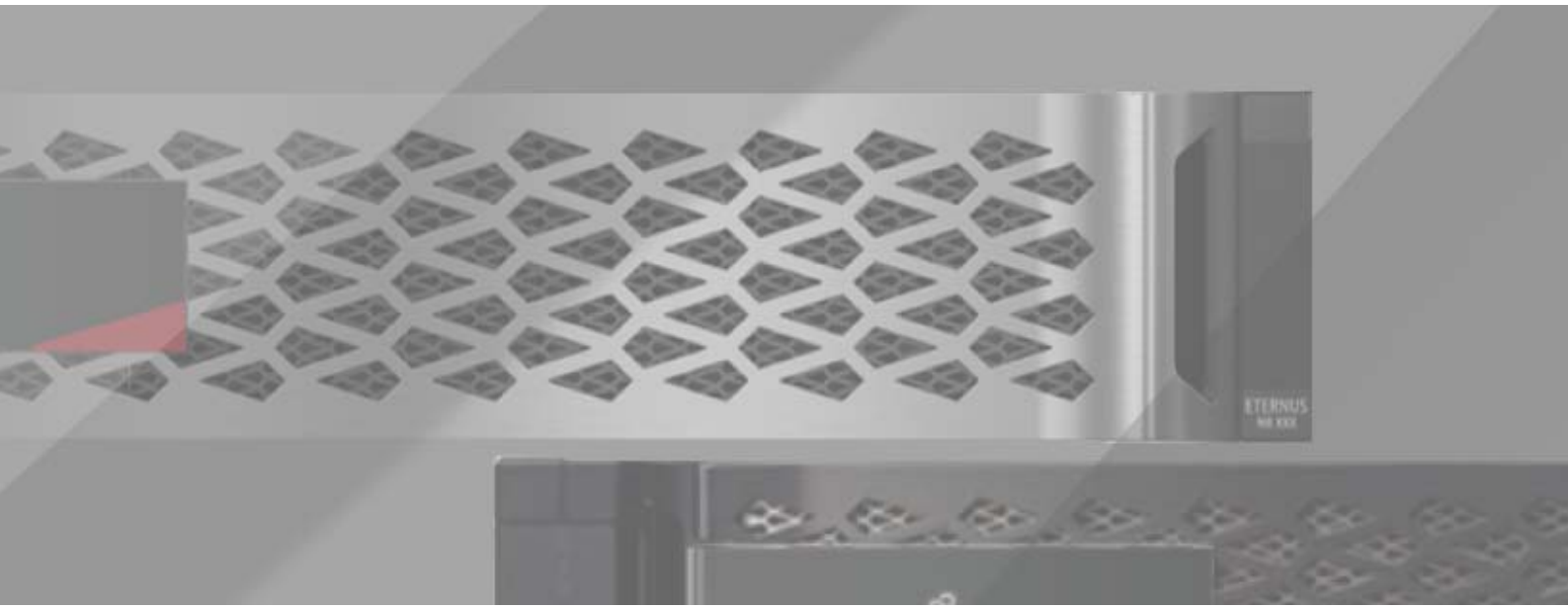
P3AG-5672-01Z0

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

SANtricity OS Dynamic Disk Pool 機能の説明とベストプラクティス



FUJITSU

目次

第 1 章　はじめに 8

1.1　概要 8

1.2　使用目的 8

1.3　単位の規則 9

第 2 章　技術概要 10

2.1　データレイアウト 10

2.2　ドライブ障害時の動作 11

2.3　複数ドライブの障害 13

2.4　ドロワー損失保護 14

2.5　管理 15

2.6　DDP とボリューム・グループの比較 17

2.7　構成のガイドライン 18

第 3 章　パフォーマンス 19

第 4 章　プール構成のベストプラクティス 22

4.1　DDP テクノロジーと従来の RAID の選択 23

4.2　プールのサイズはどのくらいにする必要がありますか？ 23

4.3　分析のベストプラクティス：小ブロック・ランダム・ワークロード 24

4.4　バックアップ / ビデオ監視のベスト・プラクティス：シーケンシャル・ワークロード 24

4.5　テクニカル・コンピューティングのベスト・プラクティス：大きなブロックのシーケンシャル・ワークロード 25

4.6　同容量ボリュームによるプールの構成 25

4.7　再構築の優先順位の設定 27

第 5 章　要約 29

付録 A　用語集 30

付録 B　シンプロビジョニング 32

図目次

図 2.1 D ピースと D ストライプの例 10
図 2.2 24 ドライブプール 11
図 2.3 1 台のドライブに障害が発生した 24 ドライブプール 12
図 2.4 HDD のリビルド時間の例 12
図 2.5 SSD のリビルド時間の例 13
図 2.6 4 U 60 ドライブシェルフ 14
図 2.7 SANtricity System Manager でのプールの作成 15
図 2.8 SANtricity System Manager のプール設定 16
図 2.9 SANtricity System Manager でのボリュームの作成 17
図 3.1 ブロックサイズが 16K/64K の場合の性能（Read/Write） 19
図 3.2 ブロックサイズが 16K の場合のレーテンシー（Read:Write=3:1） 20
図 3.3 ブロックサイズが 16K の場合のレーテンシー（Read:Write=1:3） 20
図 4.1 ボリュームグループとプールの混在 22
図 4.2 DDP テクノロジーまたは RAID の選択 23
図 4.3 SANtricity System Manager の同一ボリューム 26
図 4.4 SANtricity System Manager で同じボリュームを使用して容量をプールする 26
図 4.5 RAID のリビルドと DDP の再構築 27
図 4.6 再構築中の IOPS パフォーマンス低下 28

表目次

表 2.1 DDP とボリュームグループの比較 17
表 2.2 DDP 構成のガイドライン 18

はじめに

ダイナミック・ディスク・プール (DDP) を使用すると、ストレージ管理者は、同じディスクのセットをプール・トポロジーにグループ化できます。このトポロジーでは、プール内のすべてのドライブが I/O ワークフローに参加します。このテクノロジーにより、RAID 5 または RAID 6 よりも高速なドライブ再構築が可能になり、RAID グループ構成の複雑さが解消されるため、ストレージ管理者は容量の割り当てに専念できます。このドキュメントでは、DDP 機能について詳しく説明します。

Copyright 2020 FUJITSU LIMITED

初版
2020 年 11 月

登録商標

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

本書では、本文中の ™、® などの記号は省略しています。

本書の読み方

対象読者

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

本書の表記について

■ 本文中の記号

本文中では、以下の記号を使用しています。

注 意	お使いになるときに注意していただきたいことを記述しています。必ずお読みください。
備 考	本文を補足する内容や、参考情報を記述しています。

第 1 章
はじめに

1.1 概要

Dynamic Disk Pools (DDP) テクノロジーは、ストレージ・システムのデータ保護と管理において大きな進歩を遂げています。データ転送速度が対応して向上しないままディスク容量が増加し続けると、従来の RAID の再構築時間は長くなり、数日にもなります。再構築に時間がかかると、パフォーマンスが低下し、ディスク障害が発生する可能性が高くなります。

DDP テクノロジーは、5 件の特許を取得しており、ドライブ障害からのリカバリを含むあらゆる状況下で予測可能なパフォーマンスを維持しながら、容易な管理と自己最適化によって安心してストレージを提供します。DDP テクノロジーでは、再構築時間が従来の方法より最大 4 倍高速になるため、複数のカスケード・ディスク障害の危険性が大幅に軽減され、優れたデータ保護が提供されます。

次に、これらの利点を実現する主な DDP 属性を示します。

- 管理の合理化 :
  - 分散ホットスペア容量 (保存能力として知られている) により、専用のアイドルホット スペア ドライブが不要になります。
  - RAID を再構成しなくても、プールにドライブを追加できます。
  - 保護スキームとストライプサイズは自動的に設定されます。構成する必要はありません。
- 予測可能なパフォーマンス :
  - 決定論的アルゴリズムにより、データ、スペア容量、保護情報がドライブのプール全体に動的に分散されます。
  - ドライブに障害が発生した場合、セグメントは別の場所に再作成されるため、パフォーマンスの中断の規模と期間が短縮されます。
  - ドライブの大規模なプールにより、ホット・スポットが減少します。
- 複数のディスク障害のリスクを軽減 :
  - セグメント再配置により、システムはより早く最適な状態に戻ります。
  - 優先順位を付けた再構築により、複数のドライブ障害が発生したストライプには、最も高い優先順位が与えられます。

このテクニカルレポートでは、DDP テクノロジーの概要と、プールの使用に関するベストプラクティスのガイドラインについて説明します。

1.2 使用目的

この情報は、お客様およびパートナーを対象としています。

1.3　単位の規則

本書では、2 進法の値を参照するときは IEC 2 進法の単位を使用し、10 進法の値を参照するときは 10 進数の単位を使用します。次に、バイナリ単位の例を示します。

