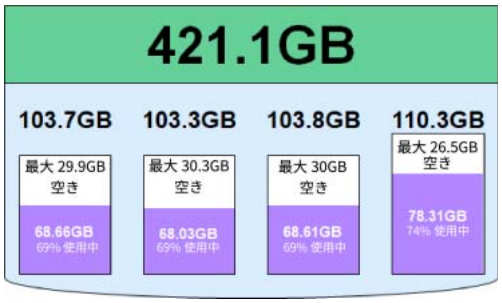
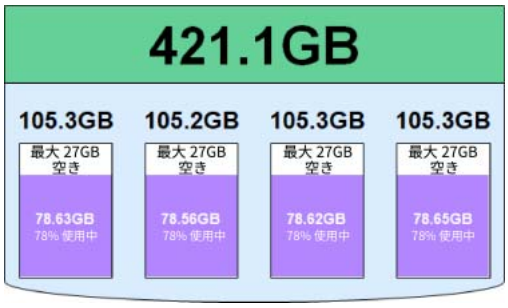


図 7.13 FlexGroup のデータバランス (使用率は最大 68%)：プロアクティブなサイズ変更、自動サイズ設定が無効



使用容量が約 70% になると、バランスのとれた空き領域を維持するために、メンバーボリュームのサイズが少し変更され始めますが、総容量は同じままです。

図 7.14 FlexGroup データバランス、ジョブ完了：プロアクティブなサイズ変更、自動サイズ設定が無効



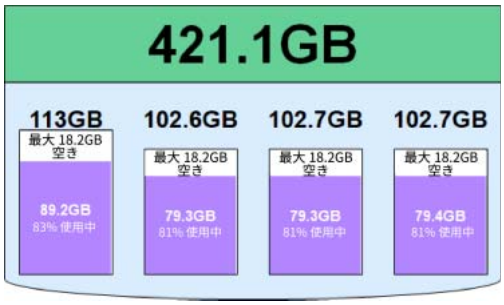
ジョブが完了すると、ONTAP は、使用されている領域がすべてのメンバーボリュームにわたって均等であることを確認し、プロアクティブなサイズ変更によってメンバーボリュームが元のサイズに縮小され、十分な空き領域があるためすべて同じサイズになります。合計 FlexGroup サイズは変更されていません。

では、この後に新しい 10GB のファイルが書き込まれるとどうなるでしょうか？

ファイルが書き込まれると、そのファイルはメンバーボリュームの 1 つに配置されます。これによりデータの不均衡が生じますが、ONTAP はそれに応じて他のメンバーボリュームのサイズを変更し、空き領域を均等に維持します。

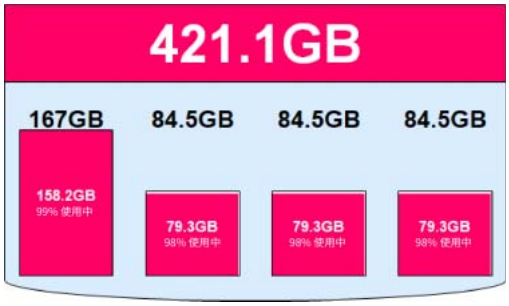
新しい 10GB ファイルが書き込まれた後のデータバランスは次のとおりです。

図 7.15 FlexGroup データバランス、新しい大容量ファイル : プロアクティブなサイズ変更、自動サイズ設定が無効



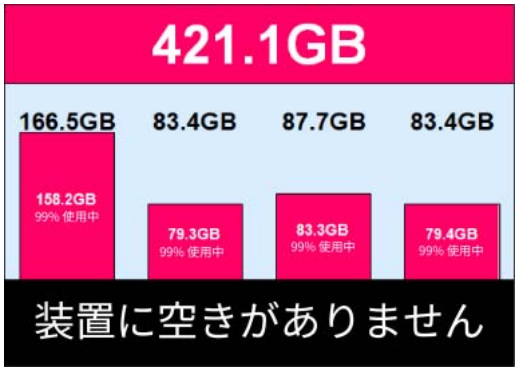
ご覧のとおり、新しいファイルは最初のメンバーボリュームに配置されました。エラスティックサイジングにより、そのメンバーボリュームは 113GB に拡張され、他のメンバーボリュームは縮小されましたが、使用可能な空き容量と FlexGroup の合計サイズはほぼ同じ量に維持されています。

図 7.16 FlexGroup データバランス、80GB ファイル：プロアクティブなサイズ変更、自動サイズ設定が無効



次に、新しいファイルを FlexGroup に再度書き込みます。今回は、ファイルが大きすぎて 1 つのメンバーボリューム (80GB) に収まりません。また、FlexGroup 自体には大きすぎます。その場合、ONTAP はプロアクティブなサイズ変更を使用しますが、すべてのメンバーボリュームに 1GB の空き容量しか残っていない状況ではありません。つまり、FlexGroup 全体の容量が不足し、自動サイズ設定が無効になっているため、次の 10GB のファイルは失敗します。

図 7.17　FlexGroup のデータバランス、領域不足 : プロアクティブなサイズ変更、自動サイズ設定が無効



その結果、次のファイルの作成は失敗しますが、領域不足エラーを回避するために、メンバーボリュームに空き領域を追加する際には、プロアクティブなサイズ変更の方がはるかに積極的になります。ただし、FlexGroup ボリューム自体の領域が不足している場合は、FlexGroup ボリュームを手動で拡張するか、ボリュームの自動サイズ設定を有効にするしかありません。

プロアクティブなサイズ変更動作 : ボリュームの自動サイズ設定が有効

次の例では、4 つの 100GB メンバーボリュームを持つ 400GB FlexGroup ボリュームが使用されています。クライアントは、4 つのフォルダにわたって 32 個の 10GB ファイルを FlexGroup ボリュームに作成します。FlexGroup ボリュームでは、デフォルト設定でボリュームの自動サイズ設定が有効になっています。これは次のことを意味します。

- FlexGroup ボリュームは、プロアクティブなサイズ変更が行われた場合でも、使用スペースが 92% のしきい値に達するまで、同じ容量を維持します。
- 使用済み領域のしきい値に達すると、デフォルト設定に従って、ボリュームの増加は 20% 以下になります。この場合、566.7GB がボリュームの最大拡張サイズになります。これは、このボリュームのサイズが拡張され、その後縮小されたため、20% より大きくなります。
- 使用済み容量が 50% を下回ると、ボリュームは元のサイズの 421.1GB に縮小されます。

FlexGroup の自動サイズ設定は次のとおりです。

```
cluster::> vol autosize -vserver DEMO -volume FG_SM_400G
Volume autosize is currently ON for volume "DEMO:FG_SM_400G".
The volume is set to grow to a maximum of 566.7g when the volume-used space is above 92%.
The volume is set to shrink to a minimum of 421.1g when the volume-used space falls below
50%.
Volume autosize for volume 'DEMO:FG_SM_400G' is currently in mode grow_shrink.
```

FlexVol ボリュームまたは FlexGroup ボリュームが小さい場合、デフォルトの増加しきい値の割合は小さくなります。
例 :

- 100GB のボリュームのデフォルトの拡張しきい値は 90%、縮小しきい値は 50% です。
- 10TB のボリュームのデフォルトの拡張しきい値は 98%、縮小しきい値は 50% です。

最初のファイルが書き込まれた後のジョブの開始時には、容量バランスは次のように表示されます。

図 7.18　初期 FlexGroup データバランス：プロアクティブなサイズ変更、自動サイズ設定が有効

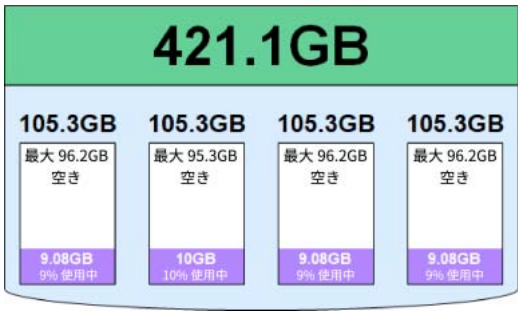
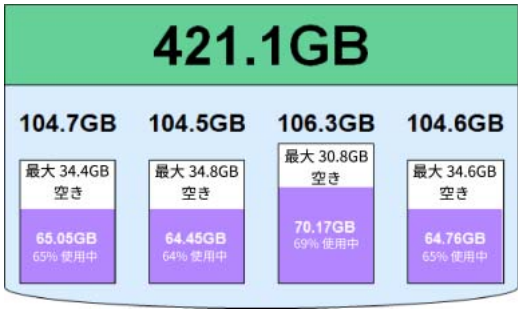
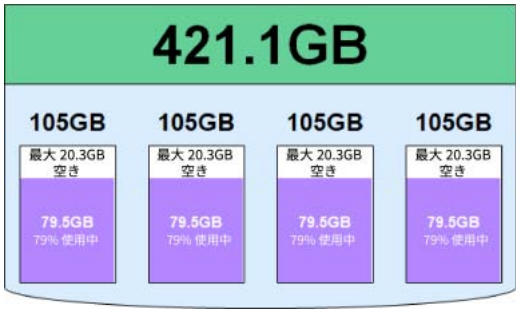


図 7.19 FlexGroup のデータバランス (使用率は最大 68%)：プロアクティブなサイズ変更、自動サイズ設定が有効



使用容量が約 66% になると、バランスのとれた空き領域を維持するためにメンバーボリュームのサイズが少し変更されますが、自動サイズ設定が無効になっている場合と同様に、総容量は同じままです。

図 7.20 FlexGroup データバランス、ジョブ完了：プロアクティブなサイズ変更、自動サイズ設定が有効



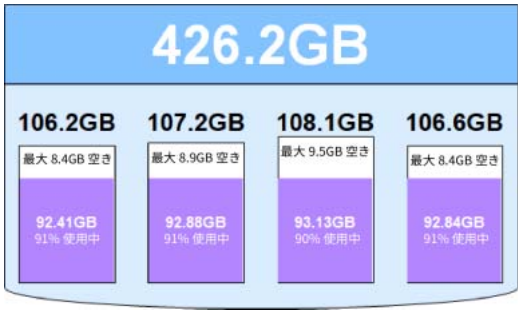
ジョブが完了すると、ONTAP は、使用されている領域がすべてのメンバーボリュームにわたって均等であることを確認し、プロアクティブなサイズ変更によってメンバーボリュームが元のサイズに縮小され、十分な空き領域があるためすべて同じサイズになります。合計 FlexGroup サイズは変更されていません。

ご覧のように、自動サイズ設定が有効になっている FlexGroup ボリュームは、自動サイズ設定が無効になっているときに、空き容量のしきい値が自動サイズ設定が有効になる値を下回ると、FlexGroup ボリュームと同じように動作します。

上の図では、FlexGroup ボリューム全体で最大約 81GB の空き容量があります。10GB のファイルを書き続けると、最終的に自動サイズ設定のしきい値に達し、ONTAP はそれに応じて反応し始めます。今回は、他のメンバーボリュームから空き領域を借りるのではなく、追加の領域を必要とするメンバーボリュームを自動サイズ設定によって拡張します。

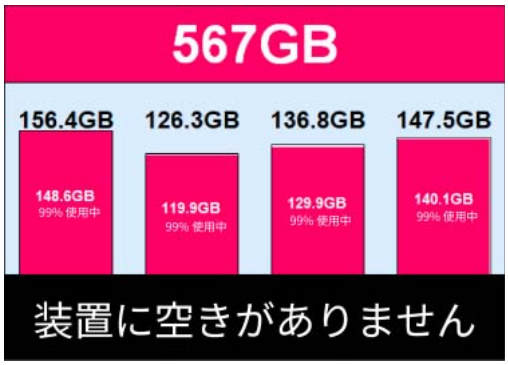
その結果、FlexGroup ボリューム全体の容量が増加します。次のテスト実行では、同じ FlexGroup に新しいフォルダを作成し、4 つのフォルダで FlexGroup ボリュームに 32 個の 10GB ファイルを作成するテストを再実行しました。

図 7.21 FlexGroup データバランス、2 回目のテスト実行：プロアクティブなサイズ変更、自動サイズ設定が有効



FlexGroup メンバーボリュームのいずれかが 92% の使用済み容量のしきい値に達すると、自動サイズ設定によりそのメンバーボリューム「のみ」増加します。他のメンバーボリュームも 92% に達した時点で増加する必要があり、その後も増加しています。これにより、FlexGroup ボリューム全体の容量が増加します。プロアクティブなサイズ変更では、他のメンバーボリュームの容量も上下に調整されるため、メンバーボリュームごとに比較的均等な空き容量を使用できます。

図 7.22　FlexGroup データバランス、自動サイズ設定制限：プロアクティブなサイズ変更、自動サイズ設定が有効



デフォルトのボリュームの自動サイズ設定では、ボリュームは総ボリュームサイズの 120% までしか拡大できません。2 回目のテスト実行では、正常に完了するために 320GB の容量が必要であり、ボリュームの自動拡張では、ボリュームの合計容量が 566.7GB までしか増やせなかったため、容量不足のためにジョブは失敗しました。

自動サイズ設定の考慮事項 : 小さい FlexGroup ボリューム

自動サイズ容量は総サイズのパーセンテージに基づいているので、小さい FlexGroup ボリューム (420GB の FlexGroup ボリュームなど) は、大きい FlexGroup ボリュームよりもデフォルトでは拡張のためのランウェイが少なくなります。デフォルトの自動拡張最大値は、合計ボリュームサイズの 120% に制限されます。ボリュームが手動で拡張され、縮小されて元に戻る場合、自動拡張の値はより大きなボリュームサイズを反映します。

表 7.7　最大サイズの自動調整の例

FlexGroup ボリュームサイズ	デフォルトの最大自動サイズ	デフォルトのサイズ差分
420GB	480GB	+80GB
100TB	120TB	+20TB
400TB	480TB	+80TB

そのため、FlexGroup ボリュームに対してボリューム自動サイズ設定を使用している場合は、次のガイダンスに従ってください。より大きな FlexGroup ボリュームを使用し、デフォルトの自動サイズ値を維持します。

- 小さい FlexGroup ボリュームを使用している場合は、停止を避けるためにデフォルトの `-max-autosize` 値を変更します。
- エンドユーザーに提供した容量以上の容量を提供したくない場合でも、qtree とクォータを使用してエンドユーザーが表示および使用する容量を制限すると、ボリュームの自動サイズ設定を使用できます。
- ボリュームの自動サイズ設定を無効にする場合は、FlexGroup ボリューム内に使用可能なスペースがなくなると、ONTAP 9.8 でプロアクティブなサイズ変更を行ってもファイルの書き込みが失敗することに注意してください。

ベストプラクティス 9: 容量管理のための ONTAP 機能の組み合わせ
容量管理にアプローチする最善の方法は、より大きな FlexGroup ボリューム、ボリュームの自動サイズ設定、ONTAP 9.8 以降、qtree、およびクォータの適用を組み合わせることです。FabricPool を使用してクラウドまたは S3 への自動階層化を実行すると、この話はさらに説得力を増します。これらの機能を使用すると、ストレージ管理者の容量管理オーバーヘッドを最小限に抑えることができます。

7.11 ネットワークに関する考慮事項

CIFS/SMB または NFS を使用する場合、各マウントポイントは、クラスタ内の単一の IP アドレスへの単一の TCP 接続を介して作成されます。ONTAP では、これらの IP アドレスは SVM における仮想ネットワークインタフェースであるデータ LIF に付加されます。

IP アドレスは、リンクアグリゲーション制御プロトコル (LACP) または別のトランク構成に参加する 1 つのハードウェア Ethernet ポートまたは複数のハードウェア Ethernet ポートに割り当てることができます。しかし、

ONTAP では、これらのポートは常に 1 つのノード上に存在します。つまり、これらのポートはそのノードの CPU や PCI バスなどを共有しています。単一ノードの潜在的なボトルネックを軽減するために、ONTAP を使用すると、クラスタ内の任意のノードに TCP 接続を確立した後、その要求をクラスタのバックエンドネットワーク経由で適切なノードにリダイレクトできます。この方法は、ハードウェアシステム全体にネットワーク接続と負荷を適切に分散するのに役立ちます。

ベストプラクティス 10: FlexGroup を使用するネットワークの設計
FlexGroup ネットワークのベストプラクティスは、FlexVol ネットワークのベストプラクティスに似ています。ONTAP で NAS ソリューションを設計する場合は、ボリュームのスタイルに関係なく、次のネットワークのベストプラクティスを考慮してください。 • 各ノードへのパスを確認するために、SVM ごとにノードごとに少なくとも 1 つのデータ LIF を作成します。 • 何らかの形式の DNS ロードバランシングを使用して、単一の完全修飾ドメイン名 (FQDN) の背後にあるクライアントに複数の IP アドレスを提示します。 • 可能な場合は、LACP ポートを使用してデータ LIF をホストし、スループットとフェイルオーバーを考慮します。 • クライアントを手動でマウントする場合は、TCP 接続をクラスタノード間で均等に分散します。それ以外の場合は、DNS の負荷分散でクライアントの TCP 接続の分散を処理できるようにします。 • 頻繁にマウントとアンマウントを行うクライアントの場合は、オンボックス DNS を使用して負荷を分散することを検討してください。クライアントが頻繁にマウントおよびアンマウントされない場合、オンボックス DNS はあまり役に立ちません。 • ワークロードが「マウントストーム」(数百または数千のクライアントが同時にマウントされる) の場合は、オフボックス DNS ロードバランシングを使用するか、FlexCache ボリュームの使用を検討します。単一ノードへのマウントストームは、クライアントへのサービス拒否またはパフォーマンスの問題を引き起こす可能性があります。 • NFSv4.1 を使用している場合は、データのローカリゼーションとファイルへの並列接続に pNFS を利用することを検討してください。pNFS はシーケンシャル I/O ワークロードに最適です。メタデータの負荷が高いと、単一のメタデータサーバー接続でボトルネックになる可能性があります。 • 最新の SUSE および Ubuntu クライアントなど、pNFS をサポートするクライアントがある場合は、Nconnect マウントオプションを使用すると、単一クライアントでの NFS マウントのパフォーマンスがさらに向上します。 • SMB3 ワークロードの場合は、CIFS サーバでマルチチャネルおよび大容量 MTU 機能を有効にすることを検討してください。 • ネットワークでジャンボフレームを使用している場合は、ネットワークアーキテクチャの各エンドポイントでジャンボフレームが有効になっていることを確認します。ジャンボフレーム構成が一致しないと、どのボリュームタイプでも、診断が困難なパフォーマンスの問題が発生する可能性があります。 • NFS クライアントは、複数のネットワークインタフェースを介して ONTAP 内の同じボリュームに接続された同じクライアントからの複数のマウントポイントを使用して、パフォーマンスを向上させることができます。ただし、この構成は複雑になる可能性があります。NFS クライアントがサポートしている場合は、Nconnect を使用します。

7.11.1 LACP の考慮事項

クライアント側のネットワークで LACP ポートを使用する理由はいくつかあります。一般的で適切な使用例は、SMB 1.0 プロトコルを介してファイルサーバーに接続するクライアントに回復力のある接続を提供することです。SMB 1.0 プロトコルはステートフルであり、OSI スタックの上位レベルでセッション情報を維持するため、LACP はファイルサーバーが HA 構成にある場合に保護を提供します。SMB プロトコルを後で実装すると、LACP ポートを設定しなくても、耐障害性のあるネットワーク接続を実現できます。

LACP はスループットとリカバリ性にメリットをもたらしますが、LACP 環境の保守の複雑さを考慮して判断する必要があります。LACP が含まれる場合でも、複数のデータ LIF を使用する必要があります。

7.11.2 DNS ロードバランシングの考慮事項

DNS ロードバランシング (オフボックスとオンボックスの両方) は、ネットワーク接続をクラスター内のノードおよびポートに分散する方法を提供します。FlexGroup ボリュームは、DNS ロードバランシングの背後にある全体的な考え方を変えるものではありません。ストレージ管理者は、NAS コンテナの種類に関係なく、ネットワーク接続をクラスタ全体に均等に分散する必要があります。ただし、FlexGroup ボリュームの設計により、FlexGroup ボリュームが複数のクラスタノードにまたがる場合、リモートクラスタトラフィックはほぼ確実です (pNFS データのローカリティは例外です)。したがって、これらの構成では、ネットワーク接続とデータのローカリティに関する考慮事項は無効になります。その結果、データのローカリティを気にする必要がなくなったため、ある種の DNS ロードバランシングは FlexGroup ボリュームとの相性が少し良くなりました。最終的に、どの DNS ロードバランシング方式を使用するかは、ストレージ管理者およびネットワーク管理者の目標によります。

ベストプラクティス 11: DNS ロードバランシングの形式を使用する
可能な場合は、FlexGroup メンバーボリュームを含むノード上で、FlexGroup ボリュームを使用して DNS ロードバランシングを行います。

オンボックス DNS かオフボックス DNS か ?

ONTAP は、オンボックス DNS サーバーを介して DNS クエリをサービスする方法を提供します。この方法では、ノードの CPU とスループットを考慮して、NAS アクセス要求に対応するのに最適なデータ LIF を決定します。

- オフボックス DNS は、DNS 管理者が、データ LIF へのラウンドロビンアクセスを提供する外部 DNS サーバー上に、同じ名前を持つ複数の「A」名前レコードを作成することによって構成されます。
- マウントストームのシナリオを作成するワークロードの場合、ONTAP オンボックス DNS サーバでは適切なバランスを維持できないため、オフボックス DNS を使用することをお勧めします。

特に FlexGroup ボリュームを使用する場合に、SVM ごとにノードごとに少なくとも 1 つのデータ LIF を作成することをベスト・プラクティスとして推奨しています。このため、DNS の負荷分散によって DNS エイリアスの背後にある IP アドレスをマスクすることをお勧めします。次に、各クライアントの複数の IP アドレスに複数のマウントポイントを作成して、クラスタと FlexGroup ボリュームのスループットを向上させる必要があります。

7.11.3 Border Gateway Protocol (BGP)

ONTAP は BGP をサポートして、ストレージシステムに最新のネットワークスタックを提供します。BGP サポートでは、ポートの使用効率を向上させるためにレイヤ 3 (L3) ルーティング、改善されたロードバランシングインテリジェンス、および仮想 IP (VIP) が提供されています。

FlexGroup ボリュームでは、この新しいネットワーク要素を使用するために構成を変更する必要はありません。

7.12 セキュリティとアクセス制御リストのスタイルに関する考慮事項

ONTAP では、NFS および SMB/CIFS を介して同じデータにアクセスできますが、ファイルの所有権は保持され、適切なファイル権限が適用されます。これは、マルチプロトコル NAS アクセスと呼ばれます。FlexVol ボリュームに適用されるマルチプロトコル NAS の一般的なガイダンスは、FlexGroup ボリュームにも適用されます。これらは機能的には認証と許可で同じように動作します。このガイドは、「CIFS, NFS, Multiprotocol Express Guides」および 「CIFS and NFS Reference Guides」に記載されています。

一般に、マルチプロトコルアクセスには次のものが必要です。

- 有効なユーザー (Windows および UNIX)
- 有効な名前マッピング規則、または LDAP や NIS などのローカルファイルやサーバーを介した 1:1 の名前マッピング。ONTAP は名前マッピングを使用して、クライアントのアクセスを調整します。
- ボリュームセキュリティのスタイル (NTFS、UNIX、または混在)。ボリュームまたは qtree に対して構成できます。
- デフォルトの UNIX ユーザー (pcuser、Windows から UNIX への名前マッピング用にデフォルトで作成)。デフォルトの Windows ユーザー (UNIX から Windows への名前マッピング) は、デフォルトでは構成されていません。

ボリュームを作成すると、セキュリティスタイルが適用されます。セキュリティスタイルを指定せずにボリュームを作成した場合、ボリュームは SVM ルートボリュームのセキュリティスタイルを継承します。ボリュームセキュリティスタイルは、NAS ボリュームに使用されるアクセス制御リスト (ACL) のスタイルを決定し、ユーザーの認証方法と SVM へのマッピング方法に影響します。FlexGroup ボリュームでセキュリティスタイルが選択されている場合、すべてのメンバーボリュームに同じセキュリティスタイル設定が適用されます。

qtree を使用して、FlexGroup ボリュームに一意のセキュリティスタイルを指定できます。

7.12.1 ボリュームセキュリティスタイルの基本的なガイダンス

次に、ボリュームのセキュリティスタイルの選択に関する一般的なガイダンスを示します。

- UNIX セキュリティスタイルでは、Windows ユーザーは有効な UNIX ユーザーにマップする必要があります。UNIX ユーザーが有効なユーザー名にマップする必要があるのは、NFSv4.x を使用している場合のみです。
- NTFS セキュリティスタイルでは、Windows ユーザーは有効な UNIX ユーザーにマップする必要があり、UNIX ユーザーは認証のために有効な Windows ユーザーにマップする必要があります。有効な UNIX ユーザー名が存在する場合、NFS クライアントはファイルおよびフォルダの適切な所有権を認識します。承認 (パーミッション) は、最初の認証後に Windows クライアントによって処理されます。有効な UNIX ユーザーが存在しない場合、デフォルトの UNIX ユーザー (pcuser) が認証 /UNIX 所有権に使用されます。
- UNIX セキュリティ・スタイルでは、一部の Windows クライアントが基本モードのビット権限 (所有権の変更、rwx) を変更できます。ただし、SMB での NFSv4.x ACL 管理は許可されず、高度な NTFS アクセス許可も認識しません。
- 混在セキュリティスタイルを使用すると、任意の種類のクライアントからアクセス許可を変更できます。ただし、ACL を変更する最後のクライアントタイプに基づいて、NTFS または UNIX の基本的な有効なセキュリティスタイルがあります。
- 有効なセキュリティスタイルが変更されるため、混在するセキュリティスタイルが正しく機能するには、適切な名前マッピングが必要です。
- FlexGroup ボリューム内の ACL スタイルの細分性が必要な場合は、qtree の導入を検討してください。qtree を使用すると、ONTAP の論理ディレクトリごとにセキュリティスタイルを設定できます。FPolicy、ウイルス対策、ネイティブファイル監査、クォータの適用などの他のホームディレクトリ機能が必要な場合は、最新のパッチ適用済みリリースを使用します。
- FlexGroup ボリュームの NFSv4.x および NFSv4 ACL サポートは、ONTAP 9.7 で追加されました。

ベストプラクティス 12: ボリュームセキュリティスタイル : 混在したセキュリティスタイルのガイダンス
クライアントが両方のスタイルのクライアントから権限を変更できる必要がある場合にのみ、混在したセキュリティ・スタイルを推奨します。それ以外の場合は、マルチプロトコル NAS 環境であっても、セキュリティスタイルとして NTFS または UNIX を選択することをお勧めします。

ユーザーマッピング、ネームサービスのベスト・プラクティスなどの詳細については、製品マニュアルを参照してください。詳細については、「ETERNUS AX/HX series ONTAP で LDAP を設定する方法 マルチプロトコル NAS ID 管理」を参照してください。

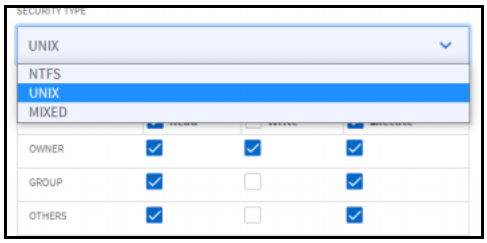
FlexGroup ボリュームのセキュリティスタイルの変更

FlexGroup ボリュームは、FlexVol ボリュームと同様に管理することを目的としています。ボリュームのセキュリティスタイルの変更は、この哲学に含まれています。ボリュームのセキュリティスタイルはライブで変更でき、クライアントを再マウントする必要はありません。ただし、その後 ACL スタイルが変更されると、アクセス権が予測できなくなる可能性があります。最善の結果を得るために、本番データセットのメンテナンスウィンドウでセキュリティスタイルを変更することを推奨しています。

FlexGroup ボリュームのセキュリティスタイルを変更するには、次のいずれかの操作を行います。

- コマンドラインから `volume modify` を使用します。
- ONTAP System Manager で FlexGroup ボリュームを最初に作成するときは、[Edit] ボタンまたは [Advanced Features] を使用します。

図 7.23　ONTAP System Manager での FlexGroup ボリュームセキュリティースタイルの変更



NFSv3 クライアントでの NFSv4.x ACL の使用

NFSv4 経由でエクスポートをマウントした管理クライアントを使用して、FlexGroup ボリュームで NFSv4 ACL をサポートする ONTAP リリースを使用している場合、ONTAP では、クライアントが NFSv3 だけを使用している場合でも、NFSv4.x の詳細 ACL サポートの利点を活用できます。そのクライアントから NFSv4 ACL を設定することによって、NFSv3 からアクセスするクライアントは、NFSv4.x マウントを使用することなく、それらの NFSv4 ACL を受け入れます。

8. FlexGroup 管理の考慮事項

この章では、FlexGroup ボリュームの表示、ボリュームの移動、FlexGroup ボリュームのサイズ変更、名前の変更など、FlexGroup ボリューム管理に関する一般的な考慮事項について説明します。

備考

このドキュメントの ONTAP System Manager の例では、ONTAP 9.7 以降で使用可能な新しいバージョンの System Manager を使用しています。

8.1 FlexGroup ボリュームの表示

FlexGroup ボリュームは、ONTAP GUI またはコマンドラインから作成でき、通常の FlexVol ボリュームと同様に、ストレージ管理者の観点から管理できるように設計されています。Snapshot コピー、サイズ変更、ストレージ効率ポリシーなどはすべて FlexGroup ボリュームレベルで管理されます。

ただし、FlexGroup ボリュームは単なる FlexVol ボリュームではありません。代わりに、FlexGroup ボリューム全体で一致して動作する一連の FlexVol メンバーボリュームで構成されます。ONTAP は、これらのメンバーボリュームをデータの取り込みに使用して、ファイルシステム全体にわたる自動ロードバランシングと並列処理を実現し、容量とパフォーマンスを向上させます。

ほとんどの場合、FlexGroup ボリュームは FlexGroup レベルで管理できます。たとえば、FlexGroup ボリュームを拡張する場合は、FlexGroup レベルで GUI を使用するか `volume size` コマンドを実行して、ボリュームの合計サイズを増やすことができます。ONTAP は、基盤となるすべてのメンバーボリュームに同等の容量が割り当てられていることを確認するため、ストレージ管理者は容量の配分方法を考慮する必要がありません。ONTAP 9.8 以降では、新しいプロアクティブなサイズ変更機能により、FlexGroup レベルでの容量管理がさらにシンプルになりました。

まれに、個々の FlexVol メンバーボリュームを表示して容量とパフォーマンスの問題を確認する場合があります。これらのタスクは、ONTAP の以前のリリースではより一般的です。

メンバーボリュームの表示が役立つ 2 つのシナリオ：

- メンバーの容量使用状況を表示する場合（1 つのメンバーボリュームで満杯に近づいているか）
- 個々のメンバーのパフォーマンスを表示する場合（`volume move` を使用する必要があるか）

次の各項では、FlexGroup ボリュームの表示について説明します。

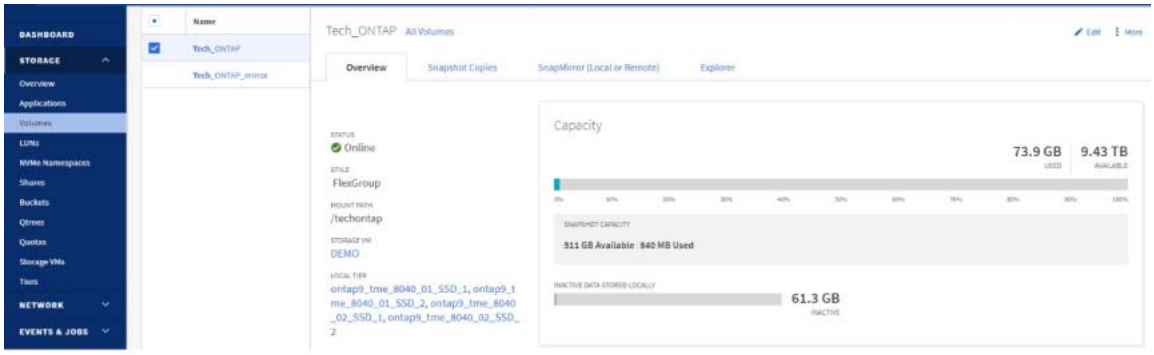
8.1.1 ONTAP System Manager

ONTAP System Manager では、[FlexGroup] タブを使用して、FlexGroup レベルで FlexGroup ボリュームを表示および管理できます。ただし、メンバーボリュームのビューはありません。これは仕様です。FlexGroup ボリュームは管理が容易である必要があります。ONTAP System Manager は、データ保護情報、リアルタイムのパフォーマンス、容量情報など、これらのビューで FlexGroup ボリュームに関する有用な情報を提供します。

備考

ONTAP System Manager は、シンプロビジョニングされた FlexGroup ボリュームのスペース割り当て情報を提供できないことに注意してください。

図 8.1　ONTAP System Manager の FlexGroup ボリュームビュー



FlexVol ボリュームと同様に、スナップショット、サイズ変更、データ保護などの基本タスクを System Manager GUI から管理できます。

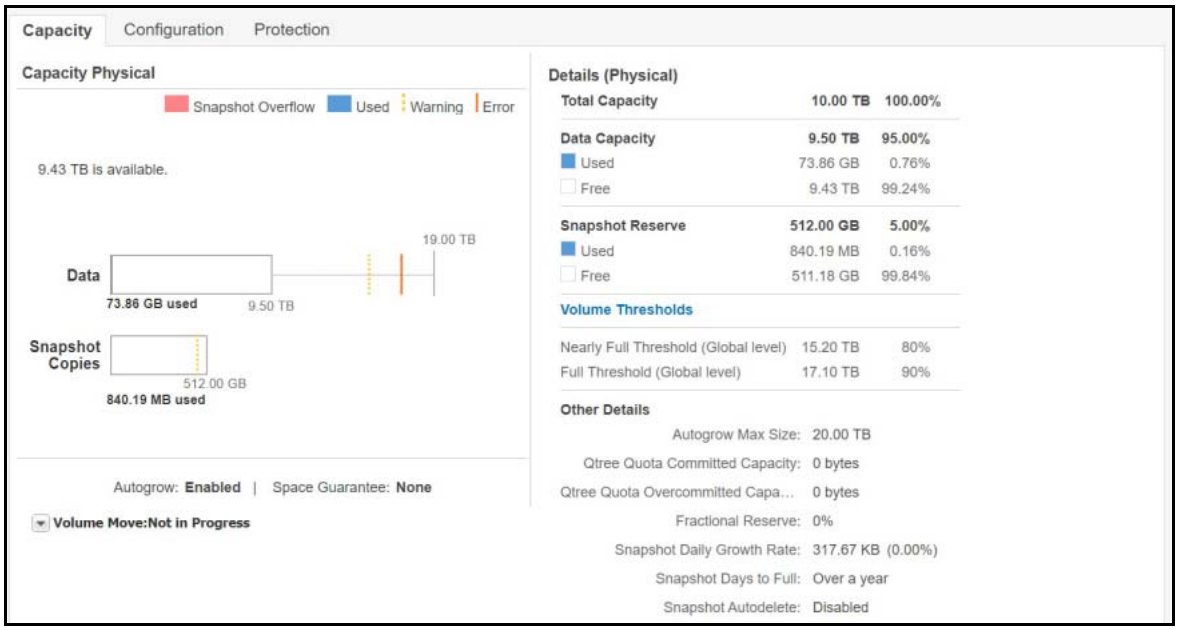
8.1.2　Active IQ Unified Manager

Active IQ Unified Manager を使用すると、ストレージ管理者は 1 つのダッシュボードで ONTAP クラスタの稼働状態とパフォーマンスを確認できます。

Active IQ Unified Manager を使用すると、FlexGroup のボリューム容量、構成、およびストレージ効率をグラフィカルな形式で確認できます。Active IQ Unified Manager の FlexGroup ボリューム容量は FlexGroup レベルで表示され、個々のメンバーのボリューム容量は表示されません。

パフォーマンスに関するメンバーボリューム情報については、Active IQ Performance Manager を使用してください。メンバーボリュームの容量情報を表示するには、コマンドラインを使用します。

図 8.2　Active IQ Unified Manager、FlexGroup 容量ビュー



Active IQ Unified Manager では、ワークロード分析を使用してボリューム容量の傾向を表示することもできます。

図 8.3　Active IQ Unified Manager での容量の傾向

備考

現在、FlexGroup ボリュームでは Active IQ ワークロードを使用できません。

パフォーマンス監視

Active IQ Unified Manager は、FlexGroup 全体とそのメンバーボリュームを含む ONTAP のパフォーマンス統計のアーカイブを収集します。ストレージ管理者は、FlexGroup ボリュームの詳細なビューを使用して、個々のメンバー FlexVol ボリュームのパフォーマンス異常を評価し、必要に応じて次のような修正措置を講じることができます。

- スペースの追加
- メンバーの追加 (`volume expand`)
- 割り当ての少ないノードへの無停止のボリューム移動

備考

これらのタスクは、Active IQ では実行できません。現在、これらのタスクを実行できるのは、コマンド行または ONTAP System Manager GUI、あるいはその両方です。

図 8.4 に、いくつかの FlexVol メンバーとそれに対応するパフォーマンスを示します。各行は FlexVol メンバーを表します。

図 8.4　Active IQ Performance Manager の FlexGroup ボリュームビュー

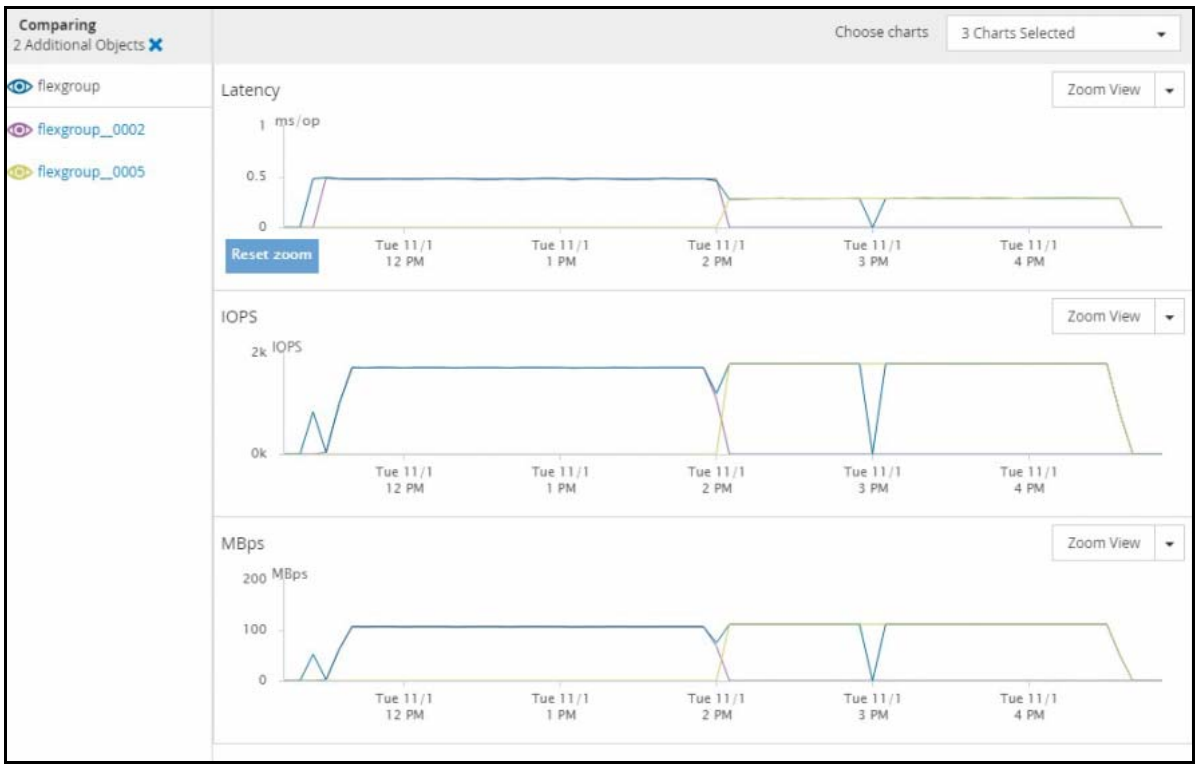
図 8.5 では、2 つの 1TB ファイルが FlexGroup ボリュームに書き込まれています。この表では、そのワークロードを引き受けたメンバーボリューム (メンバー 2 および 5) と、ワークロードのパフォーマンスのサマリーを確認できます。図 8.6 では、IOPS と MBps のグラフを確認できます。

図 8.5　　メンバーボリュームのパフォーマンス表

View and compare　Constituents of this FlexGroup　Filtering

Volume	Latency	IOPS	MBps	
flexgroup__0001	4 ms/op	< 1 IOPS	0 MBps	Add
flexgroup__0002	0.481 ms/op	1,581 IOPS	98.8 MBps	Add
flexgroup__0005	0.287 ms/op	1,743 IOPS	109 MBps	Add
flexgroup__0006	N/A	N/A	N/A	Add
flexgroup__0004	N/A	N/A	N/A	Add
flexgroup__0003	N/A	N/A	N/A	Add
flexgroup__0008	N/A	N/A	N/A	Add
flexgroup__0007	N/A	N/A	N/A	Add

図 8.6　　メンバボリュームのグラフ



8.1.3 コマンドライン

CLI は、FlexGroup ボリューム情報を表示するもう 1 つの方法です。各権限レベルには、FlexGroup ボリュームプロパティを表示するためのオプションセットがあります。

admin 権限レベル

- 総容量 (total、available、および used: すべてのメンバーボリュームから計算)、ストレージ効率
- Snapshot リザーブまたは Snapshot ポリシー
- FlexGroup ボリュームがまたがるアグリゲートとノードのリスト
- ボリュームスタイルと拡張ボリュームスタイル (ボリュームが FlexGroup ボリュームかどうかを示します)
- セキュリティのスタイル、所有者、またはグループ
- ジャンクションパス
- ファイルと inode の最大数
- メンバーボリューム情報 (`-is-constitution true` または `volume show-space` 経由)

高度な権限レベル

- ディレクトリ・サイズの最大値
- FlexGroup マスターセット ID (MSID)
- ボリュームが 7-Mode から移行されたかどうか (FlexVol から FlexGroup へのボリューム変換に重要)
- FlexGroup 最大メンバーボリュームサイズ

diag 権限レベル

- メンバーボリュームの詳細情報 (容量、使用状況など)
- FlexGroup インジェスト統計 (`flexgroup show`)

備考

メンバーボリュームスペース情報は、コマンド `volume show-space` を使用して、admin 権限レベルで表示できます。例については、「14.1 コマンド・ラインを使用した容量の監視とアラート」を参照してください。

8.2 FlexGroup ボリューム容量の表示

このセクションでは、FlexGroup ボリュームの容量を監視するためのさまざまな方法について説明します。これには、ストレージ効率での全体的な削減量の表示も含まれます。FPolicy のサポートにより、FlexGroup の容量を監視することもできます。

備考

ONTAP 9.8 以降では、プロアクティブなサイズ変更によってメンバーボリュームの空きスペースバランシングが処理されるため、容量ビューは、基礎となるメンバーボリュームではなく、FlexGroup ボリューム全体にフォーカスする必要があります。ただし、メンバーボリュームを表示する必要がある場合は、以下のセクションを参照してください。

8.2.1 FlexGroup の総容量

FlexGroup の総容量は、次から導出される数値です。

- 合計容量
  FlexGroup ボリュームの合計割り当て済み容量 (メンバーボリューム容量 × メンバー数)。
- 使用可能なスペース
  最も割り当てられたメンバーボリュームで使用可能なスペースの量。

10TB の FlexGroup ボリュームをプロビジョニングすると、クライアントには 10TB と表示されます。ONTAP では、10TB を作成されたメンバーボリュームの数で割った値が表示されます。ほとんどの場合、特に ONTAP 9.8 でのプロアクティブなサイズ変更では、このことを考慮する必要はありません。ただし、FlexGroup ボリュームが小さい場合やファイルが大きい場合は、これらの計算がより重要になります。

FlexGroup の総容量は、ONTAP System Manager、Active IQ Unified Manager、または admin 権限レベルの CLI で表示できます。

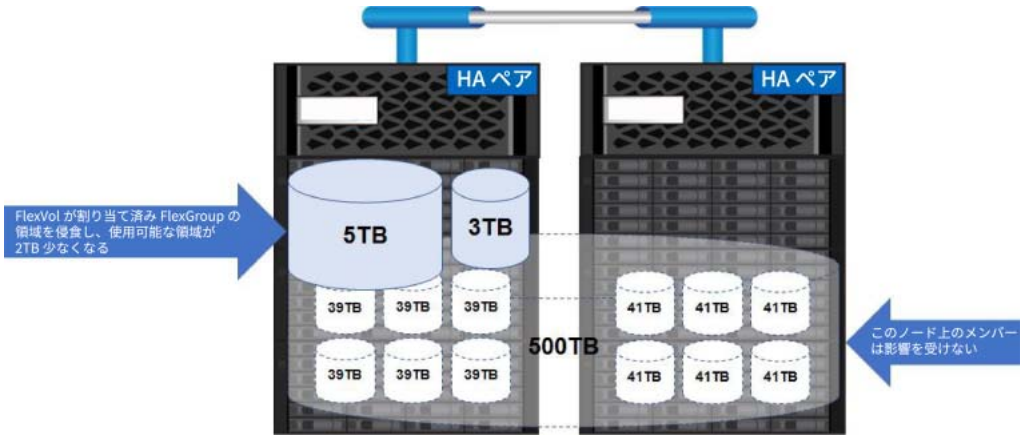
8.2.2 FlexGroup ボリュームでのオーバープロビジョニングまたはシンプロビジョニング

FlexGroup ボリュームでの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は、容量管理を最小限に抑えたいシナリオで役立ちます。データが書き込まれるまで実際の容量を消費せず、アグリゲート内の物理的に利用可能な領域に依存する、大規模な FlexGroup ボリュームを作成できます。

シンプロビジョニングは、次の点に注意して使用する必要があります。

- ボリュームが領域不足になると、物理容量が使用されているため、実際にスペース不足になります。同じノード上か、クラスタ内に新しいノードを追加するかにかかわらず、領域の問題を解決するには、ディスク容量を追加する必要があります。
- 割り当てられたスペースは、必ずしも実際に使用可能なスペースを反映しているわけではありません。物理スペースよりもはるかに大きいボリュームを割り当てることができます。
- システムが、シックプロビジョニングされたほかのボリューム (FlexGroup ボリュームまたは FlexVol ボリューム) とアグリゲートを共有している場合は、シンプロビジョニングの使用には注意が必要です。FlexGroup ボリュームと同じアグリゲート上の既存の FlexVol ボリューム、または FlexGroup ボリュームは、データの取り込み方法に影響を与える可能性があります。既存のボリュームでは、使用されている領域が他のボリューム割り当てを侵食するため、メンバーボリュームに使用できる領域が減少します。

図 8.7 シンプロビジョニングされた FlexGroup ボリュームが、スペース保証された FlexVol ボリュームとともに存在する場合の容量の影響



ボリューム容量の保証を使用すると、ボリューム容量に影響を与える他のデータセットから保護できますが、ストレージ容量を最も効率的に使用することはできません。

ベストプラクティス 13: シンプロビジョニングの使用
ONTAP でシンプロビジョニングを使用する場合は、Active IQ Unified Manager などのツールを使用するか、CLI を使用して監視をセットアップし、ストレージシステム内の使用可能な領域を追跡することが重要です。

8.3 FlexGroup ボリュームへの容量の追加

FlexGroup ボリュームの容量は非常に大きくなる可能性がありますが、データの増加に伴い、FlexGroup ボリュームのような大容量のコンテナであっても、より多くの容量が必要になる場合があります。

メンバーボリューム間のリモートハードリンク作成の増加により、メンバーボリュームの容量が満杯に近づくと、FlexGroup のパフォーマンスが悪影響を受ける場合があります。容量、空き領域、および FlexGroup ボリュームへの影響の詳細については、「7.10 容量に関する考慮事項」を参照してください。

そのため、FlexGroup メンバーボリューム内に少なくとも 10% の空き容量を維持することをお勧めします。当然のことながら、この 10% という数字は、メンバーボリュームのサイズとファイルサイズによって意味が異なります。したがって、このドキュメントの、サイズの大きいファイルまたはファイルが増加するボリュームのプロビジョニングに関する推奨事項に従ってください。

備考

ONTAP 9.8 以降では、プロアクティブなサイズ変更やその他の機能により、メンバーボリューム容量の管理オーバーヘッドが大幅に削減されるため、FlexGroup ボリュームを使用する場合は、可能であればそのリリースに移行してください。

8.3.1 容量の追加に関する推奨事項

FlexGroup ボリュームに容量を追加するには、主に次の 2 つの方法があります。推奨順に示します。

- `volume size` コマンドを使用して、既存のメンバーボリュームを拡張します。
- `volume expand` を使用して新しいメンバーボリュームを追加します。

既存の FlexGroup ボリュームに容量を追加する場合は、FlexGroup ボリュームを拡張することをお勧めします。FlexGroup に書き込まれたデータは静的であり、FlexGroup に配置された後は他のメンバーボリュームに再配分されないため、FlexGroup をインプレースで拡張する方が、ワークロードに新しいメンバーボリュームを追加するよりも、一貫したレベルのパフォーマンスと容量配分を維持できます。

物理的なアグリゲートの制限、またはメンバー FlexVol ボリュームが 100TB の制限に近づいているためにこの方法を使用できない場合、またはクラスタに新しいノードを追加してワークロードをさらに分散させたい場合は、代わりに `volume expand` を使用して新しいメンバーボリュームを追加する必要があります。

FlexGroup で使用する新しいノードまたはアグリゲートを追加する場合は、最初に `volume move` を使用してノード間でメンバーボリュームのバランスを調整してから、ボリュームを拡張するか、ノードごとに同じ倍数で新しいメンバーボリュームを追加します。ボリューム移動アプローチの視覚化については、「8.3.3 ディスク、アグリゲート、およびノードの追加」を参照してください。

ベストプラクティス 14: ボリュームサイズの増加

- 可能な場合は、`volume size` を使用して容量を増やすか、新しいメンバーを追加するのではなく ONTAP System Manager からサイズを変更します。この方法では、既存の FlexGroup データバランスが維持されます。
- FlexGroup ボリュームで使用する新しいノードまたはアグリゲートを追加する場合は、`volume move` を使用して新しいハードウェア間でメンバーを再配分してから、`volume size` を使用して FlexGroup を拡張するか、`volume expand` を使用してノードごとに同じ倍数で新しいメンバーを追加します。たとえば、各ノードに 4 つのメンバーがある場合、ノードごとに 4 つの新しいメンバーを追加します。
- 当社サポート部門のガイダンスなしで FlexGroup メンバーボリュームに対して `volume size` コマンドを個別に実行することは避けてください。FlexGroup ボリューム自体でのみ `volume size` を実行します。
- ONTAP 9.8 以降にアップグレードすると、プロアクティブなサイズ変更のメリットが得られ、ボリュームの自動サイズ設定を使用して容量を手動で管理する必要がなくなります。
- 容量の監視を使用して、メンバーボリュームがどのように満杯になるかを追跡します。問題に対処するための十分な時間を得るために、メンバーボリュームの合計容量に基づいてしきい値の割合で警告を設定します (たとえば、80% は 10TB までは問題ありませんが、100GB では問題があります)。
- シンプロビジョニングを使用して、総容量の割り当てに影響を与えずに、ボリューム上の容量の仮想上限を高く設定します。

備考

現在のところ、メンバーボリュームを削除して容量を削減する方法はありません。ボリュームの縮小のみが許可されます。

8.3.1.1 ボリュームの拡張と新しいメンバーの追加

FlexGroup ボリュームで容量やファイル数の増加が必要な場合は、主に 2 つの方法があります。

FlexGroup ボリュームの拡張 (ボリュームサイズ)

FlexGroup ボリュームの合計サイズを増やすことで、既存のメンバーボリュームに容量を追加できます。これは、CLI で `volume size` コマンドを使用するか、ONTAP System Manager で行います。追加されたサイズは、FlexGroup ボリューム内のメンバーボリューム間で均等に分割されます。たとえば、8 つのメンバーボリュームを持つ FlexGroup ボリュームに 8TB を追加する場合、各メンバーボリュームは 1TB 増加します。したがって、領域を追加するときは、FlexGroup ボリューム内のメンバーボリュームの数を把握しておくことが重要です。この数を調べるには、CLI から `volume show -name [flexgroup] -is-constituent true` を使用します。

FlexGroup ボリュームを拡張するか、メンバーボリュームを追加するかを決定する際には、希望する結果と目的を考慮してください。次の場合は、FlexGroup ボリュームを拡張します。

- 既存のノードまたはアグリゲートの容量を増やしたい。
- クラスタ内の合計ボリューム数を増やしたくない。
- FlexGroup メンバーボリュームの容量が 100TB の限界に達していない。
- メンバーボリュームのファイル数が上限の 20 億に達していない。
- メンバーボリュームが現在存在する使用可能な物理スペースがある。
- メンバーボリューム間のデータバランスを保持する場合。

これらのシナリオは決して網羅的なものではありません。他にも、新しいメンバーボリュームを追加するのではなく、FlexGroup ボリュームを拡張したい場合があります。不明な場合は、当社サポート部門までお問い合わせください。