- **KiB.** キビバイト (1,024 バイト)
- **MiB.** メビバイト ($1,024^2$ バイト)
- **GiB.** ギビバイト ($1,024^3$ バイト)
- **TiB.** テビバイト ($1,024^4$ バイト)
- **PiB.** ペビバイト ($1,024^5$ バイト)

10 進数の単位の例を次に示します。

- **KB.** キロバイト (1,000 バイト)
- **MB.** メガバイト ($1,000^2$ バイト)
- **GB.** ギガバイト ($1,000^3$ バイト)
- **TB.** テラバイト ($1,000^4$ バイト)
- **PB.** ペタバイト ($1,000^5$ バイト)

注 意

SANtricity System Manager では、バイナリ値にバイナリラベルを使用します。

第 2 章
技術概要

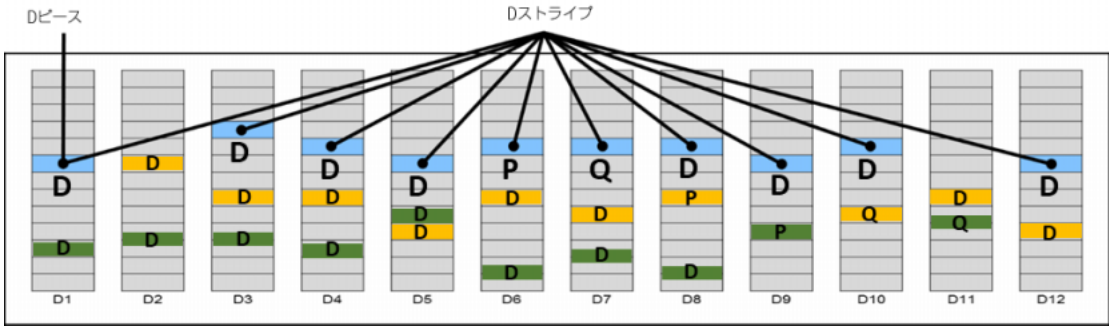
DDP テクノロジーを使用すると、SANtricity OS および管理ソフトウェアで、従来のボリューム・グループ (一般に RAID グループと呼ばれる) に加えてプールを作成できます。プールのサイズは、最小 11 台のドライブから、ストレージ・システム内のすべてのドライブ (ETERNUS HB5200 では最大 480 台のディスク) までさまざまです。プールは、ハードディスクドライブ (HDD) またはソリッドステートドライブ (SSD) で構成されます。さらに、プールとボリュームグループを同じシステムに共存させることもできます。たとえば、すべてのドライブの容量が同じである 24 ドライブのストレージシステムでは、次のような組み合わせが考えられます。その他、様々な組み合わせで構成できます。

- 8 ドライブ構成の RAID 10 (4+4 個) ボリューム・グループ ×1、ドライブ 16 台のプール ×1
- 24 ドライブ・プール ×1
- 12 ドライブ・プール ×2
- 5 ドライブ構成の RAID 5 (4+1 個) ボリューム・グループ ×1、4 ドライブ構成の RAID 10 (2+2 個) ボリューム・グループ ×1、15 ドライブ・プール ×1

2.1 データレイアウト

プール内では、プールに割り当てられているドライブの数に関係なく、ボリュームデータはすべてのドライブに分散されます。ボリュームは、D ストライプと呼ばれる多数の仮想ストライプで構成されます。各 D ストライプは 10 台のドライブに配置され、インテリジェントな最適化アルゴリズムによってプール全体に分散されます。1 つのドライブに存在する各 D ストライプの部分を D ピースと呼びます。各 D ピースは、物理ドライブの連続したセクションです。図 2.1 に、D ストライプのレイアウト例を示します。この場合、プールは 12 台のドライブで構成されますが、それ以上のドライブがあったとしても、D ストライプは 10 個の D ピースにしか分割されません。
D ピースは、必ずしも各ドライブの同じ部分に存在するとは限りません。D ストライプと D ピース、およびそれらの容量の詳細については、「付録 A 用語集」を参照してください。

図 2.1　D ピースと D ストライプの例



各 D ストライプは、ETERNUS AB6100 上の 8,192 個の RAID 6 ストライプと、他のすべてのアレイ上の 4,096 個の RAID 6 ストライプで構成されます。これらのストライプはそれぞれ 10 個の 128 KiB セグメントで構成されており、ETERNUS AB6100 では合計 1 GiB、その他の機種では合計 512 MiB になります。図に示すように、8 つのセグメントはデータ (D) 、1 つはパリティ (P) 、1 つは RAID 6 Q 値です。

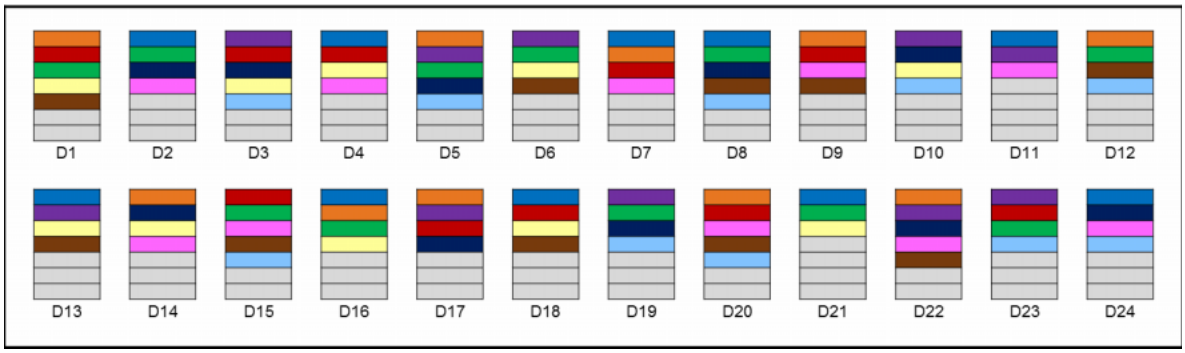
プールの作成後、プール内にボリュームを作成できます。このボリュームは、プール内のすべてのドライブにわたるいくつかの D ストライプで構成され、データ D ストライプの数は、定義されたボリューム容量を D ストライプサイズで割った値に等しくなります。たとえば、2 TiB ボリュームは、ETERNUS AB6100 では 256 個のデータ D ストライプで構成され、他のすべての機種では 512 個のデータ D ストライプで構成されます。特定のボリュームに対する D ストライプの割り当ては、特定のドライブ上の特定の D ピースに対する論理ブロックアドレス (LBA) の使用可能な最小範囲から開始して実行されます。

2.2　ドライブ障害時の動作

DDP テクノロジーの主なメリットは、専用の孤立したホット・スペアを使用するのではなく、プール自体に統合された保存機能があり、潜在的なドライブ障害に備えて再構築場所を提供できることです。この機能により、個別のホット・スペアを計画または管理する必要がなくなるため、管理が合理化されます。また、再構築の時間が大幅に短縮され、再構築中のボリューム自体のパフォーマンスも向上します。

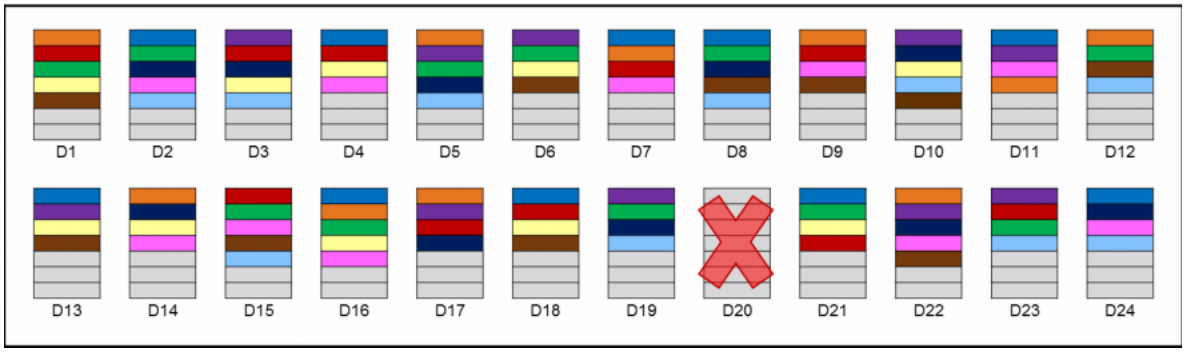
ドライブに障害が発生したときの DDP オペレーションの説明を開始するには、図 2.2 に示す 24 ドライブ・プールを検討します。図中の各色は D ストライプを表し、各ストライプには同じ色の D ピースが 10 個含まれています。前述したように、D ピースは、DDP インテリジェントアルゴリズムによってプール上に分散されています。