メンバーボリュームの追加 (ボリューム拡張)

FlexGroup ボリュームに容量やファイル数を追加するもう 1 つの方法は、メンバーボリュームを追加することです。現在、正式に最大 200 のメンバー FlexGroup ボリュームがサポートされており、最大容量は 20PB、ファイル数は 4000 億です。それ以上の容量またはメンバーボリューム数が必要な場合は、担当営業に連絡して、より大容量の FlexGroup ボリュームの認定プロセスを開始してください。

現在、メンバーボリュームをさらに追加するには、CLI で `volume expand` コマンドを使用する必要があります。このコマンドは、既存の FlexGroup メンバーボリュームとまったく同じサイズの新しい空のメンバーボリュームを追加します。新しいメンバーボリュームの数は、`volume expand` コマンドの `aggr-multiplier` および `aggr-list` オプションによって決まります。このコマンドの例については、「18.8 FlexGroup ボリュームの拡張例」を参照してください。

ほとんどの場合、新しいメンバーボリュームを追加することは、容量やファイル数を追加するための適切な方法ではありません。新しいメンバーボリュームは空であり、新しい要求の取り込みバランスが失われる可能性があるためです。新しいメンバーボリュームを追加する場合は、必ず複数 (システムにすでに存在するボリュームと同じ数) で追加してください。

ただし、一部のユースケースでは、FlexGroup ボリュームに容量を追加する最善の方法は、メンバーボリュームを追加することです。例 :

- FlexGroup のメンバーボリュームは、すでに 100TB またはそれに近い容量です。
- 新しいノードまたはアグリゲートがクラスタに追加されています。
- さらに maxfiles が必要であり、メンバーボリュームはすでにそのサイズの最大値になっています。
- クラスター容量は、メンバーボリュームを拡張できないレベルにあり、クラスター内の他のノードには、同じ容量のメンバーボリュームに対して十分なスペースしかありません。

これらのシナリオは決して網羅的なものではありません。他にも、ボリュームを拡張するのではなく、FlexGroup ボリュームにメンバーを追加する場合があります。不明な場合は、当社サポート部門までお問い合わせください。

ボリュームのサイズ変更

必要に応じてボリューム容量を拡大または縮小するには、FlexGroup ボリュームで `volume size` コマンドを実行するか、ONTAP System Manager を使用します。容量を追加すると、インジェストヒューリスティックスに好影響を与えます。これは、メンバーボリュームの空き容量が増えるため、ONTAP では、パフォーマンスを向上させるために、親フォルダへのファイルのローカル配置が優先されるためです。

このコマンドを使用すると、メンバー FlexVol ボリュームは、合計容量 / メンバーボリュームの合計数だけ増加します。たとえば、8 個のメンバーボリュームを持つ FlexGroup ボリュームが 80TB 増加すると、各メンバーボリュームは ONTAP によって自動的に 10TB 増加します。

ONTAP 9.8 より前のリリースでは、FlexGroup の合計サイズの増加を考慮に入れる場合、これらの個別の増加を考慮することが重要でした。ボリュームの自動拡張を使用したプロアクティブなサイズ変更では、個々のメンバーのボリュームサイズを考慮する重要性が低くなります。これは、ONTAP が使用可能な空きスペースを管理するためです。

8.3.1.2 ONTAP System Manager からのボリュームのサイズ変更

ONTAP System Manager 9.7 以降でボリュームのサイズを変更するには、サイズを変更するボリュームを選択し、[Edit] をクリックします。次に、新しいサイズを入力します。

8.3.2 ボリューム拡張

FlexGroup ボリュームは、ボリュームサイズを指定して無停止で拡張できます。また、admin 権限レベルで利用できる volume expand コマンドを使用して、容量を動的に追加することもできます。このコマンドは、FlexGroup ボリュームにさらに FlexVol メンバーボリュームを追加します。次のいずれかが発生した場合に使用します。

- 既存のメンバーボリュームが最大容量 (100TB または 20 億ファイル) に達している
- ノード容量の物理的な制限に達しており、クラスタにさらにアグリゲートまたはノードが追加されている

備考

FlexGroup ボリュームに容量やファイル数を追加するだけの場合は、既存のボリュームのサイズを変更するか、maxfiles の値を増やしてから、ボリューム拡張によって新しいメンバーボリュームを追加します。

8.3.2.1 ボリューム拡張に関する考慮事項

FlexGroup ボリュームに新しいメンバーボリュームを追加する場合は、次の点を考慮する必要があります。

- 既存のボリュームサイズまたはファイル数を増やすことができない場合、または FlexGroup を新しいハードウェア全体に拡張する場合にのみ、volume expand を使用します。
- volume expand は現在、CLI 経由でのみ行われます。
- 新しいメンバーを追加する必要がある場合は、必ず既存の FlexGroup ボリュームと同じ倍数で追加してください。つまり、既存の FlexGroup ボリュームに 16 個のメンバーボリュームがあり、ノードごとに 8 個ある場合、一貫したパフォーマンスを向上させるために、新しいメンバーを 16 個 (ノードごとに 8 個) 追加します。
- 既存のクラスタに新しいノードを追加し、それらのノードに新しいメンバーを追加する場合は、既存の FlexGroup ボリュームと同じように、ノードごとのメンバーボリューム数を一定に保つようにしてください。
- FlexGroup ボリュームに新しいメンバーを追加すると、インジェストヒューリスティックスが変更され、新しいデータ用の新しい空のメンバーボリュームがより頻繁に使用されるようになります。また、新しいメンバーが既存のメンバーに追いつく間に、新しいデータ取り込みの全体的なシステムパフォーマンスに影響を与える可能性があります。
- メンバーボリュームを倍数で追加します。可能であれば、メンバーボリュームのワーキング・セットと同じにしてください。たとえば、8 つのメンバーボリュームがある場合、メンバーを追加するときに 8 つの新規メンバーボリュームを追加します。
- メンバーボリュームを追加すると、FlexVols がクラスタに追加され、プラットフォームで許可されている最大ボリューム数に対してカウントされます。
- FlexGroup ボリュームに新しいメンバーを追加すると、既存の Snapshot コピーと SnapMirror の関係は、ボリュームレベルの SnapRestore 操作では無効になりますが、以前のバージョンとスナップショットディレクトリへのアクセスを介したクライアント主導のファイルリストアには使用できます。データ類に固有の機能については、「ETERNUS AX/HX series ONTAP FlexGroup ボリュームのデータ保護とバックアップ」を参照してください。
- FlexGroup 内の既存のデータにはリバランスする方法はありませんが、必ずしも問題があるわけではありません。データの不均衡とその影響の詳細は、「6.4.5 FlexGroup ボリューム内のデータ不均衡」を参照してください。

volume expand を使用する場合は、「8.3.1 容量の追加に関する推奨事項」に記載されている指示に従ってください。

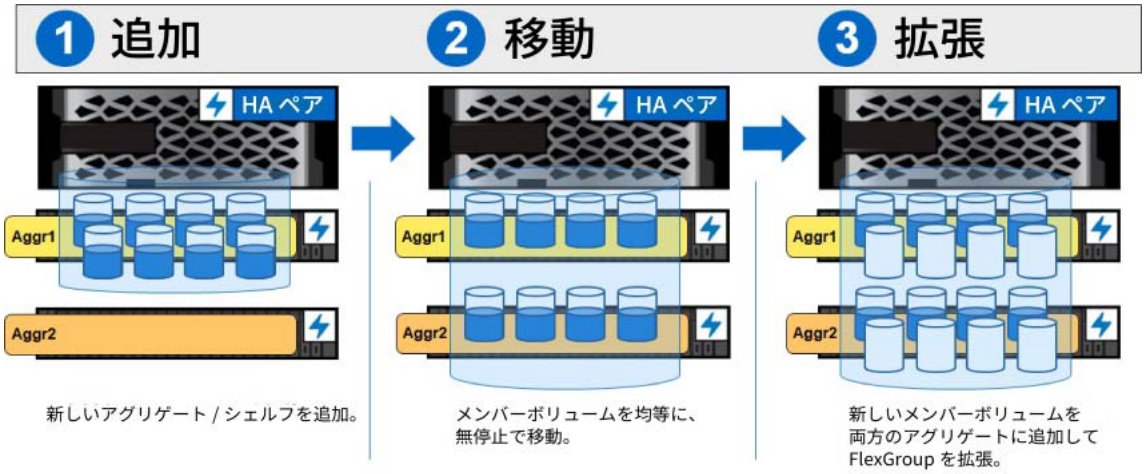
volume expand の例については、「18. コマンドの例」を参照してください。

8.3.3 ディスク、アグリゲート、およびノードの追加

FlexGroup メンバーボリュームを含む既存のアグリゲートにディスクを追加する場合は、ボリュームの合計サイズを増やす場合を除き、何もする必要はありません。

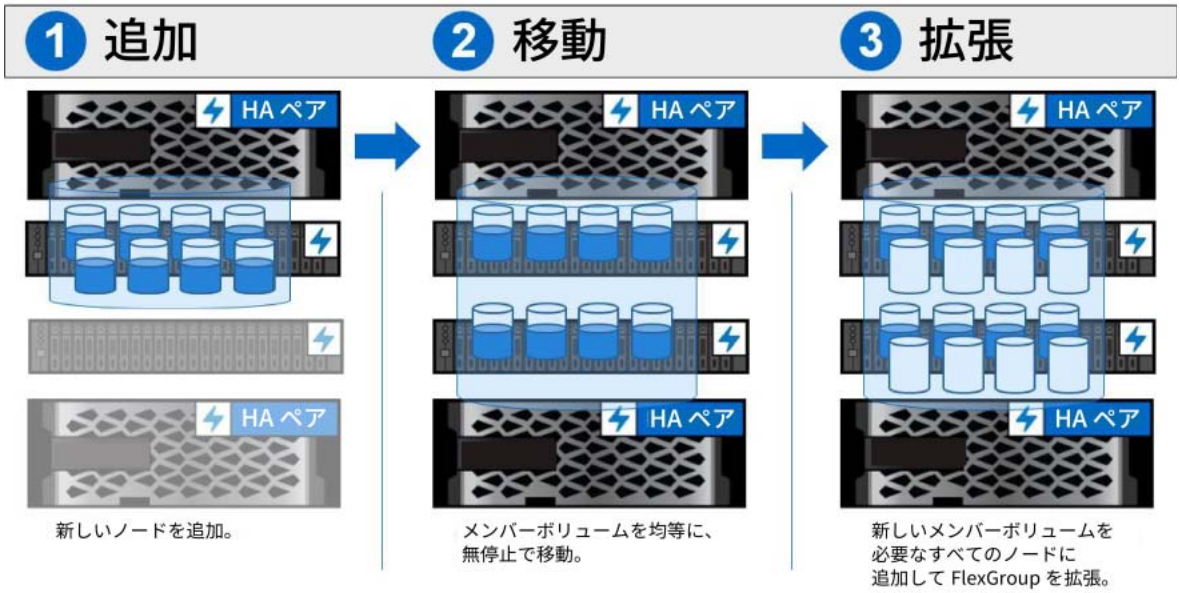
ノードにアグリゲートを追加するときに、FlexGroup ボリュームが新しいアグリゲートにまたがる必要がある場合は、メンテナンス時間を必要とせずに、無停止ボリューム移動を使用してメンバーボリュームを新しいアグリゲートに移動できます。次に、新しいアグリゲートと古いアグリゲートにまたがる FlexGroup ボリュームにメンバーボリュームを作成します。

図 8.8 FlexGroup ボリュームでのアグリゲートの追加



クラスタに新しいノードを追加するときは、クラスタにアグリゲートを追加するときと同じ手順に従います。`volume move` コマンドと `volume expand` コマンドを使用してメンバーボリュームを調整します。

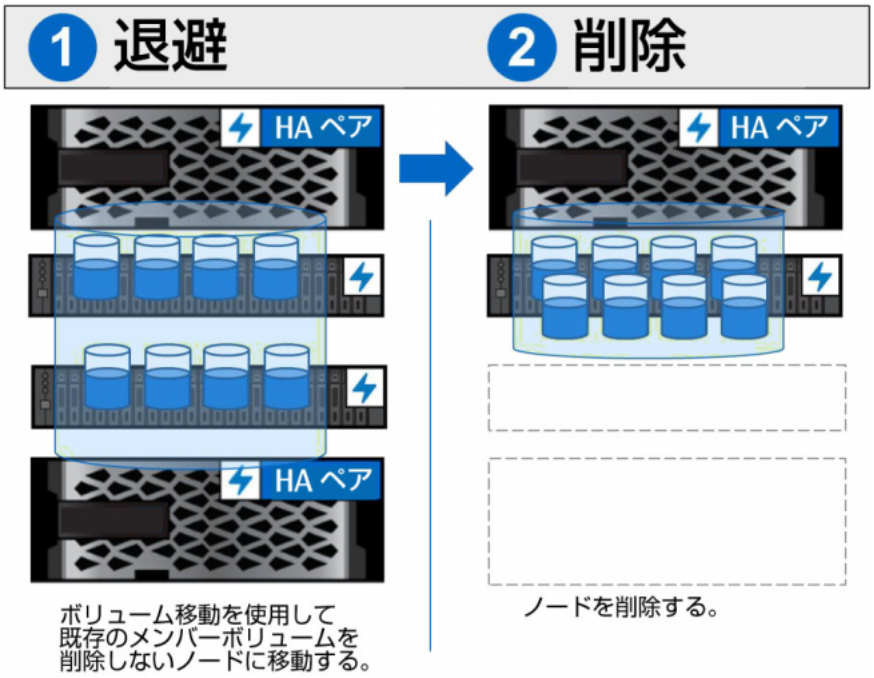
図 8.9 ノードの追加と FlexGroup ボリュームの拡張



8.3.4 クラスタからのノードの削除または退避

メンバーボリュームの無停止ボリューム移動を使用して、FlexGroup メンバーボリュームを含むクラスタからノードを削除することもできます。たとえば、8 つノードがあり各ノードに 16 個のメンバーボリュームがあるクラスタから 2 つのノードを削除する場合は、ボリューム移動を使用して、残りの 6 つのノードに 32 個のメンバーボリュームを分散します。32 個のボリュームは 6 つのノードに均等に分割されないため、次の分割可能なノード数を使用して均等に配分し、4 つのノードがノードごとに 8 個のメンバーボリュームを取得するようにします。4 つのノードがそれぞれ 8 つのメンバーボリュームを使用するのに十分なスペースがない場合は、2 つのノードに 6 つのメンバーボリュームを配置し (12)、4 つのノードに 5 つのメンバーボリュームを配置します (20)。

図 8.10 FlexGroup メンバーボリュームを含むノードの削除



備考

メンバーボリュームは、いったん追加すると FlexGroup ボリュームから削除できません。

8.4 クラスタからのノードの削除または退避

ONTAP を使用すると、同じクラスタ内のアグリゲート間またはノード間で、無停止でのボリューム移動を実行できます。この機能により、メンテナンスウィンドウで処理したり、クラスタ内のパフォーマンスや容量の割り当てのバランスをとろうとするときに柔軟に対応できます。

FlexGroup ボリュームもこの機能をサポートしていますが、さらに細分化されています。この機能を使用して、FlexGroup ボリューム内の各メンバーボリュームを移動できます。ボリュームを移動しても、1 回の移動で FlexGroup 全体が移動するわけではありません。

したがって、ストレージ管理者は、容量やパフォーマンスの問題が発生した場合に、クラスタ内でワークロードを移動します。Active IQ を使用して個々のメンバーの FlexVol ボリュームを確認すると、問題を迅速に特定して解決できます。

備考

ボリュームの移動は無停止で実行できますが、移動にかかる時間はボリュームのサイズとノードの全体的な負荷によって異なります。

8.4.1 無停止でのボリューム移動を使用する場合

FlexGroup では、次のような場合に無停止でのボリューム移動が便利です。

- メンバーボリュームが容量に近づいていますが、現在のノードには、そのボリュームを適切に拡張するための物理ストレージがありません。
- メンバーボリュームは、他の FlexVol ボリュームとアグリゲートを共有しており、FlexVol ボリュームのパフォーマンスまたは容量の影響を受けています。
- メンバーボリュームが FlexGroup ボリューム内で過負荷になっており、ノードリソースをさらに必要としている場合。
- パフォーマンスのために FlexGroup ボリュームをスピニングディスクから SSD に移行したり、アーカイブのために SSD からスピニングディスクに移行したりする場合。
- 新しいクラスタノードまたはデータアグリゲートが追加された場合。
- (ダウンタイムを最小限に抑えるために) ヘッド・スワップまたはその他の計画された保守操作を実行している場合。

8.4.2 無停止でのボリューム移動の使用

FlexGroup メンバーボリュームの無停止での `volume move` は、コマンドラインの admin 権限レベルで実行できます。ONTAP System Manager GUI を使用して FlexVol ボリュームを移動することはできますが、現時点では FlexGroup のメンバーボリュームを移動することはできません。

FlexGroup メンバーボリュームの移動は以下の手順で行います。

1. 移動するボリュームを特定します。この情報を確認するには、Active IQ Unified Manager または CLI を使用してください。
2. コマンド・ラインから、`volume move start` コマンドを実行します。このコマンドは、admin 権限レベルで実行できます。

```
cluster::> volume move start -vserver SVM -destination-aggregate aggr1_node2 -volume
flexgroup4TB__000
flexgroup4TB__0001 flexgroup4TB__0002 flexgroup4TB__0003 flexgroup4TB__0004
flexgroup4TB__0005 flexgroup4TB__0006 flexgroup4TB__0007 flexgroup4TB__0008

cluster::> volume move start -vserver SVM -volume flexgroup4TB__0003
-destination-aggregate aggr1_node2
[Job 2603] Job is queued: Move "flexgroup4TB__0003" in Vserver "SVM" to aggregate
"aggr1_node2". Use the "volume move show -vserver SVM -volume flexgroup4TB__0003"
command to view the status of this operation.

cluster::> volume move show
Vserver   Volume     State    Move Phase Percent-Complete Time-To-Complete
--------- ---------- -------- ---------- ---------------- ----------------
SVM       flexgroup4TB__0003
                      healthy replicating
                                            45%        Tue Dec 06 13:43:01 2016
```

8.4.2.1 自動バランスアグリゲート

自動バランスアグリゲート機能は、システムのパフォーマンスまたは容量がストレージ管理者によって指定されたポイントに達したときに、ONTAP が推奨する無停止ボリューム移動を提供します。この機能は、FlexGroup ボリュームでは現在サポートされていません。

8.5 FlexGroup ボリュームを削除する際の考慮事項

ボリュームリカバリキューと呼ばれる機能がサポートされており、削除されたボリュームのリカバリキューを 12 時間保持することで、ボリュームの誤った削除を防ぎます。ボリュームはクライアントからアクセスできなくなり、admin 権限レベルでは管理者ビューから非表示になりますが、領域は割り当てられたままであり、緊急時のリカバリが必要な場合に備えてボリュームの残りが残ります。FlexGroup ボリュームもこのリカバリキューを使用するため、リカバリキューが期限切れになるか、手動でパージされるまで領域は解放されません。

diag 権限レベルで `-type DEL` を指定すると、コマンドラインから削除されたボリュームを表示できます。GUI からは、ボリュームもリカバリキューも表示されません。

```
cluster::*> volume show -vserver DEMO -type DEL
Vserver   Volume       Aggregate    State      Type       Size Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      flexgroup__0001_2321
                       aggr1_node1 offline     DEL         5TB      -      -
DEMO      flexgroup__0002_2322
                       aggr1_node2 offline     DEL         5TB      -      -
```

削除されたボリュームは、volume recovery-queue コマンドや diag 権限で表示することもできます。

```
cluster::*> volume recovery-queue show
Vserver   Volume      Deletion Request Time    Retention Hours
--------- ----------- ------------------------ ---------------
DEMO      flexgroup__0001_2321
                      Tue May 01 17:14:14 2018              12
DEMO      flexgroup__0002_2322
                      Tue May 01 17:14:13 2018              12
2 entries were displayed.
```

リカバリキューからボリュームを手動でパージするには、次のコマンドを実行します。

```
cluster::*> volume recovery-queue purge -vserver DEMO -volume flexgroup__0001_2321 Queued
private job: 4660

cluster::*> volume recovery-queue purge -vserver DEMO -volume flexgroup__0002_2322 Queued
private job: 4661
```

ボリュームの削除時にリカバリキューをバイパスするには、CLI から `-force true flug` を使用し、高度な権限レベルで `volume delete` コマンドを実行します。ONTAP System Manager は、強制削除やボリュームリカバリキューの管理をサポートしていません。

8.6 ボリューム名の変更に関する考慮事項

FlexGroup によるボリューム名の変更がサポートされています。FlexGroup ボリューム名は、ストレージ管理者がクラスタ内のボリュームを識別するためのものです。ボリュームのクライアント側の名前は、クラスタ内のボリューム名の付け方ではなく、NFS の CIFS/SMB 共有およびボリュームジャンクション・パス (エクスポートパス) を介して公開されます。

たとえば、ボリュームは `vol1` という管理名を持つことができますが、`/accounting` としてクライアントにエクスポートされます。ジャンクションパスと SMB 共有は FlexGroup ボリューム内でいつでも変更できますが、変更すると、クライアントは再マウントまたは SMB 共有再接続によって新しい共有またはエクスポートパスに再接続する必要があるため、クライアントの動作が中断します。ただし、ボリューム名を変更しても、クライアント側での中断は発生しません。

9. qtree

ONTAP では、qtree と呼ばれる論理ディレクトリに対する FlexGroup ボリュームをサポートします。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、独自のエクスポートポリシー、独自のセキュリティスタイル、詳細な統計情報を有効にすることで、データ管理に柔軟性を提供します。

qtree には複数のユースケースがあり、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便利です。

FlexGroup ボリューム内の qtree に関する詳細は、次の項目を参照してください。

- qtree はクライアントに対してディレクトリとして表示されます。
- qtree はボリュームレベルで作成可能です。現在、ディレクトリの下に qtree を作成してサブディレクトリである qtree を作成することはできません。
- 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。
- SnapMirror を使用して qtree をレプリケートすることはできません。SnapMirror は現在、ボリュームレベルでのみ実行されます。FlexGroup を使用してより詳細なレプリケーションを実行する場合は、FlexGroup ボリュームとジャンクションパスを組み合わせて使用します。
- FlexGroup ボリュームごとに最大 4,995 の qtree がサポートされます。クォータの監視と適用は、qtree またはユーザーレベルで適用できます。

備考

ONTAP 9.8 では qtree QoS サポートが追加されました。

9.1 qtree とファイル移動

qtree は ONTAP では一意のファイルシステムとみなされます。NAS クライアントから見るとディレクトリのように見えますが、実際のディレクトリとは動作が異なる操作もあります。その一例は、同じボリューム内の qtree 間でファイルを移動することです。

複数のディレクトリにまたがるボリューム内でファイルの移動が行われると、ファイルは単に新しい名前に変更され、数秒以内に実行されます。これは、同じファイルシステム内での移動であるためです。

2 つの qtree 間でファイル移動が発生した場合、ファイルは名前変更されずに新しい場所にコピーされます。このため、操作に時間がかかります。

これは、qtree が FlexVol に存在するか FlexGroup に存在するかに関係なく発生する動作です。

9.1.1 qtree ID と名前変更の動作

継承されないエクスポート・ポリシーが qtree に適用されると、NFS ファイルハンドルは、qtree 間の操作を扱う際に若干変更されます。ONTAP は NFS 操作で qtree ID を検証します。これは、ソースフォルダーまたは qtree と同じボリューム内の qtree との間での移動時にファイルの名前変更や移動などに影響します。これはセキュリティ機能と考えられており、ホームディレクトリ内のシナリオのように、qtree を越えた不要なアクセスを防ぐのに役立つ。ただし、エクスポートポリシー規則とパーミッションを適用するだけでも、同様の目的を達成できます。

たとえば、同じボリューム内の qtree に対して移動または名前変更を行うと、「アクセスが拒否されました」というエラーが発生します。別のボリュームの qtree に対して同じ移動または名前変更を行うと、ファイルがコピーされます。ファイルが大きくなると、コピー動作で移動操作に非常に長い時間がかかるように見えることがあります。同じファイルシステムとボリューム内でファイル名を変更するだけなので、ほとんどの移動操作はほぼ瞬時に行われます。

この動作は高度な権限オプションによって制御されます。

これらはさまざまなオペレーションの動作です。

```
Assuming that file permissions allow and that client is allowed by export policies to access
both source and destination volume/qtree, these are the current permutations with the
'validate-qtree- export' flag enabled or disabled:

Enabled:
- Rename in same volume and qtree: SUCCESS
- Rename in same volume, different qtrees: EACCESS
- Rename between volumes where qtree IDs differ: EACCESS
- Rename between volumes where qtree IDs match: XDEV

Disabled:
- Rename in same volume and qtree: SUCCESS
- Rename in same volume, different qtrees: SUCCESS
- Rename between volumes where qtree IDs differ: XDEV
- Rename between volumes where qtree IDs match: XDEV
```

備考

NFS3ERR_XDEV および NFS3ERR_ACCESS は、RFC-1813 で定義されています。

qtree 間の名前変更 / 移動の動作を変更するには、`-validate-qtree-export` を `disabled` に変更します。詳細は、ETERNUS AX/HX Series NFS リファレンス「qtree ファイル操作の qtree ID の検証」を参照してください。

備考

`-validate-qtree-export` オプションを無効にしても、qtree 間での名前変更を許可しない限り、悪影響はありません。

9.1.2 qtree エクスポートのファイルハンドルの効果

通常、クライアントに渡される NFS エクスポートファイルハンドルのサイズは 32 ビット以下です。ただし、qtree のエクスポートでは、40 ビットのファイルハンドルを作成するために、さらにいくつかのビットが追加されます。ほとんどのクライアントではこれは問題になりませんが、古いクライアント (1996 年に発売された HPUX 10.20 など) では、これらのエクスポートのマウントで問題が発生する場合があります。qtree エクスポートを有効にする前に、別のテスト SVM で古いクライアント接続をテストしてください。qtree エクスポートを有効にした後でファイルハンドルの動作を変更する方法は現在ありません。

9.2 FlexGroup によるクォータの管理

FlexGroup ボリュームは、ユーザー / グループおよびツリークォータをサポートします。これらのサポートのレベルは、次のように分類できます。

- ONTAP 9.3 でのクォータレポートのサポート。
- ONTAP 9.4 では、FPolicy をサポートし、DefendX (以前は NTP) などのサードパーティベンダーからクォータの適用を提供できます。
- クォータの適用 (つまり、容量とファイル数のハード制限とソフト制限を設定します) は、ONTAP 9.5 以降でサポートされています。

9.2.1 ユーザー / グループのクォータに関する考慮事項

ユーザーまたはグループのクォータを実装するには、指定されたユーザー名またはグループをクラスタが解決できる必要があります。この要件は、ユーザーまたはグループが SVM 上にローカルに存在するか、Active Directory、LDAP、NIS などの解決可能なネームサービスサーバー内に存在する必要があることを意味します。SVM でユーザーまたはグループが見つからない場合、クォータの規則は作成されません。無効なユーザーが原因でユーザーまたはグループのクォータの作成に失敗した場合、コマンドラインは次のエラーを発行します。

```
Error: command failed: User name user not found. Reason: SecD Error: object not found.
```

ONTAP System Manager でも同様のメッセージが表示されます。問題をさらに調査するには、`event log show` コマンドを使用します。ONTAP での識別管理用のネーム・サービスの構成の詳細は、「ETERNUS AX/HX series ONTAP で LDAP を設定する方法 マルチプロトコル NAS ID 管理」を参照してください。

9.2.2 ユーザーまたはグループのクォータの作成

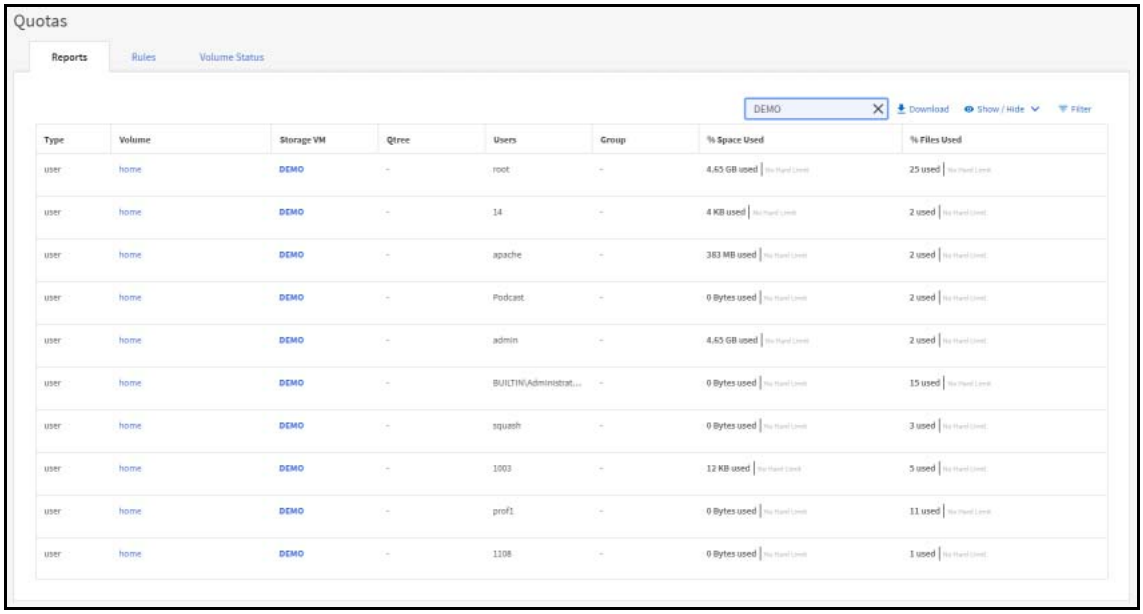
ユーザー / グループのクォータを作成して、容量またはファイル数の制限をユーザーごとに報告または適用できます。これらのクォータは、複数のユーザーまたはグループが同じ名前空間または qtree を共有するシナリオで使用されます。これらの手順は、FlexVol ボリュームと FlexGroup ボリュームでも同じです。

9.2.2.1 クォータの作成 –ONTAP System Manager

ONTAP System Manager でユーザーまたはグループのクォータを作成するには、左側のメニューから [Storage] > [Quotas] に移動します。次の 3 つのタブがあるページに移動します。[Reports]、[Rules]、[Volume Status] です。

レポートには、ユーザー、グループ、qtree の現在のクォータ追跡が表示されます。

図 9.1 クォータ・レポート：ONTAP System Manager

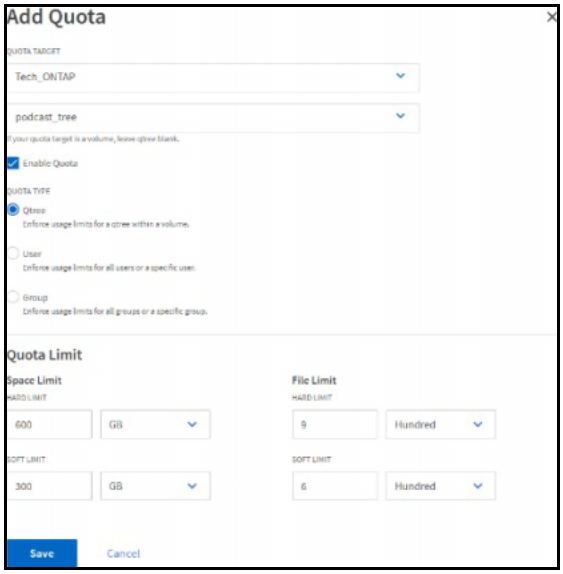


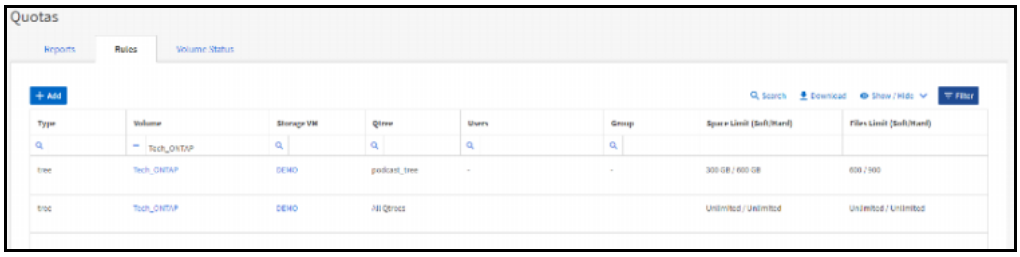
ボリュームの状態は、ボリュームのクォータがオンかオフかを示します。

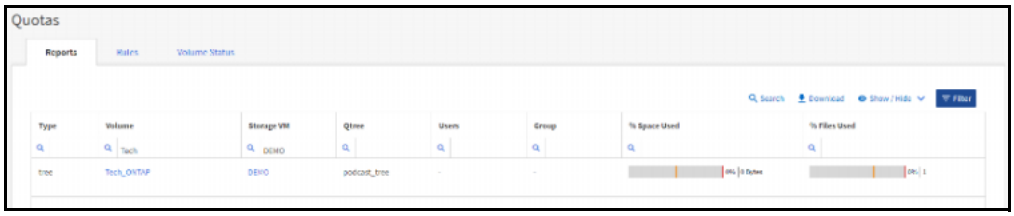
図 9.2　クォータのボリュームステータス：ONTAP System Manager

ルールは、ユーザー、グループ、または qtree の新しいクォータを作成する場所です。[Add] をクリックし、ダイアログ・ボックス画面でユーザー、グループ、qtree のクォータに関する情報を入力します。ルールが作成されると、ONTAP System Manager は必要なすべての手順を実行して、クォータを有効化およびアクティブ化します。

図 9.3　クォータ・ルール：ONTAP System Manager







9.2.2.2 ユーザーまたはグループのクォータの作成 : コマンドライン

コマンドラインで特定のユーザーまたはグループのクォータレポートを作成するには、admin 権限レベルで次のコマンドを使用します。

```
cluster::> quota policy rule create -vserver SVM1 -policy-name default -volume flexgroup
-type [user|group] -target [username or groupname] -qtree ""
```

すべてのユーザーまたはグループに対して、コマンドラインを使用してユーザーまたはグループのクォータレポートを作成するには、admin 権限レベルで次のコマンドを使用します。ターゲットは、all を示すアスタリスクとして入力します。

```
cluster::> quota policy rule create -vserver SVM1 -policy-name default -volume flexgroup
-type [user|group] -target * -qtree ""
```

9.2.3 コマンドラインからのツリーレポートクォータの作成

コマンドラインで特定のユーザーまたはグループのツリークォータレポートを作成するには、admin 権限レベルで次のコマンドを使用します。

```
cluster::> quota policy rule create -vserver DEMO -policy-name tree -volume flexgroup_local
-type tree -target qtree
```

クォータを有効にするには、quota on または quota resize を使用します。

```
cluster::> quota on -vserver DEMO -volume flexgroup_local
[Job 9152] Job is queued: "quota on" performed for quota policy "tree" on volume
"flexgroup_local" in Vserver "DEMO".

cluster::> quota resize -vserver DEMO -volume flexgroup_local
[Job 9153] Job is queued: "quota resize" performed for quota policy "tree" on volume
"flexgroup_local" in Vserver "DEMO".

cluster::> quota show -vserver DEMO -volume flexgroup_local

              Vserver Name: DEMO
               Volume Name: flexgroup_local
               Quota State: on
               Scan Status: -
          Logging Messages: -
          Logging Interval: -
          Sub Quota Status: none
  Last Quota Error Message: -
Collection of Quota Errors: -
       User Quota enforced: false
      Group Quota enforced: false
       Tree Quota enforced: true
```

次の例は、ツリークォータが指定された FlexGroup ボリュームに対する quota report コマンドを示しています。

```
cluster::> quota report -vserver DEMO -volume flexgroup_local
Vserver: DEMO
                                ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
flexgroup_local
        qtree    tree   1       0B        -       1      -  qtree
```

使用されているファイルとディスク領域は監視され、新しいファイルが作成されるたびに増加します。

```
cluster::> quota report -vserver DEMO -volume flexgroup_local
Vserver: DEMO
                                ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
flexgroup_local
        qtree    tree   1
                                13.77MB   -       4      -  qtree
```

9.2.4 クォータの適用の例

qtree またはユーザー / グループに対してクォータの適用が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。これにより、ストレージ管理者はボリュームまたは qtree に書き込まれるデータの量をより詳細に制御できるようになります。

さらに、クォータを超過すると、イベント管理システムメッセージが DEBUG 重大度レベルで記録され、ストレージ管理者にクォータ違反が通知されます。これらのメッセージは、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。

この例では、クォータに 1GB と 10 ファイルのハード制限が設定されています。

```
cluster::*> quota policy rule show -vserver DEMO
Vserver: DEMO             Policy: tree           Volume: flexgroup_local

                                          Soft          Soft
                       User        Disk    Disk  Files   Files
Type  Target   Qtree   Mapping    Limit   Limit  Limit   Limit Threshold
----- -------- ------- ------- -------- ------- ------ ------- ---------
tree  qtree    ""      -            1GB       -     10       -         -
```

ユーザーが 1.2GB のファイルを qtree にコピーしようとしたときに、ONTAP から out of space(領域不足) エラーが報告されます。

```
[root@centos7 qtree]# cp /SANscreenServer-x64-7.3.1-444.msi /FGlocal/qtree/
cp: failed to close '/FGlocal/qtree/SANscreenServer-x64-7.3.1-444.msi': No space left on
device
```

ファイルは部分的に書き込まれていますが、データが欠落しているため使用できません。

```
# ls -alh total 1.1G
drwxr-xr-x 2 root root 4.0K Jul 19 15:44 .
drwxr-xr-x 11 root root 4.0K Jun 28 15:10 ..
-rw-r--r--  1 root root0 Dec 12 2017 newfile1
-rw-r--r--  1 root root0 Dec 12 2017 newfile2
-rw-r--r-- 1 root root 1021M Jul 19 2018 SANscreenServer-x64-7.3.1-444.msi
```

ONTAP は、クォータが超過したことを報告します。

```
cluster::*> quota report -vserver DEMO
Vserver: DEMO
                                  ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
flexgroup_local
        qtree    tree   1
                                  1.01GB  1GB       5     10  qtree
```

ファイル数の制限についても同様です。この例では、ファイル数の制限は 10 で、qtree にはすでに 5 つのファイルがあります。余分の 5 つのファイルが制限に達しています。

```
[root@centos7 /]# su student1
sh-4.2$ cd ~
sh-4.2$ pwd
/home/student1
sh-4.2$ touch file1
sh-4.2$ touch file2
sh-4.2$ touch file3
sh-4.2$ touch file4
sh-4.2$ touch file5
touch: cannot touch ‘file5’: Disk quota exceeded

cluster::*> quota report -vserver DEMO
Vserver: DEMO
                                  ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
flexgroup_local
        qtree    tree   1
                                  1.01GB  1GB       5     10  qtree
home             user   student1, NTAP\student1
                                  4KB     1GB      10     10  student1
2 entries were displayed.
```

イベント・ログから、クォータ違反を確認できます。

```
cluster::*> event log show -message-name quota.exceeded
Time                Node             Severity      Event
------------------- ---------------- ------------- ---------------------------
7/19/2018 16:27:54  node02
                                     DEBUG         quota.exceeded: ltype="hard",
volname="home", app="", volident="@vserver:7e3cc08e-d9b3-11e6-85e2-00a0986b1210",
limit_item="file", limit_value="10", user="uid=1301", qtree="treeid=1", vfiler=""
7/19/2018 15:45:02  node01
                                     DEBUG         quota.exceeded: ltype="hard",
volname="flexgroup_local", app="", volident="@vserver:7e3cc08e-d9b3-11e6-85e2-
00a0986b1210", limit_item="disk", limit_value="1048576", user="", qtree="treeid=1",
vfiler=""
```

64 ビットのファイル ID を無効のままにする場合は、FlexGroup ボリュームに許可されるファイル数の制御にクォータを使用できます。詳細は、「10.2.3 クォータの適用を使用してファイル数を制限する」を参照してください。

9.2.5 クォータの使用によるパフォーマンスへの影響

機能を有効にする場合、パフォーマンスへの影響が常に懸念されます。クォータを使用する際のパフォーマンスの問題を軽減するために、クォータを有効にした場合と無効にした場合の両方で、ONTAP 9.5 の FlexGroup ボリュームに対して標準の NAS ベンチマークテストを実行しました。図 9.4 と図 9.5 に示すように、FlexGroup ボリュームでクォータを有効にした場合のパフォーマンスへの影響は無視できると結論付けました。

図 9.4 ONTAP 9.5 のパフォーマンス (オペレーション / 秒)：クォータがオンとオフ

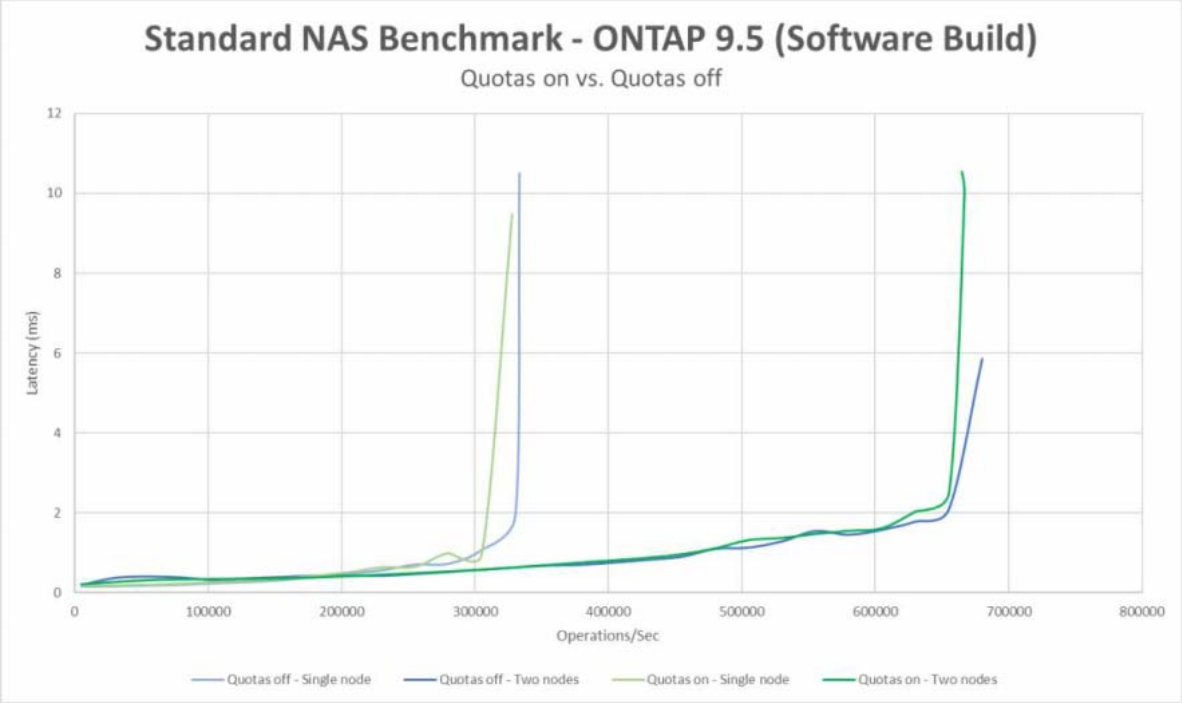
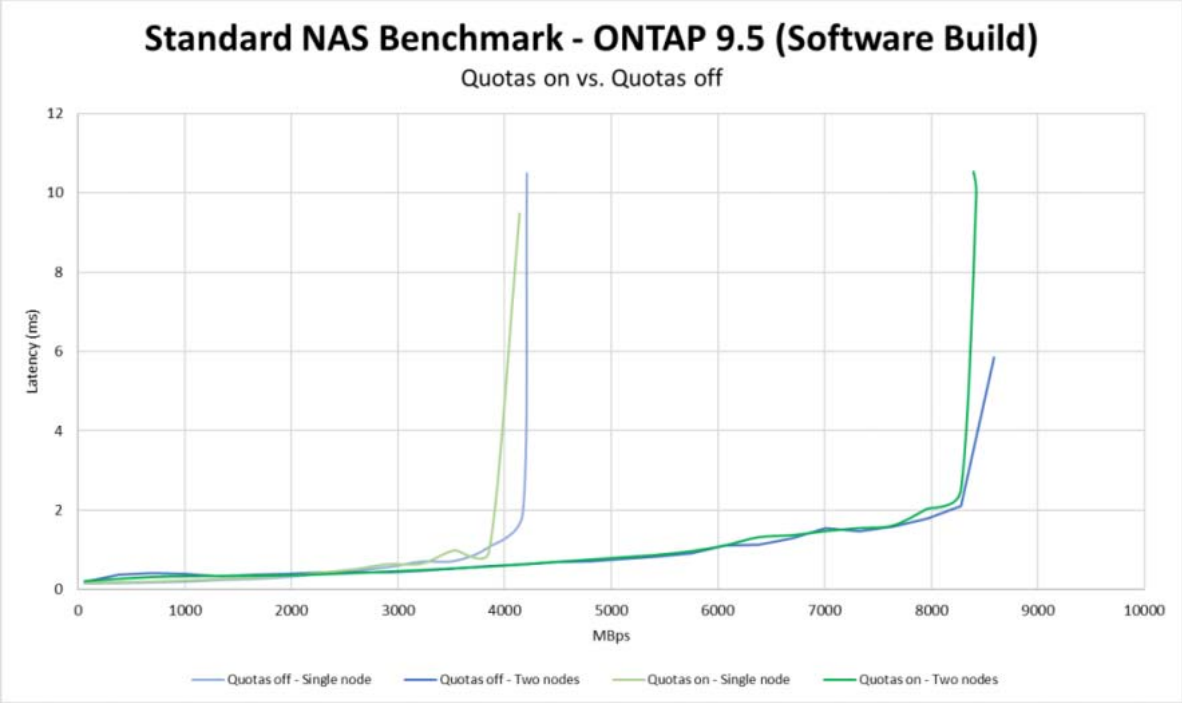


図 9.5 ONTAP 9.5 のパフォーマンス (MBps)：クォータがオンとオフ



9.2.6 クォータスキャンの完了時間

クォータの初期化またはサイズ変更が行われると、ONTAP はバックグラウンドタスクを実行して必要な作業を完了し、クォータの使用状況を正確に反映する必要があります。これらの作業には時間がかかりますが、これは以下で説明するいくつかの要因によって異なります。

9.2.6.1 初期化完了時間

ボリュームまたは qtree でのクォータの初期化にかかる時間は、次の要因によって異なります。

- ボリューム内のファイルとフォルダの数ファイル数が多いほど初期化に時間がかかりますが、ファイルサイズは初期化時間に影響しません。
- ボリュームのタイプ FlexGroup のクォータスキャンは FlexGroup が存在するノード間で並行して実行されるため、FlexVol スキャンは FlexGroup スキャンよりも時間がかかる場合があります。
- ハードウェアのタイプとシステムの負荷多くのファイルによる負荷が高いシステムでは、スキャンに数時間かかることがあります。

quota の初期化状態は、コマンド `quota show -volume volname - instance` で確認できます。

9.2.6.2 クォータサイズ変更の完了時間

クォータのサイズ変更は、クォータポリシーが変更されたときに使用されます。サイズ変更により、新しい制限でスキャンが実行されます。このプロセスにも、完了時間に関する考慮事項があります。

- サイズ変更は、新しく追加された規則を使用してスキャンするだけなので、初期化よりも速く完了します。
- サイズ変更は通常数秒で完了します。これは、クォータのオン / オフよりも少ない時間で実行する必要があるためです。
- サイズ変更は速く完了するため、クォータのオン / オフを切り替える代わりにサイズ変更を使用します。
- クォータのサイズ変更により、最大 100 個の同時ジョブを実行可能です。100 ジョブの後では、サイズ変更操作はキューで待機する必要があります。
- 同時スキャンが増えると、サイズ変更のパフォーマンスに影響し、ジョブの完了までの時間が長くなります。

9.2.7 クォータに関するユーザーマッピングの考慮事項

マルチプロトコル環境でのクォータのユーザーマッピング（SMB と NFS の両方からのデータアクセス）は、メンバーボリュームレベルで行われます。最終的には、すべてのメンバーボリュームがユーザーマッピングに同意します。ただし、ユーザーマッピングが失敗した場合や、別のメンバーで成功した名前マッピングを実行したときにタイムアウトになる場合など、不一致が発生することがあります。これは、少なくとも 1 つのメンバがユーザーをユーザーマップのペアの一部と見なし、少なくとも 1 つの他のメンバが個別のレコードと見なすことを意味します。

最悪の場合、クォータルールの適用は、問題が解決されるまで一貫性がありません。たとえば、ユーザーはクォータ制限を一時的に超過することができます。

ユーザーマッピングの結果が調整されると、イベント管理システムメッセージが送信されます。

```
cluster::*> event route show -message-name fg.quota.usermapping.result -instance

                          Message Name: fg.quota.usermapping.result
                              Severity: NOTICE
                     Corrective Action: (NONE)
                           Description: This message occurs when the quota mapper decides
whether to map the Windows quota record and the UNIX quota record of a user into a single
multiuser record.
```

9.2.8 ツリークォータに関する考慮事項

ONTAP の SVM は最大 5 つのクォータポリシーを持つことができますが、一度にアクティブにできるポリシーは 1 つだけです。SVM でアクティブなポリシーを表示するには、次のコマンドを使用します。

```
cluster::> vserver show -vserver DEMO -fields quota-policy
vserver quota-policy
------- ------------
DEMO    default
```

備考

現在、ONTAP System Manager でこの情報を表示することはできません。

ほとんどの場合、デフォルトのポリシーで十分であり、変更する必要はありません。quota on を発行すると、ボリュームに割り当てられたポリシーではなく、アクティブなポリシーが使用されます。したがって、ボリュームにクォータとルールを適用したが、quota on が失敗したと思われる状況に陥る可能性があります。

次の例では、ボリュームにクォータポリシーを適用します。

```
cluster::*> quota policy show -vserver DEMO -policy-name tree

      Vserver: DEMO
  Policy Name: tree
Last Modified: 10/19/2017 11:25:20
    Policy ID: 42949672962

cluster::*> quota policy rule show -vserver DEMO -policy-name tree -instance

                 Vserver: DEMO
             Policy Name: tree
             Volume Name: flexgroup_local
                    Type: tree
                  Target: tree1
              Qtree Name: ""
            User Mapping: -
              Disk Limit: -
             Files Limit: -
Threshold for Disk Limit: -
         Soft Disk Limit: -
        Soft Files Limit: -
```

SVM には default クォータが割り当てられており、ルールが含まれていないため、クォータをオンにするとエラーが発生します。

```
cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true

Error: command failed: No valid quota rules found in quota policy default for volume
flexgroup_local in Vserver DEMO.
```

ルールを default に追加すると、quota on コマンドは動作しますが、SVM は新しいツリーポリシーを使用しません。

```
cluster::*> quota policy rule create -vserver DEMO -policy-name default -volume
flexgroup_local - type tree -target ""

cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true
[Job 8063] Job succeeded: Successful

cluster::*> vserver show -vserver DEMO -fields quota-policy
vserver quota-policy
------- ------------
DEMO    default
```

必要なポリシーを使用するには、SVM を変更し、クォータをオフにしてからオンに戻す必要があります。

```
cluster::*> vserver modify -vserver DEMO -quota-policy tree

cluster::*> quota off -vserver DEMO *

cluster::*> quota policy rule delete -vserver DEMO -policy-name default *
1 entry was deleted.

cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true
[Job 8084] Job succeeded: Successful
```

この動作は、FlexGroup ボリュームに固有のものではありません。これは FlexVol ボリュームでも発生します。

9.2.9 クォータが有効なときにクライアントが領域を認識する方法

ONTAP 内の qtree に対してクォータが有効になっている場合、クライアントには、そのクォータによって報告された使用可能な領域のみが表示されます。

たとえば、次の例は qtree1 のクォータです。

```
cluster::*> quota report -vserver DEMO -volume flexgroupDS -tree qtree1
Vserver: DEMO
                                ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
flexgroupDS
        qtree1   tree   1       0B    500GB       1      -  qtree1
```

ボリュームの実際の容量：

```
cluster::*> vol show -vserver DEMO -volume flexgroupDS -fields size
vserver volume      size
------- ----------- ----
DEMO    flexgroupDS 10TB
```

クライアントは、このボリュームの領域を次のように認識します。

```
# df -h /mnt/nas2
Filesystem         Size Used Avail Use% Mounted on
demo:/flexgroupDS  9.5T 4.5G  9.5T   1% /mnt/nas2
```

次の例は、その qtree についてのレポートです。

```
# df -h /mnt/nas2/qtree1/
Filesystem         Size Used Avail Use% Mounted on
demo:/flexgroupDS  500G    0  500G   0% /mnt/nas2
```

10. NAS とファイル数が多い場合の一般的な考慮事項

この章では、NAS とファイル数の多い環境に関する一般的な考慮事項について説明します。

10.1 ファイル数が多い場合の考慮事項

ONTAP の inode は、Snapshot コピーを含む、ファイルシステム内の任意のファイルまたはフォルダへのポインタです。各 FlexVol ボリュームは、有限数の inode を持ち、絶対最大値は 2,040,109,451 です。

FlexVol ボリュームの作成後に inode を増やすことが可能で、まだ割り当てられていない数までしか減らすことができません。

10.1.1 デフォルトおよび最大 inode 数

ボリューム (FlexVol と FlexGroup の両方) のデフォルトおよび最大 inode 数は、ボリュームの合計割り当て容量に依存します。たとえば、100GB の FlexVol ボリュームでは、8TB の FlexVol ボリュームと同じ数の inode を保持できません。

表 10.1 に、FlexVol のサイズ、inode のデフォルト値、および最大値の例を示します。

表 10.1 FlexVol サイズに応じた inode のデフォルト値と最大値

FlexVol サイズ	既定の inode 数	最大 inode 数
20MB*	566	4,855
1GB*	31,122	249,030
100GB*	3,112,959	24,903,679
1TB	21,251,126	255,013,682
7.8TB	21,251,126	2,040,109,451
100TB	21,251,126	2,040,109,451

*FlexGroup メンバーボリュームのサイズは、100GB 以上にする必要があります。

10.1.2 最大ファイル数の増加 : 考慮事項

inode 不足状態を監視したり、それに対処したりすることを避けたい場合は、サポートされている最大ファイル数の値を使用して、ファイル数が多い FlexGroup と FlexVol をすぐに設定できます。次の点に注意してください。

FlexVol ボリューム上の inode のデフォルト数または最大数は、ボリュームサイズによって異なり、容量 4KB に対する 1 inode の比率を持ちます。つまり、1 つのボリュームに割り当てられた 4KB ごとに、1 つの inode を割り当てることができます。これらの値の例を表 10.1 に示します。

さらに、各 inode は 288 バイトの容量を使用します。つまり、1 つのボリュームに多数の inode があると、実際のデータの容量だけでなく、物理的な容量もかなり消費されます。ファイルが 64 バイト未満の場合、そのファイルは inode 自体に格納され、追加容量は使用されません。

この使用済み領域は、ONTAP のアグリゲートリザーブの 10% に相当します。20 億個のファイルで最大約 585GB の領域を使用できます。最大ファイル数の制限が多くのボリュームに設定されている場合、各ボリュームの inode 容量はそのアグリゲートリザーブに割り当てられます。この容量は、ファイルが実際にボリュームに割り当てられている場合にのみ使用され、最大ファイル値を設定するだけでは使用されません。

そのため、ファイルの値を最大にする場合は、使用される inode と使用されるアグリゲートスペースの両方に注意する必要があります。両方の値を 80% の範囲に維持すると、ファイル数が多い環境で最適な結果が得られます。

その他の考慮事項

- FlexGroup ボリュームは、制限に達したときに無停止で拡張できるため、ファイル数の多い環境に最適なボリュームです。
- 割り当てられたサイズ 4KB あたり、おおよそ最大 1 つの inode を構成できます。そのため、割り当て可能な最大ファイル数を 20 億に設定するには、FlexVol または FlexGroup のメンバーボリュームのサイズを約 7.8TB 以上にする必要があります。
  - FlexGroup ボリュームでは、各メンバーボリュームが 7.8TB 以上である必要があります。
- ご使用の環境がサポートされる最大値に達した場合でも、inode 不足状態を監視する必要があります。また、FlexVol または FlexGroup を拡張または縮小するたびに、ファイル設定を再確認する必要がある場合があります。
- ボリューム内のファイルの最大値を設定する場合は、割り当てられた inode の 80% に監視しきい値を設定して、inode が不足する前に十分な計画と対応の時間を確保することも検討してください。
- ファイルの値が FlexVol または FlexGroup の個々のメンバーボリュームの最大量に設定されていて、inode が不足している場合は、FlexGroup ボリュームを使用して新しいメンバーボリュームを追加しない限り、inode をさらに増やすことはできません。したがって、可能であれば FlexVol ボリュームを 20 億に設定しないでください。FlexGroup ボリュームを使用すると、最大数である 20 億に達した場合に、少なくともメンバーボリュームを追加するオプションがあります。
- 最後に、inode のメタデータは基になるアグリゲートに格納されることに注意してください。そのため、アグリゲートの空き領域を監視して、アグリゲートで空き領域が不足しないようにする必要があります。

10.1.3 デフォルトおよび最大 inode 数：FlexGroup ボリュームに関する考慮事項

デフォルトのボリューム inode 数が 21,251,126 に達すると、FlexVol ボリュームのサイズに関係なく、デフォルト値のままになります。この機能は潜在的なパフォーマンスの問題を軽減しますが、新しい FlexGroup ボリュームを設計する際には考慮する必要があります。FlexGroup ボリュームは、最大 4000 億個のファイル (20 億ファイル x 200 FlexVol メンバーボリューム) を処理できますが、FlexGroup ボリューム内の 200 FlexVol メンバーのデフォルトの inode 数は、わずか 4,250,225,200 です。

このカウントは、次の数式に基づいています。

```
200 member volumes * 21,251,126 default inodes per member = 4,250,225,200 total default
inodes
```

FlexGroup ボリュームがデフォルト値として表示されているよりも多くの inode を必要とする場合は、`volume modify -files` コマンドを使用して inode を増やす必要があります。前述のように、この値は、必要に応じて許容される絶対最大値まで増加させることができますが、「10.1.2 最大ファイル数の増加 : 考慮事項」のガイダンスに従う必要があります。

FlexGroup ボリュームを使用する場合、デフォルトの inode 数の合計は、FlexGroup ボリューム内の FlexVol メンバーの合計サイズと FlexVol メンバーの数の両方に依存します。

表 10.2 に、FlexGroup 構成のさまざまな例と、結果として得られるデフォルトの inode 数を示します。

表 10.2 FlexGroup メンバーサイズおよびメンバーボリューム数の結果として生じる inode のデフォルト

メンバーボリュームサイズ	メンバーボリューム数	既定の inode 数 (FlexGroup)
100GB	8	24,903,672
100GB	16	49,807,344
1TB	8	170,009,008
1TB	16	340,018,016
100TB	8	170,009,008
100TB	16	340,018,016

10.1.4 ファイル数を多くして容量を少なくするニーズ

前述のように、ONTAP はボリューム容量に基づいてデフォルトの inode と最大 inode の数を割り当てます。図 10.1 では、7.8TB 未満のメンバーボリュームは、最大 20 億の inode を達成できません。メンバーボリュームあたり 20 億個の inode を取得するには、メンバーボリュームの容量が 7.8TB 以上である必要があります。8 つのメンバーボリュームと容量が保証された FlexGroup ボリュームは、最大 160 億のファイルをサポートしますが、最大 62.4TB の予約済みストレージもプロビジョニングします。

データセットが非常に小さいファイルで構成されている場合、その予約済み容量に近づくことはなく、他のワークロードに使用される可能性のある領域を浪費することになります。たとえば、ワークロード内のすべてのファイルのサイズがそれぞれ 288 バイトである場合、160 億個のファイルが消費する容量は最大約 4.6TB にすぎず、160 億個のファイルを処理するために必要な容量を十分に下回っています。

容量をほとんど使用しない多数のファイルを導入する場合、FlexGroup ボリュームの導入には主に 2 つのオプションがあります。

- **シンプロビジョニングを使用して、7.8TB 以上のメンバーボリュームで FlexGroup ボリュームを導入**
  ボリュームのシンプロビジョニングとは、ボリュームが特定のサイズであることを ONTAP に伝えているが、そのサイズがファイルシステムで保証されていないことを意味します。これにより、ファイルシステム内で柔軟にストレージの割り当てを物理領域に制限できます。ただし、アグリゲート内の他のボリュームは、その使用スペースによって、かつ領域保証を有効にしている場合、空き容量に影響を与える可能性があるため、シンプロビジョニングを使用する場合は、使用可能なアグリゲート容量を監視することが重要です。詳細は、「8.2.2 FlexGroup ボリュームでの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」を参照してください。
- **デフォルトよりも多くのメンバーボリュームを持つ FlexGroup ボリュームを手動で作成**
  FlexGroup ボリュームの領域保証を維持したい場合は、ファイル数が多く容量が少ない環境で、FlexGroup ボリューム内にさらにメンバーボリュームを作成するという方法もあります。

inode 数は容量に応じて FlexVol メンバーボリュームごとに制限されるため、より小さいメンバーボリュームを追加すると、同じ容量でファイル数を増やすことができます。次の表に、考えられる構成を示します。FlexGroup ボリュームの手動作成の詳細については、「7.10.5 FlexGroup ボリュームを手動で作成する必要があるのはどのような場合ですか ?」を参照してください。

表 10.3 ファイル数が多い小容量のフットプリントの例 : メンバーボリューム数の増加

FlexGroup 合計サイズ	メンバーボリューム数 (サイズ)	最大 inode 数 (FlexGroup 全体)
80TB (領域保証なし)	8 (10TB)	16,320,875,608
64TB (領域保証が有効)	32 (2TB)	16,320,875,608
64TB (領域保証が有効)	64 (1TB)	16,320,875,608

10.1.5 ONTAP でファイル数が多い場合の計画

「16.6 XCP Migration Tool」などのユーティリティを使用 (スキャン機能を使用) すると、ファイル数の使用状況やその他のファイル統計を評価して、XCP を使用したファイルスキャンの詳細については、当社サポート部門までお問い合わせください。

10.1.6 使用済み inode と合計 inode の表示

ONTAP では、高度な権限で次のコマンドを使用して、ボリュームごとの inode 数を表示できます。

```
cluster::*> volume show -volume flexgroup -fields files,files-used
vserver volume    files     files-used
------- --------- --------- ----------
SVM     flexgroup 170009008 823
```

従来の df -i コマンドを使用することもできます。すべてのメンバーボリュームを表示するには、diag 権限でボリューム名とともにアスタリスクを使用します。

```
cluster::*> df -i Tech_ONTAP*
Filesystem              iused      ifree    %iused      Mounted on      Vserver
/vol/Tech_ONTAP/        10193  169998815        0%      /techontap      DEMO
/vol/Tech_ONTAP__0001/    923   21250203        0%      /techontap      DEMO
/vol/Tech_ONTAP__0002/   4177   21246949        0%      ---             DEMO
/vol/Tech_ONTAP__0003/    878   21250248        0%      ---             DEMO
/vol/Tech_ONTAP__0004/    848   21250278        0%      ---             DEMO
/vol/Tech_ONTAP__0005/    750   21250376        0%      ---             DEMO
/vol/Tech_ONTAP__0006/    972   21250154        0%      ---             DEMO
/vol/Tech_ONTAP__0007/    879   21250247        0%      ---             DEMO
/vol/Tech_ONTAP__0008/    766   21250360        0%      ---             DEMO
```

10.1.7 inode が足りなくなったらどうなりますか？

ボリュームの inode がなくなると、inode の数が増えるか、既存の inode が解放されてクラスタが EMS イベント (`callhome.no.inodes`) をトリガーするまで、そのボリュームにファイルを作成できなくなります。さらに、AutoSupport メッセージがトリガーされます。FlexGroup ボリュームは、データ取り込みに最適なメンバーボリュームを決定するときに、メンバーごとの inode 数を考慮します。例については、「17.8.1 inode 関連の EMS の例」を参照してください。

EMS メッセージは、監視やスクリプトの起動に使用できます。スクリプトは、実稼働ワークロードの問題が発生する前に領域エラーを回避するために、inode 数を自動的に増やします。

最大ファイル数を増やす方法については、「10.1.2 最大ファイル数の増加 : 考慮事項」を参照してください。

10.1.8 非同期削除

ONTAP 9.8 には、ストレージ管理者が NAS クライアントからの削除を実行する代わりに、クラスタ CLI からディレクトリ全体を削除できる新機能が導入されています。これにより、NAS プロトコルを使用するよりもはるかに高速に多数のファイルを削除できるほか、ネットワークとクライアントのパフォーマンスの競合を排除できます。このコマンドは、FlexVol ボリュームと FlexGroup ボリュームの両方で動作します。

テストでは、非同期削除は、シングルスレッドの rm コマンドよりも約 10 倍速く、FlexVol ボリュームではわずかに速く実行されました。

表 10.4 非同期削除のパフォーマンス

A300 (24,000 個のファイル / フォルダ)	rm -rf* 秒数	非同期削除秒数	速度増加
FlexVol	18.3	2	9.1 倍
FlexGroup	32.1	3	10.7 倍

非同期削除でディレクトリの削除が発生すると、ジョブが実行され、ディレクトリを削除するために並行して実行される複数のタスクが作成されます。デフォルトでは、ジョブは 5,000 の同時タスクに調整されますが、その量は最小で 50 まで減らすことも、最大で 10 万まで増やすこともできます。

削除コマンドが発行されると、ONTAP は指定されたディレクトリをスキャンします。サブディレクトリが見つかった場合、それらのディレクトリの内容が最初に削除されます。

次の注意事項が適用されます。

- CLI のみ
- SVM とボリュームは有効でなければなりません
- ボリュームはオンラインでマウントされている必要があります
- ディレクトリパスは有効でなければなりません
- 一度に実行できる非同期削除は 1 つだけです
- ディレクトリで実行する必要があります。単一ファイルでは実行できません。

Delete job を実施する場合

```
cluster::*> async-delete start -vserver DEMO -volume FlexGroup1 -path /files
[Job 34214] Job is queued: Asynchronous directory delete job.
```

進捗状況をチェックする場合

```
cluster::*> async-delete show -vserver DEMO -instance
```

10.2 64 ビットのファイル識別子

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。ファイル ID は、どのファイルがどのファイルであるかを ONTAP が追跡できるようにする、ファイルシステム内の一意の識別子です。32 ビットのファイル ID は最大 2,147,483,647 の符号付き整数に制限されていますが、これは FlexVol の 20 億 inode 制限の由来です。

FlexGroup ボリュームは、複数のメンバーボリュームをリンクすることで、単一のネームスペースで数千億のファイルをサポートできますが、32 ビット符号付き整数の制限である 20 億 を安全に超える (さらにファイル ID の競合の可能性を除去する) には、64 ビットのファイル ID を有効にする必要があります。

ファイル ID の競合が保証される前に 32 ビットのファイル ID が使用されている場合、ONTAP は FlexGroup ボリューム内で最大 4,294,967,295 個のファイル ID (32 ビット符号なし整数) を配布できます。2,147,483,647 個のファイルが存在する場合、ファイル ID の競合は数学的に不可能であるため、32 ビットのファイル ID を使用する場合に最も安全なファイル数となります。この値を超えると、ファイル数が符号なし 32 ビット整数値 4,294,967,295 に近づくにつれて、ファイル ID の競合の可能性が高くなります。maxfiles の値を大きい値に設定しても、FlexGroup ボリュームに 20 億個を超えるファイルを作成できます。ファイル ID の競合で何が起こるかの詳細については、「10.2.6 ファイル ID の競合の影響」を参照してください。

64 ビットのファイル ID を使用すると、ONTAP は最大 9,223,372,036,854,775,807 個の一意のファイル ID をファイルに割り当てることができます。ただし、FlexGroup ボリュームで公式にサポートされる最大ファイル数の制限は 4000 億です。

64 ビットのファイル識別子オプションは、デフォルトでは off/disabled 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仕様です。

備考

64 ビットのファイル ID を有効にする前に、そのサポートについてアプリケーションまたは OS のベンダーに確認するか、テスト用の SVM を作成して、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 を使用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、次の高度な権限レベルコマンドで有効にできます。また、NFSv3 および NFSv4 オプションがあります。

```
cluster::> set advanced
cluster::*> nfs modify -vserver SVM -v3-64bit-identifiers enabled -v4-64bit-identifiers
enabled
```

または、ONTAP System Manager を使用して、これらの値を有効または無効にすることもできます。

10.2.1 このオプションを変更するとどうなりますか？

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、既存のマウントで NFS 操作を試みたときに、クライアントが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。FSID 変更オプションの有効化または無効化がファイル数の多い環境の SVM に与える影響の詳細については、「10.3.1 ファイル数の多い環境の SVM での FSID の操作」を参照してください。

10.2.1.1 64 ビットのファイル ID を有効にする必要がありますか。

64 ビットのファイル ID が有効になっていない SVM で新しい FlexGroup ボリュームを作成すると、このオプションを有効にする必要があるという警告が表示されることがあります。ただし、このオプションを有効にすると、ボリュームを再マウント (および停止) しなければならなくなり、一部のアプリケーションは 64 ビットのファイル ID をサポートしないため、このオプションを有効にしない場合もあります。

FlexGroup ボリュームのファイル数が 20 億を超える場合は、この値を変更しないままにすることができます。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 の最大数も 2,147,483,647 以下に増やす必要があります。

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

備考

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。

32 ビットを必要とするボリュームと、20 億個を超えるファイルを必要とするその他のボリュームがある環境では、異なる SVM を使用してそれらのボリュームをホストし、必要に応じて 64 ビットのファイル ID を有効または無効にすることができます。

ベストプラクティス 15: 64 ビットのファイル識別子
FlexGroup ボリュームを作成する前に、NFS サーバオプションの `-v3-64bit-identifiers` を高度な権限レベルで有効にすることを強く推奨しています。特に、使用しているファイルシステムが 20 億 inode しきい値を超えるか超える可能性がある場合に有効にすることを推奨します。

10.2.2 NFSv3 と NFSv4.x: ファイル ID

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ています。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 を有効または無効にする 2 つのオプションが用意されています。

SVM で NFSv3 と NFSv4.x の両方を使用し、両方のプロトコルに 64 ビット ID オプションを適用する場合は、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る場合がありますが、NFSv4.x ではファイル ID の競合が発生するとエラーが送信されます。

オプションは次のとおりです。

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

備考

ONTAP 9.7 (FlexGroup ボリュームで NFSv4.x をサポートする最初のリリース) にアップグレードする場合は、9.7 P7 以降にアップグレードしてください。

10.2.3 クォータの適用を使用してファイル数を制限する

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、Flex Group 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。

クォータの適用ポリシーは親ボリュームの下に作成されたファイルには適用されないため (監視 / レポート作成ポリシーのみ)、FlexGroup ボリューム内に qtree を作成します。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軽減するために、その qtree に対して 20 億個のファイルを制限とするクォータツリールールを作成します。または、ボリューム内にファイルを作成するユーザー名とグループ名がわかっている場合は、特定のユーザーまたはグループのクォータ規則を作成できます。

```
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
                                ----Disk---  ---Files-----  Quota
Volume  Tree     Type   ID      Used  Limit  Used   Limit   Specifier
------- -------- ------ ------- ----- -----  ------ ------  ---------
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やポリシーのエクスポート規則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 に適用すると、NFS マウントは qtree レベルで行われます。

その後、qtree でファイルが作成されると、この制限にカウントされます。

```
[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

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

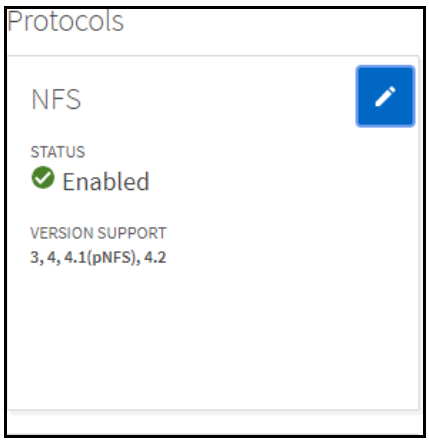
10.2.4 ONTAP System Manager: 9.7

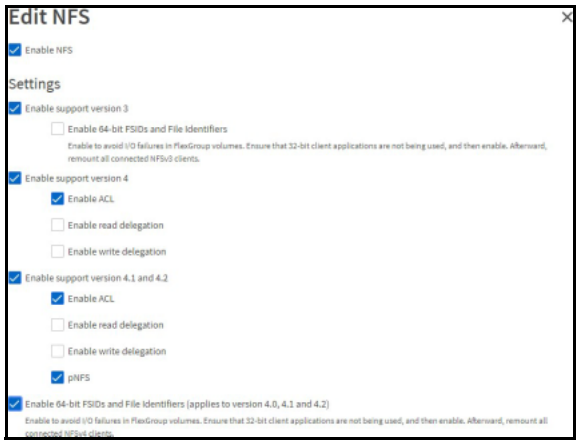
ONTAP 9.7 では、REST API 機能に基づく新しい System Manager インタフェースが導入されました。64 ビットファイル ID オプションは現在 REST API に存在しないため、System Manager でこのオプションを変更する唯一の方法は、CLI を使用することです。

10.2.5 ONTAP System Manager: 9.8 以降

ONTAP System Manager 9.8 以降には、[Storage] > [Storage VMs] メニューオプションから 64 ビットファイル ID 値を有効または無効にする GUI 方式が含まれています。目的の SVM をクリックし、NFS プロトコルメニューから [Edit] を選択します。

図 10.1 ONTAP System Manager 9.8 の 64 ビットファイル ID





10.2.6 ファイル ID の競合の影響

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 の競合の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、その影響は、クライアント上の古いファイルハンドルエラーから、ディレクトリと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にまで及びます。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。

`stat` コマンドまたは `ls -i` コマンドを使用すると、クライアントからファイルの ID をチェックできます。inode またはファイル ID の競合が発生すると、次のようになります。両方のファイルの inode は 3509598283 です。

```
# stat libs/
  File: `libs/'
  Size: 12288           Blocks: 24         IO Block: 65536 directory
Device: 4ch/76d Inode: 3509598283  Links: 3
Access: (0755/drwxr-xr-x)  Uid: (60317/ user1)    Gid: (10115/  group1)
Access: 2017-01-06 16:00:28.207087000 -0700
Modify: 2017-01-06 15:46:50.608126000 -0700
Change: 2017-01-06 15:46:50.608126000 -0700

# stat iterable/
  File: `iterable/'
  Size: 4096           Blocks:  8         IO Block: 65536 directory
Device: 4ch/76d Inode: 3509598283  Links: 2
Access: (0755/drwxr-xr-x)  Uid: (60317/ user1)    Gid: (10115/  group1)
Access: 2017-01-06 16:00:44.079145000 -0700
Modify: 2016-05-05 15:12:11.000000000 -0600
Change: 2017-01-06 15:23:58.527329000 -0700

# ls -i libs
3509598283 libs

# ls -i iterable
3509598283 iterable
```

競合が発生すると、Linux クライアント上で `find` コマンドまたは `rm` コマンドの実行中に循環ディレクトリ構造エラーが発生したり、ファイルを削除できないなどの問題が発生する可能性があります。場合によっては、古いファイルハンドルエラーが表示されることもあります。

```
rm: WARNING: Circular directory structure.
This almost certainly means that you have a corrupted file system.
NOTIFY YOUR SYSTEM MANAGER.
The following directory is part of the cycle:
'/directory/iterable'
rm: cannot remove '/directory': Directory not empty
```

備考

ファイル ID の競合は NFS にのみ影響します。SMB は、同じファイル ID 構造を使用しません。

10.3 ONTAP でのファイルシステム ID 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またがる複数のファイルシステムにまたがることができるため、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 または NFSv4 の)FSID の変更オプションを無効にする場合、NFS サーバで 64 ビットのファイル ID オプションを有効にしてください (「10.2 64 ビットのファイル識別子」を参照してください)。これは、ファイル ID の総数が SVM 内のボリューム間で共有されるため、ファイル ID の競合がより早く発生するリスクがあるためです。

この FSID 変更オプションは、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も影響する可能性があります。FSID の変更を切り替える前に、別の SVM でアプリケーションを使用して適切なテストを実行します。

10.3.1 ファイル数の多い環境の SVM での FSID の操作

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 ボリュームに固有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数はボリューム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 ビットのファイル ID によるファイル制限がすべてのボリュームに適用されます。

たとえば、SVM の 10 個の異なるボリュームに 100 億個のファイルがある場合、FSID 変更オプションを有効のままにすると、各ボリュームは固有のファイル ID のセットを持つことができます。FSID の変更オプションを無効にすると、100 億個のファイルすべてが SVM 内のファイル ID のプールを共有します。32 ビットのファイル ID を使用すると、ファイルの競合が発生する可能性があります。

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ています。

10.3.2 FSID による Snapshot コピーの処理方法

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 を組み合わせて作成されます。すべての Snapshot コピーは異なる ID を持つため、`-v3-fsid-change` オプションの設定に関係なく、ファイルのすべての Snapshot コピーは NFSv3 で異なる FSID を持ちます。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。

10.4 ディレクトリサイズに関する考慮事項 :maxdirsize

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 320MB です。つまり、ディレクトリサイズのメモリ割り当てが最大 320MB に達すると、ディレクトリはそれ以上大きくなりません。1 つのディレクトリ内のファイル数が増えると、ディレクトリサイズが大きくなります。ディレクトリ内の各ファイル項目は、そのディレクトリに割り振られたスペースに対してカウントされます。1 つのディレクトリに保存できるファイル数については、「10.4.4 デフォルトの maxdirsize で 1 つのディレクトリに収まるファイルの数」を参照してください。

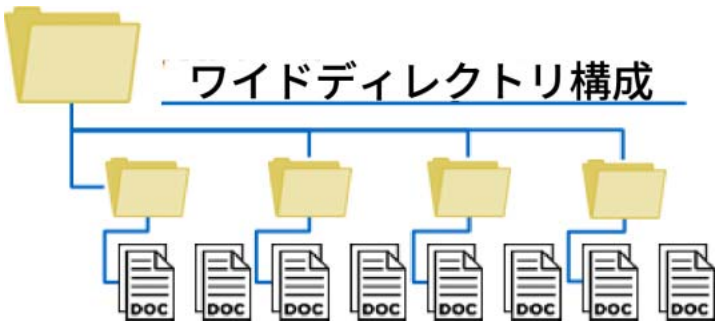
ベストプラクティス 16: ファイル数の多い環境に推奨される ONTAP バージョン
ファイル数の多い環境では、最新の ONTAP リリースを使用して、FlexGroup の機能拡張、WAFL の機能拡張、およびファイル数の多いワークロードのパフォーマンス向上のメリットを得ることができます。

10.4.1 maxdirsize はどのようなディレクトリ構造に影響されますか？

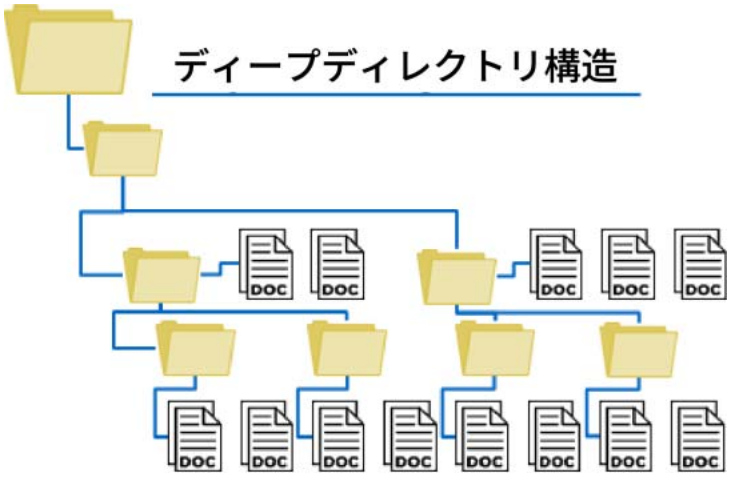
1 つのフォルダに 1 つのレベルで数百万のファイルが格納されているフラットディレクトリ構造を使用している場合は、`maxdirsize` 値が問題になることがあります。ファイル、フォルダ、サブフォルダが散在しているフォルダ構造には、`maxdirsize` への影響はほとんどありません。ディレクトリ構造の方法論はいくつかあります。

- フラットディレクトリ構造
  多数のファイルを含む単一のディレクトリ。

- ワイドディレクトリ構造
  多くの最上位ディレクトリがあり、ファイルが複数のディレクトリに分散している。



- ディープディレクトリ構造
  トップレベルディレクトリの数は少ないが、サブフォルダが多く、ファイルが複数のディレクトリに分散している。



10.4.2 フラットディレクトリ構造が FlexGroup ボリュームに与える影響

フラットディレクトリ構造 (単一または少数のディレクトリ内に多数のファイル) は、当社システムであるかどうかにかかわらず、さまざまなファイルシステムに悪影響を及ぼします。潜在的な問題には次のようなものがありますが、これらに限定されるものではありません。

- メモリ圧力
- CPU 使用率
- ネットワークパフォーマンス / レーテンシー (特にファイルの大量クエリー、GETATTR 操作、READDIR 操作など)

FlexGroup ボリュームは、maxdirsize にも影響を与えます。FlexVol ボリュームとは異なり、FlexGroup ボリュームは ONTAP 内部のリモートハードリンクを使用してトラフィックをリダイレクトします。これらのリモートハードリンクにより、FlexGroup ボリュームはクラスタ内でパフォーマンスと容量をスケールアウトできます。

ただし、フラットディレクトリでは、ローカルファイルへのリモートハードリンクの割合が高くなります。これらのリモートハードリンクは maxdirsize の合計値に対してカウントされるため、FlexGroup ボリュームは、FlexVol よりも早く maxdirsize の制限に到達する可能性があります。

たとえば、ディレクトリ内に数百万のファイルがあり、そのファイルシステムに対して約 85% のリモートハードリンクを生成する場合、maxdirsize は FlexVol の約 2 倍の量で使い果たされることが予想されます。

ベストプラクティス 17: ディレクトリ構造の推奨事項

- 最高のパフォーマンスを得るには、可能なら ONTAP でフラットなディレクトリ構造を使用しないようにしてください。ファイルまたはフォルダのパスの長さが NAS プロトコル標準を超えない限り、ワイドまたはディープディレクトリ構造が最適に機能します。
- フラットなディレクトリ構造が避けられない場合は、ボリュームの maxdirsize 値に細心の注意を払い、当社サポート部門の指示に従って必要に応じて値を増やしてください。
- NFS パスの長さは、クライアント OS によって定義されます。
- SMB パスの長さについては、この Microsoft Dev Center リンクを参照してください。

10.4.3 使用されている maxdirsize 値の問い合わせ

ONTAP で maxdirsize 割り当てを監視し、評価することは重要です。ただし、ONTAP 固有のコマンドはありません。

代わりに、クライアントから maxdirsize の割り当てを問い合わせる必要があります。

NFS クライアントから次のコマンドを実行すると、特定のマウントポイント内でディレクトリの最大数が 10 個の FlexGroup ボリューム内のフォルダのディレクトリサイズ情報を取得できます。ただし、Snapshot コピーは検索対象から除外されます。

```
# find /mountpoint -name .snapshot -prune -o -type d -ls -links 2 -prune | sort -rn -k 7 |
head
```

次の例は、数百万個のファイルが格納されているフォルダ内のデータセットで 1 秒未満しかかかりませんでした。

```
[root@centos7 /]# time find /flexgroup/manyfiles/ -name .snapshot -prune -o -type d -ls
-links 2
-prune | sort -rn -k 7 | head
787227871 328976 drwxr-xr-x     2 root  root    335544320 May 29 21:23
/flexgroup/manyfiles/folder3/topdir_8/subdir_0
384566806 328976 drwxr-xr-x     2 root  root    335544320 May 29 13:14
/flexgroup/manyfiles/folder3/topdir_9/subdir_0
3605793347 328976 drwxr-xr-x    2 root  root    335544320 May 29 21:23
/flexgroup/manyfiles/folder3/topdir_0/subdir_0
3471151639 328976 drwxr-xr-x    2 root  root    335544320 May 29 13:45
/flexgroup/manyfiles/folder3/topdir_4/subdir_0
2532103978 328976 drwxr-xr-x    2 root  root    335544320 May 29 14:16
/flexgroup/manyfiles/folder3/topdir_2/subdir_0
2397949155 328976 drwxr-xr-x    2 root  root    335544320 May 29 14:15
/flexgroup/manyfiles/folder3/topdir_1/subdir_0
1994984460 328976 drwxr-xr-x    2 root  root    335544320 May 29 13:43
/flexgroup/manyfiles/folder3/topdir_6/subdir_0
1860674357 328976 drwxr-xr-x    2 root  root    335544320 May 29 13:18
/flexgroup/manyfiles/folder3/topdir_5/subdir_0
1458235096 328976 drwxr-xr-x    2 root  root    335544320 May 29 14:25
/flexgroup/manyfiles/folder3/topdir_3/subdir_0
1325327652 328976 drwxr-xr-x    2 root  root    335544320 May 29 14:25
/flexgroup/manyfiles/folder3/topdir_7/subdir_0

real   0m0.055s
user   0m0.002s
sys    0m0.035s
```

10.4.3.1 XCP を使用した maxdirsize のチェック

XCP Migration Tool は、主に高速データムーバーと考えられていますが、堅牢なファイルスキャン機能の価値も引き出します。XCP では、find コマンドも並行して実行できるため、ストレージシステム上で前述の例をより高速に実行できるほか、指定したファイル数のディレクトリに結果をフィルタリングできます。

次の XCP コマンドの例では、2,000 を超えるエントリを持つディレクトリに対してのみ find を実行できます。

```
# xcp diag find --branch-match True -fmt "'{size} {name}'.format(size=x.digest, name=x)"
localhost:/usr 2>/dev/null | awk '{if ($1 > 2000) print $1 " " $2}'
```

次の XCP コマンドを使用すると、ディレクトリサイズの値を検索できます。

```
# xcp -match "type == d" -fmt "'{} {}'.format(used, x)" localhost:/usr | awk
'{if ($1 > 100000) print}' | sort -nr
```

XCP はディレクトリサイズの値を探すとき、最初にファイルシステムをスキャンします。次に例を示します。

```
[root@XCP flexgroup]# xcp -match "type == d" -fmt "'{} {}'.format(used, x)"
10.193.67.219:/flexgroup_16/manyfiles | awk '{if ($1 > 100000) print}' | sort -nr

660,693 scanned, 54 matched, 123 MiB in (24.6 MiB/s), 614 KiB out (122 KiB/s),  5s
1.25M scanned, 58 matched, 234 MiB in (22.1 MiB/s), 1.13 MiB out (109 KiB/s), 10s
…
31.8M scanned, 66 matched, 5.83 GiB in (4.63 MiB/s), 28.8 MiB out (22.8 KiB/s), 7m52s

Filtered: 31816172 did not match
31.8M scanned, 66 matched, 5.83 GiB in (12.6 MiB/s), 28.8 MiB out (62.4 KiB/s), 7m53s.
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_9/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_8/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_7/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_6/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_5/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_4/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_3/subdir_0
```

10.4.4 デフォルトの maxdirsize で 1 つのディレクトリに収まるファイルの数

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 は、大規模なシステムではデフォルトで 320MB に設定されるため、FlexVol ボリューム上の SMB および 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 です。

FlexGroup ボリュームは、リモートハードリンクを使用して、I/O をメンバーボリュームにリダイレクトします。これらのハードリンクはディレクトリの合計サイズにカウントされるため、320MB の maxdirsize で許容されるファイルの最大数は、作成されたハードリンクの数によって異なります。ディレクトリあたりのファイル数は、FlexGroup ボリューム内のディレクトリでは 2 から 260 万の範囲になります。

maxdirsize 値をデフォルト値のままにしておくことを強く推奨します。

10.4.5 maxdirsize を超えると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 (EMS) メッセージは、maxdirsize を超えた場合、または超えそうになった場合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 のイベント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、90% しきい値前の通知の送信が可能です (「14. 容量の監視とアラート」を参照してください)。これらのイベント管理システムメッセージは、SNMP トラップもサポートしています。

```
wafl.dir.size.max
wafl.dir.size.max.warning
wafl.dir.size.warning
```

10.4.6 maxdirsize 値の増加の効果

1 つのディレクトリに多数のファイルが含まれている場合、ルックアップ (検索操作など) は CPU とメモリを大量に消費します。ディレクトリのインデックス作成により、2MB を超えるディレクトリサイズのインデックスファイルが作成されるため、多くの検索を実行してキャッシュミスを回避する必要がありません。
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。可能であれば、ファイル数の多い環境では最新バージョンの ONTAP を使用して、WAFL の改善によるメリットを得てください。

ベストプラクティス 18: maxdirsize の最大値
`maxdirsize` の値は、4GB を超えないようにハードコードされています。パフォーマンスの問題を回避するために、`maxdirsize` の値を 1GB 以下に設定することを推奨しています。

10.4.7 FlexGroup ボリュームは maxdirsize 制限をバイパスしますか？

FlexGroup ボリュームでは、各メンバーボリュームに同じ `maxdirsize` 設定があります (これは FlexGroup レベルで設定されます)。ディレクトリ内のファイルが複数の FlexVol メンバー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、ディレクトリ自体は単一のメンバーボリューム上に存在します。その結果、FlexVol ボリュームに見られる `maxdirsize` の制限は、FlexGroup にも適用されます。これは、ボリュームではなく、ディレクトリサイズがキーコンポーネントであるためです。FlexGroup ボリュームでは、ディレクトリは単一の FlexVol メンバーボリュームに存在するため、`maxdirsize` の制限に直面している環境では問題が解決されません。

ベストプラクティス 19: maxdirsize の問題の回避
新しいプラットフォームでは、より多くのメモリと CPU 容量が提供され、ETERNUS AX series システムではファイル数の多い環境でパフォーマンスが向上します。ただし、多数のファイルがあるディレクトリの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させることです。

10.4.8 maxdirsize を超えることの影響

ONTAP で `maxdirsize` を超えると、`out of space(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。このエラーはストレージ管理者に誤解を与える可能性があります。これは、この場合の問題がファイル数に関連している場合に、実際の容量の問題を暗示しているためです。クライアントから容量の問題が報告された場合は、必ず ONTAP イベントログを確認して問題を絞り込みます。

ディレクトリサイズ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、ファイルをディレクトリから移動する必要があります。修復の詳細については、当社サポート部門にお問い合わせください。maxdirsize イベント管理システムイベントの例については、「17.8.2 maxdirsize メッセージの例」を参照してください。

10.5 ファイルシステムの分析

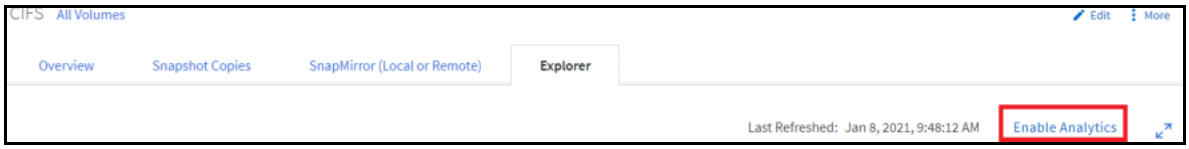
ONTAP 9.8 では、ストレージ管理者が ONTAP System Manager からファイルやディレクトリの情報に瞬時にアクセスできるようにする、ファイルシステム分析という方法を提供する新機能が導入されました。

このファイルシステム分析の最初のリリースには、以下のような情報が含まれます。

- ファイルサイズ
- フォルダサイズ
- atime および mtime のヒストグラム
- ファイルとフォルダの一覧
- 非アクティブおよびアクティブなデータレポート作成
- ファイル数とディレクトリ数

この情報は、初期スキャンの実行後にファイルシステムが更新され、使用するシステムリソースが最小限に抑えられたときに、ONTAP によって収集されます。ファイルシステム分析はデフォルトではオフになっており、使用している NAS プロトコルに関係なく、ONTAP System Manager を介してボリュームページの新しい [Explorer] タブから FlexVol ボリュームと FlexGroup ボリュームの両方について有効 (および無効) にできます。

図 10.2　ファイルシステム分析 : 有効



分析を有効にして初期スキャンが完了したら (完了時間はファイルとフォルダの数によって異なります)、ONTAP System Manager の [Explorer] タブでディレクトリツリーをクリックしてディレクトリ構造全体を参照できます。

図 10.3　ファイルシステム分析 : ディレクトリおよびファイル情報

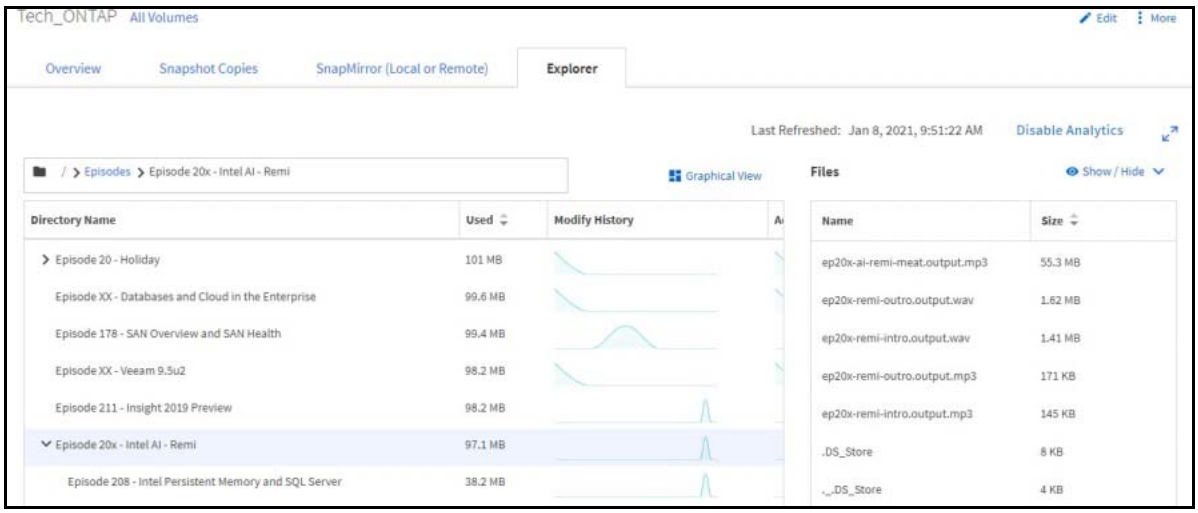
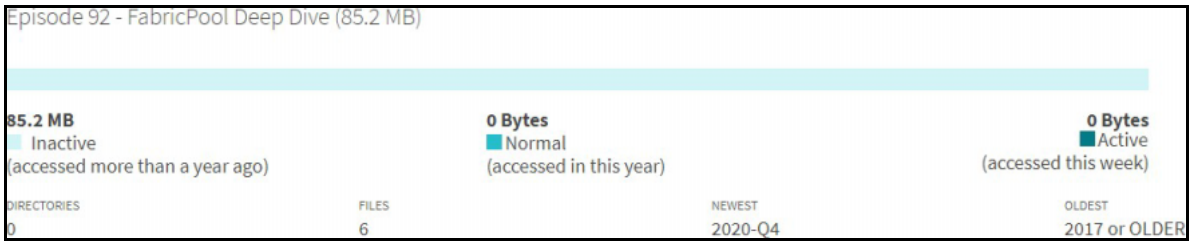


図 10.4　ファイルシステム分析 : 使用頻度の低いデータと使用頻度の高いデータ



ファイルやフォルダが作成または削除されると、ファイルシステム分析は新しい情報でツリーを数秒で更新します。ファイルシステム分析を使用すると、ストレージ管理者は、オフボックスのユーティリティや、ファイル数の多い環境で時間のかかる du、`find`、`ls` などのコマンドを使用しなくても、ファイルやフォルダの情報を取得できます。

10.6 特殊文字の考慮事項

Unicode (UTF-8 形式でエンコードされている場合) で最も一般的なテキスト文字は、3 バイト以下のエンコードを使用します。この共通テキストには、中国語、日本語、ドイツ語など、現代のすべての言語が含まれます。しかし、絵文字などの特殊文字の人気により、一部の UTF-8 文字サイズは 3 バイトを超えています。たとえば、トロフィー記号は UTF-8 エンコーディングで 4 バイトを必要とする文字です。

特殊文字には、次のようなものがあります。

- 絵文字
- 音楽記号
- 数学記号

特殊文字が FlexGroup ボリュームに書き込まれると、次の動作が発生します。

```
# mkdir /flexgroup4TB/
mkdir: cannot create directory '/flexgroup4TB/\360\237\217\206': Permission denied
```

上の例では、\ 360 \ 237 \ 217 \ 206 は 16 進数の 0xF0 0x9F 0x8F 0x86 (UTF-8) であり、トロフィー記号です。

ONTAP ソフトウェアは、NFS で 3 バイトを超える UTF-8 サイズをネイティブでサポートしていませんでした。3 バイトを超える文字サイズを処理するために、ONTAP は余分なバイトをオペレーティングシステムの bagofbits と呼ばれる領域に配置します。これらのビットは、クライアントが要求するまで格納されます。次に、クライアントは raw ビットから文字を解釈します。FlexVol および FlexGroup ボリュームは bagofbits をサポートします。

ベストプラクティス 20: FlexGroup ボリュームでの特殊文字の処理
FlexGroup ボリュームで最適な特殊文字の処理を行うには、utf8mb4 ボリューム言語を使用します。

また、ONTAP には、問題のあるファイル ID の識別方法など、bagofbits 処理の問題に関するイベント管理システムメッセージがあります。

```
Message Name: wafl.bagofbits.name Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Access the parent directory from an NFSv3 client and
rename the entry using Unicode characters.
Description: This message occurs when a read directory request from an NFSv4 client is made
to a Unicode-based directory in which directory entries with no NFS alternate name contain
non-Unicode characters.
```

FlexGroup で bagofbits 機能をテストするには、次のコマンドを使用します。

```
# touch "$(echo -e "file\xFC")"
```

10.6.1 utf8mb4 ボリューム言語のサポート

前述のように、特殊文字は、ネイティブでサポートされている UTF-8 エンコーディングの 3 バイトを超える場合があります。次に、ONTAP は bagofbits 機能を使用して、これらの文字を機能させます。

このような inode 情報の格納方法は理想的ではないため、utf8mb4 ボリューム言語がサポートされています。ボリュームがこの言語を使用する場合、サイズが 4 バイトの特殊文字は適切に格納され、bagofbits には格納されません。

ボリューム言語は、NFSv3 クライアントによって送信された名前を Unicode に変換したり、ディスク上の Unicode 名を NFSv3 クライアントが期待するエンコーディングに変換したりするために使用されます。UTF-8 以外のエンコード方式を使用するように NFS ホストが構成されている従来の状況では、対応するボリューム言語を使用する必要があります。最近では UTF-8 の使用はほぼ普遍的になっているため、ボリューム言語は UTF-8 になる可能性があります。

NFSv4 では UTF-8 を使用する必要があるため、NFSv4 ホストで非 UTF-8 エンコードを使用する必要はありません。同様に、CIFS は Unicode をネイティブで使用するため、任意のボリューム言語で動作します。ただし、utf8mb4 以外のボリュームでは、基本面より上に Unicode 名があるファイルは正しく変換されないため、utf8mb4 の使用をお勧めします。

ボリューム言語は、ボリュームの作成時に `-language` オプションを使用してのみ設定できます。ボリュームの言語を変換することはできません。新しいボリューム言語でファイルを使用するには、ボリュームを作成し、「16.6 XCP Migration Tool」などのユーティリティを使用してファイルを移行します。

ベストプラクティス 21: UTF-8 または utf8mb4?
クライアントが utf8mb4 ボリューム言語をサポートできない場合を除き、ファイル名変換の問題を防ぐために utf8mb4 ボリューム言語を使用することをお勧めします。

10.6.2 多数のファイルが存在する環境で NFS を使用して低速のディレクトリ一覧を管理

ファイル数の多い環境でのワークフローの中には、既存のデータセットに対して `find` や `ls` などの読み取りメタデータを多用する操作を実行するものがあります。このタイプのワークロードは非効率的で、完了までに長い時間がかかることがあります。これらの操作を実行する必要がある場合は、迅速に行うためにいくつかの方法を試してみてください。

一般的に、これらのタイプの操作の問題は、クライアント、プロトコル、またはネットワーク関連です。ストレージがメタデータの読み込み速度のボトルネックになることはほとんどありません。ONTAP は、メタデータ読み取り操作をマルチスレッド化できます。`ls` 操作では、`getattr` 要求は一度に 1 つずつシリアルで送信されます。つまり、数百万回の `getattr` 操作を行うと、ストレージに対して数百万回のネットワーク要求が送信される可能性があります。各ネットワークリクエストには n ミリ秒の遅延が発生し、時間の経過とともに増加します。

そのため、これらを高速化する方法がいくつかあります。

- **一度にさらに多くの `getattr` リクエストを送信**
  `ls` だけでは、要求を並列に送信できません。しかし、XCP Migration Tool のようなユーティリティーを使えば、複数のスレッドをネットワーク経由で送信して `ls` 操作を大幅に高速化することができます。`ls` の出力が後で何に使われるかによっては、XCP スキャンを使うと速度が向上します。たとえば、ファイルのユーザー権限 / 所有者が必要な場合は、`ls` を単独で使用する方が適している場合があります。しかし、ファイル名の単純なリストを作成するには、XCP スキャンが適しています。
- **ネットワークハードウェア (たとえば、10GB ではなく 100GB) を追加して、ラウンドトリップ時間 (RTT) を短縮**
  より大きなネットワークパイプでは、より多くのトラフィックがネットワーク上でプッシュされるため、負荷が軽減され、RTT 全体が削減される可能性があります。数百万のオペレーションでは、1 ミリ秒のレーテンシーを削減するだけでも、ワークロードの時間を大幅に節約できます。
- **ハイライトやカラーなどの不要なオプションを指定せずに `ls` を実行**
  `ls` を実行するときのデフォルトの動作は、読みやすくするために、並べ替え、色、および強調表示を追加することです。これらは操作のための作業を追加するので、これらの潜在的に不必要な機能を避けるために `-f` オプションを付けて `ls` を実行することは理にかなっているかもしれません。
- **クライアント上の `getattr` 操作をより積極的にキャッシュ**
  属性をクライアント側でキャッシュすると、操作のためのネットワークトラフィックを減らすのに役立つだけでなく、操作のために属性をクライアントでローカルにするのにも役立ちます。クライアントによる NFS キャッシュの管理方法は異なりますが、一般的には、ファイル数の多い環境では NFS マウントに noac を設定しないようにします。また、`actimeo` は 30 秒以上のレベルに維持してください。

- **FlexCache ボリュームを作成**
  FlexCache ボリュームは、読み取り負荷の高いワークロード用にインスタントキャッシュを作成できます。`ls` などの多数の読み取りメタデータ操作を行うワークロードに対して FlexCache ボリュームを作成すると、次のような利点があります。
  - ローカルクラスタの場合は、読み取りメタデータ操作を元のボリュームからキャッシュボリュームにオフロードすることで、元のボリュームを通常の読み取り / 書き込み用に解放できます。
  - FlexCache ボリュームはクラスタ内の任意のノードに配置できるため、FlexCache ボリュームを作成すると、元のノードからメタデータの読み取り操作を移動するだけでなく、複数のノードがこれらの操作に参加できるため、クラスタノードをより効率的に使用できます。
  - WAN を介したリモートクラスタの場合、FlexCache ボリュームは、ローカライズされた NFS キャッシュを提供して WAN の待機時間を短縮し、読み取りメタデータの多いワークロードのパフォーマンスを大幅に向上させることができます。

  FlexCache ボリュームを使用してメタデータのワークロードを読み取る場合は、FlexCache を使用するノードで `fastreaddir` を無効にしてください。

```
cluster::> node run "priv set diag; flexgroup set fastreaddir=false persist
```

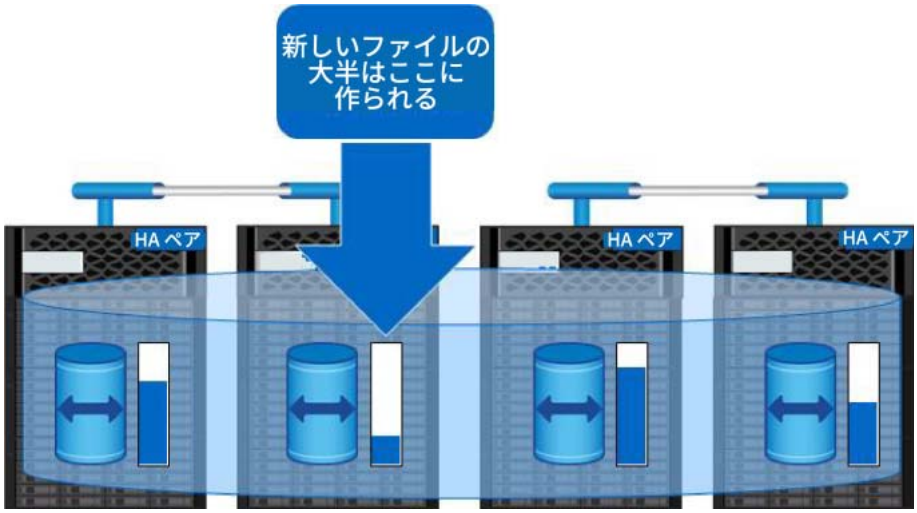
備考

- これを有効にするには、再起動またはストレージのフェイルオーバーが必要です。
- ONTAP 9.7 以降では、FlexGroup ボリュームを FlexCache ボリュームの起点にすることができます。FlexCache ボリュームの詳細については、「ETERNUS AX/HX series ONTAP での FlexCache ONTAP 9.8」を参照してください。

10.7 ファイルの削除 /FlexGroup メンバーボリュームのバランシング

FlexGroup ボリュームは、データの取り込み時に複数のメンバーボリューム間でデータを比較的均等に分散します。このデータレイアウトは、システムがより多くのハードウェアと WAFL アフィニティを使用して削除負荷をより効率的に分散し、これらの操作のためにノード当たりの CPU 使用率を少なくすることができるので、FlexVol ボリュームと比較して、FlexGroup ボリューム上でのファイル削除操作をより効率的に行うのに役立ちます。