図 2.2　24 ドライブプール



ここで、図 2.3 のように、プール内のドライブの 1 つに障害が発生したとします。

図 2.3　1 台のドライブに障害が発生した 24 ドライブプール



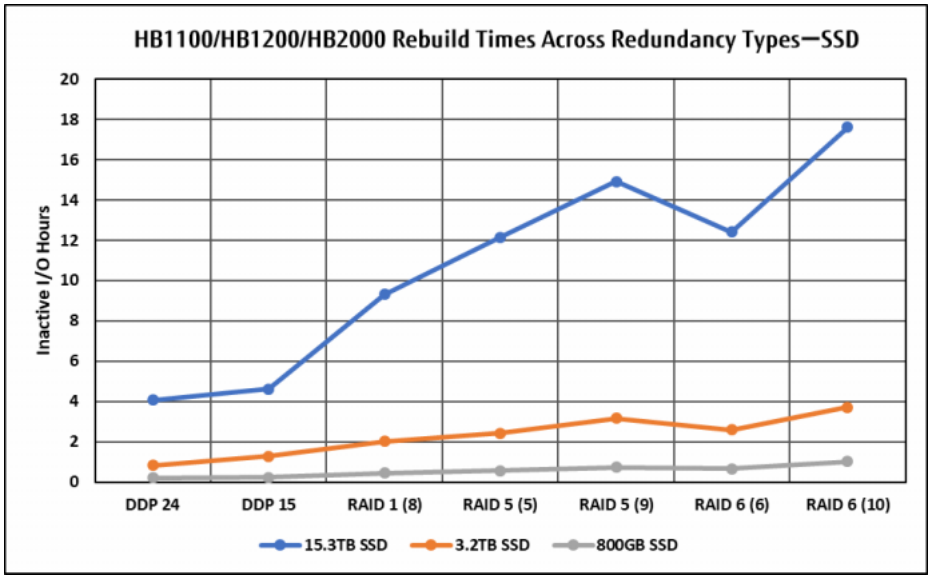
プール内のドライブに障害が発生すると、RAID 6 で通常使用されるのと同じメカニズムを使用して、障害が発生したドライブの D ピースがセグメントごとに再構築されます。インテリジェント・アルゴリズムは、再構築された D ピースを書き込むプール内の他のドライブを選択し、1 つのドライブに同じドライブの 2 つの D ピースが含まれていないことを確認します。個々の D ピースは、選択されたドライブで使用可能な LBA 範囲の最低値で再構築されます。

図 2.3 では、ドライブ D 20 に障害が発生し、D ピースが再構築されて他のドライブに書き込まれています。リビルド処理は、すべてのドライブで並行して実行されます。この作業には複数のドライブが稼働するため、この状況による全体的なパフォーマンスへの影響が軽減され、操作の完了に必要な時間も大幅に短縮されます。

図 2.4 と図 2.5 に、RAID 6 テクノロジーと DDP テクノロジーのリビルド時間の違いを示します。図 2.4 は、さまざまな数の HDD を搭載した ETERNUS HB1100/HB1200/HB2000 のデータを示していますが、図 2.5 は SSD の同様のデータを示しています。どちらも、システムの残りの部分をアイドル I/O 状態にして実行されました。これらのグラフは、プールの再構築が RAID 6 ボリューム・グループよりも高速であり、プールのスピンドル数が増加すると、DDP の再構築時間が RAID 5 および RAID 6 の再構築時間に比べ短くなることを示しています。

図 2.4　HDD のリビルド時間の例

図 2.5　SSD のリビルド時間の例



2.3　複数ドライブの障害

データ可用性のリスクを最小限に抑えるために、プール内の複数のドライブに障害が発生した場合は、2 つの D ピースが欠落している D ストライプが優先的に再構築されます。この方法はクリティカル再構成と呼ばれます。重大な影響を受ける D ストライプが再構築された後、残りの必要なデータが再構築されます。

コントローラのリソース割り当ての観点から、プール内にはユーザーが変更可能な再構築の優先順位が 2 つあります。

- 影響を受ける D ストライプに対して 1 つの D ピースのみを再構築する必要がある場合は、低い再構築優先度が割り当てられます。このインスタンスのデフォルトの優先順位は「高い」です。
- D ストライプに再構築が必要な 2 つの欠落した D ピースがある場合、クリティカル再構築優先度が割り当てられます。このインスタンスのデフォルトの優先順位は「最優先」です。

2 つのディスク障害が同時に発生する非常に大規模なプールでは、比較的少数の D ストライプだけが、2 つの D ピースを再構築する必要があるという重大な状況に遭遇する可能性があります。前述したように、これらの重要な D ピースは、識別され、最初に最優先で再構築されます。この方法では、プールは非常に短時間で縮退状態に戻り、それ以上のドライブ障害に対応できます。

たとえば、192 台のドライブのプールが作成され、2 台のドライブに障害が発生したとします。この場合、重要な D ピースは 1 分未満で再構築され、その後追加のドライブ障害に対応できる可能性があります。数学的な観点から見ると、同じ 192 ドライブのプールでは、プール内の 1 台のドライブに 1 つの D ピースを持つ D ストライプは 5.2% に過ぎず、同じドライブに 2 つの D ピースを持つ D ストライプは 0.25% に過ぎません。したがって、クリティカル・ステージを終了するには、48 GiB のデータのみを再構築する必要があります。非常に大規模なプールでは、再構築を継続するための追加の保存容量がなくなるまで、データを失うことなく複数のシーケンシャルな障害を維持できます。

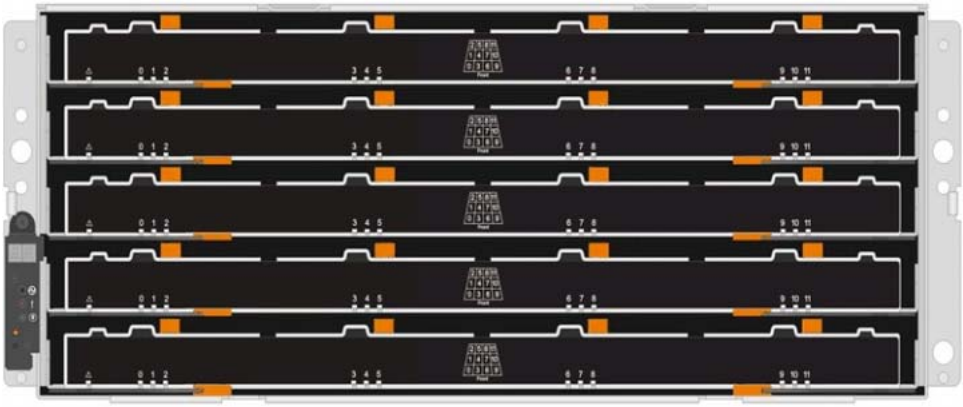
再構築後、障害が発生したドライブを交換できますが、特に交換が必要なわけではありません。基本的に、障害が発生したドライブの交換は、プールのオンライン容量拡張とほぼ同じ方法で処理されま

す。障害が発生したドライブは、プールがクリティカルまたは劣化状態から抜ける前に交換することもできます。

2.4　ドロワー損失保護

4 U 60 ドライブ・シェルフは、4 U のラック・スペースに最大 60 台のドライブを収容できます。ドライブは 5 つの引き出しに整理され、各引き出しには最大 12 台のドライブを収納できます。図 2.6 を参照してください。

図 2.6　4 U 60 ドライブシェルフ



SANtricity OS 11.25 では、1 つの 4 U 60 シェルフでドロワー損失保護 (DLP) を実現できます。ドロワー損失保護とは、プールがドロワー全体の損失に耐え、可用性を損なうことなく I/O 操作を維持できることを指します。

1 つのシェルフで DLP を有効にするには、構成に少なくとも 15 台のドライブが必要です。ドライブはドロワー間で均等に分散されていなければなりません。SANtricity System Manager は、プール作成時に DLP 候補を提示します。これらの候補のドライブ数は、常に 5 の倍数です。プール容量を増やすためにドライブを追加する場合は、新しい容量で DLP を維持するために、ドロワーごとに 1 つずつ、5 つのグループに追加する必要があります。

DLP は、複数の 4 U 60 シェルフがあるシステムでも有効にできます。この場合、同様の特性 (ドライブ・タイプ、容量、データ保証、セキュリティ) を持つドライブは、適切な候補が示されるように、すべてのドロワーに均等に分配する必要があります。管理インターフェイスを介して提示されるすべての DLP 候補は、プール内のすべてのドロワーのドロワーごとに同数のドライブで構成されます。