図 10.5 サイズの大きいファイルを削除した後の容量の不均衡



ただし、FlexGroup ボリューム間のリモートアクセスのために、ファイル削除の全体的なパフォーマンスが FlexVol ボリュームと比較して低下する場合があります。まれに、ファイル (特に大きなファイルの集合) を削除すると、容量の不均衡によって FlexGroup ボリュームに人為的なホットスポットが発生することがあります。

FlexGroup ボリュームのワークロードバランスは、次の `diag-privilege-level` コマンドで表示できます。

```
cluster::*> set diag
cluster::*> node run * flexgroup show [flexgroup name]
```

次の出力が表示されます。

- メンバーボリュームデータセット ID (DSID)
- メンバーボリューム容量 (ブロック単位の使用量および使用可能量)
- 使用済みのメンバーボリューム (%)
- 緊急度、ターゲット、および確率のパーセンテージ (インジェスト計算に用いられる)

詳細は、「14.6.2 flexgroup show」を参照してください。

10.7.1 FlexGroup ボリューム内のデータの再バランシング

FlexGroup ボリューム内の既存のデータをリバランスして容量を均等にすることは、現時点ではできませんが、ほとんどの場合、その必要はありません。ONTAP は一般的に、それほどいっぱいでないメンバーボリュームに新しい書き込みがリダイレクトされるように、取り込み負荷のバランスを適切に調整します。また、ONTAP 9.8 の新しいプロアクティブなサイズ変更機能により、ONTAP は使用可能な空き領域の均等なバッファを維持するために必要に応じてメンバーボリュームを拡張および縮小するため、リバランスは必要ありません。データの不均衡があるからといって、パフォーマンスの問題が発生するわけではありません。ただし、データの不均衡に、いっぱいになったメンバーボリュームが含まれている場合を除きます。このような場合、パフォーマンスの問題は新しいファイルの作成時にのみ発生します。

メンバーボリュームが他のメンバーボリュームよりも大幅に大きくなるというまれなケースでは、ワークロードを分析して、何か変更があったかどうかを確認する必要があります (たとえば、1MB のファイルを作成していたワークロードが、100GB のファイル作成になったなど)。XCP Migration Tool を使用してフォルダーとファイルをスキャンし、ファイルサイズと異常を識別することができます。1 つのメンバーボリュームに過剰に割り当てる可能性がある一般的なシナリオの 1 つは、エンドユーザが FlexGroup 内の大量のデータを zip 圧縮する場合です。その単一の zip ファイルが非常に大きくなり、メンバーボリュームがいっぱいになることがあります。

XCP でファイルをスキャンする例については、「16.7 XCP を使用した移行前のファイルのスキャン」を参照してください。

ファイルが識別されたら、それらを削除するか、他のボリュームに移動するか、メンバーボリュームにスペースを追加するか、新しいメンバーボリュームを追加して FlexGroup ボリュームの取り込み負荷のバランスをとります。理想的には、クラスタを ONTAP 9.8 にアップグレードして、プロアクティブなサイズ変更のメリットを活用します。これにより、メンバーボリューム容量の管理オーバーヘッドを排除できます。

10.7.1.1 FlexGroup ボリュームが既存のデータをリバランスしないのはなぜですか？

FlexGroup ボリュームがデータを取り込む際には、次の 3 つの目標があります。

- ボリュームは、すべてのメンバー FlexVol ボリュームが並行してワークロードのホストに参加するように促す必要があります。メンバーボリュームのサブセットのみがアクティブな場合、FlexGroup ボリュームはアクティブでないメンバーにより多くの新しいデータを分配します。

- FlexGroup ボリュームは、他のすべてのメンバーも空き領域が不足していない限り、すべてのメンバー FlexVol ボリュームの空き領域が不足するのを防ぎます。1 つのメンバーに他のメンバーよりも多くのデータがある場合、FlexGroup ボリュームは、十分に使用されていないメンバーに新しいデータを平均よりも高い割合で配置することによって、十分に使用されていないメンバーを調整します。
- FlexGroup ボリュームは、前の 2 つの目標を追求することによって発生するパフォーマンスの低下を最小限に抑える必要があります。FlexGroup ボリュームが、新しいファイルを最も有益な場所に慎重かつ正確に配置する場合、前の 2 つの目標は容易に達成できます。しかし、このように慎重に配置するためのコストは、サービス遅延の増加として現れます。理想的な FlexGroup ボリュームでは、パフォーマンスと容量バランスを組み合わせますが、パフォーマンスが向上します。

これらの目標の一部は矛盾しているため、ONTAP は FlexGroup ボリュームのバランスを維持するために高度なアルゴリズムとヒューリスティックスを採用しています。ただし、一部のシナリオでは、次のような不均衡が発生する可能性があります。

- 大きなファイルや時間の経過とともに増大するファイルが FlexVol メンバーボリュームに存在する場合があります。
- ワークロードは、小さいファイルから大きいファイルに変更されます (監視カメラの 4 K 解像度から 8 K 解像度への録画方法の変更など)。
- 多くのファイルは、ファイル自体と同じ FlexGroup ボリューム内の 1 つのファイルに zip 圧縮または tar 処理されます。
- 大量のデータが削除され、そのデータのほとんどが同じメンバーボリュームから削除される可能性があります (まれなケース)。

FlexGroup メンバーボリュームの容量またはファイルが不均衡になっている場合、ONTAP は、割り当てられていないメンバーボリュームがいっぱいになったメンバーに追いつくのを支援するための特別な手段を講じます。その結果、新しいファイルの作成時にパフォーマンスが影響を受ける可能性があります。既存のデータにはほとんど影響しないはずです。

メンバーボリュームの使用容量が 80% に達したときのパフォーマンスの問題

ONTAP 9.8 とプロアクティブなサイズ変更は、メンバーボリュームが容量のしきい値に達したときのパフォーマンスへの影響をさらに軽減します。これは、FlexGroup ボリュームに適した ONTAP バージョンです。

10.8 メンバーボリュームがスペース不足の場合のファイルのリスト

FlexGroup メンバーボリュームの容量が不足すると、FlexGroup ボリューム全体で容量不足が報告されます。FlexGroup メンバーの容量が不足している場合は、フォルダの内容の一覧表示などの読み取り操作でも失敗することがあります。

`ls` は読み取り専用の操作ですが、FlexGroup ボリュームを正常に動作させるには、まだ書き込み可能な領域が少し必要です。ONTAP はこのストレージを使用してメタデータキャッシュを確立します。たとえば、`foo` という名前が X プロパティを持つ inode を指し、`bar` という名前が Y プロパティを持つ inode を指しているとします。使用される容量はごくわずかであり (大規模なシステムでは数 KB、場合によっては数 MB)、この容量は繰り返し使用および解放されます。内部的には、このスペースは RAL リザーブと呼ばれます。

通常の状況では、メンバーボリュームがいっぱいになった場合でも、FlexGroup ボリュームは `ls` のような読み取り専用操作を実行するため、使用できる領域が少し残ります。ただし、ONTAP は RAL リザーブよりも他の操作を優先します。たとえば、メンバーボリュームの使用率が 100% の場合、Snapshot コピーを作成し、そのボリュームを引き続き使用しようとすると、ブロックの上書き時に WAFL Snapshot リザーブが使用されるため、より多くの領域が消費されます。ONTAP は Snapshot 領域に優先順位を付け、RAL リザーブなどから領域を確保します。このシナリオは滅多に起こりませんが、`ls` のような操作がスペース不足のために失敗する理由を説明しています。

10.9 ファイル名の変更に関する考慮事項

FlexGroup ボリュームは、メタデータの多いワークロードのほとんどを適切に処理します。ただし、一度に大量のファイル名を変更するワークロード (例えば数十万) では、これらの操作のパフォーマンスは FlexVol ボリュームと比較して低下します。これは、ファイル名を変更してもファイルシステム内のファイルは移動しないためです。代わりに、ファイル名を新しい場所に移動するだけです。FlexGroup ボリュームでは、この名前の移動はリモート操作として行われ、リモートハードリンクが作成されます。その後の名前変更では、ファイルの場所へのリモートハードリンクがさらに作成されるため、そのファイルで発生するオペレーションに遅延が生じ続けます。アプリケーションのワークフローのほとんどがファイル名の変更である場合は、FlexGroup ボリュームではなく FlexVol ボリュームの使用を検討してください。名前の変更後に最終的なランディングスポットが FlexGroup ボリュームである場合は、名前の変更プロセス後にファイルを FlexVol ボリュームから FlexGroup ボリュームに移動することを検討してください。

10.10 シンボリックリンクの考慮事項

ワークロードの 1 つの FlexGroup ボリュームに多数のシンボリックリンク (つまり、数百万単位のシンボリックリンク) が含まれている場合、その多数のシンボリックリンクを解決しようとすると、パフォーマンスに悪影響を及ぼす可能性があります。この悪影響は、ONTAP が作成するリモートハードリンクに加えて、リモートハードリンクを人為的に作成することによって発生します。

ベストプラクティス 22: FlexGroup ボリューム内のシンボリックリンク
可能であれば、FlexGroup ごとのシンボリックリンク数を数千以下にしてください。

10.11 NFS バージョンに関する考慮事項

NFS を使用するクライアントが NFS バージョン (例 :`-o nfsvers=3`) を指定せずに ONTAP にボリュームをマウントしようとすると、クライアントとサーバー間でプロトコルバージョンネゴシエーションが行われます。クライアントは、サーバーがサポートする最新バージョンの NFS を要求します。サーバ (ONTAP の場合は、NFS を提供する SVM) で NFSv4.x が有効になっている場合、クライアントはそのバージョンでマウントを試みます。

ONTAP 9.7 以降では、NFSv4.x がサポートされています。ただし、これによって別の問題が発生する可能性があります。クライアントは、NFS サーバーによって通知された最新の NFS バージョンをマウントします (この場合、ONTAP SVM)。NFSv4.x バージョンが有効になっている場合、NFSv3 が必要なとき、または必要と予測されるときに、クライアントが NFSv4.x を介してマウントすることがあります。NFSv4.x をマウントすると、パフォーマンスとアクセス権の動作が NFSv3 とは異なります。

10.11.1 ネットワーク接続の同時実行性 : NFSv3

前述の考慮事項に加えて、ONTAP には、NFSv3 オペレーションに対して TCP 接続あたり 128 の同時オペレーションという制限があることに注意してください。この制限は、各 IP アドレスに対して、システムが同時に処理できる操作の数が最大 128 に制限されることを意味します。したがって、NFSv3 クライアントが FlexGroup テクノロジーの潜在能力を最大限に引き出すほどストレージシステムを強く押し上げることはできません。NFSv3 を通じて送信される (RPC スロットテーブル使用による) 同時オペレーションの数を制御するようにクライアントを構成できるため、追跡が困難なパフォーマンスの問題を回避できます。

10.11.1.1 RPC スロットテーブルの潜在的な問題の特定

最近の NFSv3 クライアントの多くは、RPC スロットテーブルに動的な値を使用します。クライアントは、1 つの TCP スレッドで最大 65,336 個の同時オペレーションを送信します。ただし、ONTAP では TCP 接続あたり 128 の同時オペレーションしか許可されないため、クライアントが 128 を超える数を送信する場合、ONTAP は NFS オペレーションでフロー制御の形式を実行し、リソースが解放されるまで NFS オペレーション (ONTAP の exec コンテキスト) をブロックすることによって、不正なクライアントによるストレージシステムのオーバーランを防止します。このフロー制御は、一般的なストレージシステム統計からは容易に明らかな理由がない可能性がある余分な遅延やジョブ完了時間を遅延を引き起こすパフォーマンスの問題として顕在化する可能性があります。これらの問題はネットワークに関連している可能性があり、ストレージ管理者が間違ったトラブルシューティングパスをたどる可能性があります。

RPC スロットテーブルが関係しているかどうかを調べるには、ONTAP パフォーマンスカウンタを使用します。オーバーラン中の接続によってブロックされた exec コンテキストの数が増加しているかどうかをチェックできます。

これらの統計情報を収集するには、次のコマンドを実行します。

```
statistics start -object cid -instance cid
```

次に、一定期間の統計を調べて、増加しているかどうかを確認します。

```
statistics show -object cid -instance cid -counter execs_blocked_on_cid
```

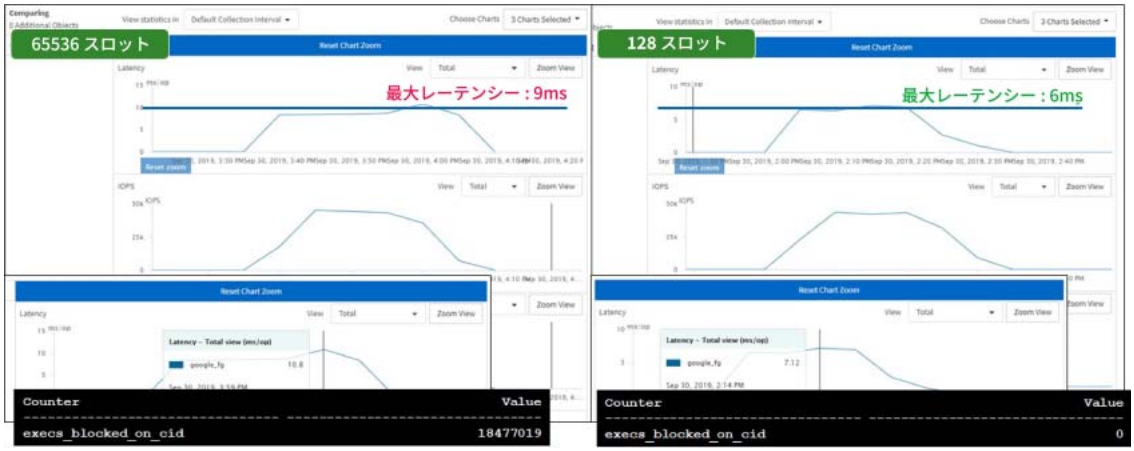
ONTAP 9.8 以降では、RPC スロットテーブルの問題の識別に役立つ新しい EMS メッセージ (`nblade.execsOverLimit`) が追加されています。この EMS は、execs_blocked_on_cid カウンタが一定の時間内に一定の値を超えたときにトリガーされます。イベントにこのメッセージが表示された場合は、当社サポート部門に問い合わせるか、NFSv3 クライアントで使用されるスロットテーブルの数を減らす方法を調べてください。

10.11.1.2 パフォーマンスに対する RPC スロットテーブルの影響の例

次の例では、18 万のサブディレクトリに 1800 万のファイルを作成するスクリプトを実行しています。この負荷生成は、3 つのクライアントから同じ NFS マウントに対して実行されました。目標は、ONTAP でフロー制御シナリオを実行するのに十分な NFS オペレーションをデフォルトの RPC スロットテーブルが設定されたクライアントで生成することでした。次に、RPC スロットテーブルを 128 に設定して、同じスクリプトを同じクライアント上で再度実行しました。

その結果、デフォルトのスロットテーブル (65,536) では 1800 万の `execs_blocked_on_cid` イベントが生成され、RPC スロットテーブルの低い設定 (128) で実行した場合と比べて、ワークロードに 3 ミリ秒の遅延が追加されました。

図 10.6 NFSv3 のパフォーマンスに対する RPC スロットテーブルの影響



3 ミリ秒はそれほど長い遅延ではないように見えるかもしれませんが、何百万もの操作を追加する可能性があり、ジョブの完了が大幅に遅くなります。

10.11.1.3 RPC スロットテーブルの問題の解決

ONTAP は、NFSv3 オペレーションに対して TCP 接続ごとにクライアントが送信するスロットテーブルの数を制御できません。したがって、NFS 経由で送信される最大スロットテーブルを 128 に制限するようにクライアントを構成する必要があります。この設定は、クライアントの OS バージョンによって異なります。詳細については、クライアントのベンダーに問い合わせてください。

同じクライアント上のクラスタ内の異なる IP アドレスに、より多くのマウントポイントを接続することで、クライアントの NFS 接続のパフォーマンスを向上させることができますが、この方法では複雑になります。たとえば、`SVM:/volumename` にボリュームをマウントするのではなく、ボリューム内の異なるフォルダおよび IP アドレスにまたがって、同じクライアント上に複数のマウントポイントを作成できます。

例：

```
LIF1:/volumename/folder1
LIF2:/volumename/folder2
LIF3:/volumename/folder3
```

また、同じ TCP 接続上で NFSv3 の多重化を実行できる一部の Linux ディストリビューションで使用できる nconnect オプションを使用する方法もあります。このオプションにより、使用可能な同時セッションが増え、全体的なパフォーマンスが向上します。

10.11.1.4 RPC スロットテーブルの制限は他の NAS プロトコルに影響しますか？

RPC スロットテーブルの制限は、NFSv3 トラフィックにのみ影響します。

- SMB クライアントは、SMB マルチチャネル、SMB 多重化、SMB クレジットなど、異なる接続方法を使用して同時実行します。SMB 接続の方法は、クライアントとサーバの構成とプロトコルのバージョンによって異なります。たとえば、SMB1.0 は SMB 多重化 (mpx) を使用し、SMB2.x は SMB クレジットを使用します。
- NFSv4.x クライアントは RPC スロットテーブルを使用せず、代わりに状態 ID とセッションテーブルを使用して、クライアントからの同時トラフィックのフローを制御します。

10.11.2 NFS 書き込みの追加

ONTAP は、これらの書き込み追加の並列処理を提供し、ファイルサイズに関係なく書き込み追加のパフォーマンスを向上します。

10.11.3 nconnect

nconnect は、一部の Linux ディストリビューションで使用可能なマウントオプションです。このオプションは、マウントごとに使用する TCP 接続の数を指定し、クライアントごとの一部のワークロードでパフォーマンスを大幅に向上させます。通常は、ネットワークスレッドがワークロードのボトルネックになっている場合のみです。これは、クライアントがマウントごとにより多くの RPC スロットテーブルを利用できるようにすることで、ONTAP にも利点をもたらします。「10.11.1 ネットワーク接続の同時実行性：NFSv3」を参照してください。

NFS クライアントもサポートしていれば、ONTAP 9.8 は NFS マウントでの nconnect の使用をサポートします。nconnect を使用する場合は、クライアントのバージョンで提供されているかどうかを確認し、ONTAP 9.8 以降を使用します。

表 10.5 は、異なる nconnect スレッド値を使用した単一の Ubuntu クライアントからの結果を示しています。

表 10.5　nconnect のパフォーマンス結果

nconnect 値	プロセスあたりのスレッド数	スループット	差
1	128	1.45GB/s	-
2	128	2.4GB/s	+66%
4	128	3.9GB/s	+169%
8	256	4.07GB/s	+181%

10.11.4 NFS 接続クライアントのボリューム名へのマッピング

クラスタからマウントされている NFS のバージョンを確認するには、ONTAP 9.7 で使用可能な `nfs connected-clients show` コマンドを使用します。

```
cluster::> nfs connected-clients show -node * -vserver DEMO

     Node: node1
  Vserver: DEMO
  Data-Ip: 10.x.x.x
Client-Ip       Volume-Name      Protocol Idle-Time     Local-Reqs Remote-Reqs
--------------- ---------------- -------- ------------- ---------- -----------
10.x.x.x        CIFS             nfs4.1   2d 0h 9m 3s   153        0
10.x.x.x        vsroot           nfs4.1   2d 0h 9m 3s   0          72
10.x.x.x        flexgroup_16__0001
                                 nfs3     0s            0          212087
10.x.x.x        flexgroup_16__0002
                                 nfs3     0s            0          192339
10.x.x.x        flexgroup_16__0003
                                 nfs3     0s            0          212491
10.x.x.x        flexgroup_16__0004
                                 nfs3     0s            0          192345
10.x.x.x        flexgroup_16__0005
                                 nfs3     0s            212289     0
```

NFSv4.x が有効になっている環境で FlexGroup ボリュームをマウントする際の問題を回避するには、`fstab` を使用して NFSv3 のデフォルトマウントバージョンを使用するようにクライアントを構成するか、マウント時に NFS バージョンを明示的に指定します。

例：

```
# mount -o nfsvers=3 demo:/flexgroup /flexgroup
# mount | grep flexgroup
demo:/flexgroup on /flexgroup type nfs (rw,nfsvers=3,addr=10.193.67.237)
```

10.11.5 FlexGroup ボリュームでの NFSv4.x の有効化と使用

環境内で NFSv4.x を設定すると、FlexGroup ボリュームは FlexVol ボリュームと同じように機能します。パフォーマンスを重視するのではなく、NFSv4.x をワークロードで使用するメリットは次のとおりです。

- **セキュリティ**
  NFSv4.x は、補助プロトコル (NLM、NSM、mountd、portmapper など) を 2049 以上の単一ポートに統合することで NFS のセキュリティを大幅に向上させています。開いているファイアウォールポートが少ないほど、利用可能な脅威ベクトルが少なくなります。
  さらに、NFSv4.x には、RFC 要件の一部として Kerberos 暗号化 (krb5、krb5i、および krb5p) が含まれています。これは、Kerberos がサポートされていない限り、クライアント / サーバが RFC に準拠していないことを意味します。
  NFSv4.x では、クライアントとサーバーの構成でドメイン ID を一致させる必要があるため、UID/GID 情報のマスキングが改善され、特に Kerberos 暗号化を使用する場合に、ユーザーのなりすましが困難になります。
  最後に、NFSv4.x は、Windows NTFS ACL の機能を模倣した詳細な ACL サポートを提供します。これにより、モードビットで提供される NFSv3 よりも多くのユーザーとグループを ACL に追加できるだけでなく、基本的な読み取り / 書き込み / 実行 (rwx) を超える ACL 機能をより多く使用できます。NFSv4.x ACL は、NFSv3 のみをマウントするデータセットにも適用できます。これにより、NFSv4.x を使用していない場合でも、ファイルおよびフォルダのセキュリティをより詳細に設定できます。
- **ロックの改善**
  NFSv3 ロックは、NSM や NLM などの補助プロトコルを使用して、NFS プロトコルの外部で実行されました。このため、クライアントまたはサーバが停止したときに古いロックが発生し、古いロックが解除されるまでファイルへのアクセスができなくなることがよくありました。
  NFSv4.x では、リースメカニズムを使用してロックが強化されています。このメカニズムでは、指定した時間だけリースが保持され、クライアント / サーバ間の通信に影響がない場合はリースが維持されます。通信に問題がある場合 (ネットワークまたはサーバーの停止)、リースは期限切れになり、ロックが再確立されるまで解放されます。
  さらに、NFSv4.x でのロックは NFS パケット内に統合されており、NFSv3 よりも信頼性が高く効率的なロックの概念を提供します。
- **データのローカリティと並列アクセス**
  NFSv4.x は、NFSv4.x 参照などのスケールアウト NAS 環境向けのデータローカリティ機能を提供します。これにより、ノード上の場所に従って ONTAP 内のボリュームにマウント要求をリダイレクトし、マウントへのローカルアクセスを確保できます。
  NFSv4.1 では、マウント時にメタデータサーバを確立し、ネームスペース間でデータ I/O をリダイレクトする並列 NFS サポートも提供されています。そのために、ノードとデータ LIF の位置に従ってデータを追跡するクライアント / サーバー通信を使用します。この概念は、SAN における非対称論理ユニットアクセス (ALUA) の概念に似ています。詳細は、「10.11.5.2 FlexGroup ボリュームによる pNFS」を参照してください。

10.11.5.1 ONTAP での NFSv4.x パフォーマンスの強化

一般に、NFSv4.x は NFSv3 よりもパフォーマンスが低くなります。これは、NFSv4.x がステートフルであるため、プロトコル操作ごとに多くの処理が必要になるためです。NFSv4.x オーバーヘッドは、ロックとリース、ACL、複合呼び出し、クライアントとサーバ間の状態 ID の通信、および各パケットの処理という形で発生します。

NFSv4.x のパフォーマンスの弱点の 1 つに、メタデータの取り込みが多いワークロードがあります。FlexGroup ボリュームは、このようなタイプのワークロードに最も適しているため、これらのワークロードに NFSv4.x を使用することを検討している場合は、FlexGroup ボリュームの使用を強く推奨しています。

NFSv4.x を使用する利点の 1 つは、操作で RPC スロットテーブルを使用しないため、RPC スロットの枯渇の影響を受けないことです。

NFS で Kerberos を使用している場合は、暗号化されたパケットのオーバーヘッドを処理する操作のパフォーマンスにもわずかな影響があります。影響は、次のようないくつかの要因によって異なります。

- ONTAP バージョン
- 使用しているハードウェア
- ネットワーク遅延、WAN 遅延、およびクラウド領域
- クラスタのパフォーマンスヘッドルーム
- 使用されている Kerberos 暗号化 (krb5、krb5i、または krb5p)

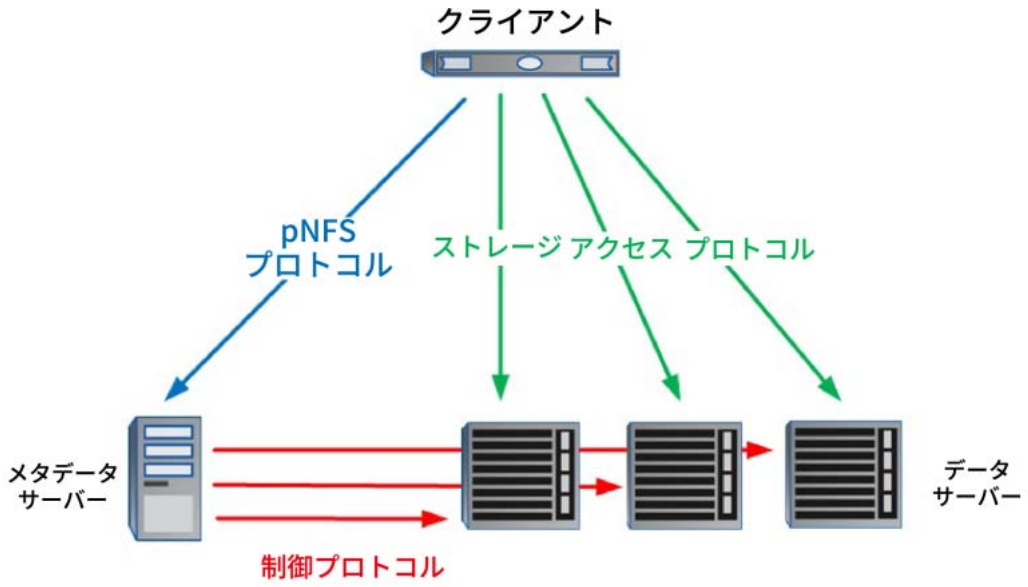
10.11.5.2 FlexGroup ボリュームによる pNFS

ONTAP は、NFSv 4.1 とその RFC 必須機能を含む NFSv4.x のサポートを提供します。これらの機能には並列 NFS (pNFS) が含まれており、クラスタ内の複数のボリュームおよびノード間で読み取りと書き込みのローカライズを行うことができます。ONTAP は、pNFS のファイルバージョンを提供しますが、ストライピングやブロックバージョンの機能は使用しません。

ONTAP での pNFS の動作

SVM の NFS サーバーで pNFS が有効になっている場合、pNFS をサポートし、NFSv 4.1 を使用してマウントするクライアントは、メタデータサーバーとして機能する単一の TCP 接続を使用して、最初にクラスター内の特定のノードに接続します。この接続は、データレイアウト用のクライアント / サーバー通信、LIF ロケーション、および I/O トラフィックをクラスター内のローカルボリュームおよびデータ LIF にリダイレクトするのに役立つ pNFS マッピングなどの pNFS 操作を処理します。メタデータサーバーは、`getattr` 操作や `setattr` 操作などの NFS メタデータ操作も行います。

図 10.7 pNFS



pNFS アーキテクチャーには、次の 3 つの主要コンポーネントがあります。

すべての非データ I/O トラフィックを処理するメタデータサーバー。GETATTR、SETATTR、LOOKUP、ACCESS、REMOVE、RENAME 操作などのすべてのメタデータ操作を担当します。メタデータサーバーは、ファイルのレイアウトに関する情報も提供します。

- **ファイルデータを格納し、クライアントの読み取り / 書き込み要求に直接応答するデータサーバ**
  データサーバは、純粋な読み取り / 書き込み I/O を処理します。
- **データサーバーに直接アクセスできる 1 つ以上のクライアント**
  このアクセスは、メタデータサーバーから受信したメタデータに基づいて行われます。

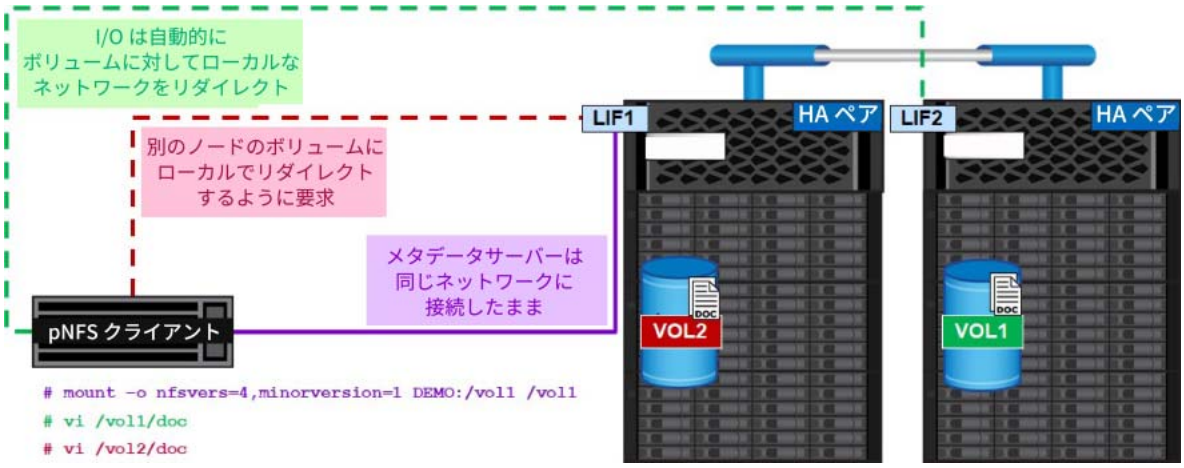
クライアント、メタデータサーバー、およびデータサーバー間で使用されるプロトコルには、次の 3 種類があります。

- **メタデータサーバーとデータサーバーの間で使用される制御プロトコル**
  このプロトコルは、ファイルシステムデータを同期化します。
- **クライアントとメタデータサーバー間で使用される pNFS プロトコル**
  これは基本的に NFSv4.1 プロトコルであり、pNFS 固有の拡張がいくつかあります。これは、多数のデータサーバに格納されているファイルにアクセスするために必要な場所とストレージアクセスプロトコルを記述するメタデータを含むレイアウトを取得および操作するために使用されます。
- **クライアントがデータサーバーに直接アクセスするために使用するストレージアクセスプロトコルのセット**
  現在、pNFS 仕様には、ファイルベース、ブロックベース、オブジェクトベースの 3 つのストレージプロトコルのカテゴリがあります。

読み取り要求または書き込み要求が pNFS を介してクライアントによって実行されると、クライアントとサーバーはデータレイアウトマッピングを使用して、これらの要求の送信先をネゴシエートします。たとえば、ファイルがクラスタ内の volume1 (これは node1 に存在します) に存在するが、メタデータサーバーが node2 に接続されている場合、データレイアウトマッピングは node1 に対してローカルなネットワーク接続を介して読取り / 書込みを実行するようにクライアントに通知します。

ボリュームが移動されると (たとえば、無停止のボリューム移動オペレーションの場合)、データレイアウトテーブルが更新され、ONTAP は次の要求でローカルトラフィックをボリュームにリダイレクトします。このプロセスは、ALUA が SAN 環境で機能する方法に似ています。SAN 環境では、ブロックデバイスのローカリティに基づいてパスを切り替えることができます。

図 10.8　　pNFS 操作



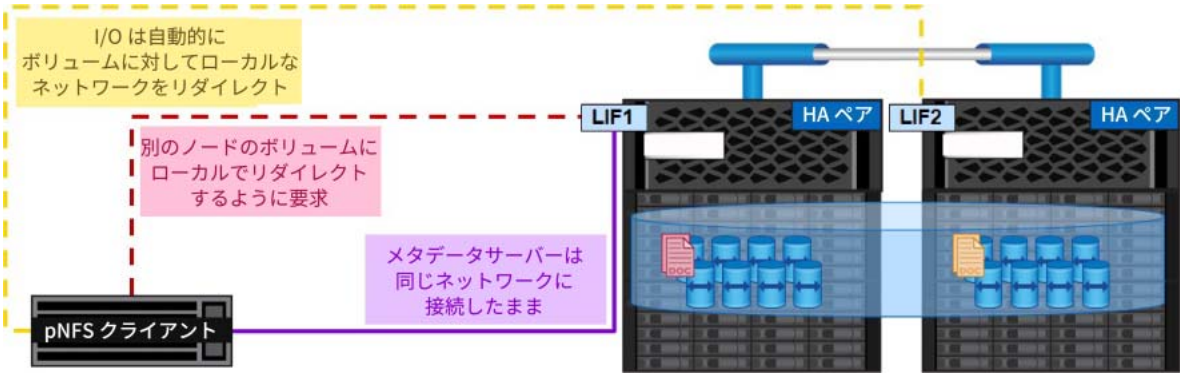
FlexGroup ボリュームでの pNFS の動作

FlexGroup ボリュームは単一のエンティティとして動作しますが、複数の FlexVol メンバーボリュームで構成されます。各メンバーボリュームには、ボリューム間でストライプ化されていない一意のファイルが含まれています。NFS 操作が FlexGroup ボリュームに接続すると、ONTAP はクラスタネットワーク経由で操作のリダイレクトを処理します。

pNFS では、データレイアウトマッピングがメンバーボリュームの場所とローカルネットワークインタフェースを追跡するため、これらのリモート操作が削減されます。クライアントが 1 つのネームスペースしか認識しない場合でも、読み取り / 書き込みは FlexGroup ボリューム内のローカルメンバーボリュームにリダイレクトされます。このアプローチにより、よりシームレスで管理が容易なスケールアウト NFS ソリューションが実現します。また、クラスタネットワークトラフィックが削減され、ノード間でのデータネットワークトラフィックがより均等に分散されます。

FlexGroup pNFS は、FlexVol pNFS とは少し異なります。ファイルを開くためのメタデータサーバー間の FlexGroup ロードバランシングは行われますが、pNFS は異なるアルゴリズムを使用します。pNFS は、ターゲットファイルがあるノードにトラフィックを送信しようとします。ノードごとに複数のデータ LIF が与えられる場合、各 LIF に接続することができますが、ネットワークインターフェースごとのボリュームにトラフィックを送るために使用されるのはセットの LIF の 1 つだけです。

図 10.9 pNFS 操作：FlexGroup ボリューム



pNFS のベストプラクティス

ONTAP における pNFS のベストプラクティスは、通常の NAS のベストプラクティスとほとんど変わりませんが、ここではいくつかの点に注意してください。一般的には、

- サポートされている最新のクライアント OS バージョンを使用します。
- サポートされている最新の ONTAP パッチリリースを使用します。
- すべてのノードのデータのローカリティを保証するために、ノードごと、SVM ごとにデータ LIF を作成します。
- NFSv4.1 はステートフルなプロトコルであり、LIF の移行では NFS の状態が再確立されるときに短時間の停止が発生する可能性があるため、メタデータサーバーのデータ LIF で LIF の移行を使用しないでください。
- 複数の NFSv 4.1 クライアントがマウントされている環境では、メタデータサーバー接続を複数のノード間でバランスさせて、メタデータ操作が 1 つのノードまたはネットワークインタフェースに集中しないようにします。
- 可能であれば、SVM 内の同じノードで複数のデータ LIF を使用しないでください。
- 一般に、NFSv3 と NFSv4.x を同じデータセットにマウントすることは避けてください。これを避けることができない場合は、ロックが正しく管理されていることをアプリケーションベンダーに確認してください。
- pNFS で NFS 参照を使用している場合は、参照によってローカルメタデータサーバーが確立されますが、データ I/O はリダイレクトされることに注意してください。FlexGroup ボリュームでは、メンバーボリュームが複数のノード上に存在する場合があるため、NFS 参照はあまり使用されません。代わりに、DNS 負荷分散を使用して接続を分散します。

10.11.5.3 NFSv4.x の一般的な考慮事項

SVM に対して NFSv4.x を検討する場合は、導入前に必ずパフォーマンス、クライアントとアプリケーションのサポート、ネームサービスのインフラストラクチャ、ロックメカニズムを考慮してください。また、アプリケーションが同じデータセット上で NFSv3 と NFSv4.x の両方を使用できるかどうかを検討してください。例えば VMware では、両方のプロトコルバージョンでデータストアのサービスを行わないことを推奨しています。

可能であれば、本番環境に導入する前に、機能とパフォーマンスのテストを実行するために別の SVM を設定します。

NFSv4.x 構成を正しく機能させるには、通常、次のものが必要です。

- NFSv4.x をサポートする NFS クライアント。
- NFSv4.x を指定した NFS マウント。
- NFS サーバーの構成 (NFSv4.x および必要な機能の有効化 : クライアントと NFS サーバで同一に構成された参照、pNFS、ACL サポート、NFSv4 ID ドメインなど)。
- クライアントとサーバーでユーザー名とグループが一致しています (大文字と小文字を区別。たとえば、user1@domain.com はサーバーとクライアントの両方に存在する必要があります。USER1 と user1 は一致しているとは見なされません)。
- オプション。UNIX ID 用のネームサービス (NIS や LDAP など) を使用すると、NFSv4.x の実装と機能を大幅に簡素化できます。

10.11.6 FlexGroup ボリュームにおける NAS メタデータの影響

メタデータ操作のオーバーヘッドはワークロードの実行方法に影響し、リモート操作のパフォーマンスは 10% から 30% の範囲で低下します。メタデータの影響のほとんどは、メタデータの書き込みに関連しています。ほとんどの読み取りメタデータはほとんど影響しません。

- `getattr`、`access`、`statfs`、`lock`、`unlock`。FlexGroup のオーバーヘッドはほとんどありません。
- `nfs create`、`unlink`、`lookup`。負荷が高い場合、FlexGroup のオーバーヘッドはほとんどないか、まったくありません。
- `nfs mkdir`、`rmdir`、`lookup dir`。50% ～ 100% のリモートアクセス、高いオーバーヘッド。
- `CIFS open/close`。高いオーバーヘッド。

10.12 CIFS/SMB に関する考慮事項

FlexGroup ボリュームは、NFS ワークロードと SMB ワークロードの両方をサポートします。ONTAP SMB サーバは、FlexGroup ボリュームと FlexVol ボリュームの両方で SMB ワークロードの全体的なパフォーマンス向上に役立つ機能をいくつか提供します。次のセクションでは、これらの機能の一部と、FlexGroup ボリュームまたはファイル数の多い環境に適用される注意事項について説明します。

10.12.1 SMB バージョンに関する考慮事項

FlexGroup ボリュームは、SMB 2.x バージョンおよび SMB 3.x バージョンのみをサポートします。SMB 1 バージョンでは、FlexGroup ボリュームを指す CIFS/SMB 共有にアクセスできません。SMB 1 は Microsoft によって非推奨にされているため、FlexGroup ボリュームに SMB 1 のサポートを追加する予定はありません。SMB サポート情報の詳細については、「5. FlexGroup 機能のサポートと最大化」を参照してください。

CIFS/SMB ワークロードを FlexGroup に移行する前に、既存のワークロードに SMB 1 クライアントが接続されていないことを確認する必要があります。ONTAP では、次のコマンドを使用してこれを実行できます。

```
cluster::> cifs session show -protocol-version SMB1
```

FlexGroup ボリュームに SMB 1 アクセスを試みると、`Nblade.flexgroupStatefulProtocolAccess` EMS イベントがログに記録されます。

10.12.2 SMB での変更通知の使用

SMB 変更通知は、SMB クライアントが、セッションを閉じたり画面を更新 (F5 キーを押すなど) したりすることなく、SMB 共有内にファイルが存在することを通知される方法です。SMB クライアントは SMB セッション中に SMB サーバと常に通信し、SMB サーバは共有内のファイルの変更に関する定期的な更新をクライアントに送信します。この機能は、ファイルを書き込み、SMB 共有内のファイルをすぐに読み取れるようにする必要があるアプリケーションに最も便利です。変更通知は、`changenotify` 共有プロパティによって制御されます。ONTAP は変更通知が不要な場合でも、新しい SMB 共有にこの共有プロパティを自動的で設定します。

ベストプラクティス 23: SMB 変更通知の推奨事項
SMB 変更通知を使用する必要がある場合は、ONTAP 9.7 以降を使用してください。

10.12.3 大容量 MTU

大容量 MTU は SMB の Maximum Transmission Unit (MTU) を 64KB から 1MB に拡張できるため、処理が必要なパケット数を削減することで、大容量ファイル転送の速度と効率が大幅に向上します。大きな MTU を有効にするには、高度な権限コマンド `cifs options modify -is-large-mtu -enabled true` を使用します。

ONTAP の CIFS/SMB サーバでこの機能を有効にした場合、クライアントと SMB プロトコルのバージョンでサポートされていれば (SMB 2.1 以降)、MTU サイズのネゴシエーションが自動的に行われます。

ONTAP SVM で使用している MTU が大きいかどうかを確認するには、次のコマンドを使用します。

```
cluster::> cifs session show -is-large-mtu-enabled true
```

備考

大規模な MTU は、SMB 2.1 以降のサーバで許可される大規模な読み取り / 書き込みを意味します。ネットワーク層による MTU サイズではありません。

10.12.4 SMB マルチチャネル

ONTAP では、SMB マルチチャネルがサポートしています。これは SMB 3.0 プロトコルの機能で、SMB 3.x クライアントが単一のネットワークインタフェースカード (NIC) または複数の NIC を介して接続プールを確立し、それらを使用して単一の SMB セッションに対する要求を送信できるようにします。これは NFS の nconnect 機能に似ています。

これにより、この機能を使用しないクライアントよりも、単一クライアントのパフォーマンスを大幅に向上させることができます。

SMB マルチチャネルは、次の高度な権限コマンドを使用して有効にすることができ、新しい SMB セッションで有効になります。

```
cluster::*> cifs options modify -is-multichannel-enabled true
```

ONTAP SVM の CIFS/SMB セッションが SMB マルチチャネルを使用しているかどうかを確認するには、次のコマンドを使用します。

```
cluster::> cifs session show -connection-count >1
```

Windows クライアントでは、Get-SmbMultichannelConnection PowerShell コマンドレットでマルチチャネルが使用されているかどうかを確認できます。

10.12.5 継続的に利用可能な共有 (CA 共有)

CA 共有は、スケールアウト、永続的なハンドル、監視、透過的なフェイルオーバーなどの SMB 3.x 機能を使用することにより、SMB 接続がストレージのフェイルオーバーに中断されることなく対応できるようにします。CA 共有は、正式には SQL および Hyper-V ワークロードでのみサポートされています。

CA 共有は、次のコマンドを使用して CIFS/SMB 共有レベルで設定します。

```
cluster::*> cifs share properties add -share-name SQL -share-properties continuously-
available
```

FlexGroup ボリュームは CA 共有をサポートし、Hyper-V および SQL ワークロードについてのみ正式に認定されています。ただし、このサポートには注意事項があります。

- 大規模なデータベースファイルが少ない SQL Server ワークロードは、FlexGroup ボリュームに適していない場合があります。ただし、多数のファイル (ログまたはデータベース) を含む SQL Server ワークロードは、FlexGroup ボリュームと CA 共有の適切な使用例です。「10.14 FlexGroup ボリューム上のデータベース」も参照してください。
- Hyper-V ワークロードは、CA 共有で公式にサポートされていると記載されていますが、ONTAP 9.8 の時点では、FlexGroup ボリュームで公式にサポートされているのは VMware 仮想化ワークロードだけです。Hyper-V ワークロードは、CA 共有と FlexGroup ボリュームで使用できますが、VMware ワークロードと同じテストおよび認定は行われていません。
- 仮想ハードディスクワークロード (FSLogix VHDx プロファイルなど) は、FlexGroup ボリュームで使用でき、CA 共有で動作しますが、公式にはテストまたは認定されていません。場合によっては、これらのワークロードをホストするために CA 共有が必要ないため、本番環境に展開する前にテストを行う必要があります。

一般に、CA 共有はメタデータ集約型の SMB ワークロード (ホームディレクトリなど) では使用しないでください。パフォーマンスの問題が発生する可能性があります。CA 共有を使用している場合は、homedirectory、branchcache、access-based-enumeration、attribute cache などの他の共有プロパティを設定しないでください。

次のコマンドを使用すると、CA 共有を使用している CIFS/SMB セッションを確認できます。

```
cluster::*> cifs session show -continuously-available Yes|Partial
```

10.12.6 その他の考慮事項

SMB 8.3 の短い名前や CIFS のシンボリックリンクを使用した名前変更などの特定のタスクの実行中に、特定のシナリオで発生する可能性のある問題は他にもいくつかあります。これらの問題のリストを次に示します。リンクは、これらの問題が解決された ONTAP リリースを示しています。最良の結果を得るには、ONTAP の最新のパッチリリースを実行することをお勧めします。

10.13 仮想化ワークロードに関する考慮事項

ONTAP 9.8 は、VMware 仮想化ワークロードを正式にサポートする最初のリリースです。つまり、FlexGroup ボリュームを使用して VMware NFS データストアをプロビジョニングし、クラスタ内の複数のノード間で拡張して、仮想マシンに 100TB 以上を提供できます。

FlexGroup ボリュームを使用した拡張性の高い VMware データストアは、FlexVol ボリュームよりも優れています。

- 1 つの NFS データストアで最大 20PB および 4000 億ファイル (VMware の制限により、この量が削減される可能性がある)
- ONTAP の SIS クローンとテンプレートキャッシュを使用した高速 VM クローン作成

ただし、FlexGroup ボリューム上の VMware データストアは、現時点では VMware VVol や VMFS ファイルシステムをサポートしていません (FlexGroup は NAS のみ)。

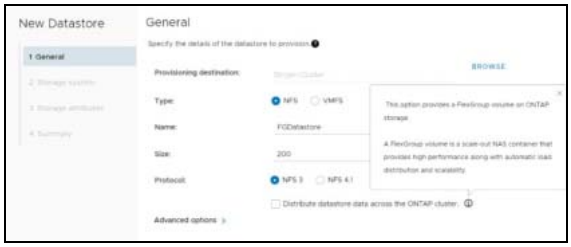
10.13.1 ONTAP tools for VMware vSphere(以前の Virtual Storage Console)

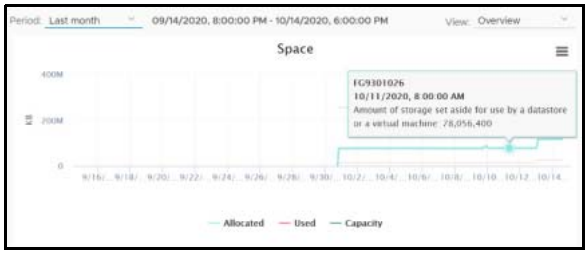
ONTAP tools for VMware vSphere の新しいリリースでは、データストアのプロビジョニングと管理の方法も提供されます。この機能には次のようなものがあります。

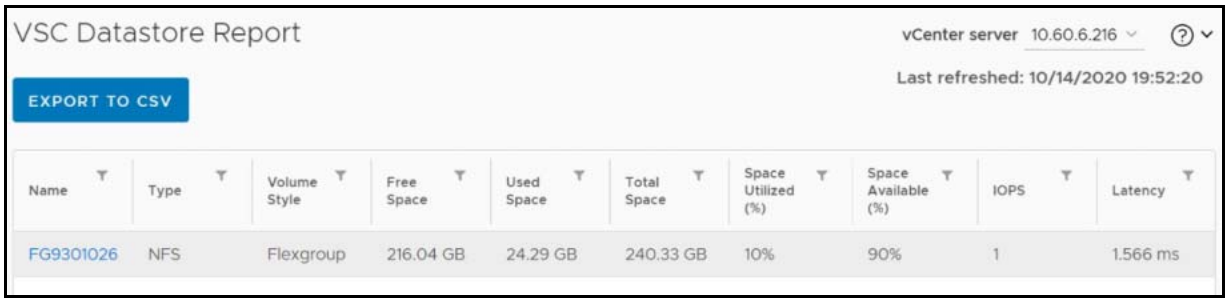
- FlexGroup または FlexVol としてのデータストアプロビジョニング
- VM レベルまでの QoS ポリシー管理
- VM レベルのパフォーマンス・メトリック
- SnapCenter による vSphere のサポート

次の図は、使用可能な機能の一部を示しています。

図 10.10 ONTAP tools for VMware vSphere：FlexGroup データストア







Name	Type	Volume Style	Free Space	Used Space	Total Space	Space Utilized (%)	Space Available (%)	IOPS	Latency
FG9301026	NFS	Flexgroup	216.04 GB	24.29 GB	240.33 GB	10%	90%	1	1.566 ms

10.13.2 コピーオフロード

ONTAP 9.8 は、クローン作成を高速にするために、FlexVol データストアよりも高速にするために、最適化されたコピーオフロード (VAAI) をサポートしていますが、留意すべきいくつかの制限があります。

- コピーオフロード操作にかかる時間は一定ではありません (ファイルサイズに比例)。
- 1 ノードあたりの並列オンデマンドジョブ数は 50 に制限されています。その上限を超えると、キューに入れられます。
- テンプレートファイルを各メンバにコピーする必要があるため、領域使用のオーバーヘッドが増加します。その結果、十分な空き領域がない FlexGroup ボリュームでのオフロード操作が失敗することがあります。
- Snapshot の作成は、コピーが完了するまで許可されません。
- データプルスキャンが完了するまでレプリケーションは実行できません。

10.13.3 考慮事項

ONTAP は仮想化ワークロードのテストと認定を行っており、ONTAP 9.8 はこのようなタイプのデータセットに対応するためのプロアクティブなサイズ変更などの機能を提供していますが、いくつかの考慮事項があります。

- FlexGroup データストアは、ONTAP tools for VMware vSphere を使用してプロビジョニングすることができます (推奨)。また、CLI または ONTAP System Manager を使用して手動で作成し、VMware vSphere を使用してマウントすることもできます。

- 仮想マシンと Snapshot コピーは、最初は小さなファイルとして作成され、時間の経過とともに増大します。ONTAP は、容量のしきい値に達すると個々のメンバーボリュームのサイズを変更して、使用可能な空きスペースのバランスを自動的に均等に維持します。この結果、一部のメンバーボリュームが他のメンバーボリュームよりも大きくなり、正常になります。
- 認定テストは、FlexGroup データストアで最大 1500 台の VM に対して実施されました。これは厳密な制限ではありませんが、テストした制限を超えると、予期しない結果になる可能性があります。
- FlexGroup データストアのサイズを設定する場合は、FlexGroup が複数の小さい FlexVol ボリュームで構成されており、大きいネームスペースが作成されることに注意してください。したがって、データストアのサイズは、最大の仮想マシンのサイズの 8 倍以上にしてください。たとえば、環境内に 6TB の VM がある場合、FlexGroup データストアのサイズは 48TB 以上にします。
- FlexGroup ボリュームは、ONTAP 9.7 以降で VAAI をサポートしています。これは、vSphere からストレージへのコピー操作の負荷を軽減するために使用されます。コピーオフロードは常にホストコピーよりも高速であるとは限らず、vSphere は NFS ストレージのコールド VM 上の操作をオフロードするだけであることに注意してください。
- 仮想化ワークロードを使用すると、同じ FlexGroup データストア内の VMDK ファイルを、クラスタ全体の複数の FlexVol メンバーボリュームに配置できます。そのため、SnapCenter for vSphere を使用してスナップショットとレプリケーションを調整します。
- FlexGroup ボリュームは、Veeam や Rubrik などの仮想化バックアップ製品とともに使用できます。SnapDiff 2.0 以降に関するサポートおよび操作のレベルについては、これらのベンダーにお問い合わせください。
- FlexGroup ボリュームは、VMware データストアについてのみテストおよび認定されています。Hyper-V、Citrix Xen、RedHat KVM などの使用はテストまたは認定されておらず、公式にはサポートされていません。
- 仮想ハードディスク (VHD) ファイルの FlexGroup ボリュームへの配置は、仮想化プロバイダーに関係なくサポートされます。
- SIS クローンのサポートは ONTAP 9.8 で追加されましたが、VMware vSphere API 経由でのみ使用できます。
- VMware サポートの制限が適用されます (例えば、pNFS なし)。
- VMware と当社は現在、一般的なマルチパスネットワーキングのアプローチをサポートしていません。NFSv4.1 では、当社は pNFS をサポートしますが、VMware はセッショントランキングをサポートします。NFSv3 では、ボリュームへの複数の物理パスはサポートされていません。ONTAP 9.8 を使用する FlexGroup の場合、間接的なアクセスの影響は通常最小限であるため (マイクロ秒)、ONTAP tools for VMware vSphere を単一のマウントにすることを推奨します。ラウンドロビン DNS を使用して、FlexGroup 内の異なるノード上の LIF 間で ESXi ホストを分散することは可能ですが、そのためには、ONTAP tools for VMware vSphere を使用せずに FlexGroup を作成してマウントする必要があり、パフォーマンス管理機能は使用できません。
- ONTAP tools for VMware vSphere 9.8 を使用して、ONTAP メトリック (ダッシュボードと VM レポート) から FlexGroup VM のパフォーマンスを監視し、個々の VM の QoS を管理します。これらのメトリックは、現在のところ ONTAP コマンドまたは API では使用できません。
- SnapCenter for vSphere リリース 4.4 は、プライマリストレージシステム上の FlexGroup データストア内の VM のバックアップとリカバリをサポートしています。SnapMirror を手動で使用して FlexGroup をセカンダリシステムに複製できますが、SCV 4.4 はセカンダリコピーを管理しません。

10.14 FlexGroup ボリューム上のデータベース

通常、データベース (Oracle など) は配備時に小さなファイルをいくつか作成します。FlexGroup ボリュームでは、少数の小さいファイルは親フォルダにローカルに配置される傾向があります。つまり、8 つのデータベースを配置した Oracle 環境は、すべて同じ FlexGroup メンバーボリューム内に配置される可能性があります。この方法では、クラスタ内のノード間での負荷分散のメリットが得られないだけでなく、時間の経過とともにファイルが増大するために問題が発生する可能性があります。最終的には、ファイルがメンバーボリュームの容量いっぱいになり、データを移動するための修復手順が必要になります。

理論的には、データベースのワークロードは 1 つのクラスタにまたがる単一のネームスペースでうまく動作します。ただし、ファイルは時間の経過とともに増大する可能性が高く、遅延の影響を受けやすいデータベースはクラスタネットワークを横断するボリュームで実行される可能性があるため、現在、データベースファイルを FlexVol ボリュームに配置することを推奨しています。

備考

ONTAP 9.8 はプロアクティブなサイズ変更機能を備えており、大きなファイルや増大するファイルをホストする際の問題を軽減し、FlexGroup ボリュームのデータベースワークロードをより現実的なものにします。

10.15 FlexCache ボリュームの考慮事項

ONTAP では、FlexCache をサポートしています。この機能により、NAS ワークロードのパフォーマンスを向上させ、クラスタ内または WAN 内のボリュームホット・スポットを防止することができるスパースボリュームが提供されます。FlexCache のキャッシュボリュームは FlexGroup ボリュームにより強化され、ポインタとブロックをリダイレクトする基本プロトコルはリモートアクセスレイヤー (RAL) です。RAL は、FlexGroup ボリュームを FlexGroup ボリュームにするものでもあります。ONTAP 9.7 では、FlexCache の FlexGroup 元ボリュームのサポートが追加されました。

ONTAP 9.8 では、FlexCache ボリュームに次のような機能が追加されています。

- SMB およびマルチプロトコル NAS のサポート
- 1:100 の元ボリューム対キャッシュボリュームの割合
- SnapMirror のセカンダリ元ボリューム
- ブロックレベルの無効化
- FlexCache の事前設定

FlexCache の詳細については、「ETERNUS AX/HX series ONTAP での FlexCache ONTAP 9.8」を参照してください。

10.16 FlexClone

ONTAP 9.7 以降、FlexClone は FlexGroup ボリュームでの使用がサポートされています。この機能により、ストレージ管理者は、テスト、開発、バックアップの検証、その他のさまざまな用途に使用するボリュームの、スペース効率に優れたインスタント・コピー (Snapshot テクノロジーによるバックアップ) を作成できます。FlexGroup ボリュームで使用する場合、特別な考慮事項はありません。ただし、FlexGroup ボリュームの FlexClone コピーでは、FlexGroup の親ボリュームと同じ数のメンバーボリュームが使用されます。その結果、FlexClone コピーが作成されると、ノード上のボリューム数が加算され始めます。

たとえば、16 個のメンバーボリュームを含む FlexGroup ボリュームがあり、その FlexGroup ボリュームの FlexClone コピーを作成すると、システムで 32 個のボリュームが使用されることになります。ボリュームの新しい各クローンは、16 個のメンバー FlexVol ボリュームも使用します。

```
cluster::*> volume clone create -vserver DEMO -flexclone FGclone -type RW -parent-vserver
DEMO - parent-volume flexgroup_16

cluster::*> vol show -vserver DEMO -volume flexgroup_16*,FGclone* -fields name -sort-by
name
vserver volume             name-ordinal
------- ------------------ ------------
DEMO    FGclone            -
DEMO    flexgroup_16       -
DEMO    FGclone__0001      base
DEMO    FGclone__0002      base
DEMO    FGclone__0003      base
DEMO    FGclone__0004      base
DEMO    FGclone__0005      base
DEMO    FGclone__0006      base
DEMO    FGclone__0007      base
DEMO    FGclone__0008      base
DEMO    FGclone__0009      base
DEMO    FGclone__0010      base
DEMO    FGclone__0011      base
DEMO    FGclone__0012      base
DEMO    FGclone__0013      base
DEMO    FGclone__0014      base
DEMO    FGclone__0015      base
DEMO    FGclone__0016      base
DEMO    flexgroup_16__0001 base
DEMO    flexgroup_16__0002 base
DEMO    flexgroup_16__0003 base
DEMO    flexgroup_16__0004 base
DEMO    flexgroup_16__0005 base
DEMO    flexgroup_16__0006 base
DEMO    flexgroup_16__0007 base
DEMO    flexgroup_16__0008 base
DEMO    flexgroup_16__0009 base
DEMO    flexgroup_16__0010 base
DEMO    flexgroup_16__0011 base
DEMO    flexgroup_16__0012 base
DEMO    flexgroup_16__0013 base
DEMO    flexgroup_16__0014 base
DEMO    flexgroup_16__0015 base
DEMO    flexgroup_16__0016 base
```

10.16.1 異なるストレージ仮想マシン (SVM) への FlexClone

ONTAP を使用すると、親ボリュームとは異なる SVM にまたがる FlexClone ボリュームを作成できます。これには、-vserver および -parent-vserver コマンドオプションを使用します。これにより、マウントパスを維持する必要がある場合に、クライアントに同じエクスポートパスを使用できます。

例 :

```
cluster::> vol clone create -vserver NFS -flexclone clone -type RW -parent-vserver DEMO
-parent- volume flexgroup -junction-path /flexgroup

cluster::> vol show -junction-path /flexgroup -fields junction-path,volume,size
vserver volume       size junction-path
------- ------------ ---- -------------
DEMO    flexgroup    1PB  /flexgroup
NFS     clone        1PB  /flexgroup
```

10.16.2 ボリュームのリホスト

ONTAP には、volume rehost コマンドを使用して、ボリュームの所有 SVM をすばやく変更する方法が用意されています。現在、FlexGroup ボリュームでの使用はサポートされていません。

10.16.3 FlexClone の削除

FlexGroup でもある FlexClone を削除すると、複数のボリュームが同時に削除されます。ほとんどの場合、これは問題ではありませんが、バグ 1368356 が原因で、条件がクリアされるまで新しいボリューム作成が失敗する可能性があります。これによってデータの停止やアクセスの問題が発生することはありませんが、FlexGroup ボリュームで FlexClone を使用している場合に最適な結果を得るには、どの ONTAP リリースでバグが修正されているかを確認し、そのリリースを使用します。

11. 保存中の暗号化

ONTAP 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 ボリューム向け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。既存の FlexGroup ボリュームのキーを更新できます。詳細は、「11.1 FlexGroup ボリュームのキー更新または既存の FlexGroup ボリュームの暗号化」を参照してください。

一般的に、VE には次のものが必要です。

- 有効な VE ライセンス
- 鍵管理サーバー
- 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 series

FlexGroup ボリュームおよび FlexVol ボリュームを使用した VE の実装と管理については、「ETERNUS AX/HX series Encryption パワー ガイド」および「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 では Aggregate Encryption (AE) がサポートされており、アグリゲートレベルでの暗号化が可能です。同じ FlexGroup ボリュームに属する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。

11.1 FlexGroup ボリュームのキー更新または既存の FlexGroup ボリュームの暗号化

ONTAP では、FlexGroup ボリュームのキー更新と、まだ暗号化されていない FlexGroup ボリュームの暗号化の両方をサポートしています。プロセスは、FlexVol ボリュームの場合と同じです。

11.2 ドライブレベルの暗号化 (NSE と SED)

FlexGroup ボリュームは、暗号化されたドライブだけにまたがる場合、NSE ディスクと SED ディスクを使用できます。

MetroCluster の考慮事項

「7.8 MetroCluster」で説明されているように、MetroCluster で VE/AE を使用する場合は、VE/AE を設定する前に MetroCluster 構成を完了する必要があります。

12. FlexGroup サンプル設計

ONTAP FlexGroup には複数のメリットがあり、通常の FlexVol ボリュームと同様に管理できます。次の設計バリエーションは、FlexGroup ボリュームで許可される機能の例です。

FlexGroup ボリュームでは、次のことが可能です。

- SVM を FlexVol ボリュームとして共有
- FlexVol ボリュームと同じ物理ディスクとアグリゲートを共有
- 他の FlexGroup ボリュームまたは FlexVol ボリュームにマウントされる
- FlexGroup レベルの下にマウントされる (FlexVol ボリュームと同様)
- FlexVol ボリュームとのエクスポートポリシーおよびルールの共有
- クォータを適用する

FlexGroup ボリュームは、次の条件を満たさないことが理想的です。

- 複数のディスクタイプまたはアグリゲートタイプ (たとえば、SATA および SSD 上の同じ FlexGroup ボリュームのメンバーボリューム) にまたがるように構成する
- 異なるハードウェアタイプのノードにまたがる
- 不均一な空き容量がある複数のアグリゲートにまたがる

12.1 ボリュームアフィニティと CPU の飽和

並行処理をサポートするために、ONTAP は起動時に使用可能なハードウェアを評価し、アグリゲートとボリュームをアフィニティと呼ばれる個別のクラスに分割します。一般に、1 つのアフィニティに属するボリュームは、他のアフィニティに属するボリュームと並行してサービスすることができます。これとは対照的に、同じアフィニティの 2 つのボリュームは、ノードの CPU 上でスケジューリング時間 (逐次処理) を待機する必要があります。

ノードのアフィニティは、高度な権限 nodeshell コマンド `waffinity_stats -g` で確認できます。

```
cluster::> set -privilege advanced

cluster::*> node run * waffinity_stats -g

   Waffinity configured with:
                                    # AGGR affinities :     2
      # AGGR_VBN_RANGE affinities / AGGR_VBN affinity :     4
                     # VOL affinities / AGGR affinity :     4
        # VOL_VBN_RANGE affinities / VOL_VBN affinity :     4
           # STRIPE affinities / STRIPEGROUP affinity :     9
              # STRIPEGROUP affinities / VOL affinity :     1
                    # total AGGR_VBN_RANGE affinities :     8
                               # total VOL affinities :     8
                     # total VOL_VBN_RANGE affinities :    32
                            # total STRIPE affinities :    72
                                   # total affinities :   149
                                            # threads :    19
```

上記の NetApp FAS8080 EX ノードの例では、8 つの個別のボリューム上で同時に完全な同時オペレーションをサポートできることが報告されています。また、その最大の可能性に到達するためには、それぞれ 4 つの構成要素をホストする、少なくとも 2 つの別々のアグリゲートで最もうまく動作することを示しています。したがって、このノードが処理する新しい FlexGroup ボリュームを構築する場合、新しい FlexGroup ボリュームには、このノード上の 8 つの構成要素が 2 つのローカルアグリゲートに均等に分散されている必要があります。ONTAP System Manager などのプロビジョニングツールは、新しい FlexGroup ボリュームを作成するときに、FlexGroup のサイズが利用可能なアフィニティにまたがり、最小 100GB のメンバーボリュームサイズを超える場合に、これらのアフィニティを考慮しようとします。

ETERNUS AX series などのハイエンドプラットフォーム用に、使用可能なアフィニティの数がアグリゲートあたり 8(2 つのアグリゲート、ノード当たり 16) に増加しました。

```
cluster::*> node run * waffinity_stats -g

Waffinity configured with:
                                  # AGGR affinities :       2
   # AGGR_VBN_RANGE affinities / AGGR_VBN affinity :       8
                  # VOL affinities / AGGR affinity :       8
     # VOL_VBN_RANGE affinities / VOL_VBN affinity :       4
        # STRIPE affinities / STRIPEGROUP affinity :       3
           # STRIPEGROUP affinities / VOL affinity :       3
                 # total AGGR_VBN_RANGE affinities :      16
                            # total VOL affinities :      16
                  # total VOL_VBN_RANGE affinities :      64
                         # total STRIPE affinities :     144
                                # total affinities :     325
                                         # threads :      18
                                          # pinned :       0
                                # leaf sched pools :      18
                                     # sched pools :      21
```

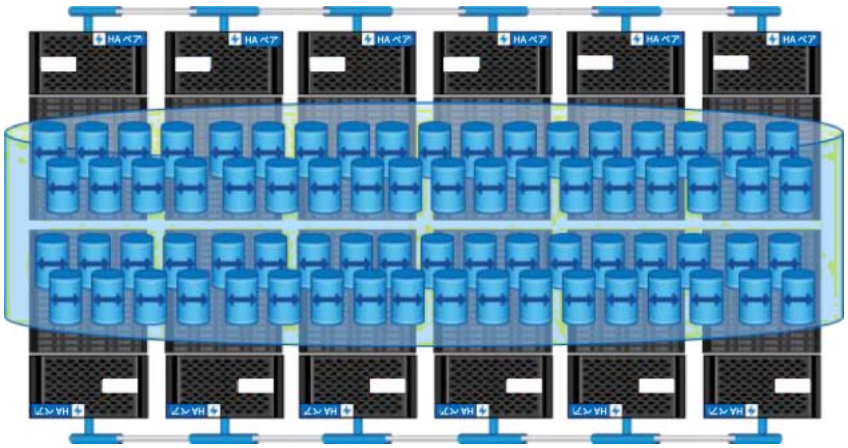
ただし、ONTAP はほとんどの使用例のベストプラクティスに従って FlexGroup ボリュームを導入するため、ストレージ管理者は通常、ボリュームアフィニティについて心配する必要はありません。FlexGroup ボリュームを手動で作成する必要がある場合については、上記のセクションを参照してください。

操作を簡単にするために、`vol create -auto-provision-as flexgroup` コマンド、`flexgroup deploy` コマンド、および ONTAP System Manager GUI を使用して、ストレージ管理者がこの設定を行うことができます。

12.1.1 FlexGroup サンプル設計 1: FlexGroup ボリューム、クラスタ全体 (24 ノード)

FlexGroup ボリュームは、24 ノードのクラスタ全体にまたがることができるため、クラスタ内の利用可能なすべてのハードウェアを単一の分散ネームスペースで使用できるという利点があります。使用可能なすべてのハードウェアを使用できるだけでなく、より多くの潜在的な容量とワークロード内のより多くのボリュームアフィニティを得ることができます。

図 12.1 FlexGroup ボリューム、クラスタ全体 (24 ノード)



考慮事項

クラスタ全体を使用して FlexGroup ボリュームをホストする場合は、「7.1 クラスタに関する考慮事項」の情報に注意してください。

使用例

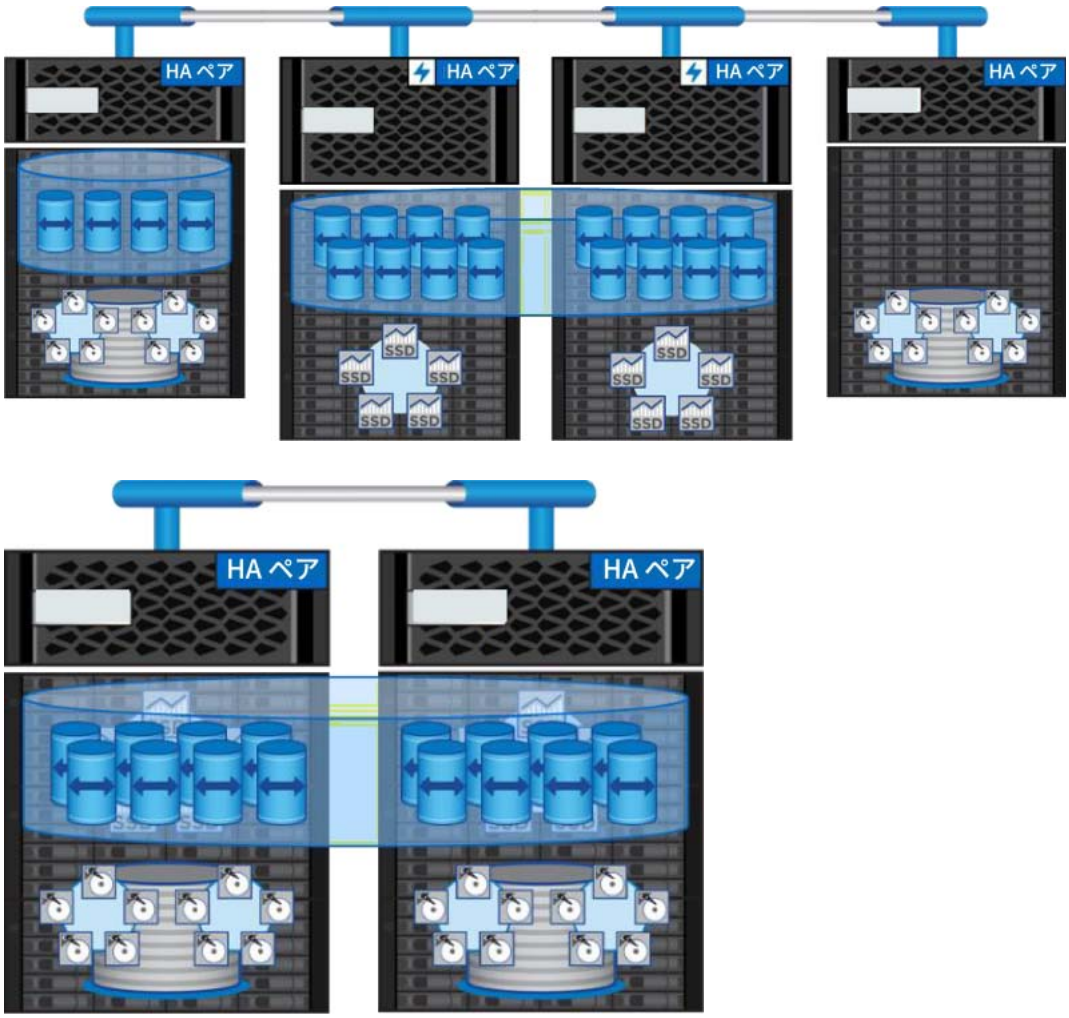
- 膨大な容量 (アーカイブ、スクラッチスペース、およびメディアリポジトリー)
- ストレージに加えて膨大な処理能力を必要とするワークロード (EDA)

12.1.2 FlexGroup サンプル設計 2: 複数ノード、アグリゲート、部分クラスタ

ストレージ管理者は、クラスタ全体のノードに FlexGroup ボリュームを分散させたくない場合があります。その理由には、次のようなものがあります。

- ハードウェアまたは ETERNUS HX series の混在 (一部のノードは ETERNUS AX series)
- アグリゲートタイプまたはディスクタイプの混在 (つまり、同じノード上のハイブリッドアグリゲート)
- ノードを特定のタスク、ストレージ階層、またはテナント専用にする必要があります。
- このようなシナリオでは、FlexGroup ボリュームを作成して、同じノード上または複数のノード上の特定のアグリゲートのみを使用できます。FlexGroup ボリュームがすでに作成されている場合は、メンバー FlexVol ボリュームを目的のノードおよびアグリゲートに無停止で移動できます。詳細は、「8.4.1 無停止でのボリューム移動を使用する場合」を参照してください。

図 12.2 複数のノード、部分的なクラスタ



考慮事項

ノードとアグリゲートの組み合わせで FlexGroup ボリュームを作成しようとすると、自動化されたコマンドはあまり役に立ちません。代わりに、FlexGroup 作成時にアグリゲートを指定できる `volume create` または GUI を使用します。すでに作成されている FlexGroup ボリュームの場合、コマンドラインが唯一のオプションです。

使用例

- 混在ワークロード (高パフォーマンスとアーカイブ)
- 混在クラスタハードウェア
- ハイブリッドアグリゲートがあるノード

12.1.3 FlexGroup サンプル設計 3: FlexGroup、単一ノード

ONTAP クラスタは、堅牢なバックエンドクラスタネットワークを使用して、データ LIF 上の I/O 要求を受信したノードから、物理データを所有するノードに読み取りと書き込みを渡します。トラフィックがリモートの場合は、これらのパケットが処理されるため、リモート I/O の遅延ペナルティがわずかに発生します (約 5% から 10%)。トラフィックがデータを所有するノードに対してすべてローカルである場合、クラスタバックエンドは使用されません。また、NAS オペレーションでは、要求をより高速にディスクに送信するための特別なバイパスが考慮されるため、ノードにローカルに送信するメリットがあります。

FlexGroup では、データ要求の送信先を手動で制御する必要はありません。ONTAP は単純化のためにこの部分を制御します。このため、FlexGroup ボリュームがクラスタ内の複数のノードにまたがる場合、クラスタインターコネクト上に間接的なトラフィックが発生します。

FlexGroup の同時実行性は、リモートトラフィックのパフォーマンスの低下を上回る場合がありますが、FlexGroup ボリュームを 1 つのノードに分離することで、パフォーマンスをある程度向上させることができます。さらに、クラスタの全体的な障害ドメインを減らすために、1 つのノードでいくつかのデプロイメントが実行されます。

図 12.3 は、AX4100 ノードでローカルに 100% アクセスされる単一の FlexVol ボリュームと、同じくローカルに 100% アクセスされる 8 つの FlexVol メンバーを持つ単一の FlexGroup ボリュームを示しています。使用したテストは GCC ライブラリのコンパイル時の Git クローンでした。「6.3.1 AFF A700 テスト」 に記載されているものと同じ試験装置およびデータを使用しました。

このテストでは、より多くのハードウェアを使用できるため、クラスタ全体の FlexGroup ボリュームの方が完了時間がわずかに短くなります。追加のスレッドがローカル FlexGroup ボリュームに追加されると、ハードウェアも対応できなくなるため、完了時間が長くなります。ただし、どちらの FlexGroup ボリュームもローカルの FlexVol ボリュームよりも 2 ～ 3 倍高速で、パフォーマンス曲線も緩やかです。

図 12.3　Git クローン完了時間の比較

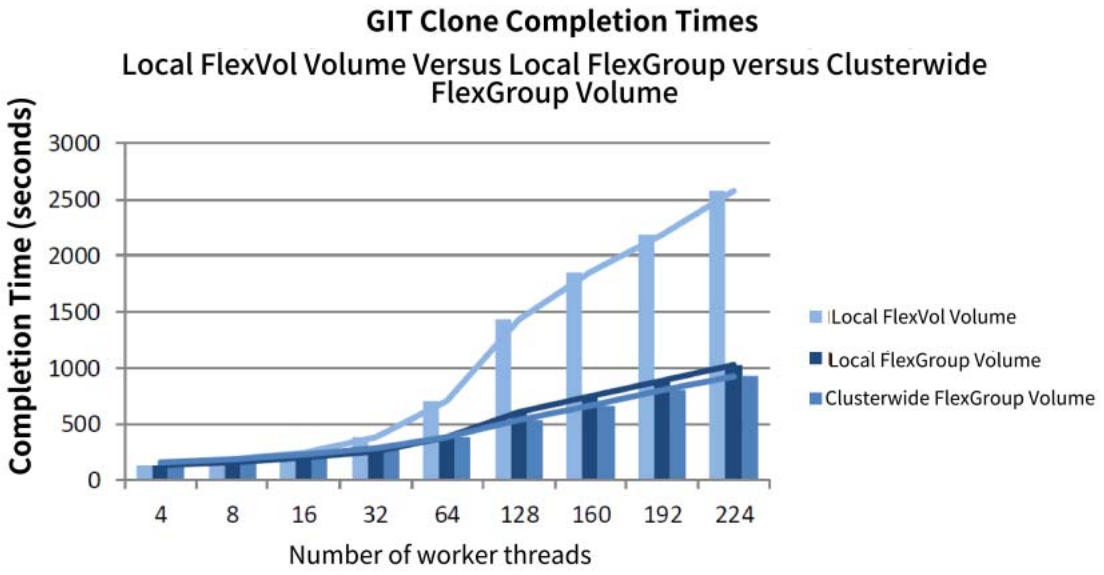


図 12.4 に、ローカル FlexVol ボリュームとローカル FlexGroup ボリュームの平均および最大スループットを示します。さらに、比較のためにクラスタ全体の FlexGroup ボリュームも追加しました。ローカル FlexGroup ボリュームは、16 スレッドに達するまで、クラスタ全体の FlexGroup ボリュームよりも全体的なスループットが優れています。その後、すべてのローカル FlexGroup ボリュームはわずかに遅れて開始されます。これは、ハードウェアを追加することで、ワークロードが 1 つのノードの制限を超えるためです。

図 12.4　平均スループットと最大スループットの比較

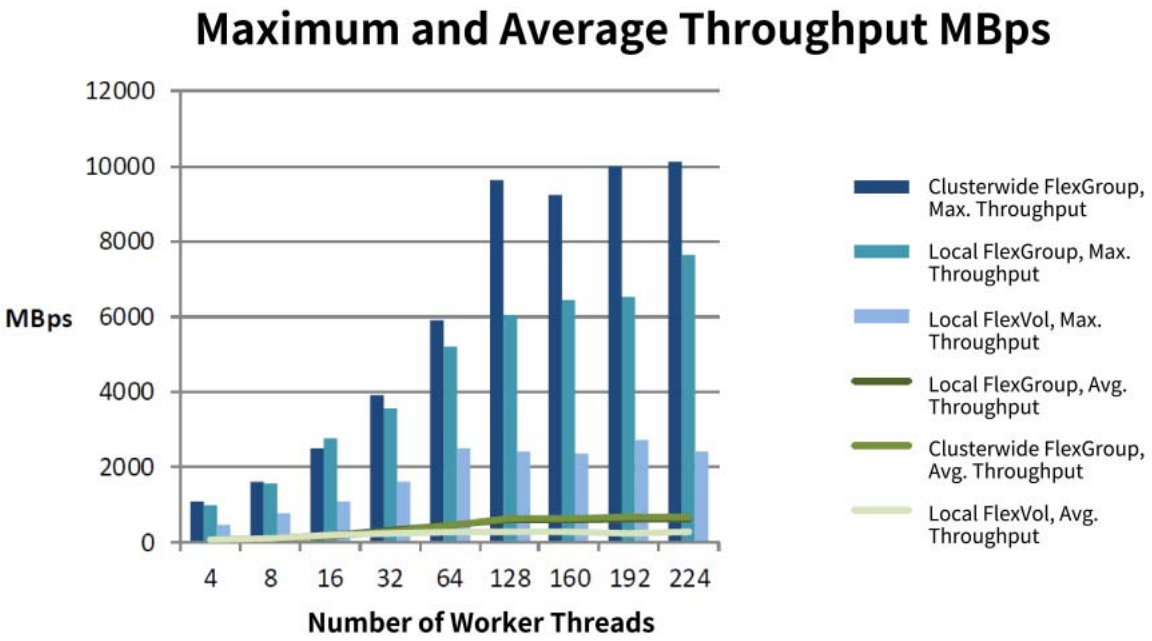


図 12.5 と図 12.6 では、読み取りと書き込みのスループットを、それぞれローカル FlexGroup ボリュームとクラスタ全体の FlexGroup ボリュームで比較しています。64 スレッドの転換点で、ローカル FlexGroup ボリュームはシフトを示し始めます。読み取りスループットは向上しますが、書き込みスループットは低下します。クラスタ全体の FlexGroup ボリュームは、反対の傾向を示します。

図 12.5　最大読み取りスループットの比較

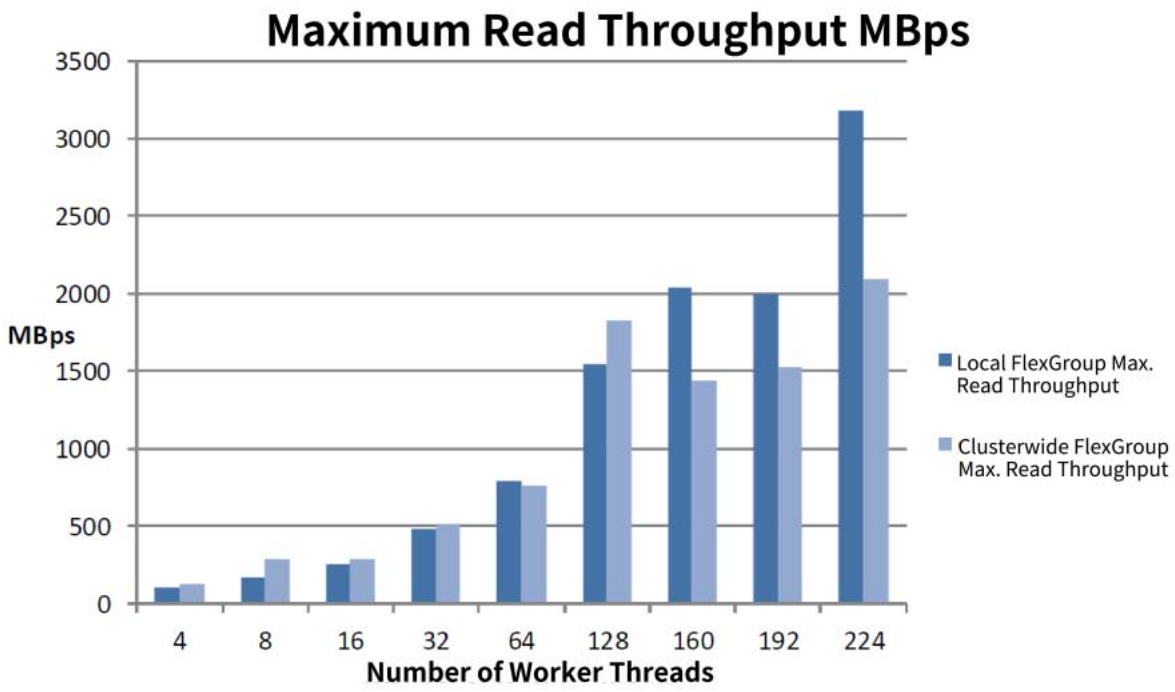


図 12.6　最大書き込みスループットの比較

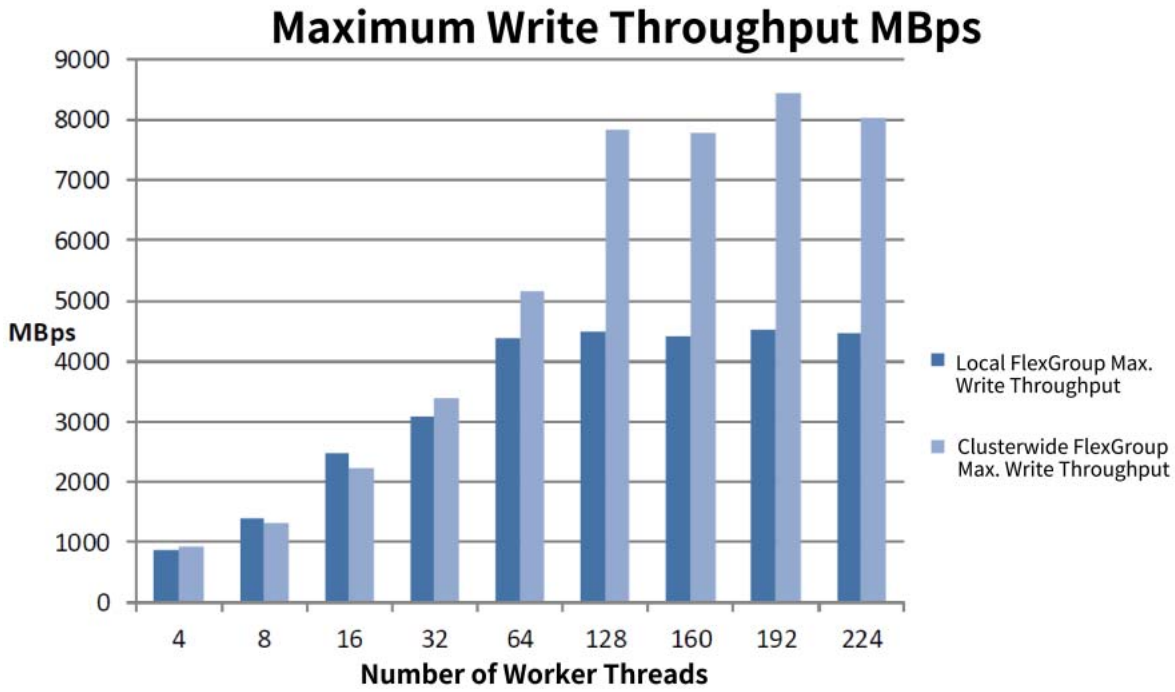
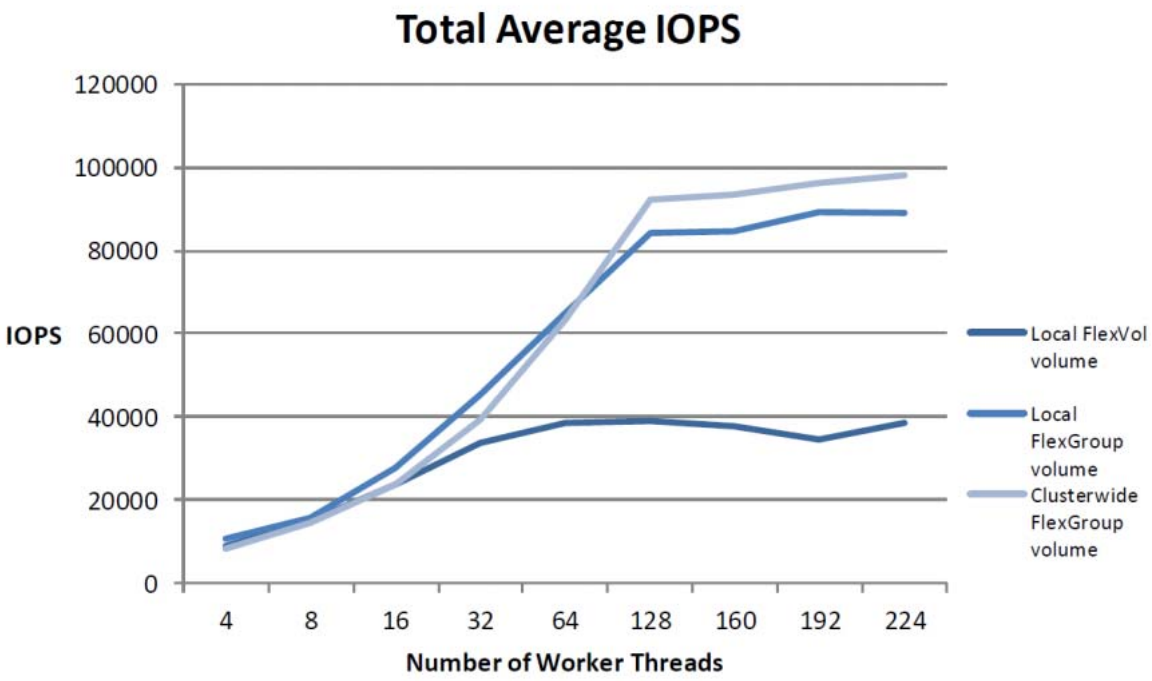


図 12.7 に、ローカル FlexVol ボリュームとローカルおよびクラスタ全体の FlexGroup 構成の合計平均 IOPS を示します。FlexGroup 構成では、FlexVol ボリュームの 2 倍の IOPS が生成されます。ローカル FlexGroup ボリュームのパフォーマンスは、64 スレッドの転換点までクラスタ全体の FlexGroup ボリュームを上回ります。

図 12.7　合計平均 IOPS の比較

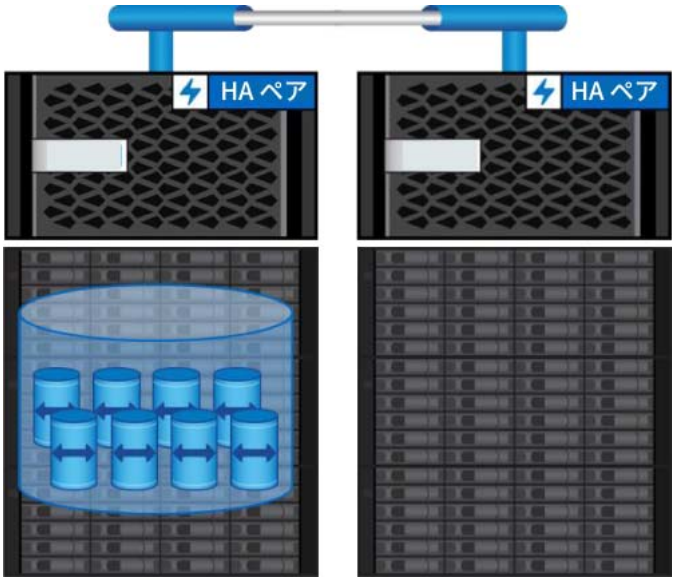


このテストでは、64 個のワーカースレッドが最高の結果をもたらすようです。シングルノードの FlexGroup ボリュームと、64 スレッドをわずかに上回る HA ペアにまたがる FlexGroup ボリュームの平均 CPU 使用率を比較してみましょう。より多くの CPU を使用することは良いことであることを覚えておいてください。作業が実行されていることを意味します。この作業は、同じワークロードで複数のノードにまたがる FlexGroup ボリュームの IOPS 数の増加とスループットの向上によって証明されています。

図 12.8　FlexGroup ボリュームの平均 CPU 使用率、スループット、および IOPS (ETERNUS AX4100 HA ペア、128 スレッド)

図 12.9 FlexGroup ボリュームの平均 CPU 使用率、スループット、および IOPS (シングルノード ETERNUS AX4100、128 スレッド)

図 12.10 FlexGroup ボリューム、単一ノード



考慮事項

FlexGroup ボリュームに単一のノードを使用すると、クラスタインターコネクト横断を除外した場合に得られるメリットは、比較的すぐに消えます。ノードに負荷が追加され、CPU、RAM、ネットワーク帯域幅、ディスク使用率が問題になると、これらのメリットは消滅します。通常、最小のクラスタインターコネクト帯域幅を節約するよりも、FlexGroup ボリュームを複数のノードに分散する方が理にかなっています。

使用例

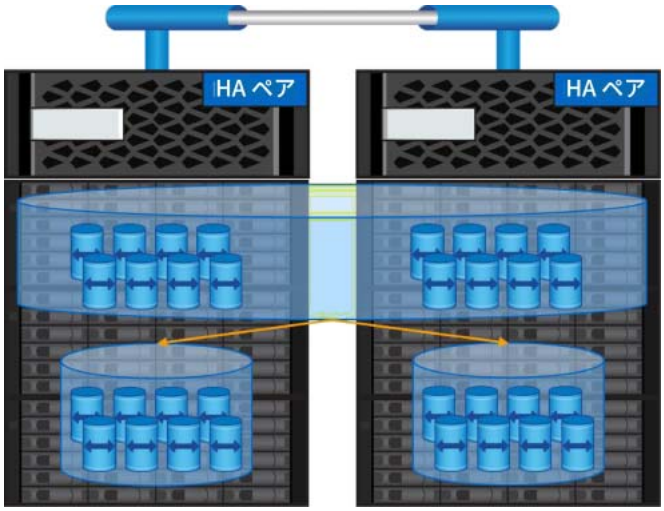
- 高い読み取りワークロード
- ワークロードをノードに分離する必要がある
- クラスターネットワークからトラフィックを遮断する必要がある

12.1.4 FlexGroup サンプル設計 4: FlexGroup ボリュームを FlexGroup ボリュームにマウント

ONTAP で FlexVol ボリュームを使用すると、ボリュームを他のボリュームにマウントしてクラスタにまたがることができ、
1 つの FlexVol ボリュームでは実現できなかった 100TB を超える容量を実現できます。このファイルシステムの設計方法は、パフォーマンスの点で FlexGroup に勝るものです。ただし、複数のノード間で複数の FlexVol ボリュームを作成し、それらをネームスペース内で相互にマウントする管理オーバーヘッドには、多大な時間がかかります。さらに、容量とパフォーマンスの拡張でも、同様の管理上の問題が生じる可能性があります。

FlexGroup ボリュームは、FlexGroup を別の FlexGroup にマウントして、より詳細なデータ管理が可能なフォルダ構造を作成することで、FlexVol ボリュームと同様に管理できます。

図 12.11 FlexGroup ボリュームが FlexGroup ボリュームにマウントされている場合



考慮事項

FlexGroup ボリュームを他の FlexGroup ボリュームにマウントすると、柔軟性が得られますが、管理オーバーヘッドと追加のメンバーボリューム数が犠牲になります。

使用例

- エクスポートポリシーと規則のより詳細な管理
- データの物理的な場所をより詳細に制御
- ボリュームレベルで細分化された SnapMirror により、小規模なデータセットに対応

12.1.5 FlexGroup ボリュームにマウントされている FlexVol ボリューム

FlexVol ボリュームは、FlexGroup ボリュームにマウントすることも、逆にマウントすることもできます。FlexGroup ソリューションでは、この構成も可能です。

使用例

- エクスポートポリシーと規則のより詳細な管理
- データの物理的な場所をより詳細に制御
- ワークロードによってサイズの大きいファイルや時間の経過とともに増大する小さいファイルが作成される可能性がある場合 (たとえば、一連のファイルがより大きな zip ファイルに圧縮された場合)、FlexGroup ではサポートされていない FlexVol の機能

13. 一般的なトラブルシューティングと修復

13.1 障害シナリオ

このセクションでは、障害のシナリオと FlexGroup の対応方法について説明します。

13.1.1 ストレージのフェイルオーバー

FlexGroup ボリュームは FlexVol ボリューム上に構築されるため、ストレージフェイルオーバオペレーションは機能的には FlexVol ボリュームの場合と同じです。テイクオーバーに目立った混乱はありません。無停止アップグレード、ヘッド・スワップ、ローリング・アップグレードなどはすべて正常に実行されます。SMB のようなステートフルなプロトコルのギブバックは、ロック状態の転送のために、少し混乱を招きます。

注意すべき点は、アグリゲートがホームにない (部分的なギブバック状態やアグリゲートが再配置された場合など、所有するノード上にない) 場合、アグリゲートがホームに来るまで FlexGroup ボリュームを作成できないことです。

備考

まれに、4 つ以上のノードがあるクラスタでノードの電源がオフ (ダーティシャットダウン) になっていると、FlexGroup ボリュームをホストする他のノードのメンバーボリュームキャッシュエントリが正常にフラッシュされず、正常に動作しているノードが他のメンバーボリュームの位置が変更されたことを認識しないため、FlexGroup がハングしたように見えることがあります。この問題は、フェイルオーバーおよび回復性テスト中に最もよく発生します。これは ONTAP 9.7P9 以降で修正されています。

13.1.2 ネットワーク障害

FlexGroup ボリュームにアクセスしているネットワーク接続に中断や障害がある場合、FlexGroup ボリュームの動作は FlexVol ボリュームの動作と同じになります。クラスターは、ネットワークに正常にアクセスできるポートまたはノードにデータ LIF を移行しようとします。ネットワークの問題が発生し、使用しているプロトコルのバージョンによっては、クライアントで短時間の中断が発生する場合があります。

13.1.3 スナップショット障害

FlexGroup Snapshot コピーに障害が発生した場合、ONTAP はその Snapshot コピーを部分的なコピーとみなし、SnapRestore 操作では無効にします。Snapshot セットは ONTAP によってクリーンアップされ、EMS イベントが記録されます (`mgmtgwd.snapshot.partCreate`)。

13.1.4 ハードウェア障害

FlexGroup ボリュームをホストするアグリゲート上のディスク障害は、FlexVol の場合と同じように動作します。ONTAP はディスクに障害を起こし、再構築操作で使用するスペアを選択します。RAID 構成で許可されているよりも多くのアグリゲート内ディスクに障害が発生した場合、そのアグリゲートはオフラインと見なされ、オフラインのアグリゲート上に存在するメンバーボリュームにはアクセスできません。

ここでの FlexVol と FlexGroup の主な違いは、複数のアグリゲートにまたがる FlexGroup ボリュームでは、ハードウェアの問題が解決され、他のメンバーボリュームがオンラインに戻るまで、データの不整合を防ぐために他のメンバーボリュームへのアクセスが遮断されることです。

ノードに障害が発生すると、ストレージフェイルオーバイベントが発生します。このイベントでは、ノードが所有するアグリゲートが失敗して HA パートナーノードに所有権が移動し、FlexGroup ボリュームは正常に動作を継続します。同じ HA ペアで 2 つのノードに障害が発生した場合、FlexGroup ボリュームにはオフラインと見なされるメンバーボリュームがあり、ノードが修復されて正常な状態に戻るまで、データアクセスは制限されます。

13.1.5 時刻同期

ONTAP FlexGroup のメンバーをホストするノード間にタイムスキューがあると、FlexGroup ボリュームで中断が発生する場合があります。タイムスキューは、FlexGroup をホストするノード間のローカル時間の相対的な差です。

ディレクトリまたはシンボリックリンクの名前変更やシンボリックリンクのリンク解除などの SMB/NFS プロトコル操作は、内部キャッシュの無効化を引き起こし、結果としてノード間で時間ベースの計算が行われます。タイムスキューが十分に大きい場合、キャッシュ無効化操作の誤った反復トリガが発生する可能性がある。キャッシュ無効化操作を繰り返すと、名前変更およびリンク解除操作の完了が妨げられます。メッセージは再試行され続け、FlexGroup 全体の操作に影響を与え、中断を引き起こす可能性があります。

この問題は ONTAP 9.8 で修正されています。ONTAP 9.8 より前のリリースの ONTAP を実行している場合は、クラスタノードの時刻が同期していることを確認します。理想的には Network Time Protocol (NTP) 構成と同期しています。ONTAP 9.7P6 では、この問題が発生している場合にストレージ管理者に警告する EMS 通知が導入されています。ONTAP リリースを入手するには、FTI のダウンロードサイトを使用します。

14. 容量の監視とアラート

この章では、FlexGroup ボリュームの容量を監視するためのさまざまな方法について説明します。これには、ストレージ効率での全体的な削減量の表示も含まれます。FPolicy のサポートにより、FlexGroup の容量を監視することもできます。

容量の監視とアラートは、ONTAP 9.8 のプロアクティブなサイズ変更機能ではあまり問題になりません。これは、FlexGroup の合計空き容量が個々のメンバーのボリューム空き容量をより正確に反映するためです。

14.1 コマンド・ラインを使用した容量の監視とアラート

シンプロビジョニングを使用する場合は、コマンド `storage aggregate show-space` を `volume show -is-constitution true`、`volume show-space`、および `storage aggregate show` とともに使用すると、FlexGroup のボリュームスペースの使用状況を全体的に把握することができます。コマンドラインでは、`-sort-by` オプションを使用してリストを整理することもできます。

備考

使用中のスペースを正確に描写するには、`volume show-space` コマンドの `Physical Used` 部分に注意してください。例については、「18. コマンドの例」を参照できます。

14.1.1 イベント管理システムメッセージ

イベント管理システムメッセージは、ONTAP のボリュームの容量についてストレージ管理者に警告します。このセクションでは、メッセージを一覧表示します。コマンドラインで `event route show-messagename [ メッセージ ]-instance` と入力すると表示されます。これらのメッセージの例については、「17.8.3 容量関連のイベント管理システムメッセージの例」を参照してください。

- 変更できない値
  - 重大度レベル
  - 修正措置
  - 説明
  - SNMP サポート
- 変更可能な値
  - 送信先
  - 許容される送信間のドロップまたはインターバル

SNMP をサポートするイベント管理システムメッセージがトリガーされると、SNMP トラップが設定済みの SNMP サーバーに対して起動します。このアクションは、`destinations` の値によって指定されます。イベント管理システムの送信先の設定の詳細については、お使いの ONTAP バージョンの Express Guide を参照してください。

Nearly Full (警告) および Full (エラー) のデフォルト値は、次のとおりです。

```
cluster::*> vol show -vserver SVM -volume flexgroup -fields space-nearly-full-threshold-
percent,space-full-threshold-percent
vserver volume    space-nearly-full-threshold-percent space-full-threshold-percent
------- ------    ----------------------------------- ----------------------------
SVM     flexgroup 95%                                 98%
```

`volume.full` のイベント管理システムメッセージは、次のようになります。

```
11/28/2016 18:26:34 cluster-01
DEBUG  monitor.volume.full: Volume flexgroup@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7
is full (using or reserving 99% of space and 0% of inodes).
```

前述の例では、次の値が指定されています。

- オブジェクトのタイプ
- ボリュームの名前
- SVM (CLI で vserver と呼ばれる) 汎用一意識別子 (UUID)
- 使用されているスペースの割合
- inode の使用率

これらの値は、イベント管理システムメッセージをテストするときに使用できます。エラーの影響を受ける SVM を探す場合は、高度な権限レベルで UUID 文字列を使用します。

```
cluster::*> vserver show -uuid 05e7ab78-2d84-11e6-a796-00a098696ec7
                                Admin      Operational Root
Vserver     Type    Subtype    State      State       Volume     Aggregate
----------- ------- ---------- ---------- ----------- ---------- ----------
SVM         data    default    running    running     vsroot     aggr1_node1
```

イベント管理システムメッセージのテスト

イベント管理システムメッセージをテストするには、event generate コマンド (diag 権限レベルで使用が可能) を使用します。各メッセージには、一意の値の文字列があります。volume.full と volume.nearlyFull の値については、前のセクションで説明しました。次の例は、volume.nearlyFull イベントに対するテストメッセージと、その結果のイベント管理システムメッセージを作成する方法を示しています。

```
cluster::*> event generate -message-name monitor.volume.nearlyFull -values Volume flexgroup
@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7 95 0

cluster::*> event log show -message-name monitor.volume.nearlyFull
Time                Node             Severity      Event
------------------- ---------------- ------------- ---------------------------
11/28/2016 18:36:35 cluster-01

                                     ALERT         monitor.volume.nearlyFull: Volume
flexgroup@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7 is nearly full (using or reserving
95% of space and 0% of inodes).
```

ボリュームフルおよびほぼフルのしきい値の変更

FlexVol ボリュームでは、ボリュームが単一のコンテナに分離されているため、full と nearlyFull のデフォルト値で十分です。FlexGroup ボリュームでは、メンバーの FlexVol ボリュームが最大しきい値またはほぼ最大しきい値に達するまでに、アプリケーションまたはエンドユーザーはパフォーマンスの低下をすでに確認している可能性があります。このようにパフォーマンスが低下するのは、リモートファイルの割り当てが増加しているか、メンバーボリュームが満杯またはほぼ満杯であるために空き容量が不足していることをすでに報告している FlexGroup ボリュームが原因です。
ONTAP 9.8 以降のバージョンでは、容量管理で推測作業の大半を占めるプロアクティブなサイズ変更などの機能のため、同じレベルの積極性は必要ありません。

このような状況を監視するには、ボリュームが一杯になる前に警告やエラーが発生するように、ボリュームが満杯またはほぼ満杯のしきい値を調整する必要があります。ボリュームには、admin 権限レベルでこれらのしきい値を調整するオプションがあります。

```
-space-nearly-full-threshold-percent
-space-full-threshold-percent
```

これらのしきい値を調整するには、volume modify コマンドを使用します。

ベストプラクティス 24: FlexGroup のボリューム領域しきい値に関する推奨事項
一般的に、FlexGroup ボリュームのほぼ満杯のしきい値は 80% に設定し、満杯のしきい値は 90% に設定します。これらの設定を使用すると、FlexGroup ボリュームのサイズを増やしたり、ボリューム拡張によってメンバーボリュームを追加する形で容量を追加したりして、容量の問題を解決するのに十分な時間がありますこれらの値は、ファイルの平均サイズと FlexGroup メンバーボリュームのサイズによって異なります。 たとえば、1TB の FlexGroup メンバーは、平均ファイルサイズが 800GB の場合、即座に 80% に達することができますが、100TB の FlexGroup メンバーは、このしきい値に達するまでに時間がかかります。大容量ファイルワークロードのガイダンスについては、「6.4.3 理想的でないワークロード：大容量ファイル」を参照してください。

Active IQ を使用して容量と inode を監視および警告する例については、「17.4 Active IQ Unified Manager での容量の監視とアラートの例」を参照してください。

14.2 シンプロビジョニングにおけるクライアント側の容量に関する考慮事項

FlexGroup ボリュームを使用する場合、クライアントは通常、ストレージ管理者がプロビジョニングした内容を反映した方法で、使用可能な容量や使用済み容量などを報告します。このレポートは、ボリューム領域保証が volume に設定されている場合に特に有効です。これは、ONTAP が予想される容量をクライアントに返すためです。

ただし、物理ストレージにシンプロビジョニングとオーバープロビジョニングを使用する場合、クライアントの値は FlexGroup ボリュームの予想使用容量を反映しません。代わりに、物理アグリゲートの使用済み容量が反映されます。この方法は、FlexVol ボリュームの動作と同じです。

次の例では、3 つの FlexGroup ボリュームがあります。

- flexgroup には 80TB が割り当てられており、2 つのアグリゲートにわたってシンプロビジョニングされ、約 10TB が使用可能です。
- flexgroup4TB には 4TB が割り当てられており、volume の領域が保証されています。
- flexgroup4TB_thin には 4TB が割り当てられており、2 つのアグリゲートにわたってシンプロビジョニングされ、約 4TB が使用可能です。