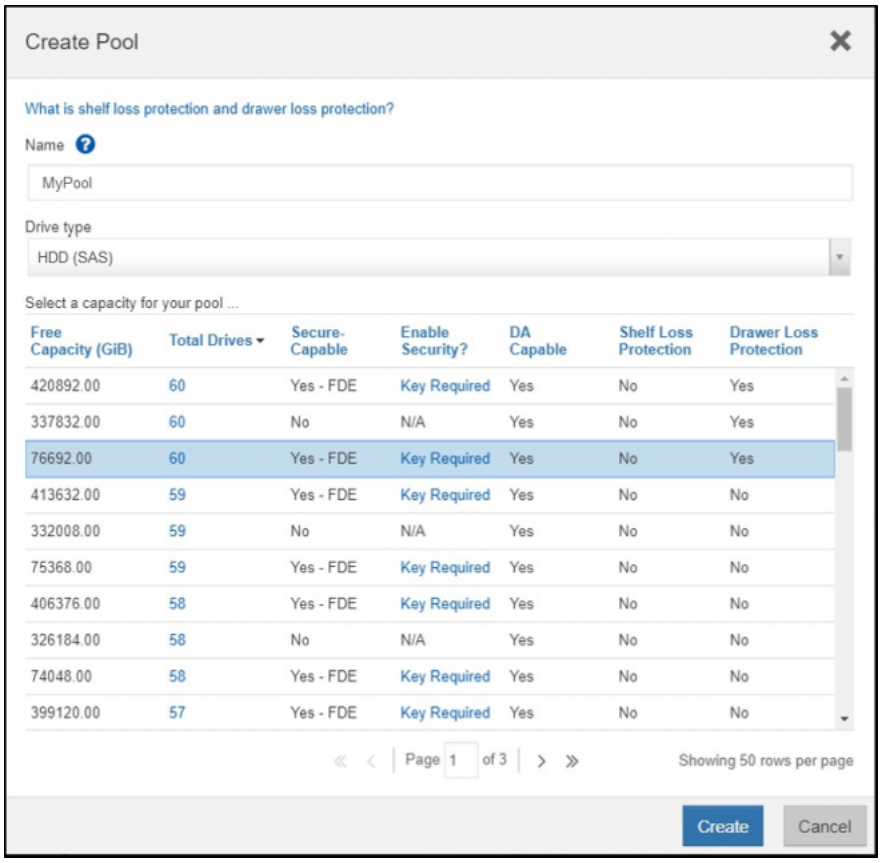
複雑な構成では、最初に小さい DLP プールを作成する必要がある場合があります。この状況は、一部のドロワーには他の種類のドライブが含まれているため、同様のドライブがドロワー間で均等に分散されていない場合に発生します。この場合、DLP を維持するために一度に 5 台のドライブを追加することで、プールを拡張できます。

2.5　管理

プールの構成は、従来のボリュームグループの構成よりも簡単です。RAID レベル、セグメントサイズ、またはグローバルホットスペアは、システムのデフォルトによって決定されるため、選択する必要はありません。管理者は、ドライブのタイプ、ドライブの数、プールにセキュリティと DLP のどちらを適用するかを決定します。

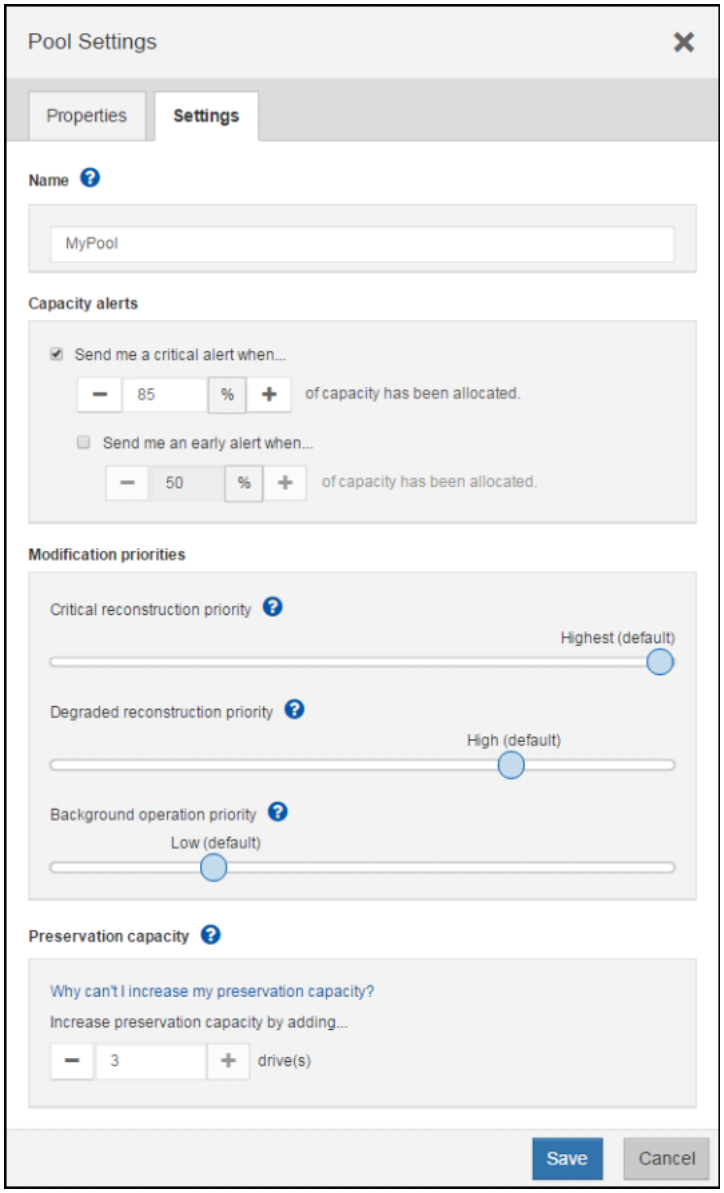
図 2.7 は、SANtricity System Manager でのプール作成を示しています。この場合、DLP とフルディスク暗号化 (FDE) の両方のセキュリティを提供する 60 台の HDD を持つ候補が管理者によって選択されています。

図 2.7　SANtricity System Manager でのプールの作成



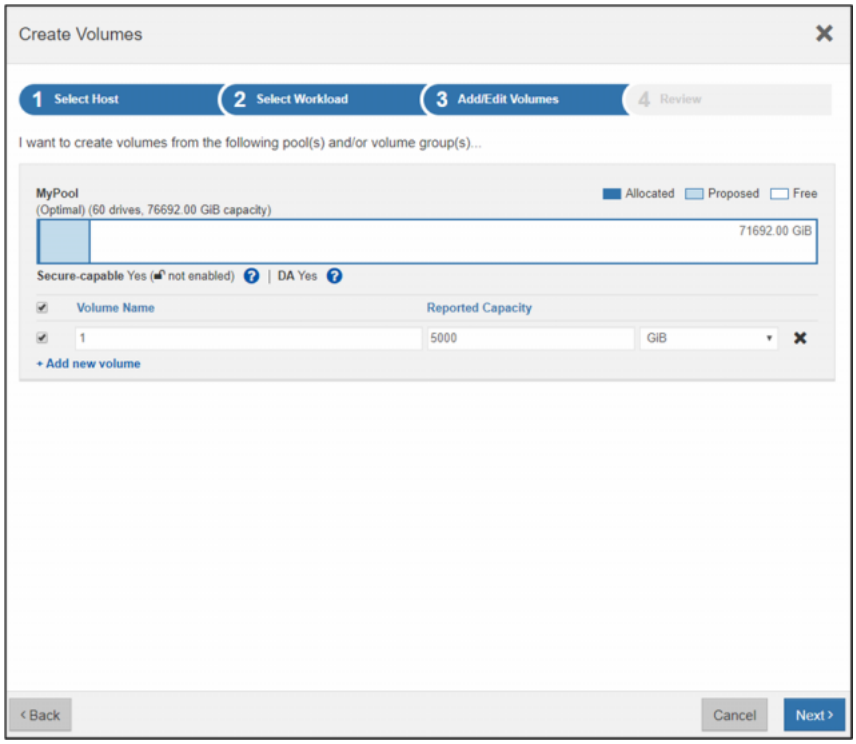
プールの設定は変更できますが、通常は変更する必要はありません。図 2.8 は、SANtricity System Manager のさまざまなプール設定を示しています。再構築およびその他のバックグラウンド操作のデフォルト設定に注意してください。この場合、管理者は保存容量に相当するドライブ数も変更できます。この場合、デフォルトは 3 ですが、ユーザーは 0 から最大 10 の任意のドライブ数、またはドライブの 20% のいずれか小さい方を選択できます。