次の出力は、クラスタがボリューム内の適切な使用済み領域を認識していることを示しています。

```
cluster::> vol show -fields size,used,percent-used,space-guarantee,available -vserver SVM
-volume flexgroup*,!*__0* -sort-by size
vserver volume            size available used    percent-used space-guarantee
------- ----------------- ---- --------- ------- ------------ ---------------
SVM     flexgroup4TB      4TB  3.77TB    30.65GB 5%           volume
SVM     flexgroup4TB_thin 4TB  3.80TB    457.8MB 5%           none
SVM     flexgroup         80TB 10.13TB   5.08GB  87%          none
3 entries were displayed.
```

ただし、クライアントでは、flexgroup という名前のオーバープロビジョニングされた FlexGroup ボリュームの使用済み容量が、クラスタ上の 5GB ではなく 66TB として認識されます。この合計には、物理的なアグリゲートの使用可能な合計サイズ (5.05TB+5.08TB=~10TB) が含まれ、合計サイズからそのサイズが差し引かれます。

オーバープロビジョニングされていないボリュームは、通常、容量を報告します。

```
# df -h
Filesystem                        Size   Used  Avail  Use%  Mounted on
10.193.67.220:/flexgroup          76T    66T   11T     87%  /flexgroup
10.193.67.220:/flexgroup4TB       3.9T   31G   3.8T     1%  /flexgroup4TB
10.193.67.220:/flexgroup4TB_thin 3.9T   230M   3.8T    1%  /flexgroup4TB_thin


cluster::*> aggr show -aggregate aggr1* -fields usedsize,availsize,percent-used,size
aggregate   availsize percent-used size   usedsize
----------- --------- ------------ ------ --------
aggr1_node1 5.05TB    36%          7.86TB 2.80TB
aggr1_node2 5.08TB    35%          7.86TB 2.78TB
2 entries were displayed.
```

約 11TB の空き容量は、Linux クライアントが容量を計算する方法に基づいています。このクライアントは 1 K ブロックを実行するため、10881745216 という数は 1,000 の因数に分割されます。ONTAP では、係数 1,024 を使用して空間が計算されます。

```
# df | grep flexg
10.193.67.220:/flexgroup         85899345920 75017600704 10881745216 88% /flexgroup
10.193.67.220:/flexgroup4TB       4080218944    32143296  4048075648  1% /flexgroup4TB
10.193.67.220:/flexgroup4TB_thin  4080218944      468736  4079750208  1% /flexgroup4TB_thin
```

また、出力の size 部分では、Snapshot 領域に割り当てられているデフォルトの 5% が考慮されます。これが、前述の df 出力で 80TB が 76TB になる理由です。

```
cluster::> vol show -fields size,percent-snapshot-space -vserver SVM -volume flexgroup*,!*__0*
-sort-by size
vserver volume            size percent-snapshot-space
------- ----------------- ---- ----------------------
SVM     flexgroup4TB      4TB  5%
SVM     flexgroup4TB_thin 4TB  5%
SVM     flexgroup         80TB 5%
3 entries were displayed.
```

Snapshot の領域割り当てを 0 に減らすと、df はより正規化された実際のサイズを報告します (used 領域に奇妙な値がまだ残っています)。

```
cluster::*> vol modify -vserver SVM -volume flexgroup -percent-snapshot-space 0
[Job 2502] Job succeeded: volume modify succeeded

# df -h | grep flexgroup
Filesystem                    Size Used Avail Use% Mounted on
10.193.67.220:/flexgroup       80T  70T   11T  88% /flexgroup
```

14.2.1 Windows 容量レポート

Windows は、Linux クライアントとほぼ同じ方法でレポートを作成します。違いは、Windows では因数 1,024 が使用されるため、数値は ONTAP 値に近くなります。

14.3 ONTAP コマンド行からの FlexVol メンバー容量の表示

FlexGroup ボリュームが作成されると、各メンバーは総容量と FlexVol メンバーの数に従って均等に分割されます。たとえば、80TB の FlexGroup ボリュームの場合、FlexVol メンバーはそれぞれ 10TB です。メンバーボリュームの容量を表示するには、diag 権限レベルで `volume show` コマンドを使用します。`volume show -is-consolution true` を使用するか、admin 権限レベルで `volume show-space` コマンドを使用します。

FlexGroup ボリューム内の実際の空き容量を確認する場合は、FlexVol メンバー容量を表示すると便利です。FlexGroup ボリュームは、使用可能な容量の合計を報告するときに、すべてのメンバーボリューム上の使用可能な容量の合計を考慮します。ただし、個々のメンバーボリュームが容量いっぱいになると、他のメンバーボリュームが使用可能な領域を示していても、FlexGroup ボリューム全体が領域不足として報告されます。このシナリオを緩和するために、FlexGroup インジェストアルゴリズムは、他のボリュームよりも頻繁に使用されるボリュームからトラフィックを誘導しようとします。

14.4 FlexGroup 容量ビューア

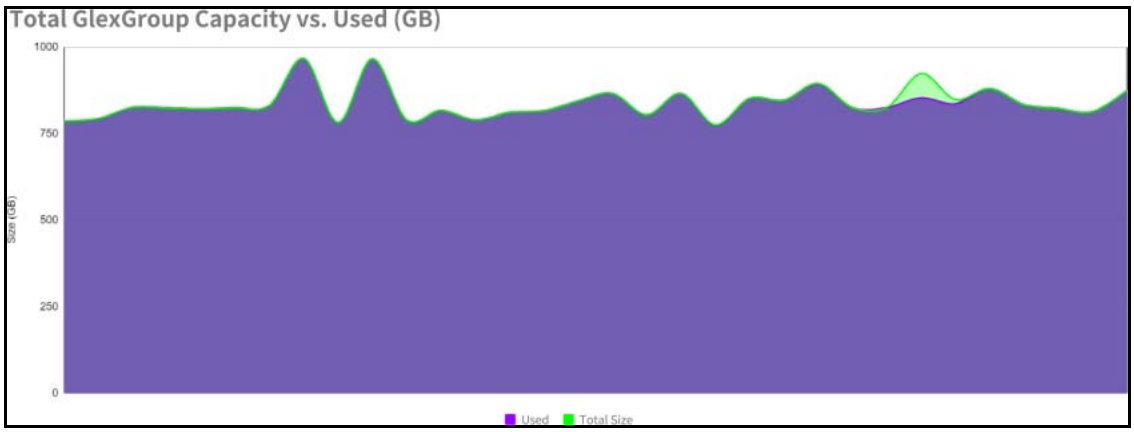
FlexGroup ボリュームの容量と inode の使用状況は、カスタムの Google シートを使用して表示することもできます。当社サポート部門に対して、クラスタのシリアル番号と、グラフ化したい FlexGroup ボリュームの名前を含めた電子メールを送り、自分のコピーにアクセスするよう要求することができます。

Google シートのグラフには、次のものが含まれます。

- 使用されている FlexGroup と合計スペースの全体的なビュー
- スナップショットリザーブと使用済みのグラフ
- 使用済み容量と総容量のメンバーボリュームレベルのビュー
- inode と使用済み容量のトレンド線

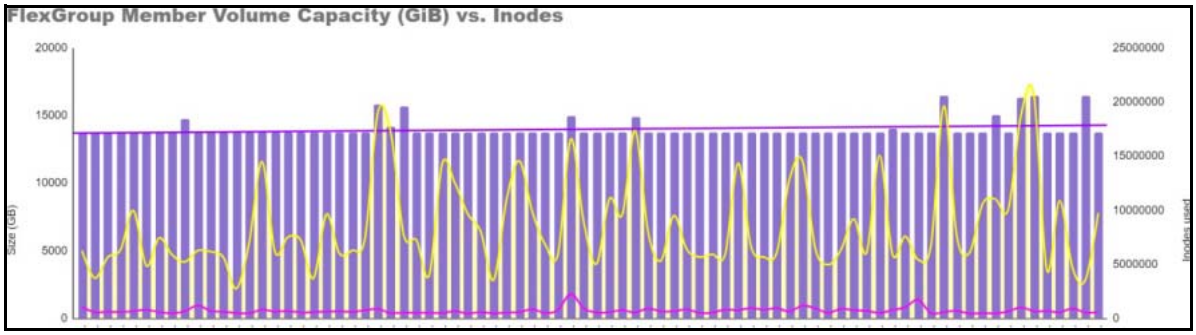
次のビューでは、容量の使用率が不均等な FlexGroup ボリュームを示しています。このボリュームでは、エラスティックサイジングも実行されています。この場合、FlexGroup は総容量を増やす必要があります。

図 14.1 Google シート：FlexGroup の容量ビュー



このビューには、メンバーボリュームレベルでの FlexGroup ボリューム容量が表示され、使用中の inode のトレンド・ラインが表示されます。理想的には、使用される inode には、使用される容量と平行した傾向が見られます。この場合、使用済み inode の数に対応しない、メンバーボリュームあたりの容量の急増が見られます。これは、一部のメンバーボリュームが他のボリュームよりも大きなファイルを持つ可能性があることを意味します。

図 14.2　　Google シート：FlexGroup の容量ビュー



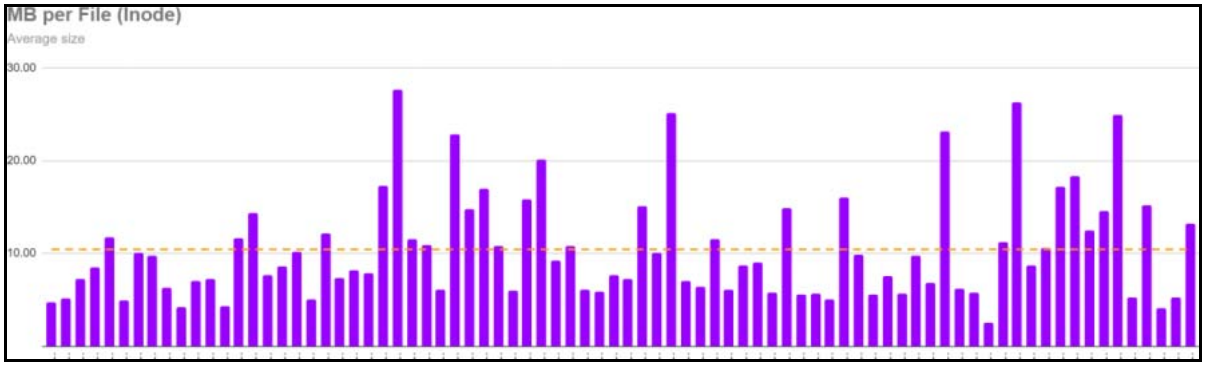
また、使用済み容量の合計を使用済み inode で割って平均ファイルサイズを表示しようとするグラフもあります。正確な方法ではありませんが、かなり近い値が得られます。

次の例では、紫色の棒グラフにはさまざまな値が表示されています。つまり、平均的なファイルサイズには大きな範囲があります。オレンジ色の線は、FlexGroup ボリューム全体の平均 inode サイズです。平均 inode サイズの範囲がこれほど広い理由を発見するには、この FlexGroup に対して、さらに調査が必要になる場合があります。

備考

ファイルサイズの不一致や容量の不均衡は、必ずしも問題を示しているわけではありませんが、問題があり特定する必要がある場合は、これらのグラフを使用できます。詳細は、「6.4.5 FlexGroup ボリューム内のデータ不均衡」を参照してください。

図 14.3　　Google シート：平均 inode サイズ



備考

クラスタのシリアル番号と FlexGroup ボリューム名をメールで当社サポート部門に送信して、FlexGroup Capacity Viewer のコピーを要求します。

14.5　論理スペース計算

論理スペース計算によって、ストレージ管理者は、ストレージ効率による削減をマスクして、エンド・ユーザーが指定されたストレージクォータを過剰に割り当てないようにすることができます。

たとえば、ユーザーが 10TB のボリュームに 6TB を書き込み、ストレージ効率によって 2TB を節約した場合、論理スペース計算によって、ユーザーが 6TB と 4TB のどちらを表示するかを制御できます。

図 14.4　論理スペース計算の動作

現在、FlexGroup ボリュームはこの機能をサポートしていません。FlexVol ボリュームでのみ使用できます。

14.6　FlexGroup パフォーマンスの監視

FlexGroup のパフォーマンスは、通常の FlexVol ボリュームのパフォーマンスを監視する場合と同じ方法で監視できます。CPU 使用率、ディスク飽和、NVRAM ボトルネック、およびその他の WAFL 関連のパフォーマンス特性についても、同じ概念が適用されます。また、NAS のパフォーマンス監視は変更されません。CIFS/SMB および NFS の基本的な統計情報は、従来どおり使用できます。

FlexGroup パフォーマンスの監視との主な違いは、複数のノード、アグリゲート、メンバ FlexVol 構成ボリューム、およびファイルとフォルダのリモート配置を考慮する必要があることです。これらの要素により、パフォーマンスの問題を監視して切り分けたい場合に考慮すべき別のレイヤーが追加されます。

14.6.1 コマンドラインからのパフォーマンスの監視

コマンドラインからパフォーマンス統計を表示する方法はいくつかあります。

14.6.1.1 リアルタイムのパフォーマンス監視

システムのパフォーマンスをリアルタイムで監視するには、`statistics show-periodic` コマンドを使用します。

```
cluster::*> statistics show-periodic ?
   [[-object] <text>]          *Object
   [ -instance <text> ]        *Instance
   [ -counter <text> ]         *Counter
   [ -preset <text> ]          *Preset
   [ -node <nodename> ]        *Node
   [ -vserver <vserver name> ] *Vserver
   [ -interval <integer> ]     *Interval in Seconds (default: 2)
   [ -iterations <integer> ]   *Number of Iterations (default: 0)
   [ -summary {true|false} ]   *Print Summary (default: true)
   [ -filter <text> ]          *Filter Data
```

このコマンドを使用すると、システムパフォーマンスを最新の状態で確認できます。デフォルト値をそのままにしておくと、クラスタ全体が表示されます。SVM を指定すると、より詳細に表示されますが、主に CPU やディスクではなく、NAS カウンタなどの SVM に固有のカウンタが表示されます。SVM 固有の統計情報を使用する場合、オブジェクトに提供されるカウンタを定義すると、CLI のノイズを減らすのに役立ちます。ローカルとリモートのトップレベルディレクトリ (tld)、高レベルディレクトリ (hld)、通常のディレクトリ、およびファイルの比率に関する FlexGroup 統計情報をリアルタイムで取得することもできます。

これらのコマンドの例については、「18. コマンドの例」を参照してください。

FlexGroup 統計情報では、その他のさまざまな情報も表示できます。また、statistics start -object flexgroup コマンドを開始すると、一定期間にわたって収集できます。このコマンドは、Perfstat や perfarchives などの自動化されたツールを使用して、反復で取得できる時間の経過に伴う統計情報を収集します。

```
cluster::*> statistics start -object flexgroup
Statistics collection is being started for sample-id: sample_69197
```

統計を表示するには、次のコマンドを使用します。

```
cluster::*> statistics show -object flexgroup -instance 0

Object: flexgroup
Instance: 0
Start-time: 11/30/2016 16:44:42
End-time: 11/30/2016 17:42:57
Elapsed-time: 3495s
Scope: cluster-01

Counter                                                     Value
-------------------------------- --------------------------------
cat1_tld_remote                                                 2
cat2_hld_local                                                180
cat2_hld_remote                                              1292
cat3_dir_local                                             146804
cat3_dir_remote                                               283
cat4_fil_local                                             734252
cat4_fil_remote                                              1124
groupstate_analyze                                          12232
groupstate_update                                           86242
instance_name                                                   0
node_name                                       cluster-01
process_name                                                    -
refreshclient_create                                         5241
refreshclient_delete                                         5241
refreshserver_create                                         5244
refreshserver_delete                                         5244
```

統計情報のキャプチャは、FlexGroup ボリュームが複数のノードにまたがる場合に、ボリューム内のリモートファイルとディレクトリの配置の割合をわかりやすくまとめたものです。次の例では、値は 14% のリモートディレクトリと 1% のリモートファイルです。

```
    remote_dirs                    14
    remote_files                    1
```

14.6.1.2　プロトコル統計

個々の NAS プロトコルがパフォーマンスにどのように影響しているかを確認することもできます。statistics start コマンドを使用するだけで、NFS または SMB のパフォーマンスカウンタをキャプチャに含めることができます。diag 権限を使用すると、より多くのオプションを利用できます。

```
cluster::*> statistics start -object nfs
   nfs_credstore           nfs_exports_access_cache
   nfs_exports_cache       nfs_exports_match
   nfs_file_session_cache  nfs_file_session_cache:constituent
   nfs_generic             nfs_idle_conn
   nfs_idle_total_conn     nfs_qtree_export
   nfs_server_byname       nfserr
   nfsv3                   nfsv3:constituent
   nfsv3:cpu               nfsv3:node
   nfsv4                   nfsv4:constituent
   nfsv4:cpu               nfsv4:node
   nfsv4_1                 nfsv4_1:constituent
   nfsv4_1:cpu             nfsv4_1:node
   nfsv4_1_diag            nfsv4_1_error
   nfsv4_diag              nfsv4_error
   nfsv4_spinnp_errors

cluster::*> statistics start -object smb
   smb1   smb1:node        smb1:vserver smb1_ctx    smb1_ctx:node
   smb2   smb2:node        smb2:vserver smb2_ctx    smb2_ctx:node

cluster::*> statistics start -object cifs
   cifs                    cifs:node
   cifs:vserver            cifs_cap
   cifs_cap:constituent    cifs_client
   cifs_client:constituent cifs_ctx
   cifs_ctx:node           cifs_shadowcopy
   cifs_unsupp_ioctl       cifs_unsupp_ioctl:constituent
   cifs_watch
```

14.6.2 flexgroup show

FlexGroup I/O 中に、nodeshell コマンドの flexgroup show を使用して、メンバー構成要素の使用状況とバランスを表示することもできます。このコマンドは、新しいファイルでメンバーボリュームを回避する頻度など、有用なその他の情報も提供します。FlexGroup の問題が発生し、サポートケースを開く必要がある場合は、必ずこのコマンド出力をキャプチャしてください。

```
cluster::*> node run * flexgroup show Tech_ONTAP
FlexGroup 0x80F03868 (Tech_ONTAP)
* next snapshot cleanup due in 9334 msec
* next refresh message due in 334 msec (last to member 0x80F0386E)
* spinnp version negotiated as 10.13, capability 0x3F7F
* Ref count is 8
* ShouldEnforceQuotas true
* IsAnyMemberInNvfailedState false
* reaction +0.0, workload +0.0, activity level 0, cv 0%
Idx    Member L          Used       Avail Urgc  Target          Probabilities  D-Ingest Alloc  F-Ingest Alloc
---  -------- - ------------  ---------- ---- -------  --------------------- --------- ----- --------- -----
  1     4503  L 2238647   0%  318698244    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  2     4369  R 3239783   1%  318638088    0%  12.49%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  3     4674  L 2011415   0%  318697586    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  4     4477  R 2334885   0%  318694396    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  5     4329  L 2250619   0%  318697596    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  6     4370  R 2255368   0%  318697148    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  7     4675  L 2252390   0%  318697125    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
  8     4478  R 2201995   0%  318698611    0%  12.50%   [100% 100% 87% 87%]      0+  0     0     0+  0     0
```

14.6.2.1 flexgroup show コマンドの出力内訳

`flexgroup show` コマンドには、一見しただけでは直感的でない一連の値があります。表 14.1 に、これらの値とその解釈方法を示します。

表 14.1 flexgroup show 出力列の定義

列	定義
Idx	メンバーボリュームのインデックス番号。
Member	FlexGroup メンバーの DSID。
L	ノードに対してローカルまたはリモート。
Used	使用されている 4 K ブロックの数と全体の割合。
Urgc	緊急度 : ファイルまたはディレクトリの作成が、メンバーボリュームの ENOSPC を早めることなくリモートメンバーボリュームに割り当てられる可能性。ボリュームの容量の使用率が 100% に近いほど大きくなります。
Targ	ターゲット : ピアに関連してメンバーボリュームに配置される新しいコンテンツの割合。すべての目標パーセンテージの合計は最大 100% に等しい。
Probabilities	メンバーボリュームが使用されない可能性。この数は、メンバーボリュームが他のメンバーボリュームと比較してどの程度いっぱいになるか (許容範囲) に応じて増加します
D-Ingest および D-Alloc	ディレクトリの取り込みとディレクトリの割り当て。ローカルメンバーボリュームに割り当てられているディレクトリ数。
F-Ingest および F-Alloc	ファイル取り込みとファイル割り当て。ローカルメンバーボリュームに割り当てられたファイルの数。

FlexGroup ボリュームの I/O アクティビティの期間中に、`flexgroup show` コマンドを実行する必要があります。このコマンドは、次の有用な情報を提供します。

- メンバー間でトラフィックを均等に分散する方法
- メンバー上でのスペースの均等な配分
- メンバーボリュームを取り込みに使用する可能性
- ワークロードでのディレクトリとファイルの作成の比率
- メンバーボリュームのノード・ローカリティ

14.6.3　パフォーマンスアーカイバ

サポート上の問題を解決するために、ONTAP でデフォルトで実行されるパフォーマンスアーカイバを介してパフォーマンスデータがキャプチャされます。

14.6.4　パフォーマンスの監視 (Active IQ Unified Manager)

FlexGroup ボリュームのパフォーマンスを監視するツールとして、Active IQ Unified Manager の方が適しており、幅広く利用できます。このツールは無料で利用できます。ova ファイルをダウンロードするか、FTI のダウンロードサイトから Linux インストールとして入手します。

Active IQ Unified Manager は、リアルタイムのパフォーマンス情報と履歴のパフォーマンス情報の両方を提供し、単一の監視ポイントを提供します。Active IQ Unified Manager では、FlexGroup ボリューム全体または個々のメンバー構成要素の FlexVol ボリュームの詳細なパフォーマンスビューを表示できます。

図 14.5 は、単一の Linux VM 上の単純なファイル作成スクリプトのキャプチャーであるため、FlexGroup のパフォーマンス上の利点はここでは説明しません。ただし、この図は、Unified Manager が提供できる機能を示しています。

図 14.5　Active IQ Performance Manager のグラフ

15. FlexGroup データ保護のベストプラクティス

FlexGroup データ保護のベストプラクティスについては、「ETERNUS AX/HX series ONTAP FlexGroup ボリュームのデータ保護とバックアップ」を参照してください。

16. ONTAP FlexGroup への移行

多数のファイルまたは大容量を保持する場合の課題の 1 つは、データをできる限り迅速かつ無停止で効果的に移動する方法を決定することです。この課題は、ファイル数が多く、メタデータ操作が多いワークロードで最も大きくなります。ファイルレベルでデータをコピーするには、属性とファイルリストをファイルシステムでクロールする必要があります。これは、ある場所から別の場所にファイルをコピーする時間に大きく影響します。この期間は、ネットワーク・レーテンシー、WAN、システムパフォーマンスのボトルネック、データ移行を困難にするその他の要因を考慮していません。

ONTAP FlexGroup を使用すると、パフォーマンス、拡張性、管理性のメリットが明らかになります。
FlexGroup を扱う場合、データ移行には次の 3 つの一般的な形式があります。

- 他社製 (サードパーティ) のストレージから FlexGroup への移行
- 7-Mode で動作する Data ONTAP から FlexGroup への移行
- ONTAP 内の FlexVol ボリュームまたは SAN LUN から FlexGroup への移行

FlexGroup ボリュームへのデータ移行は、移行に最適な方法です。FlexGroup によるボリューム移行は、現時点では次の方法では実行できません。

- FlexVol から FlexGroup へのボリューム移動
- FlexVol と FlexGroup 間の SnapMirror または SnapVault
- 7-Mode Transition Tool (CBT および CFT)

以降のセクションでは、さまざまな移行の使用例とそのアプローチ方法について説明します。

16.1 NDMP を使用した移行

ONTAP 9.7 以降では、FlexGroup ボリュームが NDMP オペレーションをサポートするようになりました。これには、FlexVol から FlexGroup ボリュームへのデータ移行に使用できる ndmpcopy コマンドが含まれます。
ndmpcopy の設定については、次を参照してください。

How to run ndmpcopy in Clustered Data ONTAP

次の例では、ndmpcopy を使用して、約 500 万のフォルダとファイルを FlexVol から FlexGroup ボリュームに移行しています。このプロセスには約 51 分かかりました。

```
cluster::*> system node run -node ontap9-tme-8040-01 ndmpcopy -sa ndmpuser:AcDjtsU827tputjN -da
ndmpuser:AcDjtsU827tputjN 10.x.x.x:/DEMO/flexvol/nfs 10.x.x.x:/DEMO/flexgroup_16/ndmpcopy
Ndmpcopy: Starting copy [ 2 ] ...
Ndmpcopy: 10.x.x.x: Notify: Connection established
Ndmpcopy: 10.x.x.x: Notify: Connection established
Ndmpcopy: 10.x.x.x: Connect: Authentication successful
Ndmpcopy: 10.x.x.x: Connect: Authentication successful
Ndmpcopy: 10.x.x.x: Log: Session identifier: 12584
Ndmpcopy: 10.x.x.x: Log: Session identifier: 12589
Ndmpcopy: 10.x.x.x: Log: Session identifier for Restore : 12589
Ndmpcopy: 10.x.x.x: Log: Session identifier for Backup : 12584
Ndmpcopy: 10.x.x.x: Log: DUMP: creating "/DEMO/flexvol/../snapshot_for_backup.1" snapshot.
Ndmpcopy: 10.x.x.x: Log: DUMP: Using subtree dump
Ndmpcopy: 10.x.x.x: Log: DUMP: Using snapshot_for_backup.1 snapshot
Ndmpcopy: 10.x.x.x: Log: DUMP: Date of this level 0 dump snapshot: Thu Jan 9 11:53:18 2020.
Ndmpcopy: 10.x.x.x: Log: DUMP: Date of last level 0 dump: the epoch.
Ndmpcopy: 10.x.x.x: Log: DUMP: Dumping /DEMO/flexvol/nfs to NDMP connection
… (output omitted for length)
Ndmpcopy: 10.x.x.x: Notify: dump successful
Ndmpcopy: 10.x.x.x: Log: RESTORE: RESTORE IS DONE
Ndmpcopy: 10.x.x.x: Notify: restore successful
Ndmpcopy: Transfer successful [ 0 hours, 50 minutes, 53 seconds ]
Ndmpcopy: Done
```

NFS で cp を使用した同じデータセットの所要時間は 316 分で、ndmpcopy の 6 倍でした。

```
# time cp -R /flexvol/nfs/* /flexgroup/nfscp/

real   316m26.531s
user   0m35.327s
sys    14m8.927s
```

XCP Migration Tool を使用した場合、このデータセットにかかる時間は 20 分弱で、ndmpcopy よりも約 60% 高速でした。

```
# xcp copy 10.193.67.219:/flexvol/nfs 10.193.67.219:/flexgroup_16/xcp Sending statistics...
5.49M scanned, 5.49M copied, 5.49M indexed, 5.60 GiB in (4.81 MiB/s), 4.55 GiB out (3.91 MiB/s),
19m52s.
```

備考

この XCP コピーは、RAM や CPU の少ない 1GB のネットワークを持つ VM 上で実行されました。より堅牢なサーバは、さらに優れたパフォーマンスを発揮します。

16.2 FlexVol から FlexGroup への変換

ONTAP 9.7 以降では、中断時間が 40 秒未満で、単一の FlexVol を単一のメンバーボリュームを含む FlexGroup ボリュームに変換できます。これは、ボリューム内のデータ容量やファイル数には関係ありません。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、メンテナンス画面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、変換した FlexGroup ボリュームに新しいメンバーボリュームを追加して容量を拡張できます。

16.2.1 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 100TB 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

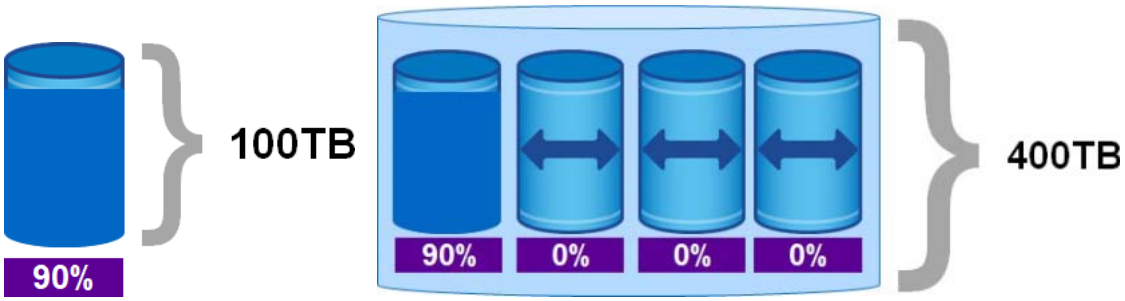
たとえば、ワークロードが急速に増大しており、データを移行する必要はないが、容量を増やす必要があるとします。あるいは、FlexVol ボリュームではワークロードのパフォーマンスが十分ではないため、FlexGroup ボリュームでパフォーマンス処理を向上させたい場合もあります。ここでは変換が役立ちます。

16.2.2 FlexVol ボリュームを変換しない場合

FlexVol ボリュームを FlexGroup ボリュームに変換することは、必ずしも最適なオプションとは限りません。FlexGroup ボリュームでは使用できない FlexVol 機能が必要な場合は、保留してください。たとえば、SVM-DR とカスケード SnapMirror の関係は ONTAP 9.7 ではサポートされていないため、必要な場合は FlexVol ボリュームのままにしておく必要があります。

また、すでに非常に大きい FlexVol ボリューム (80~100TB) と、すでに非常にいっぱいになっている FlexVol ボリューム (80~90%) がある場合は、変換後の FlexGroup ボリュームのメンバーボリュームが非常に大きいため、変換せずにデータをコピーすることをお勧めします。これにより、パフォーマンスの問題が発生する可能性があり、特に、時間の経過とともに増大するファイルがデータセットに含まれている場合、容量の問題は完全に解決されません。

図 16.1 最大容量に達しそうな FlexVol ボリュームの変換



この 90% フル・ボリュームを FlexGroup ボリュームに変換すると、90% フル・メンバーボリュームになります。新しいメンバーボリュームを追加すると、それぞれ 100TB になり、0% フルであるため、新しいワークロードの大部分を処理できます。データはリバランスされず、元のファイルが時間の経過とともに増大した場合、どこにも移動できず容量が不足する可能性があります (最大メンバーボリュームサイズは 100TB であるため)。

変換をブロックできるもの

ONTAP は、次の理由で FlexVol の変換をブロックします。

- ONTAP のバージョンがすべてのノードで 9.7 以降ではない。
- ONTAP アップグレードの問題により、変換できない。
- 7MTT (ONTAP 9.7) を用いて 7-Mode から FlexVol ボリュームを移行した。
  - 移行ボリュームは、ONTAP 9.8 以降で変換できます。
- ボリューム上で FlexGroup でサポートされていない機能がまだ有効になっている (SAN LUN、Windows NFS、SMB 1、ファン・アウト / カスケード SnapMirror の一部、SVM-DR、Snapshot の命名 / 自動削除、`vmalign` セット、SnapLock、スペース SLO、論理スペースの適用 / レポート作成など)
- FlexClone ボリュームが存在し、FlexVol が親ボリュームである (変換されるボリュームは親またはクローンであってはいけない)。
- ボリュームが FlexCache の元のボリュームである。
- スナップ ID が 255 より大きい Snapshot コピー (ONTAP 9.7)。
  - ONTAP 9.8 では 1023 個のスナップショットのサポートが追加されたため、この制限はそのリリースでは適用されません。
- ストレージ効率が有効になっている (後に再度有効にすることがでる)。
- ボリュームが SnapMirror 関係のソースであり、デスティネーションはまだ変換されていない。
- ボリュームがアクティブな (静止しない) SnapMirror 関係の一部である。
- クォータが有効になっている (これらは最初に無効にしてから、後で再度有効にする必要がある) 。
- ボリューム名が 197 文字を超えている。
- ボリュームはアプリケーションに関連付けられている。
  - ONTAP 9.7 のみ。ONTAP 9.8 では、この制限がなくなりました。
- ONTAP プロセスが実行されている (ミラー、ジョブ、wafliron、NDMP バックアップ、処理中の inode 変換など)。
- Storage Virtual Machine (SVM) ルートボリューム。
- ボリュームが満杯。

次のコマンドを使用して、アップグレードの問題をチェックできます。

```
cluster::*> upgrade-revert show
cluster::*> system node image show-update-progress -node *
```

次のコマンドを使用して、移行されたボリュームをチェックできます。

```
cluster::*> volume show -is-transitioned true
There are no entries matching your query.
```

次のコマンドを使用して、255 より大きいスナップ ID を持つ Snapshot コピーをチェックできます。

```
cluster::*> volume snapshot show -vserver DEMO -volume testvol -logical-snap-id >255
-fields logical-snap-id
```

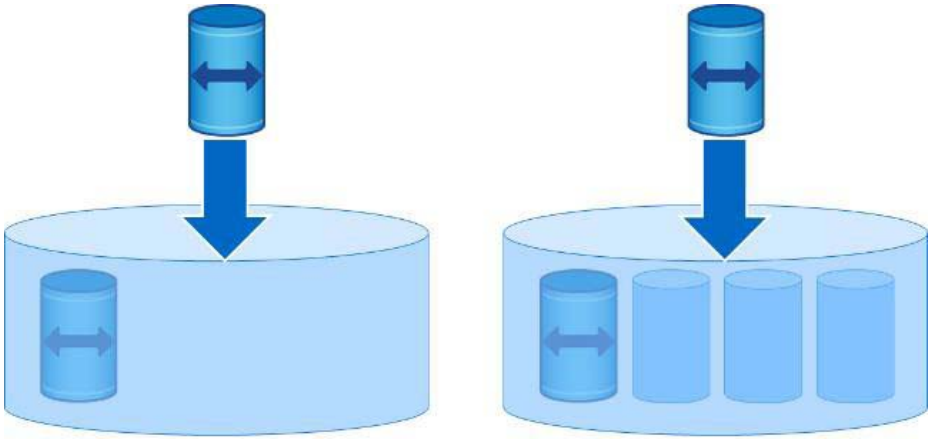
16.2.3 仕組み

ONTAP 9.7 以降で FlexVol ボリュームを FlexGroup ボリュームに変換するには、高度な権限レベルで次の簡単なコマンドを 1 つ実行します。

```
cluster::*> volume conversion start ?
-vserver <vserver name> *Vserver Name
[-volume] <volume name> *Volume Name
[ -check-only [true] ] *Validate the Conversion Only
[ -foreground [true] ] *Foreground Process (default: true)
```

このコマンドを実行すると、ONTAP は 1 つの FlexVol ボリュームを 1 つのメンバーを持つ FlexGroup ボリュームに変換します。実際に変換を行う前に、変換の検証を実行することもできます。

図 16.2 FlexVol ボリュームの FlexGroup への変換とメンバーボリュームの追加



プロセスは 1:1 であるため、現時点では複数の FlexVol ボリュームを単一の FlexGroup ボリュームに変換することはできません。変換が完了すると、単一メンバーの FlexGroup ボリュームが作成され、同じサイズのメンバーボリュームをさらに追加して容量とパフォーマンスを向上させることができます。

16.2.4 その他の考慮事項および注意事項

実際の変換プロセスは単純ですが、変換前に考慮すべき点がいくつかあります。これらの考慮事項のほとんどは、機能のサポートが追加された将来の ONTAP リリースではなくなりますが、ここでそれらを特定することは賢明です。

初期変換の実行後、ONTAP はボリュームを内部的にアンマウントして再マウントし、新しい FlexGroup 情報を適切な場所に取得します。クライアントは再マウントまたは再接続する必要はありませんが、この処理が行われている間、1 分未満の中断が発生します。例については、「17.5 FlexVol から FlexGroup への変換例」を参照してください。データはまったく変更さず、ファイルハンドルはすべて同じままです。

- FabricPool には何も必要ありません。単に動作します。オンプレミスでデータをリハイドレートする必要はありません。
- Snapshot コピーは、クライアントがデータにアクセスするために使用できますが、`snaprestore` コマンドを使用してボリュームをリストアするために使用することはできません。これらの Snapshot コピーは変換前としてマークされます。
- ソースボリュームとデスティネーションボリュームの両方が変換されている場合、SnapMirror 関係は、ベースラインを変更することなく中断した場所から再開されます。ただし、ボリュームの SnapMirror リストアは行われず、クライアントからのファイルの取得のみが行われます。SnapMirror デスティネーションは、最初に変換する必要があります。
- 変換するボリュームから FlexClone ボリュームを削除または分割する必要があります。
- 変換中はストレージ効率を無効にする必要がありますが、変換後もスペースの節約は維持されます。
- 変換されるオリジナルボリュームを持つ FlexCache インスタンスは削除する必要があります。
- 容量保証は、FlexGroup ボリュームがどれだけ大きくなるかに影響します (ボリューム保証の場合) 。新しいメンバーボリュームは、既存のメンバーと同じサイズである必要があるため、それらを維持するための十分なスペースが必要です。
- クォータは FlexGroup ボリュームでサポートされますが、FlexVol ボリュームとは若干異なります。したがって、変換の実行中は、クォータを無効にし (`quota off`)、後で再度有効にする (`quota on`) 必要があります。

FlexGroup ボリュームへの変換は、拡張した後に一方通行になるため、必ずジャンプする準備ができていることを確認してください。変換プロセス中に問題が発生した場合は、問題が発生した場合でもデータを安全に保護するために当社サポート部門が支援する救済方法があります。

FlexGroup ボリュームを拡張して新しいメンバーボリュームを追加する場合は、変換されたメンバーボリュームと同じサイズになるため、十分な空き容量があることを確認してください。さらに、元のボリュームに存在する既存のデータはそのメンバーボリュームに残ります。データは再配布されません。代わりに、FlexGroup ボリュームは、新しいファイル用に新しく追加されたメンバーボリュームを優先します。

16.2.4.1 変換するのが不安ですか？

本番の FlexVol ボリュームをすぐに FlexGroup ボリュームに変換することに不安がある場合は、オプションがあります。

まず、ONTAP を使用すると、`-check-only true` を指定して変換コマンドのチェックを実行できます。これにより、不足している可能性のある前提条件がわかります。

例：

```
cluster::*> volume conversion start -vserver DEMO -volume flexvol -foreground true -check-only true
Error: command failed: Cannot convert volume "flexvol" in Vserver "DEMO" to a FlexGroup.
Correct the following issues and retry the command:
* The volume has Snapshot copies with IDs greater than 255. Use the (privilege: advanced) "volume
snapshot show -vserver DEMO -volume flexvol -logical-snap-id >255 -fields logical-snap-id" command to
list the Snapshot copies with IDs greater than 255 then delete them using the "snapshot delete -vserver
DEMO -volume flexvol" command.
* Quotas are enabled. Use the 'volume quota off -vserver DEMO -volume flexvol' command to disable
quotas.
* Cannot convert because the source "flexvol" of a SnapMirror relationship is source to more than one
SnapMirror relationship. Delete other Snapmirror relationships, and then try the conversion of the
source "flexvol" volume.
* Only volumes with logical space reporting disabled can be converted. Use the 'volume modify - vserver
DEMO -volume flexvol -is-space-reporting-logical false' command to disable logical space reporting.
```

FlexVol から FlexGroup への変換の例については、「17.5 FlexVol から FlexGroup への変換例」を参照してください。

16.2.4.2 変換サンドボックスの作成 – データの移行

ONTAP は複数の SVM を作成でき、ネットワークアクセスから保護できます。この方法を使用して、ボリューム変換などをテストできます。唯一の方法は、データのコピーを転送することですが、実際にはそれほど難しいことではありません。

オプション 1: SnapMirror

SnapMirror を使用して、変換するボリュームを同じ SVM または新しい SVM に複製できます。次に、ミラーを解除して関係を削除します。これで、変換、拡張、パフォーマンスをテストするための Snapshot コピーを含む、ボリュームのサンドボックス・コピーが完成しました。

オプション 2: FlexClone とボリュームのリホスト

SnapMirror がインストールされていない場合、またはネットワークへの負荷が少ない方法を試す場合は、FlexClone (Snapshot コピーによってバックアップされたボリュームのインスタント・コピー) および `volume rehost` (ある SVM から別の SVM へのボリュームの瞬時の移動) を組み合わせて使用できます。FlexClone コピーはリホストできませんが、クローンを分割してからリホストできます。

基本的に、プロセスは次のとおりです。

1. `flexclone create` を使用します。
2. `flexclone split` を使用します。
3. 新しい SVM に `volume rehost` を発行します（または既存の SVM で変換します）。

備考

または、ソース SVM からデスティネーション SVM に FlexClone ボリュームを作成し、「10.16.1 異なるストレージ仮想マシン (SVM) への FlexClone」の説明に従って FlexClone を分割することもできます。

16.2.4.3 SnapMirror 関係での FlexVol ボリュームの変換

また、既存の SnapMirror 関係の一部である FlexVol ボリュームを無停止で変換することもできます。
基本的な手順は次のとおりです。

1. SnapMirror を解除します。
2. SnapMirror デスティネーション FlexVol ボリュームを FlexGroup ボリュームに変換します。
3. ソース SnapMirror FlexVol ボリュームを FlexGroup ボリュームに変換します。
4. SnapMirror を再同期します。

新たに変換された FlexGroup ボリュームを拡張してメンバーボリュームを追加すると、SnapMirror を再ベースラインすることなく、ONTAP によってデスティネーションボリュームが自動的に拡張されます。
このプロセスの例については、「17.6 既存の SnapMirror 関係における FlexVol の変換例」を参照してください。

16.2.4.4 ファイル数が多いと、変換プロセスに影響しますか？

答え：いいえ。
「17.5 FlexVol から FlexGroup への変換例」に示す変換例では、30 万個のファイルを持つボリュームが変換されています。ただし、1 つのボリューム内に 30 万個のファイルがあるということは、実際のファイル数としては多くありません。5 億個のファイルを持つ FlexVol ボリュームの変換例については、「17.6 既存の SnapMirror 関係における FlexVol の変換例」を参照してください。

備考

ビデオの例については、「Statistics show-periodic during FlexVol-FlexGroup convert」を参照してください。

16.3 サード・パーティ製ストレージから FlexGroup への移行

ETERNUS 以外のストレージ (SAN または NAS) から移行する場合、移行パスはファイルベースのコピーです。この移行を実行するには、さまざまな方法があります。無料のものもあれば、サードパーティーベンダーから支払われるものもある。
NFSv3 のみのデータの場合、XCP Migration Tool を強く推奨します。XCP は無料のライセンス・ベースのツールであり、ファイル数の多い環境のデータ移行速度を大幅に向上させることができます。また、XCP は堅牢なレポート作成機能も提供します。XCP 1.5 以降のバージョンでは、NetApp によって公式にサポートされているだけでなく、NFSv4.x および NFSv4.x ACL もサポートしています。

備考

XCP は、ETERNUS ストレージシステムへの移行でのみサポートされます。

CIFS/SMB データの場合は、XCP for SMB が利用可能です。Robocopy はフリーツールですが、転送速度はマルチスレッド機能の使用に依存します。サードパーティのプロバイダも、この種類のデータ転送を実行できます。

16.4 7-Mode で動作する Data ONTAP からの移行

7-Mode で動作する Data ONTAP から FlexGroup にデータを移行するには、次の 2 つの方法のいずれかを使用します。

- コピー・ベースまたはコピー・フリーの移行方法を使用した、7-Mode システムから ONTAP システムへの完全移行。コピーフリー移行を使用する場合、FlexVol ボリューム内のデータを FlexGroup ボリュームにコピーベースで移行します。
- ETERNUS 以外のストレージから FlexGroup に移行するための前述のツールを使用した、FlexVol からのコピー・ベースの移行または LUN からのホスト・ベースのコピー。

FlexVol から FlexGroup ボリュームに移行する場合は、FlexVol から FlexGroup への変換を使用できます。ONTAP 9.8 以降では、これは 7-Mode システムから移行したボリュームで機能します。

16.5 ONTAP での SAN LUN からの移行

SAN ベースの LUN などの既存の ONTAP オブジェクトから移行する場合、現在の移行パスはコピー・ベースです。ETERNUS 以外のストレージから FlexGroup に移行するための前述のツールは、ONTAP オブジェクトからの移行にも使用できます。

16.6 XCP Migration Tool

XCP Migration Tool は無料で、構造化されていない大規模な NAS データセットの範囲設定、移行、管理に特化して設計されています。初期バージョンは NFSv3 のみでしたが、CIFS バージョンが利用可能になりました。ツールを使用するには、ツールをダウンロードしてフリーライセンス (ソフトウェア追跡専用) を要求します。

XCP は、多数の要求を並列に処理できるマルチコア、マルチチャネル I/O ストリーミング・エンジンを使用することにより、メタデータ操作とデータ移行のパフォーマンスに関してファイル数の多い環境が抱える課題に対応します。

これらの要求には、次のものが含まれます。

- データ移行
- ファイルまたはディレクトリの一覧 (ls に代わる、パフォーマンスが高く柔軟性のある選択肢)
- スペースレポート (du に代わる、パフォーマンスが高く柔軟性のある選択肢)

ファイル数の多い環境では、XCP によりデータ移行の時間が 20~30 倍短縮されることがあります。さらに、XCP により、競合他社のシステムでは 9 日間かかっていた 1 億 6500 万ファイルのファイルリスト時間が NetApp テクノロジーでは 30 分に短縮されました。パフォーマンスが 400 倍に向上します。XCP 1.5 の時点で、このツールは NetApp サポートによって正式にサポートされています。

XCP 1.6 には File Systems Analytics 機能も追加されている。これは ONTAP 9.8 に追加された機能に似ていますが、ONTAP を実行していないシステムもスキャンできます。

備考

最良の結果を得るには、入手可能な最新の XCP リリースを使用してください。

XCP には、図 16.3 に示すような便利なレポート作成グラフもあります。

図 16.3　XCP レポート作成グラフ

詳しくは、XCP の公式 Web サイト http://xcp.netapp.com を参照してください。

16.7　XCP を使用した移行前のファイルのスキャン

FlexGroup ボリュームを導入する場合は、ファイルシステムと構造を評価して、初期サイズ設定の考慮事項と、メンバーボリュームを配置する最適な方法を決定します。ファイル数が多い環境では、これは時間がかかり、面倒な作業になります。XCP を使用すると、ファイルをスキャンして CSV または XML 形式にエクスポートし、ファイルシステムを簡単に確認できます。

次の例は、100 万を超えるファイルを持つ FlexGroup ボリュームを示しています。理想的には、これらのファイルの分析に多くの時間を費やしたくありません。

```
Cluster::> vol show -vserver DEMO -fields files,files-used -volume flexgroup_16
vserver volume       files     files-used
------- ------------ --------- ----------
DEMO    flexgroup_16 318766960 1103355
```

このプロセスを合理化するには、xcp scan を使用してファイル情報を取得します。コマンドの例を次に示します。

```
C:\> xcp scan -stats \\demo\flexgroup > C:\destination.csv
```

この場合、クライアントはファイルをスキャンし、コンマ区切り値 (CSV ファイル) 文書に情報を追加します。このドキュメントには、次のような情報が表示されます。

- size、depth of directory、dirsize の最大値と平均値

```
== Maximum Values ==
   Size  Depth  Namelen  Dirsize
 340MiB      9       86      500

== Average Values ==
   Size  Depth  Namelen  Dirsize
1.61KiB      4        6       11
```

- 上位ファイル拡張子

```
== Top File Extensions ==
         .docx  .png  .pptx  .pdf  .css  other
1000038    260   175    128    91    33    219
```

- サイズ範囲別のファイル数

```
== Number of files ==
empty    <8KiB  8-64KiB  64KiB-1MiB  1-10MiB  10-100MiB  >100MiB
    8  1000215      156         288      265         10        2
```

- サイズ範囲で使用されるスペース

```
== Space used ==
empty    <8KiB  8-64KiB  64KiB-1MiB  1-10MiB  10-100MiB  >100MiB
    0  28.7MiB  3.94MiB      124MiB   695MiB     272MiB   453MiB
```

- ディレクトリ・エントリ、ファイルカウント別

```
== Directory entries ==
    empty    1-10  10-100  100-1K  1K-10K  >10k
        7  100118      30     200
```

- ディレクトリの深さの範囲

```
== Depth ==
    0-5  6-10  11-15  16-20  21-100  >100
1100966   333
```

- 変更および作成された日付範囲

```
== Modified ==
   >1 year  >1 month  1-31 days  1-24 hrs  <1 hour  <15 mins  future
       579   1100559         11                150

== Created ==
   >1 year  >1 month  1-31 days  1-24 hrs  <1 hour  <15 mins  future
             1100210                          1089
```

- 合計ファイル数、合計ディレクトリ、シンボリックリンク、ジャンクション、および合計使用スペースを含む、ファイル構造の要約

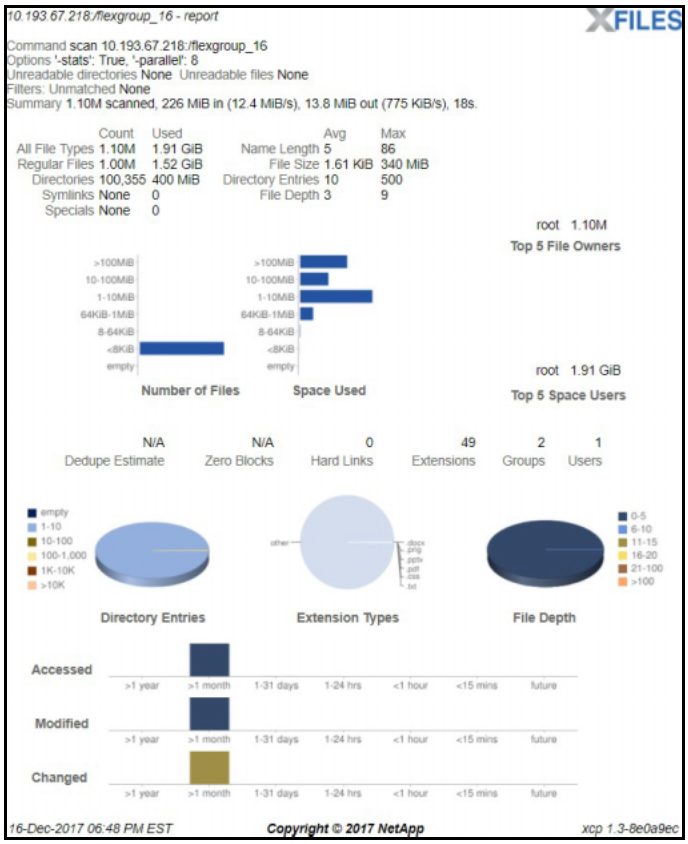
```
Total count: 1101299
Directories: 100355
Regular files: 1000944
Symbolic links:
Junctions:
Special files:
Total space for regular files: 1.54GiB
Total space for directories: 0
Total space used: 1.54GiB
1,101,299 scanned, 0 errors, 26m34s
```

また、NFS 経由の XCP を使用して CIFS ボリュームをスキャンし、より堅牢なレポート作成機能と、データをグラフィカル形式で表示する HTML へのエクスポート機能を利用することもできます。

たとえば、次のコマンドを実行すると、図 16.4 に示すレポートが作成されます。

```
xcp scan -stats -html demo:/flexgroup_16 > /flexgroup.html
```

図 16.4　XCP レポート



XCP を使用してファイルシステムをスキャンすると、ファイルの平均サイズ情報、ファイルサイズの最大値、容量、上位 5 つのファイル所有者のファイルカウントなどを取得できます。これらの統計情報は NFS バージョンの XCP でのみ使用できますが、NFS を使用できる仮想マシンを設定することによって、SMB トラフィックのみを実行するデータセットに対して NFS スキャンを実行することもできます。

16.7.1　XCP を使用したディスク使用量 (du) スキャンの実行

よくある不満の 1 つは、ファイル数の多い環境では、du のようなコマンドの実行に非常に長い時間がかかることです。たとえば、この du コマンドは、1,101,002 個のファイルとフォルダを含む FlexGroup ボリュームで実行され、21 分 22.600 秒かかりました。

XCP を使用する場合、このコマンドは同じクライアントで同じデータセットを 22.852 秒でスキャンします。

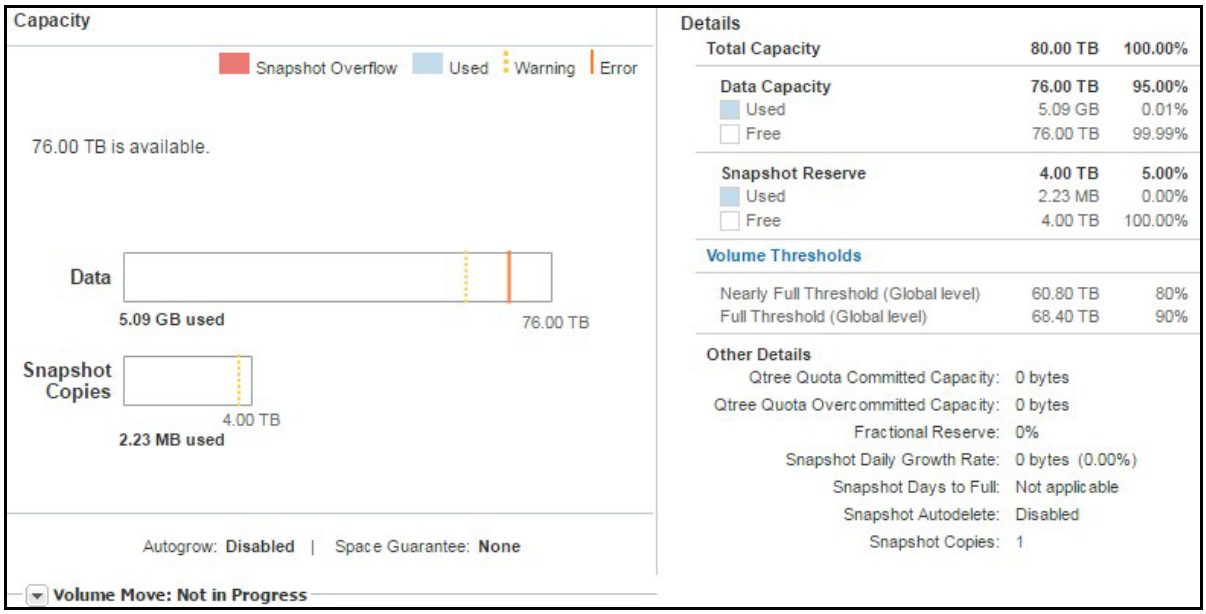
```
[root@centos7 ~]# xcp -duk DEMO:/FGlocal 2>/dev/null | egrep -v '.*?/.*?/'
```

17. 例

17.1 シンプロビジョニングの例

次の図は、FlexGroup ボリュームの合計容量が 80TB、使用容量が 5GB であることを示しています。また、Snapshot コピー用に 4TB が予約されています (5%)。使用可能な容量は 76TB です。

図 17.1 FlexGroup の容量の内訳 :Active IQ Unified Manager



ただし、次の CLI 出力では、いくつかの例外が目立っています。

- Active IQ Unified Manager では、FlexGroup ボリュームは 76TB が使用可能であると表示されますが、CLI では 11.64TB しか使用できません。
- FlexGroup ボリュームの空き容量は 11.64TB と表示されますが、メンバーの FlexVol ボリュームはすべて約 5.8TB と表示されます。
- FlexGroup ボリュームに使用されている割合は 85% ですが、使用している容量は 5GB のみで、80TB と比較するとごくわずかです (81920GB のうちの 5GB は 1% 未満)。
- FlexGroup ボリュームは 85% 使用済みと表示されますが、メンバーの FlexVol ボリュームは、FlexVol メンバーごとに容量が異なるにもかかわらず、すべて 41% 使用済みと表示されます。

例 :

```
cluster::> volume show -is-constituent true -fields size,used,percent-used,available -vserver SVM -volume
flexgroup* -sort-by volume
vserver volume             size  available used    percent-used
------- ------------------ ----- --------- ------- ------------
SVM     flexgroup          80TB  11.64TB   5.08GB  85%
SVM     flexgroup__0001    10TB  5.81TB    147.5MB 41%
SVM     flexgroup__0002    10TB  5.83TB    145.2MB 41%
SVM     flexgroup__0003    10TB  5.81TB    144.9MB 41%
SVM     flexgroup__0004    10TB  5.83TB    148.0MB 41%
SVM     flexgroup__0005    10TB  5.81TB    4.08GB  41%
SVM     flexgroup__0006    10TB  5.83TB    147.6MB 41%
SVM     flexgroup__0007    10TB  5.81TB    145.3MB 41%
SVM     flexgroup__0008    10TB  5.83TB    146.5MB 41%
9 entries were displayed.
```

この異常は、ONTAP がアグリゲートの使用可能な領域に対して計算を行うために発生します。FlexVol メンバーボリュームは、それらが配置されているアグリゲートに応じて、同等の使用可能な値を示します。

```
cluster::> volume show -is-constituent true -fields available,aggregate -vserver SVM -volume
flexgroup* -sort-by aggregate
vserver volume             aggregate   available
------- ------------------ ----------- ---------
SVM     flexgroup__0001    aggr1_node1 5.81TB
SVM     flexgroup__0003    aggr1_node1 5.81TB
SVM     flexgroup__0005    aggr1_node1 5.81TB
SVM     flexgroup__0007    aggr1_node1 5.81TB
SVM     flexgroup__0002    aggr1_node2 5.83TB
SVM     flexgroup__0004    aggr1_node2 5.83TB
SVM     flexgroup__0006    aggr1_node2 5.83TB
SVM     flexgroup__0008    aggr1_node2 5.83TB

cluster::> storage aggregate show -aggregate aggr1* -fields availsize
aggregate   availsize
----------- ---------
aggr1_node1 5.81TB
aggr1_node2 5.83TB
2 entries were displayed.
```

シンプロビジョニングを使用するということは、容量の監視時にアグリゲートの容量とボリュームのフットプリントを考慮する必要があるということです。

17.2 ボリュームの自動サイズ設定の例

次の例では、メンバーボリュームが容量のしきい値に達したときに、ボリュームの自動サイズ設定がどのように機能するかを示します。

この場合、メンバーボリュームのサイズはすべて 1GB です (推奨されない)。これは、単一のファイルでボリュームを簡単に埋めるために行われました。

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available
vserver volume           size
------- ---------------- ------
DEMO    fgautogrow__0001 1GB
DEMO    fgautogrow__0002 1GB
DEMO    fgautogrow__0003 1GB
DEMO    fgautogrow__0004 1GB
DEMO    fgautogrow__0005 1GB
DEMO    fgautogrow__0006 1GB
DEMO    fgautogrow__0007 1GB
DEMO    fgautogrow__0008 1GB
```

備考

1GB はメンバーボリュームの推奨サイズではありません。メンバーボリュームの最小サイズは 100GB 以上でなければなりません。ONTAP は、REST API および ONTAP System Manager を使用して、100GB 未満のメンバーボリュームを持つ FlexGroup ボリュームの作成をプログラムで禁止し、CLI で警告します。

volume autosize を使用すると、書き込みが行われるメンバーボリュームは書き込みを実行するのに適切なサイズに拡大されるため、書き込みは成功します。この場合、ファイルはメンバーボリューム fgautogrow 0003 に書き込まれています。

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available,size
vserver volume           size   available used
------- ---------------- ------ --------- -------
DEMO    fgautogrow__0004 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0005 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0006 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0007 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0008 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0002 1GB    915.5MB   57.26MB
DEMO    fgautogrow__0001 1GB    915.5MB   57.27MB
DEMO    fgautogrow__0003 1.60GB 498.7MB   1.03GB
```

これが発生すると、イベント管理システムでイベントがトリガされ、event log show で確認できます。

```
INFORMATIONAL wafl.vol.autoSize.done: Volume Autosize: Automatic grow of volume
'fgautogrow__0003@vserver:7e3cc08e-d9b3-11e6-85e2-00a0986b1210' by 611MB complete.
```

このイベントは、SNMP、イベント送信先経由でのアラート送信、または Active IQ Unified Manager を使用して監視することもできます。

```
cluster::> event route show -message-name wafl.vol.autoSize.done -instance

                                     Message Name: wafl.vol.autoSize.done
                                         Severity: INFORMATIONAL
                                Corrective Action: (NONE)
                    Description: This message occurs on successful autosize of volume.
                                Supports SNMP trap: true
                                      Destinations: -
             Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

17.3 Snapshot Spill の例

たとえば、400GB の FlexGroup ボリュームでスナップリザーブが 5% に設定されている場合、スナップショットリザーブの合計は 20GB になります。メンバーボリュームが 4 つある場合、メンバーボリュームごとのスナップショットリザーブは 5GB までです。

```
cluster::*> vol show -vserver DEMO -volume FG_SM_400G* -fields size,used,
size-used-by-snapshots,snapshot-reserve-available
vserver volume     size    used   size-used-by-snapshots snapshot-reserve-available
------- ---------- ------- ------ ---------------------- --------------------------
DEMO    FG_SM_400G 420.9GB 2.01GB 3.16MB                 21.04GB
DEMO    FG_SM_400G__0001
                   105.2GB 513.7MB
                                   860KB                  5.26GB
DEMO    FG_SM_400G__0002
                   105.2GB 513.8MB
                                   432KB                  5.26GB
DEMO    FG_SM_400G__0003
                   105.2GB 513.8MB
                                   828KB                  5.26GB
DEMO    FG_SM_400G__0004
                   105.2GB 513.7MB
                                   1.09MB                 5.26GB
```

そのボリュームに 4GB のファイルを書き込むと、スナップショットでは何も使用されません。これは、上書きされたブロックがないためです。

```
cluster::*> vol show -vserver DEMO -volume FG_SM_400G* -fields size,used,size-used-by-snapshots,
snapshot-reserve-available
vserver volume     size    used   size-used-by-snapshots snapshot-reserve-available
------- ---------- ------- ------ ---------------------- --------------------------
DEMO    FG_SM_400G 420.9GB 6.70GB 3.17MB                 21.04GB
DEMO    FG_SM_400G__0001
                   105.2GB 5.19GB 868KB                  5.26GB
DEMO    FG_SM_400G__0002
                   105.2GB 513.9MB
                                  432KB                  5.26GB
DEMO    FG_SM_400G__0003
                   105.2GB 513.9MB
                                  828KB                  5.26GB
DEMO    FG_SM_400G__0004
                   105.2GB 513.8MB
                                  1.09MB                 5.26GB
```

スナップショットを今作成すると、後で上書きが発生した場合に備えて既存のブロックがロックされます。ただし、スナップショットが使用している領域はまだありません。

```
cluster::*> snapshot create -vserver DEMO -volume FG_SM_400G -snapshot file1

cluster::*> vol show -vserver DEMO -volume FG_SM_400G* -fields size,used,size-used-by-snapshots,
snapshot-reserve-available
vserver volume     size    used    size-used-by-snapshots snapshot-reserve-available
------- ---------- ------- ------- ---------------------- --------------------------
DEMO    FG_SM_400G 420.9GB 6.69GB  3.78MB                  21.04GB
DEMO    FG_SM_400G__0001
                   105.2GB 5.18GB  1.02MB                  5.26GB
DEMO    FG_SM_400G__0002
                   105.2GB 514.0MB
                                   580KB                   5.26GB
DEMO    FG_SM_400G__0003
                   105.2GB 514.0MB
                                   976KB                   5.26GB
DEMO    FG_SM_400G__0004
                   105.2GB 513.9MB
                                   1.25MB                  5.26GB
```

データを上書きする場合はスペースが使用されます。これは、ファイルを削除したとき、ファイルを上書きしたとき、または単にファイルデータを変更したときに発生します。ただし、スペースの変更は、ファイルが存在するメンバーボリュームでのみ行われることに注意してください。これは FlexGroup 自体にも反映されていますが、スナップショットの実際の使用済み容量は組み込まれていません。これは、それが予約済み容量を下回っているためです。次の例では、ファイルを削除するとスナップショット内の 4.4GB の領域がすべて使用され、ボリュームで使用される領域が同じ量だけ削減されます。

```
cluster::*> vol show -vserver DEMO -volume FG_SM_400G* -fields size,used,size-used-by-snapshots,
snapshot-reserve-available
vserver volume     size    used    size-used-by-snapshots snapshot-reserve-available
------- ---------- ------- ------- ---------------------- --------------------------
DEMO    FG_SM_400G 420.9GB 772.7MB 4.41GB                  16.63GB
DEMO    FG_SM_400G__0001
                   105.2GB 596.7MB 4.41GB                  874.4MB
DEMO    FG_SM_400G__0002
                   105.2GB 58.65MB 624KB                   5.26GB
DEMO    FG_SM_400G__0003
                   105.2GB 58.69MB 1012KB                  5.26GB
DEMO    FG_SM_400G__0004
                   105.2GB 58.71MB 1.29MB                  5.26GB
```

使用可能なスナップショットリザーブを超過し始めると、使用済みスナップショットスペースが AFS にあふれ始めます。たとえば、スナップリザーブを 1% に減らすと、既存のボリュームでそれが発生することがわかります。

```
cluster::*> vol show -vserver DEMO -volume FG_SM_400G* -fields size,used,size-used-by-snapshots,
snapshot-reserve-available
vserver volume     size    used   size-used-by-snapshots snapshot-reserve-available
------- ---------- ------- ------ ---------------------- --------------------------
DEMO    FG_SM_400G 420.9GB 3.86GB 4.64GB                 3.15GB
DEMO    FG_SM_400G__0001
                   105.2GB 3.69GB 4.64GB                 0B
DEMO    FG_SM_400G__0002
                   105.2GB 58.72MB
                                  624KB                  1.05GB
DEMO    FG_SM_400G__0003
                   105.2GB 58.84MB
                                  1012KB                 1.05GB
DEMO    FG_SM_400G__0004
                   105.2GB 58.71MB
                                  1.29MB                 1.05GB
```

これで、ほぼすべてのスナップショットスペースが used に反映され、使用可能なスナップショットリザーブが 0 になりました。これは、df 出力に表示されるスナップリザーブ使用率にも影響します。

```
cluster::*> df -g FG_SM_400G
Filesystem                total used avail capacity  Mounted on             Vserver
/vol/FG_SM_400G/          416GB  3GB 412GB       0%  /FG_SM_400G            DEMO
/vol/FG_SM_400G/.snapshot 4GB    4GB   0GB     110%  /FG_SM_400G/.snapshot  DEMO
```

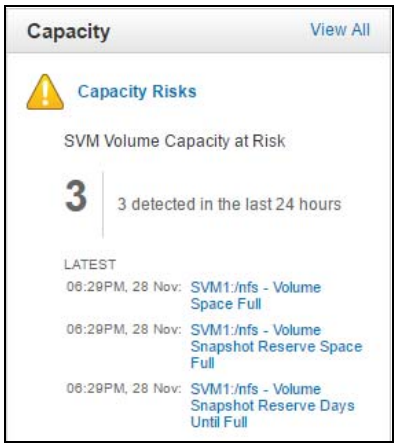
現在、スナップショットリザーブの 110% を使用しています。この容量はどこかに移動する必要があるため、AFS に移行して 3GB を使用していますが、スナップショットリザーブが調整される前は約 597MB しか使用していませんでした。これは Snapshot Spill と呼ばれ、容量レポートに悪影響を及ぼす可能性があります。また、使用可能な空き領域が現在レポートされている場合でも、FlexGroup ボリュームの空き領域が不足する可能性があります。

Snapshot Spill の量は、`volume show-space` コマンドで確認できます。

```
cluster::> volume show-space -vserver DEMO -volume FG_SM_400G* -fields snapshot-spill
vserver volume           snapshot-spill
------- ---------------- --------------
DEMO    FG_SM_400G__0001 3.58GB
DEMO    FG_SM_400G__0002 -
DEMO    FG_SM_400G__0003 -
DEMO    FG_SM_400G__0004 –
```

17.4 Active IQ Unified Manager での容量の監視とアラートの例

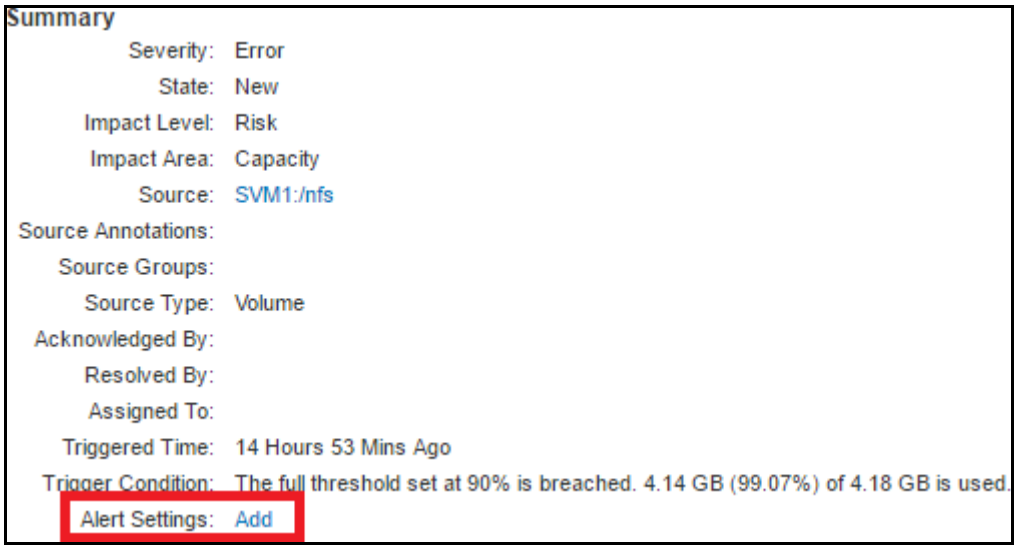
Active IQ Unified Manager は、ストレージシステムのさまざまな機能 (使用容量や空き容量など) を監視および警告する方法を提供します。メインの [Health] ページでは、Active IQ OnCommand によって、容量に関するアクティブな警告とエラーが表示されます。



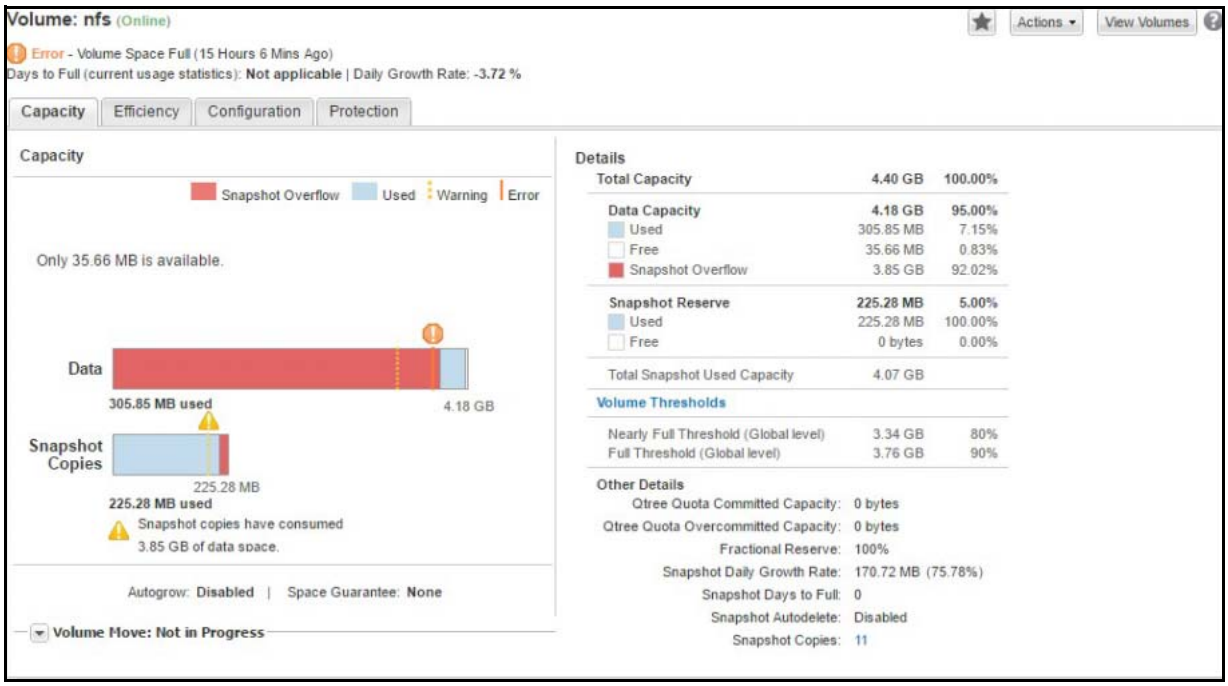
また、Active IQ Unified Manager は、容量関連のイベントをより詳細に表示します。

いずれかのイベントをクリックすると、問題の完全なレポートが表示されます。

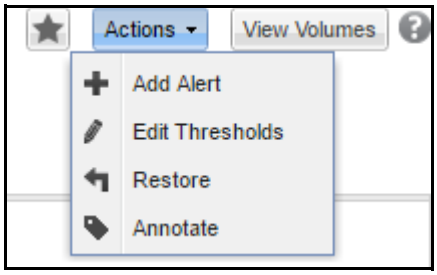
この詳細ビューでは、アラートを構成することもできます。これを行うには、[Alert Settings] の横にある [Add] リンクをクリックします。



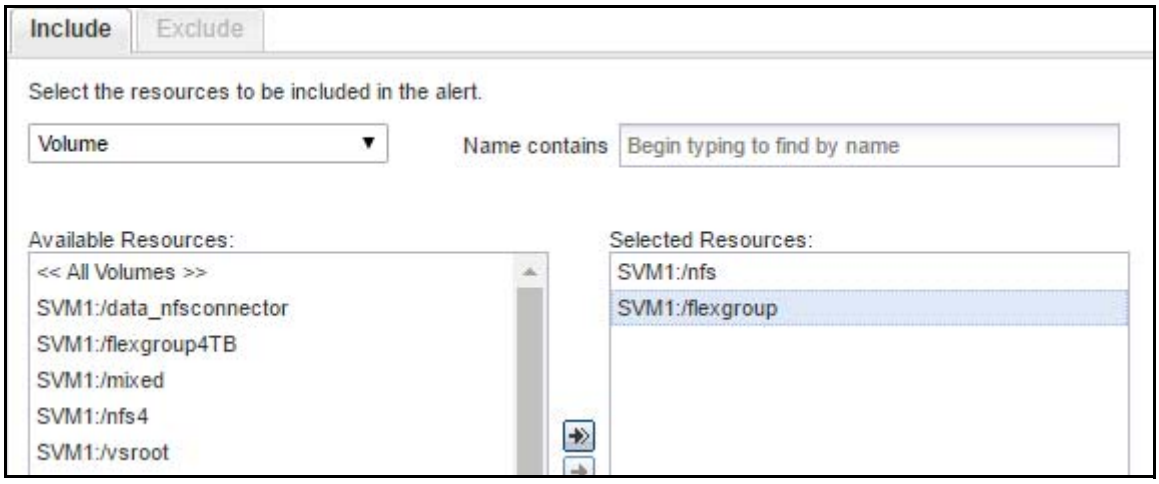
ボリューム画面からボリュームの容量を表示することもできます。[Storage] > [Volumes] をクリックしてボリュームを選択すると、次のような画面が表示されます。



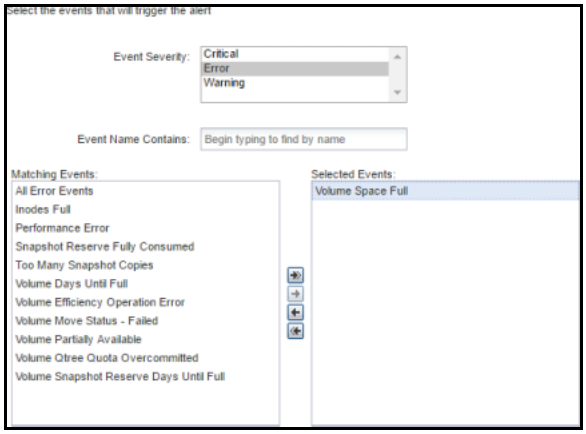
[Actions] ボタンをクリックして、ボリュームに固有のアラートを作成します。

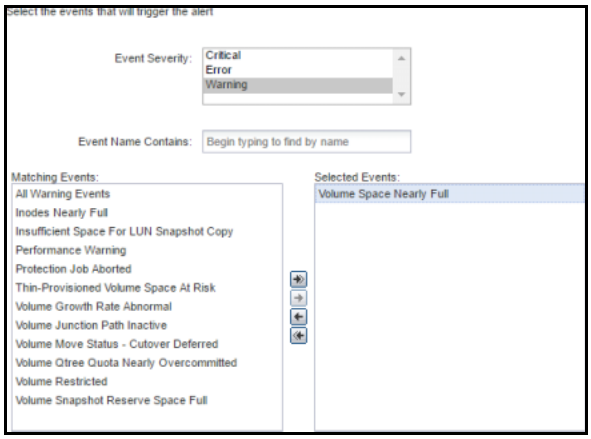


このアラートを使用すると、さまざまなイベントに 1 つ以上のボリュームを追加 (または除外) できます。



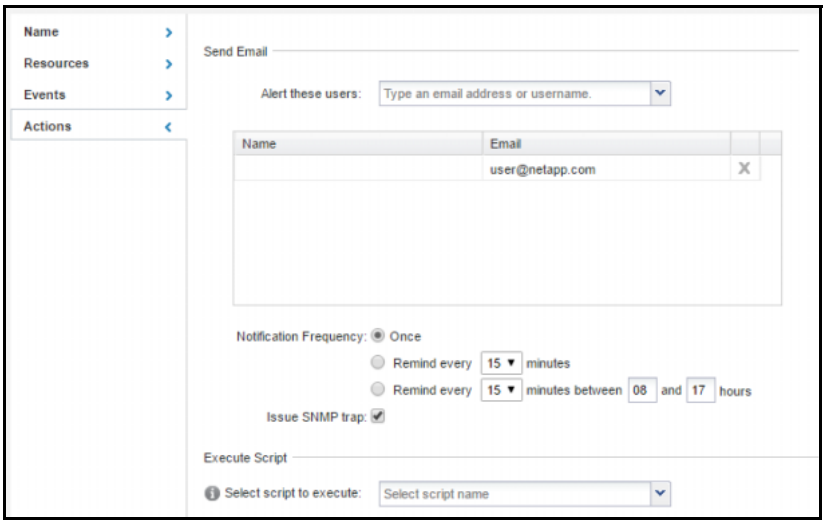
イベントは重大度別に分類され、[Critical]、[Error]、[Warning] の各レベルが含まれます。[Volume Space Full] は [Error] レベルに含まれ、[Volume Space Nearly Full] は [Warning] レベルに含まれます。





イベントがトリガーされると、Active IQ Unified Manager のアラートメカニズムは次の処理を実行できます。

- ユーザー、ユーザーのリスト、および配布リストに電子メールを送信する
- SNMP トラップをトリガする
- リマインダーを送信する
- スクリプトの実行 (自動ボリューム拡張スクリプトなど)

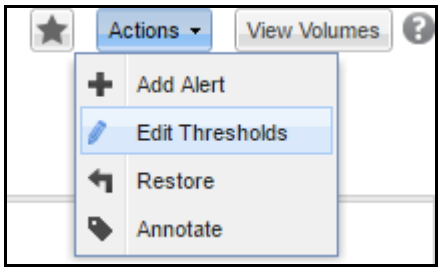


17.4.1 Active IQ Unified Manager でのボリュームしきい値の編集

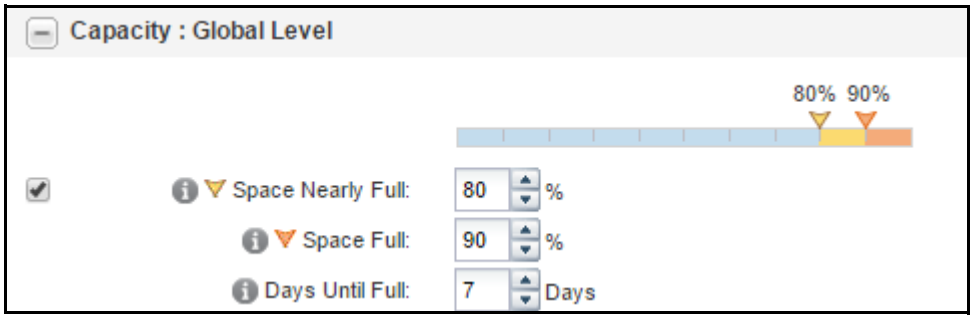
イベント管理システムイベントがクラスタによってトリガされたときの、ボリュームがほぼいっぱいになった場合とボリュームがいっぱいになった場合のしきい値。この制御により、ストレージ管理者はボリューム容量を常に把握し、ボリュームが容量不足になるのを防ぐことができます。FlexGroup では、このアプローチにより、ファイルとフォルダのリモート割り当ても行われます。これは、ボリュームが満杯に近づくにつれて、取り込みのリモート性が増すためです。すでに説明したように、FlexGroup ボリュームのボリュームがほぼいっぱいになった場合とボリュームがいっぱいになった場合のしきい値を変更して、ストレージ管理者が容量の潜在的な問題についてデフォルトよりも早く通知されるようにする必要があります。
詳細については、「ベストプラクティス 24: FlexGroup のボリューム領域しきい値に関する推奨事項」を参照してください。

コマンドラインでは、ONTAP System Manager と同様にしきい値を変更できます。ボリューム詳細の [Actions] ボタンで、[Edit Thresholds] を選択してボリュームごとにボリュームしきい値を変更します。FlexGroup ボリュームでは、設定は FlexGroup ボリューム全体に適用され、しきい値は各メンバーボリュームに個別に設定されます。

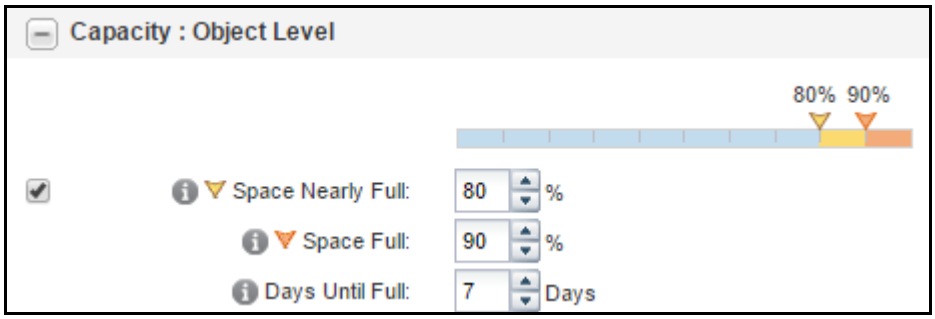
図 17.2　ボリュームしきい値の編集



[Capacity: Global Level] の下のチェックボックスを最初に選択した場合、デフォルトは次のとおりです。これらのデフォルトは、ONTAP イベント管理システムのボリュームしきい値とは関係ありません。むしろ、Active IQ Unified Manager に固有のものです。



値を変更すると、しきい値がオブジェクトレベルに変更されます。



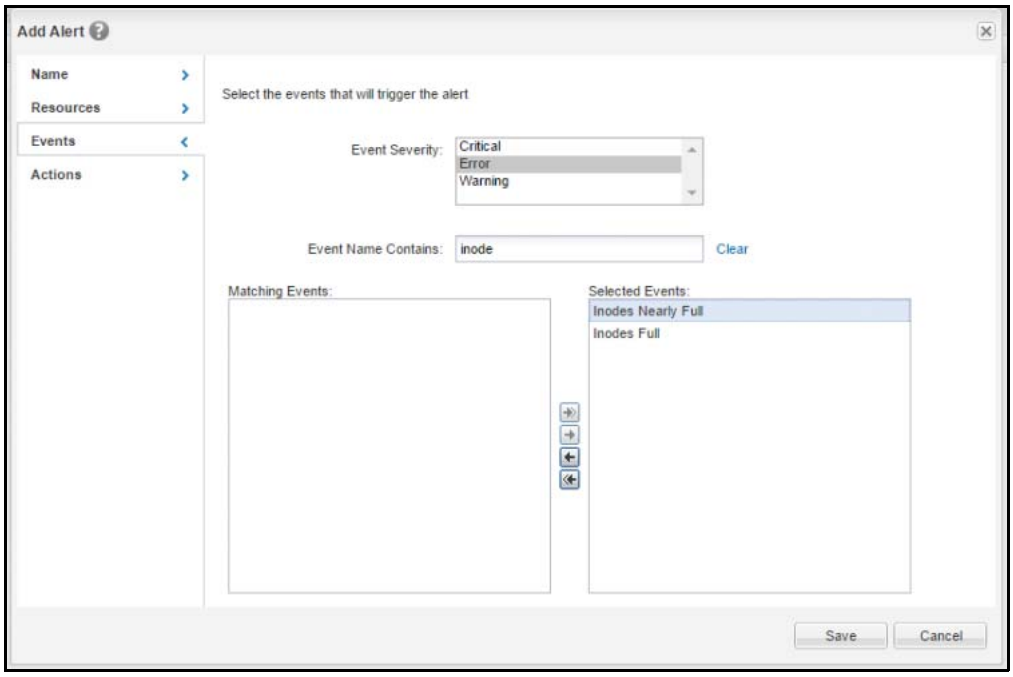
クラスタでは、ボリュームレベルのしきい値オプションは変更されません。

```
cluster::*> vol show -fields space-nearly-full-threshold-percent,space-full-threshold-percent
-sort-by space-nearly-full-threshold-percent -volume flexgroup
vserver volume    space-nearly-full-threshold-percent space-full-threshold-percent
------- --------- ----------------------------------- ----------------------------
SVM     flexgroup 95%                                 98%
```

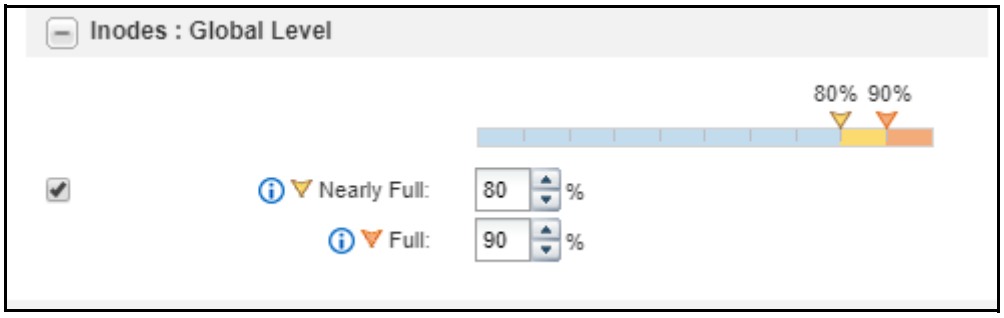
Active IQ Unified Manager アラートは、クラスタのイベント管理システムアラートおよびイベント送信先ロジックとともに使用することも、このロジックとは別に使用することもできます。

17.4.2 inode 監視

また、Active IQ Unified Manager では、[Inodes Nearly Full] （警告）および [Inodes Full] （エラー）イベントを使用して、FlexGroup ボリューム内の inode 数に関するアラートも生成できます。inode のアラートは、容量のアラートと同様に構成されます。



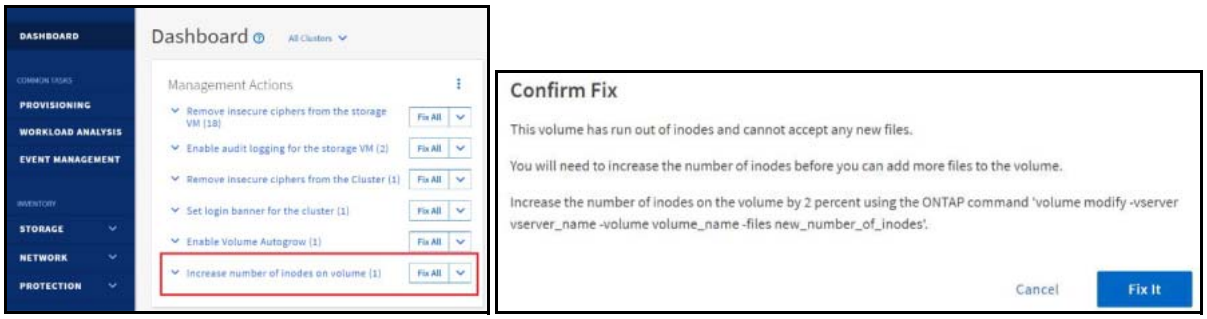
[Edit Thresholds] 画面から inode のしきい値を編集して、アラートをより細かく制御することもできます。



17.4.3 Active IQ の「Fix It」

Active IQ Unified Manager 9.8 以降では、ストレージ管理者がシングルクリックで問題のリストを解決できる「Fix It」機能が導入されています。これらの問題の 1 つは、ボリュームが inode しきい値に達したときです。これらの管理アクションは、Active IQ Unified Manager にログインするとダッシュボードに表示されます。次の例では、使用済み inode のしきい値を超えたボリュームがあります。

図 17.3 Active IQ Unified Manager:inode の修正



[Fix] をクリックすると、Active IQ Unified Manager によって合計ファイル数が 2% 増加します。

17.5 FlexVol から FlexGroup への変換例

この変換例では、ボリュームを変換する前に約 30 万個のファイルを追加して、多数のファイルが存在する場合に処理にどのくらいの時間がかかるかを確認しました。

```
cluster::*> df -i lotsafiles
Filesystem iused ifree %iused Mounted on Vserver
/vol/lotsafiles/ 330197 20920929 1% /lotsafiles DEMO

cluster::*> volume show lots*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      lotsafiles   aggr1_node1  online     RW         10TB     7.33TB    0%
```

最初に、検証を試行します。

```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true -check-only true
Error: command failed: Cannot convert volume "lotsafiles" in Vserver "DEMO" to a FlexGroup.
Correct the following issues and retry the command:
* SMB1 is enabled on Vserver "DEMO". Use the 'vserver cifs options modify -smb1-enabled false -vserver
DEMO' command to disable SMB1.
* The volume contains LUNs. Use the "lun delete -vserver DEMO -volume lotsafiles -lun *" command to
remove the LUNs, or use the "lun move start" command to relocate the LUNs to other FlexVols.
* NFSv3 MS-DOS client support is enabled on Vserver "DEMO". Use the "vserver nfs modify -vserver DEMO
-v3-ms-dos-client disabled" command to disable NFSv3 MS-DOS client support on the Vserver. Note that
disabling this support will disable access for all NFSv3 MS-DOS clients connected to Vserver "DEMO".
```

ご覧のとおり、SMB 1 や (意図的に変換を中断するために) 作成した LUN などの一部のブロック機能があります。したがって、推奨事項をクリアし、検証を再度実行します。注意点がいくつかあります。

```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true -check-only true
Conversion of volume "lotsafiles" in Vserver "DEMO" to a FlexGroup can proceed with the following
warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to a flexible
volume.
* Converting flexible volume "lotsafiles" in Vserver "DEMO" to a FlexGroup will cause the state of all
Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot copies cannot be
restored.
```

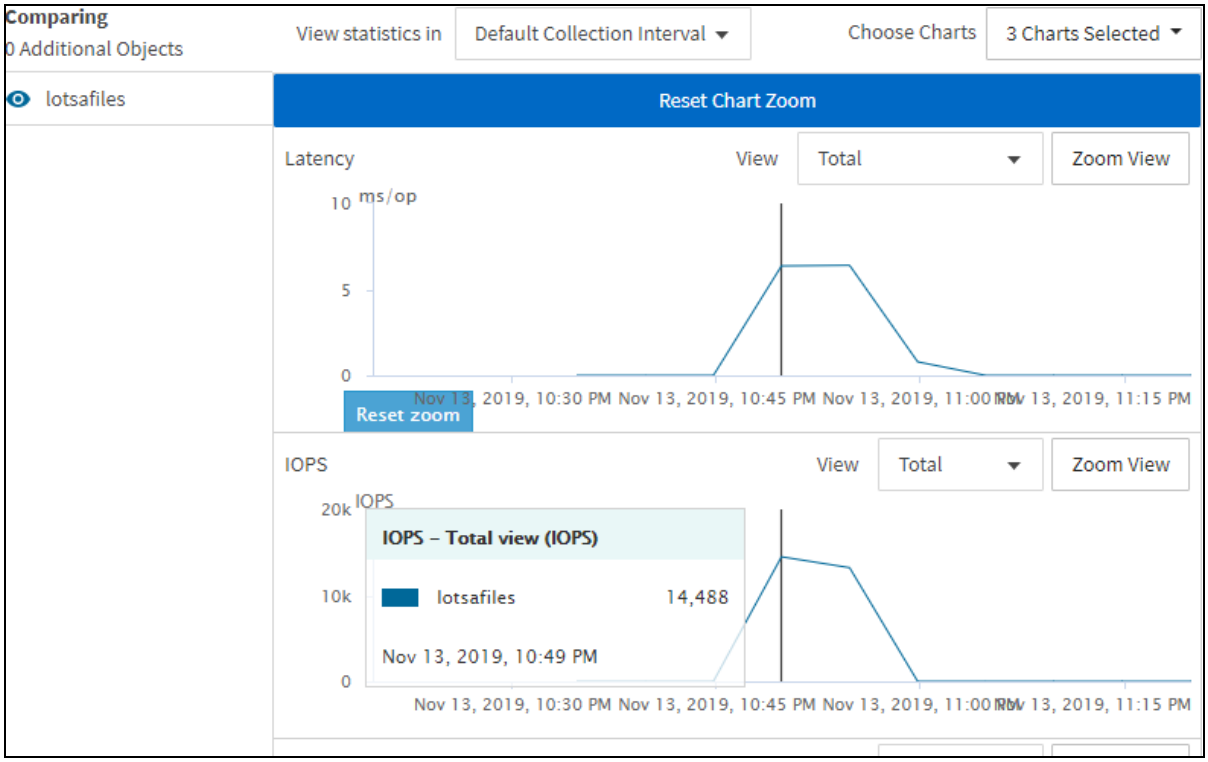
ここで、変換します。最初に、Active IQ Performance Manager を使用して変換中のパフォーマンスを監視しながら、完了までに時間がかかるスクリプトを開始します。
ボリュームの変換にかかる時間は 1 分未満で、中断されるのは IOPS のわずかな低下だけです。

```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back to a
flexible volume.
Do you want to continue? {y|n}: y
Warning: Converting flexible volume "lotsafiles" in Vserver "DEMO" to a FlexGroup will cause the state
of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot copies
cannot be restored.
Do you want to continue? {y|n}: y [Job 23671] Job succeeded: success cluster::*> statistics show-
periodic
cpu cpu total fcache total total data data data cluster cluster cluster disk disk pkts pkts avg busy
ops nfs-ops cifs-ops ops spin-ops recv sent busy recv sent busy recv sent read write recv sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------ -------- -------- -------- -------- -------- --------
34% 44% 14978   14968   10      0 14978 14.7MB 15.4MB 0% 3.21MB 3.84MB 0% 11.5MB 11.6MB 4.43MB 1.50MB
49208 55026
40% 45% 14929   14929   0               14929   15.2MB  15.7MB  0%      3.21MB  3.84MB  0%      12.0MB
11.9MB?3.93MB 641KB 49983
55712
36% 44% 15020   15020   0               15019   14.8MB  15.4MB  0%      3.24MB  3.87MB  0%      11.5MB
11.5MB?3.91MB 23.9KB 49838
55806
30% 39% 15704 15694 10 0 15704 15.0MB 15.7MB 0% 3.29MB 3.95MB 0% 11.8MB 11.8MB 2.12MB 4.99MB
50936 57112
32% 43% 14352 14352 0 0 14352 14.7MB 15.3MB 0% 3.33MB 3.97MB 0% 11.3MB 11.3MB 4.19MB 27.3MB 49736
55707
37% 44% 14807 14797 10 0 14807 14.5MB 15.0MB 0% 3.09MB 3.68MB 0% 11.4MB 11.4MB 4.34MB 2.79MB
48352 53616
39% 43% 15075   15075   0       0       15076   14.9MB  15.6MB  0%      3.24MB  3.86MB  0%      11.7MB
11.7MB  3.48MB 696KB 50124
55971
32% 42% 14998   14998   0       0       14997   15.1MB  15.8MB  0%      3.23MB  3.87MB  0%      11.9MB
11.9MB  3.68MB 815KB 49606
55692
38% 43% 15038 15025 13 0 15036 14.7MB 15.2MB 0% 3.27MB 3.92MB 0% 11.4MB 11.3MB 3.46MB 15.8KB
50256 56150
43% 44% 15132 15132 0 0 15133 15.0MB 15.7MB 0% 3.22MB 3.87MB 0% 11.8MB 11.8MB 1.93MB 15.9KB 50030
55938
34% 42% 15828 15817 10 0 15827 15.8MB 16.5MB 0% 3.39MB 4.10MB 0% 12.4MB 12.3MB 4.02MB 21.6MB
52142 58771
28% 39% 11807   11807   0               11807   12.3MB  13.1MB  0%      2.55MB  3.07MB  0%      9.80MB
9.99MB  6.76MB 27.9MB 38752
43748
33% 42% 15108   15108   0               15107   15.1MB  15.5MB  0%      3.32MB  3.91MB  0%      11.7MB
11.6MB  3.50MB 1.17MB 50903
56143
32% 42% 16143   16133   10      0 16143 15.1MB 15.8MB 0% 3.28MB 3.95MB 0% 11.8MB 11.8MB 3.78MB 9.00MB
50922 57403
24% 34% 8843 8843 0     0       8861 14.2MB 14.9MB 0%   3.70MB 4.44MB 0%        10.5MB 10.5MB 8.46MB
10.7MB 46174
53157
27% 37% 10949 10949     0       0 11177 9.91MB 10.2MB   0% 2.45MB 2.84MB        0% 7.46MB 7.40MB 5.55MB
1.67MB 31764
35032
```

```
28% 38% 12580 12567     13 0 12579 13.3MB 13.8MB 0% 2.76MB 3.26MB 0% 10.5MB 10.6MB 3.92MB 19.9KB
44119 48488
30% 40% 14300 14300     0 0 14298 14.2MB 14.7MB 0% 3.09MB 3.68MB 0% 11.1MB 11.1MB 2.66MB 600KB 47282
52789
31% 41% 14514 14503     10 0 14514 14.3MB 14.9MB 0% 3.15MB 3.75MB 0% 11.2MB 11.2MB 3.65MB 728KB 48093
53532
31% 42% 14626 14626     0 0 14626 14.3MB 14.9MB 0% 3.16MB 3.77MB 0% 11.1MB 11.1MB 4.84MB 1.14MB 47936
53645
cluster: cluster.cluster: 11/13/2019 22:44:39
cpu cpu total fcache total total data data data cluster cluster cluster disk disk pkts pkts avg busy
ops nfs-ops cifs-ops ops spin-ops recv sent busy recv sent busy recv sent read write recv sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------ -------- -------- -------- -------- -------- --------
30% 39% 15356 15349 7 0 15370 15.3MB 15.8MB 0% 3.29MB 3.94MB 0% 12.0MB 11.8MB 3.18MB 6.90MB 50493
56425
32% 42% 14156 14146 10 0 14156 14.6MB 15.3MB 0% 3.09MB 3.68MB 0% 11.5MB 11.7MB 5.49MB 16.3MB
48159 53678
```

Active IQ のパフォーマンスは次のようになります。



これで、FlexGroup ボリュームのメンバーが 1 つになりました。

```
cluster::*> volume show lots*
Vserver   Volume           Aggregate   State    Type Size    Available  Used%
--------- ---------------- ----------- -------- ---- ------- ---------- -----
DEMO      lotsafiles       -           online   RW      10TB     7.33TB    0%
DEMO      lotsafiles__0001 aggr1_node1 online   RW      10TB     7.33TB    0%
2 entries were displayed.
```

Snapshot コピーはまだ存在しますが、変換前としてマークされています。

```
cluster::> set diag
cluster::*> snapshot show -vserver DEMO -volume lotsafiles -fields is-convert-recovery,state
vserver volume     snapshot                        state          is-convert-recovery
------- ---------- ------------------------------- -------------- -------------------
DEMO    lotsafiles base                            pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_1705          pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_1805          pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_1905          pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_2005          pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_2105          pre-conversion false
DEMO    lotsafiles hourly.2019-11-13_2205          pre-conversion false
DEMO    lotsafiles clone_clone.2019-11-13_223144.0 pre-conversion false
DEMO    lotsafiles convert.2019-11-13_224411       pre-conversion true
9 entries were displayed.
```

Snapshot コピーが変換前の状態の場合、SnapRestore オペレーションでそのコピーを使用すると失敗します。

```
cluster::*> snapshot restore -vserver DEMO -volume lotsafiles -snapshot convert.2019-11-13_224411

Error: command failed: Promoting a pre-conversion Snapshot copy is not supported.
```

ただし、Snapshot コピーを使用してクライアントからファイルを取得することはできます。

```
[root@centos7 scripts]# cd /lotsafiles/.snapshot/convert.2019-11-13_224411/pre-convert/ [root@centos7
pre-convert]# ls
topdir_0 topdir_14 topdir_2 topdir_25 topdir_30 topdir_36 topdir_41 topdir_47 topdir_52
topdir_58 topdir_63 topdir_69 topdir_74 topdir_8 topdir_85 topdir_90 topdir_96
topdir_1 topdir_15 topdir_20 topdir_26 topdir_31 topdir_37 topdir_42 topdir_48 topdir_53
topdir_59 topdir_64 topdir_7 topdir_75 topdir_80 topdir_86 topdir_91 topdir_97
topdir_10 topdir_16 topdir_21 topdir_27 topdir_32 topdir_38 topdir_43 topdir_49 topdir_54
topdir_6 topdir_65 topdir_70 topdir_76 topdir_81 topdir_87 topdir_92 topdir_98
topdir_11 topdir_17 topdir_22 topdir_28 topdir_33 topdir_39 topdir_44 topdir_5 topdir_55
topdir_60 topdir_66 topdir_71 topdir_77 topdir_82 topdir_88 topdir_93 topdir_99 topdir_12
topdir_18 topdir_23 topdir_29 topdir_34 topdir_4 topdir_45 topdir_50 topdir_56 topdir_61
topdir_67 topdir_72 topdir_78 topdir_83 topdir_89 topdir_94
topdir_13 topdir_19 topdir_24 topdir_3 topdir_35 topdir_40 topdir_46 topdir_51 topdir_57
topdir_62 topdir_68 topdir_73 topdir_79 topdir_84 topdir_9 topdir_95
```

新たに変換された FlexGroup ボリュームの拡張は簡単です。volume expand を使用してさらにメンバーボリュームを追加できます。

```
cluster::*> volume expand -vserver DEMO -volume lotsafiles -aggr-list aggr1_node1,aggr1_node2
-aggr-list-multiplier 2

Warning: The following number of constituents of size 10TB will be added to FlexGroup "lotsafiles": 4.
Expanding the FlexGroup will cause the state of all Snapshot copies to be set to "partial".
Partial Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y

Warning: FlexGroup "lotsafiles" is a converted flexible volume. If this volume is expanded, it will no
longer be able to be converted back to being a flexible volume.
Do you want to continue? {y|n}: y
[Job 23676] Job succeeded: Successful
```

しかし、覚えておいてほしいのは、データは再配布できないということです。元のメンバーボリュームは、ファイルを所定の位置に保持します。

```
cluster::*> df -i lots*
Filesystem iused ifree %iused Mounted on Vserver
/vol/lotsafiles/ 3630682 102624948 3% /lotsafiles DEMO
/vol/lotsafiles__0001/ 3630298 17620828 17% /lotsafiles DEMO
/vol/lotsafiles__0002/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0003/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0004/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0005/ 96 21251030 0% --- DEMO
6 entries were displayed.

cluster::*> df -h lots*
Filesystem total used avail capacity Mounted on Vserver
/vol/lotsafiles/ 47TB 2735MB 14TB 0% /lotsafiles DEMO
/vol/lotsafiles/.snapshot
2560GB 49MB 2559GB 0% /lotsafiles/.snapshot DEMO
/vol/lotsafiles__0001/ 9728GB 2505MB 7505GB 0% /lotsafiles DEMO
/vol/lotsafiles__0001/.snapshot
512GB 49MB 511GB 0% /lotsafiles/.snapshot DEMO
/vol/lotsafiles__0002/ 9728GB 57MB 7505GB 0% --- DEMO
/vol/lotsafiles__0002/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0003/ 9728GB 57MB 7766GB 0% --- DEMO
/vol/lotsafiles__0003/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0004/ 9728GB 57MB 7505GB 0% --- DEMO
/vol/lotsafiles__0004/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0005/ 9728GB 57MB 7766GB 0% --- DEMO
/vol/lotsafiles__0005/.snapshot 512GB 0B
512GB 0% --- DEMO
12 entries were displayed.
```

17.6 既存の SnapMirror 関係における FlexVol の変換例

次に、既存の SnapMirror がある FlexVol を変換する例を示します。

1. これは、SnapMirror 関係にあるボリュームです。

```
cluster::*> snapmirror show -destination-path data_dst -fields state
source-path destination-path state
----------- ---------------- ------------
DEMO:data   DEMO:data_dst    Snapmirrored
```

2. ソースを変換しようとすると、エラーが発生します。

```
cluster::*> vol conversion start -vserver DEMO -volume data -check-only true

Error: command failed: Cannot convert volume "data" in Vserver "DEMO" to a FlexGroup. Correct
the following issues and retry the command:
    * Cannot convert source volume "data" because destination volume "data_dst" of the
SnapMirror relationship with "data" as the source is not converted. First check if the source
can be converted to a FlexGroup volume using "vol conversion start -volume data -convert-to
flexgroup -check-only true". If the conversion of the source can proceed then first convert the
destination and then convert the source.
```

3 そのため、最初にデスティネーションを変換する必要があります。そのためには、SnapMirror 関係を停止する必要があります。

```
cluster::*> vol conversion start -vserver DEMO -volume data_dst -check-only true

Error: command failed: Cannot convert volume "data_dst" in Vserver "DEMO" to a FlexGroup.
Correct the following issues and retry the command:
* The relationship was not quiesced. Quiesce SnapMirror relationship using "snapmirror quiesce
-destination-path data_dst" and then try the conversion.
```

4 これでボリュームを変換できます。

```
cluster::*> snapmirror quiesce -destination-path DEMO:data_dst
Operation succeeded: snapmirror quiesce for destination "DEMO:data_dst".

cluster::*> vol conversion start -vserver DEMO -volume data_dst -check-only true
Conversion of volume "data_dst" in Vserver "DEMO" to a FlexGroup can proceed with the following
warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to a
flexible volume.
* Converting flexible volume "data_dst" in Vserver "DEMO" to a FlexGroup will cause the state of
all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot
copies cannot be restored.
```

5 ボリュームを変換すると、次のステップが通知されます。