図 2.8　SANtricity System Manager のプール設定



プールを作成したら、従来のボリューム・グループとほぼ同じ方法でプール内にボリュームを作成できます。

図 2.9　SANtricity System Manager でのボリュームの作成



ボリュームグループと同様に、プールおよびプールボリュームも CLI または REST API を使用して構成できます。

2.6　DDP とボリューム・グループの比較

プールとその中のボリュームでは、従来のボリューム・グループのオペレーションと同様の複数のオペレーションを実行でき、DDP テクノロジーに固有の機能を提供します。

表 2.1 に、DDP とボリューム・グループの機能の簡単な比較を示します。

表 2.1　DDP とボリュームグループの比較

機能	DDPおよびプール・ボリューム	ボリューム・グループとボリューム・グループ・ボリューム
Snapshot technology	可	可
ボリュームコピー	可	可
同期ミラーリング	ETERNUS AB6100 および ETERNUS HB1x00 を除いて可	ETERNUS AB6100 を除いて可
非同期ミラーリング	ETERNUS AB6100 を除いて可	ETERNUS AB6100 を除いて可
動的なボリューム拡張	可	可
オンライン容量の拡張と縮小	一度に最大 60 台のドライブを追加または削除が可	部分的に可 : 最大 2 台のドライブを追加し、容量を削減しない
動的なRAID移行	不可 : 常に RAID 6	可
動的セグメントサイズ	不可 : セグメントは常に 128 KiB	可

機能	DDPおよびプール・ボリューム	ボリューム・グループとボリューム・グループ・ボリューム
ホット・スペア	分散保存容量	専用のグローバル・ホット・スペア
動的再配分	可 : 無停止でのバックグラウンド・オペレーション	不可 : 断片化と、欠失による孤立した容量
ドライブ排気装置	可	可
シェルフ及びドロワーの損失防止	可	可
SSDのサポート	可	可

2.7　構成のガイドライン

表 2.2 に、プールを構成する際の重要な考慮事項を示します。

表 2.2　DDP 構成のガイドライン

説明	構成
システムあたりの最大プール数	20
プールあたりの最小ドライブ数	11
1 つの 4 U 60 シェルフでのドロワー損失保護 (DLP) のためのプールあたりの最小ドライブ数 (*1)	15
プールの最大容量 (*2) (システム内のすべてのプールの容量の合計)	AB6100-12PiB
最大ボリュームサイズ	すべてのシステム :4 PiB
プール・サイズ別のデフォルト保存容量 (同等ドライブ数)	11 ドライブ : 1 12-31 台のドライブ : 2 32-63 台のドライブ : 3 64-127 台のドライブ : 4 128-191 台のドライブ : 6 192-255 台のドライブ : 7 256-384 台のドライブ : 8 385-480 台のドライブ : 10
サポートされるドライブタイプ 注 意 プール内のすべてのドライブは、同じタイプで、同じ特性 (データの保証 , セキュリティ) を持つ必要があります。 大容量ドライブの容量が失われないように、すべてのドライブの容量を同じにする必要があります。	SAS、NL-SAS、SSD、NVMe
オンラインでのプールへの追加またはプールからの削除	一度に最大 60 台のドライブ

*1:　ディスク・プールに容量を追加したときに DLP を維持するには、ドライブを 5 台ずつ (4 U シェルフのドロワーごとに 1 台ずつ) 追加する必要があります。

*2:　最大プール容量には、RAID 保護オーバーヘッド、プール保存容量、有効容量、プールのサイズに基づく小規模な DDP 固有のリザーブが含まれます。

第 3 章
パフォーマンス

ただし、RAID 5 および RAID 6 と比較した場合の差分は、読み取りワークロードが 100% の場合は無視できる程度であり、書き込みワークロードが 100% の DDP を実行した場合はパフォーマンスがわずかに低下する程度です。図 3.1 は、ワークロードが 100% の書き込みから 100% の読み取りに変更された場合の、16 KB および 64 KB の I/O を実行する ETERNUS AB5100 を使用した一般的なパフォーマンス比較を示しています。

図 3.1　ブロックサイズが 16K/64K の場合の性能（Read/Write）

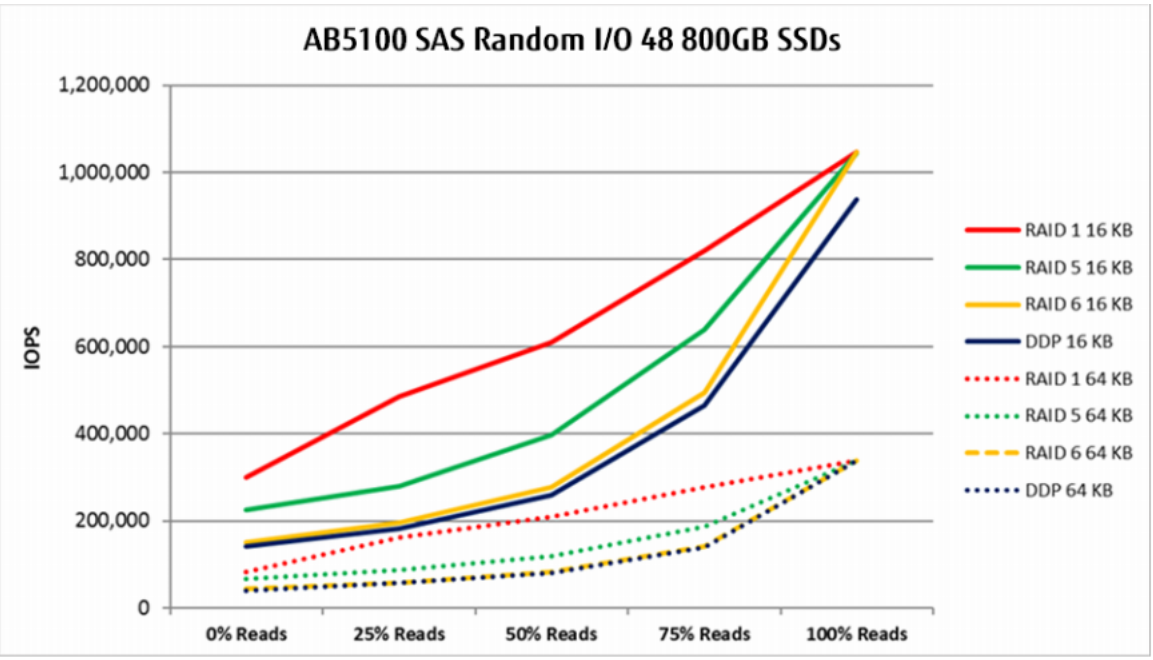
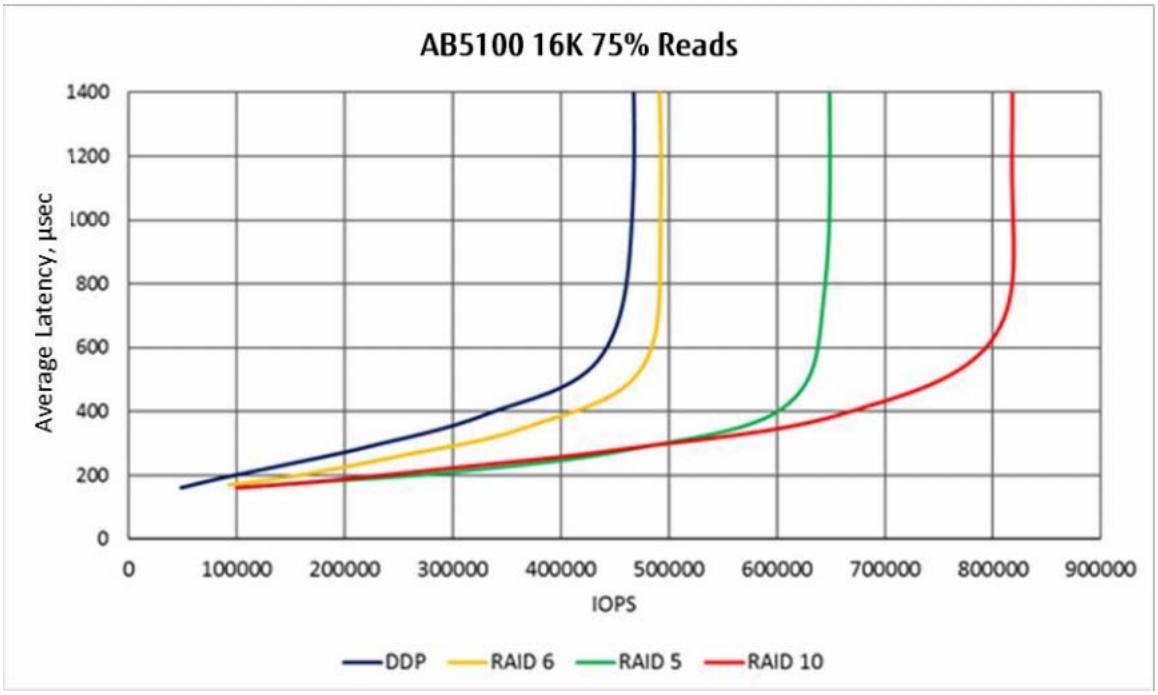