```
cluster::*> vol conversion start -vserver DEMO -volume data_dst

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
Warning: Converting flexible volume "data_dst" in Vserver "DEMO" to a FlexGroup will cause the
state of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion
Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23710] Job succeeded: SnapMirror destination volume "data_dst" has been successfully
converted to a FlexGroup volume. You must now convert the relationship's source volume,
"DEMO:data", to a FlexGroup. Then, re-establish the SnapMirror relationship using the
"snapmirror resync" command.
```

6 次に、ソースボリュームを変換します。

```
cluster::*> vol conversion start -vserver DEMO -volume data

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
Warning: Converting flexible volume "data" in Vserver "DEMO" to a FlexGroup will cause the state
of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot
copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23712] Job succeeded: success
```

7 ミラーを再同期します。

```
cluster::*> snapmirror resync -destination-path  DEMO:data_dst
Operation is queued: snapmirror resync to destination "DEMO:data_dst".

cluster::*> snapmirror show -destination-path DEMO:data_dst -fields state source-path
destination-path state
----------- ---------------- ------------
DEMO:data   DEMO:data_dst    Snapmirrored
```

変換は正常に機能しますが、SnapMirror 関係で最も重要なのはリストアです。したがって、デスティネーションボリュームの Snapshot コピーからファイルにアクセスできるかどうかを確認する必要があります。

8 まず、ソースとデスティネーションをマウントし、ls の出力を比較します。

```
# mount -o nfsvers=3 DEMO:/data_dst /dst
# mount -o nfsvers=3 DEMO:/data /data
```

これがソースボリュームの内容です。

```
# ls -lah /data
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
dr-xr-xr-x. 54 root root 4.0K Nov 15 10:08 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
drwxr-xr-x 12 root root 4.0K Nov 16 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
drwxrwxrwx 5 root root 4.0K Nov 15 10:06 .snapshot
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

デスティネーションボリュームが、必要に応じて正確に一致しています。

```
# ls -lah /dst
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
dr-xr-xr-x. 54 root root 4.0K Nov 15 10:08 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
dr-xr-xr-x 2 root root 0 Nov 15 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
drwxrwxrwx 4 root root 4.0K Nov 15 10:05 .snapshot
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

9 デスティネーションボリューム内の Snapshot コピーに対して ls コマンドを実行すると、目的のファイルが表示されます。

```
# ls -lah /dst/.snapshot/snapmirror.7e3cc08e-d9b3-11e6-85e2-
00a0986b1210_2163227795.2019-11- 15_100555/
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
drwxrwxrwx 4 root root 4.0K Nov 15 10:05 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
dr-xr-xr-x 2 root root 0 Nov 15 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

10 FlexGroup ソースを拡張して容量を増やします。

```
cluster::*> volume expand -vserver DEMO -volume data -aggr-list
aggr1_node1,aggr1_node2 -aggr- list-multiplier

Warning: The following number of constituents of size 30TB will be added to FlexGroup
"data": 4. Expanding the FlexGroup will cause the state of all Snapshot copies to be
set to "partial".
Partial Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y [Job 23720] Job succeeded: Successful
```

これで、ソースボリュームに 5 つのメンバーボリュームができました。デスティネーションボリュームには 1 つしかありません。

```
cluster::*> vol show -vserver DEMO -volume data*
Vserver  Volume  Aggregate  State  Type  Size  Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO data - online RW 150TB 14.89TB 0%
DEMO data__0001 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0002 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0003 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0004 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0005 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data_dst - online DP 30TB 7.32TB 0%
DEMO data_dst__0001
aggr1_node1 online DP 30TB 7.32TB 0%
8 entries were displayed.
```

11 ミラーを更新すると、ONTAP によって自動的に修正されます。

```
cluster::*> snapmirror update  -destination-path DEMO:data_dst
Operation is queued: snapmirror update of destination "DEMO:data_dst".
```

最初に更新が失敗し、次のエラーメッセージが表示されます。

```
Last Transfer Error: A SnapMirror transfer for the relationship with destination FlexGroup
"DEMO:data_dst" was aborted because the source FlexGroup was expanded. A SnapMirror AutoExpand
job with id "23727" was created to expand the destination FlexGroup and to trigger a SnapMirror
transfer for the SnapMirror relationship. After the SnapMirror transfer is successful, the
"healthy" field of the SnapMirror relationship will be set to "true". The job can be monitored
using either the "job show -id 23727" or "job history show -id 23727" commands.
```

12 ジョブによってボリュームが拡張されるので、再度更新できます。

```
cluster::*> job show -id 23727
Owning
Job ID Name Vserver Node State
------ -------------------- ---------- -------------- ----------
23727 Snapmirror Expand cluster
node1 Success
Description: SnapMirror FG Expand data_dst


cluster::*> snapmirror show -destination-path DEMO:data_dst -fields state source-path
destination-path state
----------- --------------- ------------
DEMO:data DEMO:data_dst Snapmirrored
```

これで、両方の FlexGroup ボリュームのメンバーボリューム数が同じになりました。

```
cluster::*> vol show -vserver DEMO -volume data*
Vserver Volume Aggregate State Type Size Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO data - online RW 150TB 14.88TB 0%
DEMO data__0001 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0002 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0003 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0004 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0005 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data_dst - online DP 150TB 14.88TB 0%
DEMO data_dst__0001
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0002
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0003
aggr1_node2 online DP 30TB 7.57TB 0%
DEMO data_dst__0004
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0005
aggr1_node2 online DP 30TB 7.57TB 0%
12 entries were displayed.
```

17.7 FlexVol から FlexGroup への変換例：5 億ファイル

この例では、5 億個のファイルを持つ FlexVol ボリュームを FlexGroup ボリュームに変換します。

```
cluster::*> vol show -vserver DEMO -volume fvconvert -fields files,files-used,is-flexgroup
vserver volume    files      files-used is-flexgroup
------- --------- ---------- ---------- ------------
DEMO    fvconvert 2040109451 502631608  false
```

その多くのファイルの作成に時間がかかったため、FlexClone ボリュームを作成して分割しました。このアプローチでは、元のボリュームを変更することなく、リスクなしでテストできます。

この例では、クローン作成プロセスに約 30 分かかりました。

```
cluster::*> vol clone split start -vserver DEMO -flexclone fvconvert -foreground true

Warning: Are you sure you want to split clone volume fvconvert in Vserver DEMO ? {y|n}: y
[Job 24230] 0% inodes processed.

cluster::*> job history show -id 24230 -fields starttime,endtime
node                record  vserver         endtime        starttime
------------------- ------- --------------- -------------- --------------
node1               2832338 cluster         12/09 10:27:08 12/09 09:58:16
```

クローンの分割後、チェックを実行しました。共有 FlexClone ブロックを削除するために、`volume clone sharing-by-split undo` を実行する必要がありました。これには時間がかかりましたが、チェックによって次の出力が生成されました。

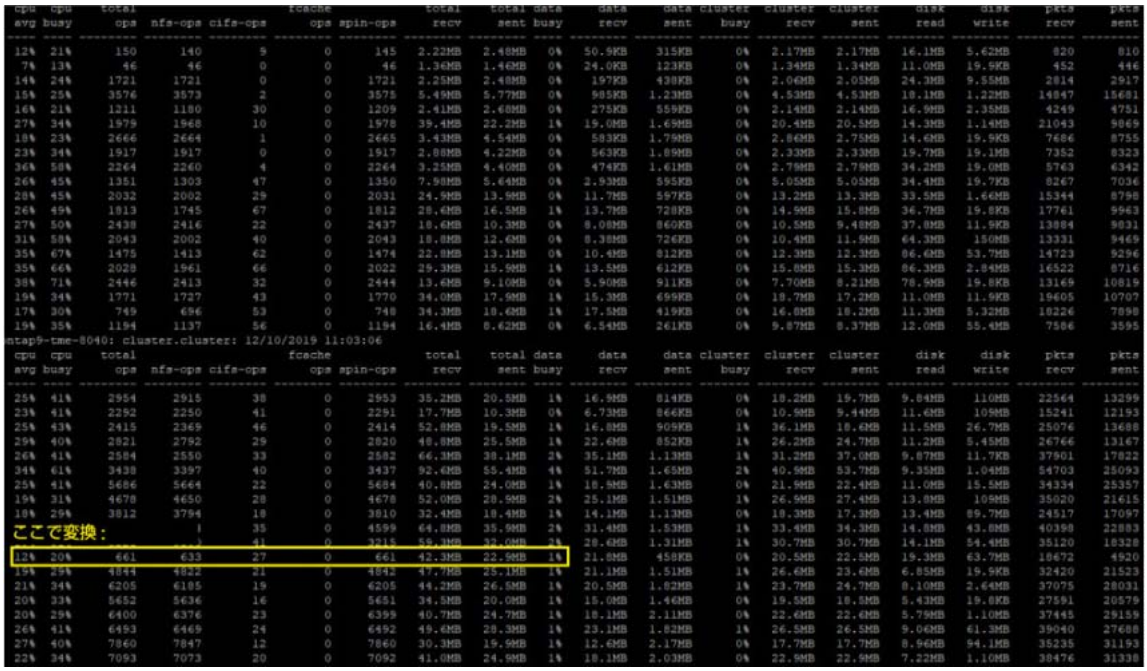
```
cluster::*> volume conversion start -vserver DEMO -volume fvconvert -foreground true
-check-only true
Conversion of volume "fvconvert" in Vserver "DEMO" to a FlexGroup can proceed with the
following warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to
a flexible volume.
```

次に、前に実行したスクリプトを実行して負荷を生成し、クラスタの統計を監視して停止が発生したかどうかを確認しました。この場合も、変換には (5 億ファイルで) 数秒かかり、ほんの少し、ほとんど気づかれない程度の遅れがありました。

```
cluster::*> volume conversion start -vserver DEMO -volume fvconvert -foreground true

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it
back to a flexible volume.
Do you want to continue? {y|n}: y
[Job 24259] Job succeeded: success
```

図 17.4　　変換プロセスのサンプル統計



cpu avg	cpu busy	total ops	nfs-ops	cifs-ops	fcache ops	spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	cluster read	disk write	pkts recv	pkts sent
12%	21%	150	140	9	0	145	2.22MB	2.48MB	0%	50.9KB	315KB	0%	2.17MB	2.17MB	16.1MB	5.62MB	820	810
7%	13%	46	46	0	0	46	1.36MB	1.46MB	0%	24.0KB	123KB	0%	1.34MB	1.34MB	11.0MB	19.9KB	452	446
14%	24%	1721	1721	0	0	1721	2.25MB	2.48MB	0%	197KB	438KB	0%	2.06MB	2.05MB	24.3MB	9.55MB	2814	2917
15%	25%	3576	3573	2	0	3575	5.49MB	5.77MB	0%	985KB	1.23MB	0%	4.53MB	4.53MB	18.1MB	1.22MB	14847	15681
16%	21%	1211	1180	30	0	1209	2.41MB	2.68MB	0%	275KB	559KB	0%	2.14MB	2.14MB	16.9MB	2.35MB	4249	4751
27%	34%	1979	1968	10	0	1978	39.4MB	22.2MB	1%	19.0MB	1.69MB	0%	20.4MB	20.5MB	14.3MB	1.14MB	21043	9869
18%	23%	2666	2664	1	0	2665	3.43MB	4.54MB	0%	583KB	1.79MB	0%	2.86MB	2.75MB	14.6MB	19.9KB	7686	8755
23%	34%	1917	1917	0	0	1917	2.88MB	4.22MB	0%	563KB	1.89MB	0%	2.33MB	2.33MB	19.7MB	19.1MB	7352	8323
36%	58%	2264	2260	4	0	2264	3.25MB	4.40MB	0%	474KB	1.61MB	0%	2.79MB	2.79MB	34.2MB	19.0MB	5763	6342
26%	45%	1351	1303	47	0	1350	7.98MB	5.64MB	0%	2.93MB	595KB	0%	5.05MB	5.05MB	34.4MB	19.7KB	8267	7036
28%	45%	2032	2002	29	0	2031	24.9MB	13.9MB	0%	11.7MB	597KB	0%	13.2MB	13.3MB	33.5MB	1.66MB	15344	8798
26%	49%	1813	1745	67	0	1812	28.6MB	16.5MB	1%	13.7MB	728KB	0%	14.9MB	15.8MB	36.7MB	19.8KB	17761	9963
27%	50%	2438	2416	22	0	2437	18.6MB	10.3MB	0%	8.08MB	860KB	0%	10.5MB	9.48MB	37.8MB	11.9KB	13884	9831
31%	58%	2043	2002	40	0	2043	18.8MB	12.6MB	0%	8.38MB	726KB	0%	10.4MB	11.9MB	64.3MB	150MB	13331	9469
35%	67%	1475	1413	62	0	1474	22.8MB	13.1MB	0%	10.4MB	812KB	0%	12.3MB	12.3MB	86.6MB	53.7MB	14723	9296
35%	66%	2028	1961	66	0	2022	29.3MB	15.9MB	1%	13.5MB	612KB	0%	15.8MB	15.3MB	86.3MB	2.84MB	16522	8716
38%	71%	2446	2413	32	0	2444	13.6MB	9.10MB	0%	5.90MB	911KB	0%	7.70MB	8.21MB	78.9MB	19.8KB	13169	10819
19%	34%	1771	1727	43	0	1770	34.0MB	17.9MB	1%	15.3MB	699KB	0%	18.7MB	17.2MB	11.0MB	11.9KB	19605	10707
17%	30%	749	696	53	0	748	34.3MB	18.6MB	1%	17.5MB	419KB	0%	16.8MB	18.2MB	11.3MB	5.32MB	18226	7898
19%	35%	1194	1137	56	0	1194	16.4MB	8.62MB	0%	6.54MB	261KB	0%	9.87MB	8.37MB	12.0MB	55.4MB	7586	3595

ntap9-tme-8040: cluster.cluster: 12/10/2019 11:03:06

cpu avg	cpu busy	total ops	nfs-ops	cifs-ops	fcache ops	spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	cluster read	disk write	pkts recv	pkts sent
25%	41%	2954	2915	38	0	2953	35.2MB	20.5MB	1%	16.9MB	814KB	0%	18.2MB	19.7MB	9.84MB	110MB	22564	13299
23%	41%	2292	2250	41	0	2291	17.7MB	10.3MB	0%	6.73MB	866KB	0%	10.9MB	9.44MB	11.6MB	109MB	15241	12193
25%	43%	2415	2369	46	0	2414	52.8MB	19.5MB	1%	16.8MB	909KB	1%	36.1MB	18.6MB	11.5MB	26.7MB	25076	13688
29%	40%	2821	2792	29	0	2820	48.8MB	25.5MB	1%	22.6MB	852KB	1%	26.2MB	24.7MB	11.2MB	5.45MB	26766	13167
26%	41%	2584	2550	33	0	2582	66.3MB	38.1MB	2%	35.1MB	1.13MB	1%	31.2MB	37.0MB	9.87MB	11.7KB	37901	17822
34%	61%	3438	3397	40	0	3437	92.6MB	55.4MB	4%	51.7MB	1.65MB	2%	40.9MB	53.7MB	9.35MB	1.04MB	54703	25093
25%	41%	5686	5664	22	0	5684	40.8MB	24.0MB	1%	18.5MB	1.63MB	0%	21.9MB	22.4MB	11.0MB	15.5MB	34334	25357
19%	31%	4678	4650	28	0	4678	52.0MB	28.9MB	2%	25.1MB	1.51MB	1%	26.9MB	27.4MB	13.8MB	109MB	35020	21615
18%	29%	3812	3794	18	0	3810	32.4MB	18.4MB	1%	14.1MB	1.13MB	0%	18.3MB	17.3MB	13.4MB	89.7MB	24517	17097
ここで変換：																		
		1		41	0	3215	59.3MB	32.0MB	2%	28.6MB	1.31MB	1%	30.7MB	30.7MB	14.1MB	54.4MB	35120	18328
12%	20%	661	633	27	0	661	42.3MB	22.9MB	1%	21.8MB	458KB	0%	20.5MB	22.5MB	19.3MB	63.7MB	18672	4920
19%	29%	4844	4822	21	0	4842	47.7MB	25.1MB	1%	21.1MB	1.51MB	1%	26.6MB	23.6MB	6.85MB	19.9KB	32420	21523
21%	34%	6205	6185	19	0	6205	44.2MB	26.5MB	1%	20.5MB	1.82MB	1%	23.7MB	24.7MB	8.10MB	2.64MB	37075	28031
20%	33%	5652	5636	16	0	5651	34.5MB	20.0MB	1%	15.0MB	1.46MB	0%	19.5MB	18.5MB	5.43MB	19.8KB	27591	20579
20%	29%	6400	6376	23	0	6399	40.7MB	24.7MB	1%	18.1MB	2.11MB	0%	22.6MB	22.6MB	5.79MB	1.10MB	37445	29159
26%	41%	6493	6469	24	0	6492	49.6MB	28.3MB	1%	23.1MB	1.82MB	1%	26.5MB	26.5MB	9.06MB	61.3MB	39040	27688
27%	40%	7860	7847	12	0	7860	30.3MB	19.9MB	1%	12.6MB	2.17MB	0%	17.7MB	17.7MB	8.96MB	94.1MB	35235	31193
22%	34%	7093	7073	20	0	7092	41.0MB	24.9MB	1%	18.1MB	2.03MB	0%	22.9MB	22.9MB	7.22MB	1.10MB	38476	31338

次に、スクリプトの実行中に、FlexGroup ボリュームに新しいメンバーボリュームを追加しました。ここでも、中断はありませんでした。

```
cluster::*> volume expand -vserver DEMO -volume fvconvert -aggr-list aggr1_node1
-aggr-list- multiplier 3 -foreground true

Warning: The following number of constituents of size 40TB will be added to FlexGroup
"fvconvert": 3.
Do you want to continue? {y|n}: y [Job 24261] Job succeeded: Successful
```

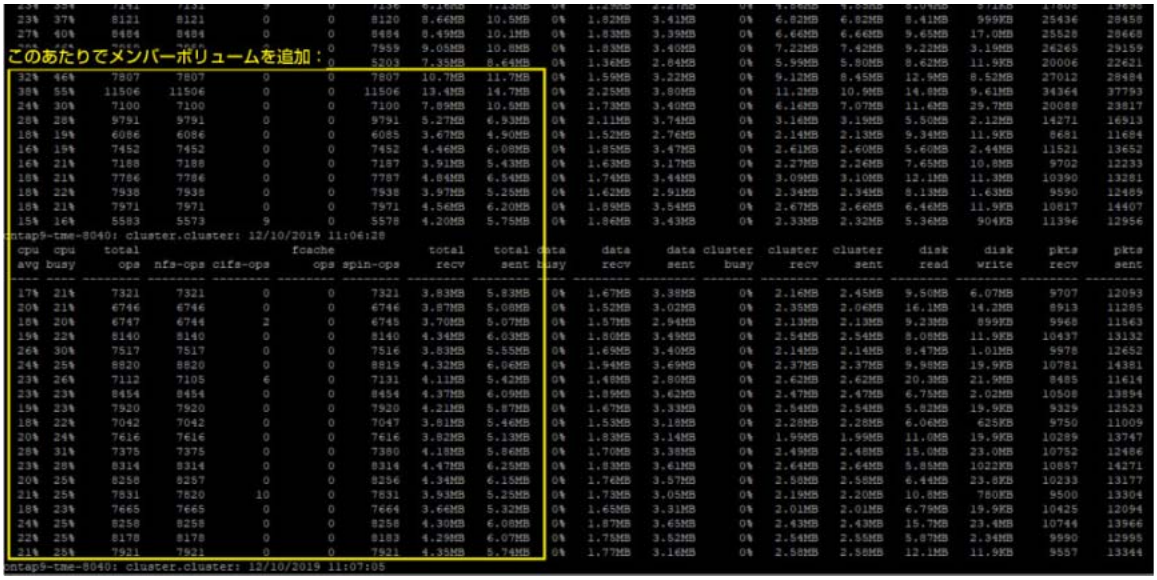
その後、さらに 4 つのメンバーボリュームを追加しました。

```
cluster::*> volume expand -vserver DEMO -volume fvconvert -aggr-list aggr1_node2
-aggr-list- multiplier 4

Warning: The following number of constituents of size 40TB will be added to FlexGroup
"fvconvert": 4.
Do you want to continue? {y|n}: y
[Job 24264] Job succeeded: Successful
```

さらに、ワークロードの合計 IOPS が増加しています。ジョブ自体は、FlexVol ボリューム上で実行した場合よりも全体的に時間がかかりませんでした。これは、FlexGroup ボリュームのパラレル取り込みが開始され、スクリプトの実行速度が向上したためです。

図 17.5　変換プロセス中の統計の例 : メンバーボリュームの追加



このキャプチャのビデオを見ることができます。

各ジョブの完了時間も記録しました。

```
This was the job on the FlexVol before it was converted:
# python file-create.py /fvconvert/files
Starting overall work: 2019-12-09 10:32:21.966337
End overall work: 2019-12-09 12:11:15.990707
total time: 5934.024611
```

FlexVol ボリュームを (追加されたメンバーボリュームを持つ)FlexGroup ボリュームに変換することで、時間を節約できました。

```
# python file-create.py /fvconvert/files2
Starting overall work: 2019-12-10 11:02:28.621532
End overall work: 2019-12-10 12:22:48.523772
total time: 4819.95753193
```

これにより、約 1100 秒 (18 分) の時間が節約され、完了までの合計時間の約 20% を節約できました。

次の出力は、スクリプト実行前のファイル配布を示しています。最初のメンバーボリュームは、以前は FlexVol ボリュームであったため、ファイル数が最も多いことに注意してください。

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume           files      files-used
------- --------------- ---------- ----------
DEMO    fvconvert__0001 2040109451 502848737
DEMO    fvconvert__0002 2040109451 12747
DEMO    fvconvert__0003 2040109451 12749
DEMO    fvconvert__0004 2040109451 12751
```

ジョブの最後に、ファイルが均等に分散していることがわかります。

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume           files      files-used
------- --------------- ---------- ----------
DEMO    fvconvert__0001 2040109451 506770209
DEMO    fvconvert__0002 2040109451 3345330
DEMO    fvconvert__0003 2040109451 3345330
DEMO    fvconvert__0004 2040109451 3345319
DEMO    fvconvert__0005 2040109451 3331657
DEMO    fvconvert__0006 2040109451 3331635
DEMO    fvconvert__0007 2040109451 3331657
DEMO    fvconvert__0008 2040109451 3331657
```

新しく変換された FlexGroup ボリュームでスクリプトを再度実行しました。今回は、ジョブの実行速度と、空の FlexVol メンバーボリューム上でファイルがどのように分散されるかを確認しました。

最初に説明したように、新しいメンバーボリュームでは、使用されるファイルはすべて 1% 未満でした (20 億個のファイルのうち 330 万個)。FlexVol ボリュームから変換されたメンバーボリュームは、全ファイルの 25% を使用していました (20 億のうち 5 億)。

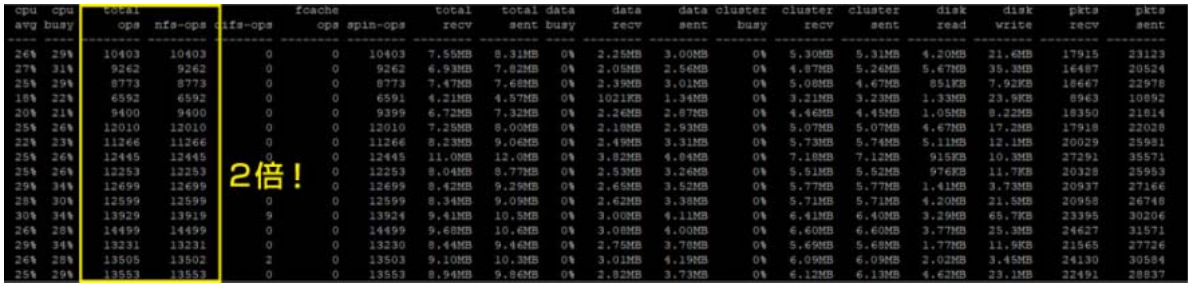
ジョブの実行後、元のメンバーボリュームのファイルカウントデルタが約 320 万、他のすべてのメンバーのファイルカウントデルタが約 358 万になりました。すべてのメンバーボリューム間のバランシングは維持されますが、新しいファイルやフォルダの作成には、より少ないボリュームが優先されます。

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume          files      files-used
------- --------------- ---------- ----------
DEMO    fvconvert__0001 2040109451 509958440
DEMO    fvconvert__0002 2040109451 6808792
DEMO    fvconvert__0003 2040109451 6809225
DEMO    fvconvert__0004 2040109451 6806843
DEMO    fvconvert__0005 2040109451 6798959
DEMO    fvconvert__0006 2040109451 6800054
DEMO    fvconvert__0007 2040109451 6849375
DEMO    fvconvert__0008 2040109451 6801600
```

新しい FlexGroup ボリュームを FlexVol ボリュームから変換すると、ジョブ時間が 5900 秒から 4656 秒に短縮されました。また、IOPS を 2 倍にすることもできました。

```
# python file-create.py /fvconvert/files3
Starting overall work: 2019-12-10 13:14:26.816860
End overall work: 2019-12-10 14:32:03.565705
total time: 4656.76723099
```

図 17.6　変換プロセスの統計の例：2 倍のパフォーマンス



cpu avg	cpu busy	total ops	total nfs-ops	cifs-ops	fcache ops	fcache spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	disk read	disk write	pkts recv	pkts sent
26%	29%	10403	10403	0	0	10403	7.55MB	8.31MB	0%	2.25MB	3.00MB	0%	5.30MB	5.31MB	4.20MB	21.6MB	17915	23123
27%	31%	9262	9262	0	0	9262	6.93MB	7.82MB	0%	2.05MB	2.56MB	0%	4.87MB	5.26MB	5.67MB	35.3MB	16487	20524
25%	29%	8773	8773	0	0	8773	7.47MB	7.68MB	0%	2.39MB	3.01MB	0%	5.08MB	4.67MB	851KB	7.92KB	18667	22978
18%	22%	6592	6592	0	0	6591	4.21MB	4.57MB	0%	1021KB	1.34MB	0%	3.21MB	3.23MB	1.33MB	23.9KB	8963	10892
20%	21%	9400	9400	0	0	9399	6.72MB	7.32MB	0%	2.26MB	2.87MB	0%	4.46MB	4.45MB	1.05MB	8.22MB	18350	21814
25%	26%	12010	12010	0	0	12010	7.25MB	8.00MB	0%	2.18MB	2.93MB	0%	5.07MB	5.07MB	4.67MB	17.2MB	17918	22028
22%	23%	11266	11266	0	0	11266	8.23MB	9.06MB	0%	2.49MB	3.31MB	0%	5.73MB	5.74MB	5.11MB	12.1MB	20029	25981
25%	26%	12445	12445	0	0	12445	11.0MB	12.0MB	0%	3.82MB	4.84MB	0%	7.18MB	7.12MB	915KB	10.3MB	27291	35571
25%	26%	12253	12253	0	0	12253	8.04MB	8.77MB	0%	2.53MB	3.26MB	0%	5.51MB	5.52MB	976KB	11.7KB	20328	25953
29%	34%	12699	12699	0	0	12699	8.42MB	9.29MB	0%	2.65MB	3.52MB	0%	5.77MB	5.77MB	1.41MB	3.73MB	20937	27166
28%	30%	12599	12599	0	0	12599	8.34MB	9.09MB	0%	2.62MB	3.38MB	0%	5.71MB	5.71MB	4.20MB	21.5MB	20958	26748
30%	34%	13929	13919	9	0	13924	9.41MB	10.5MB	0%	3.00MB	4.11MB	0%	6.41MB	6.40MB	3.29MB	65.7KB	23395	30206
26%	28%	14499	14499	0	0	14499	9.68MB	10.6MB	0%	3.08MB	4.00MB	0%	6.60MB	6.60MB	3.77MB	25.3MB	24627	31571
29%	34%	13231	13231	0	0	13230	8.44MB	9.46MB	0%	2.75MB	3.78MB	0%	5.69MB	5.68MB	1.77MB	11.9KB	21565	27726
26%	28%	13505	13502	2	0	13503	9.10MB	10.3MB	0%	3.01MB	4.19MB	0%	6.09MB	6.09MB	2.02MB	3.45MB	24130	30584
25%	29%	13553	13553	0	0	13553	8.94MB	9.86MB	0%	2.82MB	3.73MB	0%	6.12MB	6.13MB	4.62MB	23.1MB	22491	28837

ご覧のように、これらのメンバーボリュームには (元の FlexVol ボリュームよりももっと多くの) ファイルとデータの不均衡がありますが、複数のノードにわたる作業がより効率的であるため、パフォーマンスは以前の FlexVol パフォーマンスよりも大幅に向上しています。これが FlexGroup ボリュームのパワーです。

17.8 イベント管理システムの例

17.8.1 inode 関連の EMS の例

```
Message Name: callhome.no.inodes
Severity: ERROR

Corrective Action: Modify the volume's maxfiles (maximum number of files) to increase the
inodes on the affected volume. If you need assistance, contact Fujitsu technical support.

Description: This message occurs when a volume is out of inodes, which refer to individual
files, other types of files, and directories. If your system is configured to do so, it
generates and transmits an AutoSupport (or 'call home') message to Fujitsu technical
support and to the configured destinations. Successful delivery of an AutoSupport message
significantly improves problem determination and resolution.
```

```
Message Name: fg.inodes.member.nearlyFull
Severity: ALERT

Corrective Action: Adding capacity to the FlexGroup by using the "volume modify -files +X"
command is the best way to solve this problem. Alternatively, deleting files from the
FlexGroup might work, although it can be difficult to determine which files have landed on
which constituent.

Description: This message occurs when a constituent within a FlexGroup is almost out of
inodes. This constituent will receive far fewer new create requests than average, which
might impact the FlexGroup's overall performance, because those requests are routed to
constituents with more inodes.
```

```
Message Name: fg.inodes.member.full
Severity: ALERT

Corrective Action: Adding capacity to the FlexGroup by using the "volume modify -files +X"
command is the best way to solve this problem. Alternatively, deleting files from the
FlexGroup may work, but it is difficult to determine which files have landed on which
constituent.

Description: This message occurs when a constituent with a FlexGroup has run out of inodes.
New files cannot be created on this constituent. This might lead to an overall imbalanced
distribution of content across the FlexGroup.
```

```
Message Name: fg.inodes.member.allOK
Severity: NOTICE

Corrective Action: (NONE)

Description: This message occurs when conditions that led to previous
"fg.inodes.member.nearlyFull" and "fg.inodes.member.full" events no longer apply for any
constituent in this FlexGroup. All constituents within this FlexGroup have sufficient
inodes for normal operation.
```

17.8.2 maxdirsize メッセージの例

```
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. However, 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. However, 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. However, doing so could
impact system performance. If you need to increase the maximum directory size, work with
technical support.

Description: This mesaage occurs when a directory surpasses 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
```

17.8.3 容量関連のイベント管理システムメッセージの例

```
Message Name: monitor.volume.full
Severity: DEBUG
Corrective Action: (NONE)

Description: This message occurs when one or more file systems are full, typically
indicating at least 98% full. This event is accompanied by global health monitoring
messages for the customer. The space usage is computed based on the active file system size
and is computed by subtracting the value of the "Snapshot Reserve" field from the value of
the "Used" field of the "volume show- space" command. The volume/aggregate can be over 100%
full due to space used or reserved by metadata. A value greater than 100% might cause
Snapshot(tm) copy space to become unavailable or cause the volume to become logically
overallocated. See the "vol.log.overalloc" EMS message for more information.
                                Supports SNMP trap: true
                                      Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: monitor.volume.nearlyFull
Severity: ALERT
Corrective Action: Create space by increasing the volume or aggregate sizes, or by deleting
data or deleting Snapshot(R) copies. To increase a volume's size, use the "volume size"
command. To delete a volume's Snapshot(R) copies, use the "volume snapshot delete" command.
To increase an aggregate's size, add disks by using the "storage aggregate add-disks"
command. Aggregate Snapshot(R) copies are deleted automatically when the aggregate is full.
                                   Description: This message occurs when one or more file
systems are nearly full, typically indicating at least 95% full. This event is accompanied
by global health monitoring messages for the customer. The space usage is computed based on
the active file system size and is computed by subtracting the value of the "Snapshot
Reserve" field from the value of the "Used" field of the "volume show-space" command.
                                Supports SNMP trap: true
                                      Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: monitor.volume.ok
Severity: DEBUG
Corrective Action: (UNKNOWN)

Description: The previously-reported volume full condition is fixed. * We log this event,
as well as the other monitor.volume events, at LOG_DEBUG level to avoid spamming the
messages file with events which are already being reported as part of the global health
messages.
                                Supports SNMP trap: true
                                      Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: monitor.volumes.one.ok
Severity: DEBUG
Corrective Action: (NONE)

Description: This message occurs when one file system that was nearly full (usually this
means >= 95% full) is now OK. This event and other "monitor.volume" events are logged at
LOG_DEBUG  level to avoid spamming the messages file with events that are already being
reported as part of the global health messages. The space usage is computed based on the
active file system size and is computed by subtracting the value of the "Snapshot Reserve"
field from the value of the "Used" field of the "volume show-space" command.
                              Supports SNMP trap: true
                                    Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: vol.log.overalloc
Severity: ALERT
Corrective Action: Create space by increasing the volume or aggregate size, deleting data,
deleting Snapshot(R) copies, or changing the provisioning from thick to thin. To increase a
volume's size, use the "volume size" command. To delete a volume's Snapshot(R) copies, use
the "volume snapshot delete" command. To change provisioning in a volume, reserved files
can be unreserved by using the "volume file reservation" command. To increase an
aggregate's size, add disks by using the "storage aggregate add-disks" command. Aggregate
Snapshot(R) copies are deleted automatically when the aggregate is full. To change
provisioning of a volume in an aggregate, change the volume guarantee from "volume" to
"none" by using the "space-guarantee" field of the "volume modify" command.

Description: This message occurs when the volume or aggregate allocates more space than it
can honor by way of reservations, or the aggregate has allocated more space than it can
honor by way of guarantees. If the reserved or guaranteed space is consumed, there is
insufficient physical space, which can cause the volume or aggregate to be taken offline.
                              Supports SNMP trap: true
                                    Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: fg.member.elastic.sizing
Severity: NOTICE
Corrective Action: (NONE)
Description: This message occurs when a FlexGroup constituent undergoes elastic sizing,
either to restore balance among constituents or to resize constituents to accommodate space
needs.
Supports SNMP trap: false
Destinations: -
Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

18. コマンドの例

18.1 FlexGroup 容量コマンド

```
cluster::*> aggr show-space -instance -aggregate aggr1_node1

                          Aggregate Name: aggr1_node1
                       Volume Footprints: 2.05TB
               Volume Footprints Percent: 26%
Total Space for Snapshot Copies in Bytes: 0B
      Space Reserved for Snapshot Copies: 0%
                     Aggregate Metadata: 15.20MB
              Aggregate Metadata Percent: 0%
                              Total Used: 2.05TB
                     Total Used Percent: 26%
                                   Size: 7.86TB
               Snapshot Reserve Unusable: -
       Snapshot Reserve Unusable Percent: -
                Total Physical Used Size: 143.7GB
                Physical Used Percentage: 2%

                          Aggregate Name: aggr1_node2
                       Volume Footprints: 2.02TB
               Volume Footprints Percent: 26%
Total Space for Snapshot Copies in Bytes: 0B
      Space Reserved for Snapshot Copies: 0%
                     Aggregate Metadata: 8.63MB
              Aggregate Metadata Percent: 0%
                              Total Used: 2.02TB
                     Total Used Percent: 26%
                                   Size: 7.86TB
               Snapshot Reserve Unusable: -
       Snapshot Reserve Unusable Percent: -
                Total Physical Used Size: 69.71GB
                Physical Used Percentage: 1%
2 entries were displayed.
```

```
cluster::*> volume show-space -vserver SVM -volume flexgroup__*
      Vserver : SVM
      Volume : flexgroup__0001
      Feature                                     Used  Used%
      ------------------------------- ---------- ------
      User Data                            57.06MB     0%
      Filesystem Metadata                   3.51MB     0%
      Inodes                               87.26MB     0%
      Snapshot Reserve                       512GB     5%
      Deduplication                           12KB     0%
      Performance Metadata                    48KB     0%

      Total Used                           512.1GB     5%
      Total Physical Used                  148.3MB     0%

      Vserver : SVM
      Volume : flexgroup__0002
      Feature                                     Used  Used%
      ------------------------------- ---------- ------
      User Data                            57.03MB     0%
      Filesystem Metadata                   4.66MB     0%
      Inodes                               83.66MB     0%
      Snapshot Reserve                       512GB     5%
      Deduplication                           20KB     0%
      Performance Metadata                    44KB     0%

      Total Used                           512.1GB     5%
      Total Physical Used                  145.7MB     0%

      Vserver : SVM
      Volume : flexgroup__0003
      Feature                                     Used  Used%
      ------------------------------- ---------- ------
      User Data                            57.02MB     0%
      Filesystem Metadata                   3.66MB     0%
      Inodes                               84.55MB     0%
      Snapshot Reserve                       512GB     5%
      Deduplication                           12KB     0%
      Performance Metadata                    44KB     0%

      Total Used                           512.1GB     5%
      Total Physical Used                  145.6MB     0%

      Vserver : SVM
      Volume : flexgroup__0004
      Feature                                     Used  Used%
      ------------------------------- ---------- ------
      User Data                            57.19MB     0%
      Filesystem Metadata                   8.93MB     0%
      Inodes                               82.09MB     0%
      Snapshot Reserve                       512GB     5%
      Deduplication                           12KB     0%
      Performance Metadata                    44KB     0%

      Total Used                           512.1GB     5%
      Total Physical Used                  148.5MB     0%
```

```
Vserver : SVM
Volume : flexgroup__0005
Feature                                  Used  Used%
-------------------------------- ---------- ------
User Data                             3.99GB     0%
Filesystem Metadata                   4.88MB     0%
Inodes                               83.54MB     0%
Snapshot Reserve                       512GB     5%
Deduplication                           12KB     0%
Performance Metadata                    52KB     0%

Total Used                           516.1GB     5%
Total Physical Used                   4.08GB     0%

Vserver : SVM
Volume : flexgroup__0006
Feature                                  Used  Used%
-------------------------------- ---------- ------
User Data                            57.04MB     0%
Filesystem Metadata                   3.50MB     0%
Inodes                               87.26MB     0%
Snapshot Reserve                       512GB     5%
Deduplication                           12KB     0%
Performance Metadata                    44KB     0%

Total Used                           512.1GB     5%
Total Physical Used                  148.2MB     0%

Vserver : SVM
Volume : flexgroup__0007
Feature                                  Used  Used%
-------------------------------- ---------- ------
User Data                            57.02MB     0%
Filesystem Metadata                   3.50MB     0%
Inodes                               85.03MB     0%
Snapshot Reserve                       512GB     5%
Deduplication                           12KB     0%
Performance Metadata                    44KB     0%

Total Used                           512.1GB     5%
Total Physical Used                  145.9MB     0%

Vserver : SVM
Volume : flexgroup__0008
Feature                                  Used  Used%
-------------------------------- ---------- ------
User Data                            57.03MB     0%
Filesystem Metadata                   3.52MB     0%
Inodes                               86.12MB     0%
Snapshot Reserve                       512GB     5%
Deduplication                           12KB     0%
Performance Metadata                    44KB     0%

Total Used                           512.1GB     5%
Total Physical Used                  147.0MB     0%
```

```
cluster::> vol show -is-constituent true -volume flexgroup__*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup__0001
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0002
                       aggr1_node2  online     RW         10TB      5.08TB  49%
SVM       flexgroup__0003
                       aggr1_node1  online     RW         10TB      5.05TB  49%
SVM       flexgroup__0004
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0005
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0006
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0007
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0008
                       aggr1_node2  online     RW         10TB     5.08TB   49%
8 entries were displayed.

cluster::*> storage aggregate show -aggregate aggr1* -fields usedsize,size,percent-used
-sort-by percent-used
aggregate   percent-used size   usedsize
----------- ------------ ------ --------
aggr1_node1 26%          7.86TB 2.05TB
aggr1_node2 26%          7.86TB 2.02TB
2 entries were displayed.
```

18.2　クラスタ全体に対する statistics show-periodic コマンドの例

```
cluster::*> statistics  show-periodic
cluster: cluster.cluster: 11/30/2016 11:49:46
cpu  cpu        total   fcache  total   total data      data    data
cluster cluster cluster disk    disk    pkts    pkts
avg busy        ops nfs-ops cifs-ops    ops spin-ops    recv    sent busy       recv    sent
busy    recv     sent     read     write    recv     sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------- -------- -------- -------- -------- -------- --------
5%   5%   0        0        0        0        0        65.3KB   64.4KB   0%   2.22KB   1.13KB
0%      62.7KB   63.2KB   489KB    407KB    91       83
5%   5%   0        0        0        0        0        62.5KB   61.6KB   0%   1.28KB   767B
0%      61.0KB   60.9KB   23.8KB   23.8KB   64       60
4%   5%   0        0        0        0        0        62.3KB   61.3KB   0%   1.43KB   708B
0%      60.7KB   60.7KB   15.8KB   15.8KB
cluster: cluster.cluster: 11/30/2016 11:49:53
cpu  cpu        total   fcache  total   total data      data    data
cluster cluster cluster disk    disk    pkts    pkts
avg busy        ops nfs-ops cifs-ops    ops spin-ops    recv    sent busy       recv    sent
busy    recv     sent     read     write    recv     sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------- -------- -------- -------- -------- -------- --------
Minimums:
4%   5%   0        0        0        0        0        62.3KB   61.3KB   0%   1.28KB   708B
0%      60.7KB   60.7KB   15.8KB   15.8KB   64       58
4%   5%   0        0        0        0        0        63.4KB   62.4KB   0%   1.64KB   877B
0%      61.5KB   61.6KB   176KB    149KB    74       67
Maximums:
5%   5%   0        0        0        0        0        65.3KB   64.4KB   0%   2.22KB   1.13KB
0%      62.7KB   63.2KB   489KB    407KB  91  83
```

18.3 NFSv3 の読み取り / 書き込み操作に関するリアルタイムの SVM レベルの statistics show-periodic

```
cluster::*> statistics show-periodic -instance SVM -interval 2 -iterations 0 -summary true
-vserver SVM -object nfsv3 -counter nfsv3_ops|nfsv3_read_ops|nfsv3_write_ops
cluster: nfsv3.SVM: 11/30/2016 13:29:57
            nfsv3    nfsv3
   nfsv3     read    write    Complete    Number of
     ops      ops      ops Aggregation Constituents
-------- -------- -------- ----------- ------------
    2360        0      697     Yes          16
    2245        0      652     Yes          16
    2126        0      629     Yes          16
cluster: nfsv3.SVM: 11/30/2016 13:30:04
            nfsv3    nfsv3
   nfsv3     read    write    Complete    Number of
     ops      ops      ops Aggregation Constituents
-------- -------- -------- ----------- ------------
Minimums:
    2126        0      629      -           -
Averages for 3 samples:
    2243        0      659      -           -
Maximums:
    2360        0      697      -           -
```

18.4 リアルタイムの FlexGroup のローカルおよびリモート統計情報

```
cluster::*> statistics show-periodic -instance 0 -interval 2 -iterations 0 -summary true -
object flexgroup -counter
cat1_tld_local|cat1_tld_remote|cat2_hld_local|cat2_hld_remote|cat3_dir_local|cat3_dir_re-
mote|cat4
_fil_local|cat4_fil_remote
cluster: flexgroup.0: 11/30/2016 13:34:55
 cat1   cat1  cat2   cat2  cat3   cat3  cat4   cat4
  tld    tld   hld    hld   dir    dir   fil    fil    Complete    Number of
local remote local remote local remote local remote Aggregation Constituents
----- ------ ----- ------ ----- ------ ----- ------ ----------- ------------
    1      0    17    113     0      0   619      0     n/a         n/a
    0      1    17    114     0      0   654      0     n/a         n/a
    0      2    17    112     0      0   647      0     n/a         n/a
cluster: flexgroup.0: 11/30/2016 13:35:02
 cat1   cat1  cat2   cat2  cat3   cat3  cat4   cat4
  tld    tld   hld    hld   dir    dir   fil    fil    Complete    Number of
local remote local remote local remote local remote Aggregation Constituents
----- ------ ----- ------ ----- ------ ----- ------ ----------- ------------
Minimums:
    0      0    17    112     0      0   619      0      -           -
Averages for 3 samples:
    0      1    17    113     0      0   640      0      -           -
Maximums:
    1      2    17    114     0      0   654      0      -           -
```

18.5 FlexGroup ボリュームを作成し、デフォルト値より少ないメンバーボリュームを指定する例

このコマンドは、2 つのノードにまたがる 2 つの 5TB メンバーボリュームを持つ 10TB FlexGroup ボリュームを作成します。

```
cluster::> volume create -vserver DEMO -volume flexgroup -aggr-list aggr1_node1,aggr1_node2
-aggr-list-multiplier 1 -junction-path /flexgroup -size 10t

Warning: The FlexGroup "flexgroup" will be created with the following number of
constituents of size 5TB: 2.

Do you want to continue? {y|n}: y
```

備考

ボリュームが FlexGroup ボリュームであることを確認するには、`-aggr-list` フラグを使用する必要があります。

18.6 FlexGroup ボリューム作成用のサンプル REST API

次の REST API の例では、1 つのアグリゲートに対して 2TB、8 メンバーのシンプロビジョニングされた FlexGroup ボリュームを作成します。

```
{
   "aggregates": [
      {
         "name": "aggr1_node1"
      }
   ],
   "constituents_per_aggregate": 8,
   "efficiency": {
      "compaction": "inline",
      "compression": "inline",
      "cross_volume_dedupe": "inline",
      "dedupe": "inline"
   },
   "guarantee": {
      "type": "none"
   },
   "name": "RESTAPI_FG",
   "nas":
      { "export_policy":
      {
         "id": 42949672961,
         "name": "default"
      },
      "gid": 0,
      "path": "/RESTAPI_FG",
      "security_style": "unix",
      "uid": 0,
      "unix_permissions": 755
   },
   "size": "2T",
   "style": "flexgroup",
   "svm": {
      "name": "DEMO",
      "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
   }
}
```

作成後の FlexGroup は次のようになります。

```
cluster::*> vol show -vserver DEMO -volume REST*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG   -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG__0001
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0002
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0003
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0004
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0005
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0006
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0007
                       aggr1_node1  online     RW         256GB   243.1GB    0%
DEMO      RESTAPI_FG__0008
                       aggr1_node1  online     RW         256GB   243.1GB    0%
9 entries were displayed.
```

リストに複数のアグリゲートを含めるには、次の REST API を例として使用します。

```
{
   "aggregates": [
      { "name": "aggr1_node1" },{ "name": "aggr1_node2" }
   ],
   "efficiency":
      { "compaction": "inline",
      "compression": "inline",
      "cross_volume_dedupe": "inline",
      "dedupe": "inline"
   },
   "guarantee": {
      "type": "none"
   },
   "name": "RESTAPI_FG3",
   "nas":
      { "export_policy":
      {
         "id": 42949672961,
         "name": "default"
      },
      "gid": 0,
      "path": "/RESTAPI_FG3",
      "security_style": "unix",
      "uid": 0,
      "unix_permissions": 755
   },
   "size": "2T",
   "style": "flexgroup",
   "svm": {
      "name": "DEMO",
      "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
   }
}
```

次のようになります。

```
cluster::*> vol show -vserver DEMO -volume *FG3*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG3  -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG3__0001
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0002
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0003
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0004
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0005
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0006
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0007
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0008
                       aggr1_node2  online     RW        256GB    243.1GB    0%
9 entries were displayed.
```

次の REST API は、style オプションを使用して 4 つのメンバーからなる FlexGroup ボリューム y を作成しますが、constituents_per_aggregate オプションは指定しません。

```
{
   "aggregates": [
      {
         "name": "aggr1_node1"
      }
   ],
   "efficiency":
      { "compaction": "inline",
      "compression": "inline",
      "cross_volume_dedupe": "inline",
      "dedupe": "inline"
   },
   "guarantee": {
      "type": "none"
   },
   "name": "RESTAPI_FG2",
   "nas":
      { "export_policy":
      {
         "id": 42949672961,
         "name": "default"
      },
      "gid": 0,
      "path": "/RESTAPI_FG2",
      "security_style": "unix",
      "uid": 0,
      "unix_permissions": 755
   },
   "size": "2T",
   "style": "flexgroup",
   "svm": {
      "name": "DEMO",
      "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
   }
}
```

これが結果の FlexGroup です。

```
cluster::*> vol show -vserver DEMO -volume RESTAPI_FG2*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG2  -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG2__0001
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0002
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0003
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0004
                       aggr1_node1  online     RW        512GB    486.3GB    0%
5 entries were displayed.
```

18.7 FlexGroup ボリュームのサイズを増やす例

```
cluster::*> volume show -vserver SVM -volume flexgroup*
SVM       flexgroup    -            online     RW      70.20TB   10.14TB    85%
SVM       flexgroup__0001
                       aggr1_node1  online     RW         10TB    5.06TB    49%
SVM       flexgroup__0002
                       aggr1_node2  online     RW         10TB    5.08TB    49%
SVM       flexgroup__0003
                       aggr1_node1  online     RW         10TB    5.06TB    49%
SVM       flexgroup__0004
                       aggr1_node2  online     RW         10TB    5.08TB    49%
SVM       flexgroup__0005
                       aggr1_node1  online     RW         10TB    5.06TB    49%
SVM       flexgroup__0006
                       aggr1_node2  online     RW         10TB    5.08TB    49%
SVM       flexgroup__0007
                       aggr1_node1  online     RW         10TB    5.06TB    49%
SVM       flexgroup__0008
                       aggr1_node2  online     RW         10TB    5.08TB    49%

cluster::*> vol size -vserver SVM -volume flexgroup -new-size 100t vol size: Volume
"SVM:flexgroup" size set to 100t.

cluster::*> volume show -vserver SVM -volume flexgroup*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup    -            online     RW        100TB   10.14TB    89%
SVM       flexgroup__0001
                       aggr1_node1  online     RW      12.50TB    5.06TB    59%
SVM       flexgroup__0002
                       aggr1_node2  online     RW      12.50TB    5.08TB    59%
SVM       flexgroup__0003
                       aggr1_node1  online     RW      12.50TB    5.06TB    59%
SVM       flexgroup__0004
                       aggr1_node2  online     RW      12.50TB    5.08TB    59%
SVM       flexgroup__0005
                       aggr1_node1  online     RW      12.50TB    5.06TB    59%
SVM       flexgroup__0006
                       aggr1_node2  online     RW      12.50TB    5.08TB    59%
SVM       flexgroup__0007
                       aggr1_node1  online     RW      12.50TB    5.06TB    59%
SVM       flexgroup__0008
                       aggr1_node2  online     RW      12.50TB    5.08TB    59%
```

18.8 FlexGroup ボリュームの拡張例

```
cluster::*> volume show -vserver SVM -volume flexgroup4*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup4TB -            online     RW          4TB     3.78TB    5%
SVM       flexgroup4TB__0001
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0002
                       aggr1_node2  online     RW        512GB    481.2GB    6%
SVM       flexgroup4TB__0003
                       aggr1_node1  online     RW        512GB    481.5GB    5%
SVM       flexgroup4TB__0004
                       aggr1_node2  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0005
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0006
                       aggr1_node2  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0007
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0008
                       aggr1_node2  online     RW        512GB    485.5GB    5%

cluster::*> volume expand -vserver SVM -volume flexgroup4TB -aggr-list aggr1_node1,aggr1_node2
- aggr-list-multiplier 4

cluster::*> volume show -vserver SVM -volume flexgroup4*
Vserver   Volume       Aggregate    State      Type Size       Available  Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup4TB -            online     RW          8TB     7.78TB    1%
SVM       flexgroup4TB__0001
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0002
                       aggr1_node2  online     RW        512GB    481.2GB    1%
SVM       flexgroup4TB__0003
                       aggr1_node1  online     RW        512GB    481.5GB    1%
SVM       flexgroup4TB__0004
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0005
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0006
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0007
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0008
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0009
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0010
                       aggr1_node2  online     RW        512GB    481.2GB    1%
SVM       flexgroup4TB__0011
                       aggr1_node1  online     RW        512GB    481.5GB    1%
SVM       flexgroup4TB__0012
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0013
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0014
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0015
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0016
                       aggr1_node2  online     RW        512GB    485.5GB    1%
```

18.9 その他のコマンドライン例

18.9.1 flexgroup deploy を使用した FlexGroup ボリュームの作成

```
cluster::> flexgroup deploy -size 20PB -space-guarantee volume -vserver SVM -volume flexgroup
```

18.9.2 volume create を使用した複数ノード間での 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
```

18.9.3 FlexGroup Snapshot ポリシーの変更

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

18.9.4 ストレージ QoS の適用

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

18.9.5 ボリュームの自動拡張の適用

```
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.
```

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 ベストプラクティス / インプリメンテーションガイド

P3AG-6202-02Z0

発行年月　2025 年 3 月
発行責任　エフサステクノロジーズ株式会社

- 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願います。
- 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。

Fsas Technologies