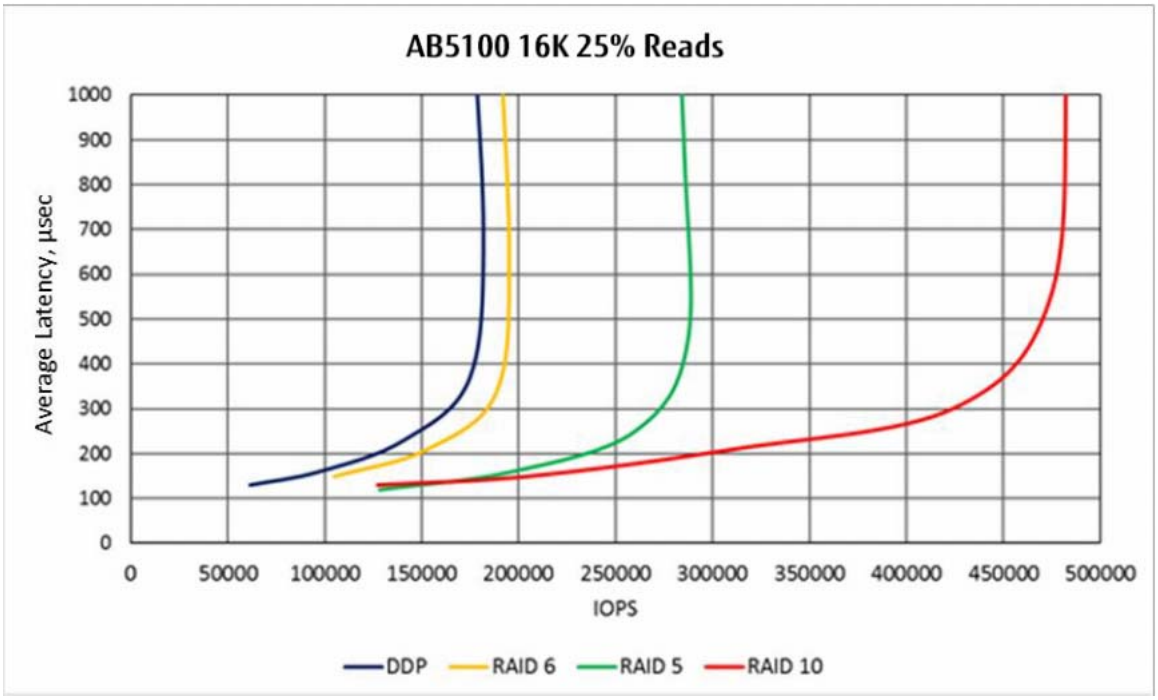
図 3.2 は、16 KB の I/O サイズと 75% の読み取りワークロードを使用した場合の、DDP、RAID 5、RAID 6、および RAID 10 のレーテンシーの比較を示しています。

図 3.2　ブロックサイズが 16K の場合のレーテンシー（Read:Write=3:1）



ワークロードの書き込みコンポーネントが増加すると、IOPS は予想どおり低下しますが、図 3.3 に示すように、レーテンシーは低いままです。

図 3.3　ブロックサイズが 16K の場合のレーテンシー（Read:Write=1:3）



これらの結果は、DDP テクノロジーを使用する場合、特にパフォーマンス範囲の下限での IOPS と遅延のパフォーマンスは、他の標準 RAID の選択肢に非常に近くなります。パフォーマンス範囲の上限では、標準 RAID の方がパフォーマンスは優れています。

ほとんどの場合、パフォーマンスは、30 台以上のドライブのプールによって提供されるドライブの再構築時間の大幅な改善と、パフォーマンスの上限とのわずかなトレードオフになります。

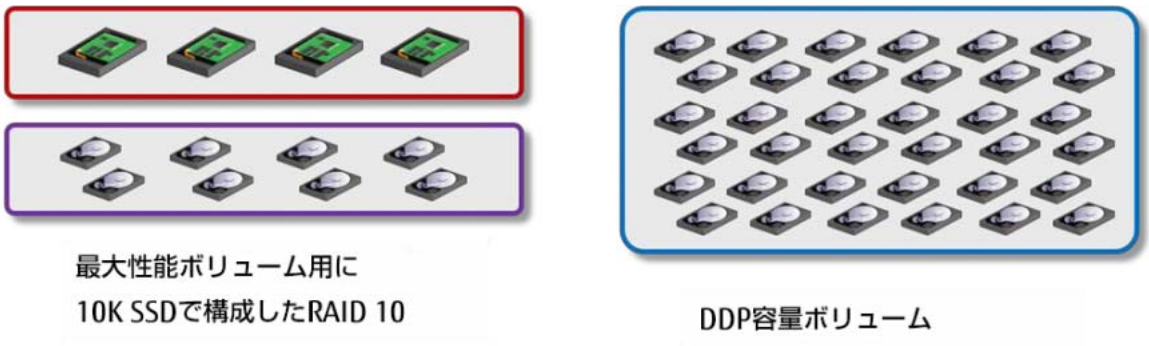
第 4 章
プール構成のベストプラクティス

一般に、DDP は同種の環境で使用するか、1 つのストレージ・システムで複数のアプリケーションを提供する場合に使用することを推奨します。DDP テクノロジーは、ランダムなワークロード環境で最高のパフォーマンスを発揮するように設計されており、単一のストレージシステム上の複数のアプリケーションが、ストレージシステムに対してランダムなプロファイルを生成します。

ストリーミング・メディアなど、帯域幅（シーケンシャル性能）を必要とするアプリケーションがストレージシステムを使用する環境でのベストプラクティスは、アプリケーションごとに従来のボリューム・グループを 1 つ構成することです。

ハイパフォーマンスと大容量の両方のワークロードに使用される SSD、ハイパフォーマンス HDD、および NL-SAS ドライブが混在するストレージシステムでは、ボリュームグループとプールの混在を構成できます。図 4.1 に例を示します。

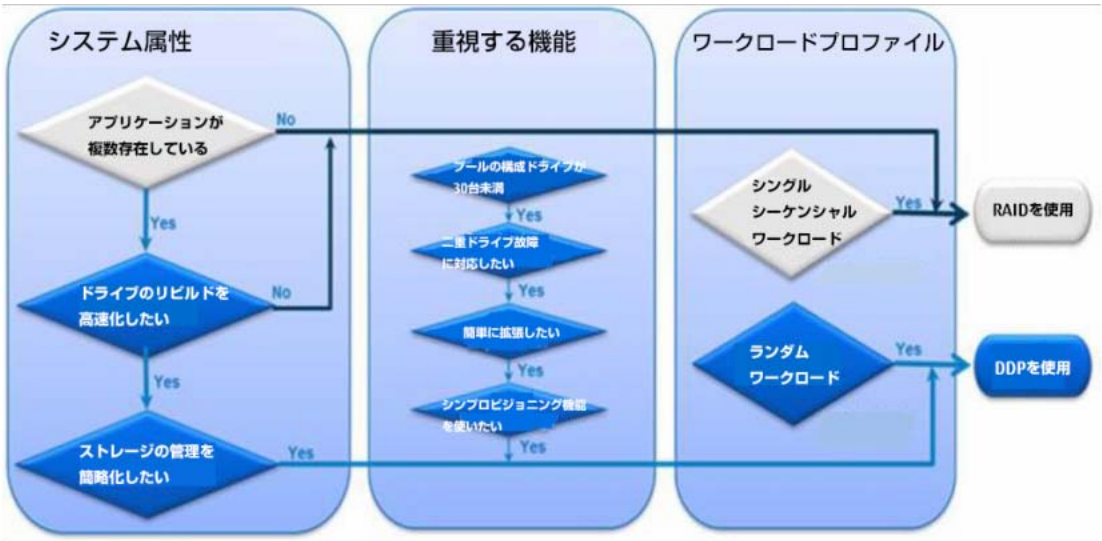
図 4.1　ボリュームグループとプールの混在



4.1　DDP テクノロジーと従来の RAID の選択

DDP テクノロジーには、ストレージ仮想化テクノロジーを選択する際に考慮すべき重要な機能がいくつかあります。図 4.2 では、DDP 機能と従来の RAID 機能を比較しています。

図 4.2　DDP テクノロジーまたは RAID の選択



注 意

二重ドライブ障害が発生した場合、DDP テクノロジーはまず重要なセグメントを再構築し、最初の二重ドライブ障害から数分以内に別のドライブ障害に対応できるようにします。

4.2　プールのサイズはどのくらいにする必要がありますか？

アーキテクチャの観点からは、すべてのドライブが同じタイプであると仮定すると、プールをシステム内のディスク総数まで拡張できます。拡張は、特定のモデルとソフトウェア・バージョンの最大プール容量 (表 2.2 に示す) によってのみ制限されます。ただし、マルチワークロード環境でのベスト・プラクティスは、ミッドレンジ・システムのパフォーマンス要件には最大 90 台のドライブ、ローレンジ・システムのパフォーマンス要件には最大 120 台のドライブからなる複数のプールを作成することです。

ETERNUS AB/HB series の RAID の選択肢はすべて、使用するドライブ・タイプによってパフォーマンスが異なります。たとえば、SSD は NL-SAS ドライブよりも高速です。このセクションでは、IOPS、スループット、およびドライブ再構築のパフォーマンス目標を達成するために ETERNUS AB/HB series ストレージシステムをプロビジョニングするための、管理者または概念検証を計画しているすべての人を支援するベストプラクティスのガイドラインについて説明します。より大きなプールを作成することもできますが、NL-SAS HDD の場合、1 プールあたり 60 ～ 120 台のドライブを使用することで、パフォーマンスと信頼性の最適なバランスを維持できます。

長期的なアーカイブの使用例には、1 つの大きなプールの方が適している場合があります。このアプローチは、管理オーバーヘッドを低く抑えながら、データを非常に信頼性の高いストレージ・プラットフォームに維持します。

4.3　分析のベストプラクティス：小ブロック・ランダム・ワークロード

従来のリレーショナル・データベースと NoSQL データベースはどちらも、ほとんどランダムで小さなブロックの読み取り / 書き込みワークロードを生成します。OLTP およびハイパーバイザワークロードの場合も同様です。DDP テクノロジーは、このようなタイプのランダムなワークロードで非常に優れたパフォーマンスを発揮するように調整されています。このような環境では、すべてのボリュームを 1 つのプールに配置することをお勧めします。このアプローチは、システム管理とデータベース管理を簡素化し、データベースログとデータストレージの両方のパフォーマンスニーズを満たします。ディスクの容量が大きく、フロアスペースが限られているため、ワークロードをセグメント化し、データベースログファイルを単一の RAID1 ボリュームに分離することはもはや効率的ではありません。このようなセグメント化は孤立したストレージを作成し、管理オーバーヘッドを増加させます。

もう 1 つのベスト・プラクティスは、ボリュームの数だけでなく、ボリュームの機能についても、2 つの RAID コントローラ間のバランスを確保することです。SANtricity 管理ソフトウェアを使用してボリュームを作成すると、コントローラ間でボリュームのバランスが自動的に調整されます。また、自動負荷分散機能では、コントローラのワークロードに基づいてボリュームの所有権を動的に調整することもできます。自動ロード・バランシング機能を使用するには、Windows、VMware、または Linux (カーネル 3.10 以上) で適切なホスト・タイプを選択する必要があります。

4.4　バックアップ / ビデオ監視のベスト・プラクティス：シーケンシャル・ワークロード

一度に複数のカメラで連続的に録画するビデオ監視ワークロードについて考えます。複数のストリームがボリュームの異なる部分に書き込まれるため、ワークロードはアレイに対してランダムに見えます。このシナリオでは、Analytics Best Practices: Small-Block Random Workloads で示されるように、DDP は優れたパフォーマンスを発揮します。

パフォーマンスが重視されるシーケンシャルなワークロードで、一般的にブロック転送サイズが大きい場合は、プールを構成するのではなく、これらのワークロードを分離して独自の従来のRAIDボリュームグループに配置することをお勧めします。このタイプのワークロードには、多くのバックアップおよび監視アプリケーションがあります。ただし、DDP テクノロジーの再構築のメリットを考慮して、ボリューム・グループが要件であるかどうかを判断するには、パフォーマンス・サイジング・ツールを使用する必要があります。一部のバックアップ実装では、DDP 構成がパフォーマンス要件を満たしながら、DDP テクノロジーの再構築のメリットも提供することが示されています。

パフォーマンスを重視する場合は、1 つのボリュームを持つ従来のボリュームグループを使用することをお勧めします。再構築時間の短縮、デグレードモード時のパフォーマンス改善、管理の容易さが目標である場合は、12 ～ 30 台のドライブのプールを使用することで要件を満たすことができます。

4.5 テクニカル・コンピューティングのベスト・プラクティス：大きなブロックのシーケンシャル・ワークロード

Lustre や IBM Spectrum Scale などの技術的なコンピューティング・ワークロードでは、大容量のブロックのシーケンシャルな読み取り / 書き込みが発生し、基盤となるメディアからの要求が最も高くなります。このタイプの環境のベスト・プラクティスは、ボリューム・グループごとに 1 つのボリュームを持つ従来の RAID ボリューム・グループを使用することです。

4.6 同容量ボリュームによるプールの構成

SANtricity System Manager および CLI では、管理者は作成するボリュームごとに必要な容量を入力する必要があるため、ユーザーはシステムを構成する前にこれらの容量を計算する必要があります。

この計算を開始するには、プールの有効容量が常に D ストライプサイズの倍数であることを知っておくと便利です。この計算では、ボリュームあたりの D ストライプ数を決定し、D ストライプ・サイズを掛けてボリュームあたりの GiB 数を求める必要があります。たとえば、プールの有効容量が 161,456 GiB で、ユーザーが 5 つの等しいボリュームに分割するとします。計算は次のようになります。

1. プール内の D ストライプの数を確認します。合計容量 (GiB)/D ストライプサイズ =161,456/4=40,364
2. ボリュームあたりの D ストライプ数を決定します。40,364/5=8,072.8
3. ボリュームあたりの D ストライプ数を 8,072 に丸めます。
4. D ストライプの数に 4 を掛けて、ボリュームあたりの GiB を求めます。8,072×4=32288 GiB

注 意

この例では、ステップ 2 の余りのために 4 つの D ストライプが残されます。D ストライプの総数が必要なボリューム数で割り切れない場合は、すべてのプール容量を使用する同じボリュームを作成することはできません。

図 4.3 は、SANtricity System Manager のこの例を示しており、プール内の 5 つの等しいボリュームを示しています。図 4.4 は、プール容量と割り当て済み容量を示しています。この例で説明した 16.00 GiB の空き容量に注意してください。また、この単純な計算を正しく実行するには、プリファレンスを設定して GiB で容量値を表示する必要があります。

図 4.3　SANtricity System Manager の同一ボリューム

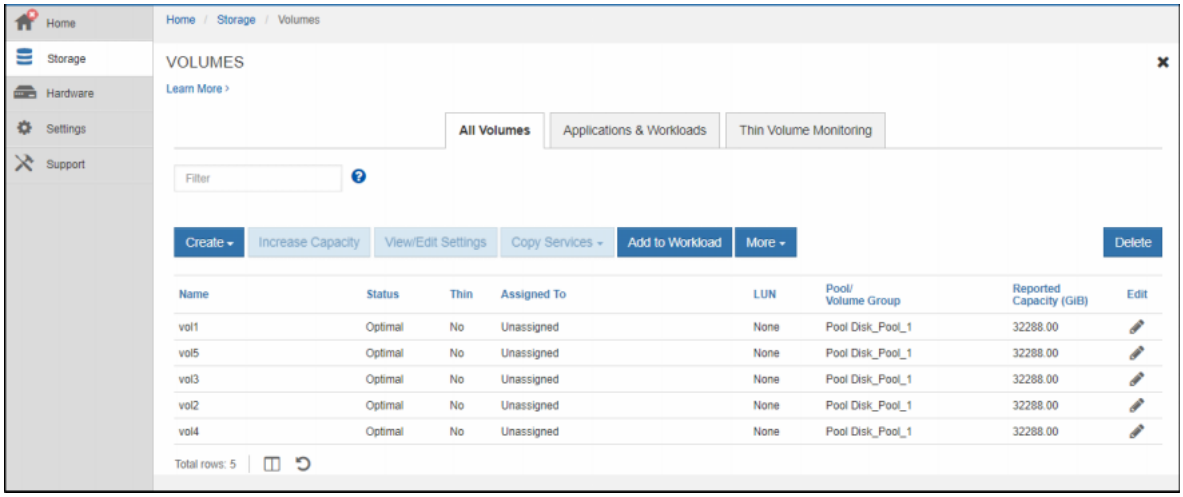
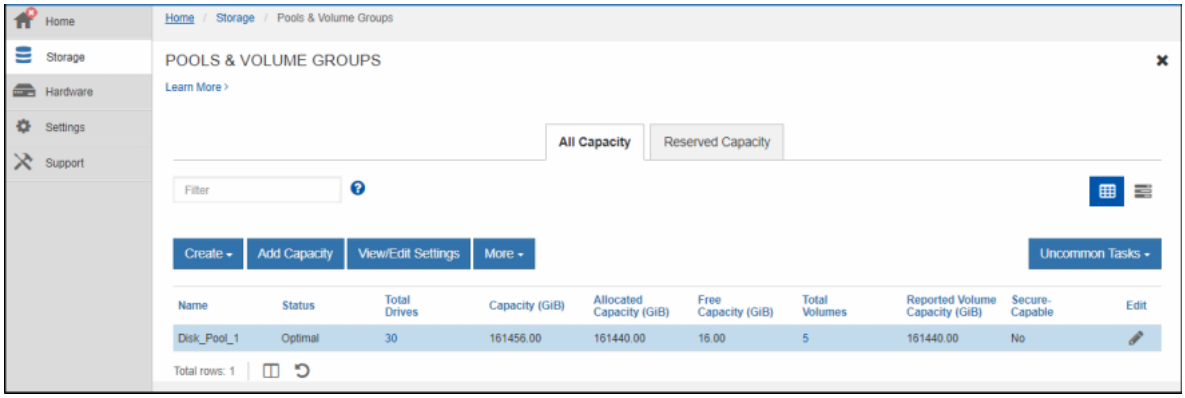


図 4.4　SANtricity System Manager で同じボリュームを使用して容量をプールする



CLI を使用した等量ボリュームの作成は、SANtricity System Manager を使用した場合と同様です。この例を説明するために、有効容量 229.996 TB、9 つの等しいボリュームを持つディスク・プールの例を考えてみましょう。次のコマンドは、プール属性を表示します。

```
show diskPool ["Disk_Pool_1"];
```

結果の容量を示す部分は次のとおりです。

```
Total capacity:                                    234.346 TB
         Preservation capacity:                          4,455.000 GB(3 Drives)
         Usable capacity:                                229.996 TB
```

9 つの等しいボリュームを作成するには、異なるボリューム名で次のコマンドを 9 回入力します。

```
create volume diskPool="Disk_Pool_1"
userLabel="vol1"
capacity=26168GB;
```

結果を確認するには、プール属性を再度表示します。

```
show diskPool ["Disk_Pool_1"];
```

容量に対して返される値は次のとおりです。

```
Total capacity:                                    234.346 TB
         Preservation capacity:                          4,455.000 GB(3 Drives)
         Usable capacity:                                229.996 TB
         Unusable capacity:                              0.000 MB

      Used capacity:                                     229.992 TB
         Volumes:                                        (9), 229.992 TB
         Repositories:                                   (0), 0.000 MB
      Free Capacity:                                     (1), 4.000 GB
      Percent full:                                      99%
```

4.7　再構築の優先順位の設定

再構築の優先度の設定を変更して、再構築に要する時間を最短にする (最優先設定)、またはドライブの再構築中にストレージ・システムのパフォーマンスが受ける影響を最小限に抑える (最低優先度設定) ことができます。既定の設定とベストプラクティスは高い優先度で、迅速な再構築時間と許容可能なシステムパフォーマンスの維持のバランスをとります。

縮退プールの再構築は、図 4.5 に示すように、従来のボリューム・グループの再構築よりも高速です。また、この縮退状態の間、ホスト・サーバへの I/O パフォーマンスは、従来のボリューム・グループと比べてほとんど影響を受けません。

図 4.5　RAID のリビルドと DDP の再構築

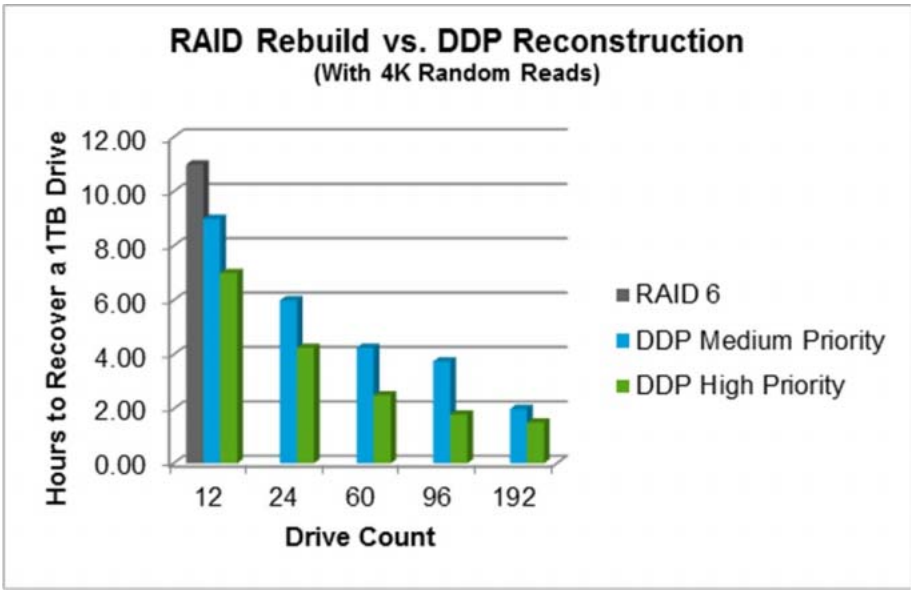
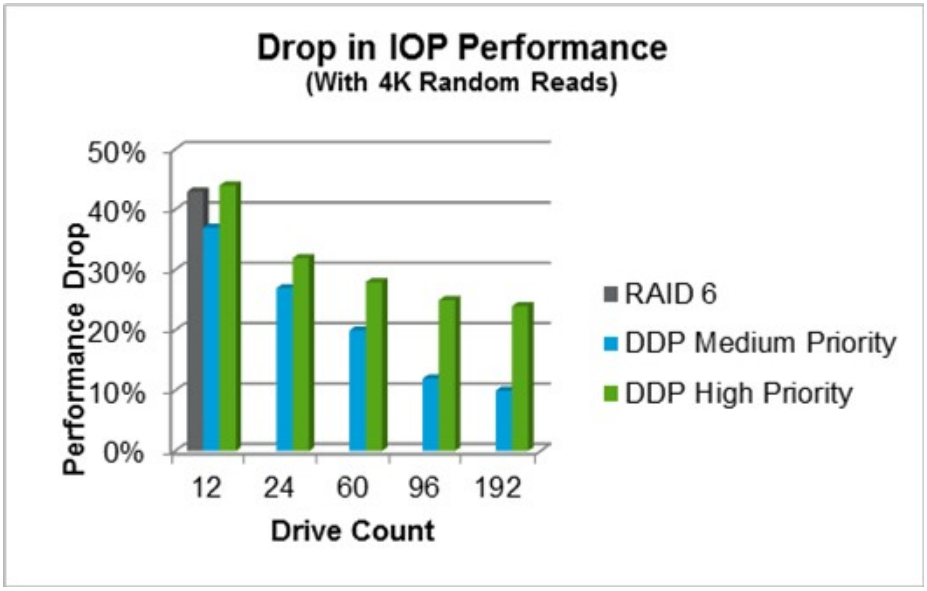


図 4.6 は、さまざまなドライブ障害シナリオにおけるパフォーマンスの例を示しています。このグラフは、ディスク・プール・サイズが大きいほど、DDP テクノロジーによるパフォーマンスへの影響が小さいことを示しています。再構築と、縮退状態またはクリティカル状態からの移行は、従来のボリューム・グループよりも高速に行われます。

図 4.6　再構築中の IOPS パフォーマンス低下



このパフォーマンスの差は、DDP テクノロジーがドライブ全体の再構築からドライブの損失保護を切り離すために発生します。DDP 機能は、残りのドライブ間でデータを再配分し、パリティの計算に必要なアレイ内のドライブの一部のみに影響を与えます。従来の RAID では、ボリューム内のすべてのドライブが影響を受け、全体的なパフォーマンスが低下します。予想されるパフォーマンスの低下と、必要な再構築操作を完了するまでの時間は、ワークロード・パターンや再構築の優先順位設定などの多くの要因によって異なります。

第 5 章
要約

ドライブ容量が増加していくにしたがってディスク故障時のリビルド時間も延長していくため、ストレージ・システムは常にデータ消失のリスクにさらされています。ドライブの再構築が長時間に及ぶと、システムのパフォーマンスに悪影響が及ぶため、管理者が SLA を満たすことが困難になり、ビジネス･アプリケーションのレスポンス･ニーズに影響が及びます。ETERNUS AB/HB series の SANtricity Dynamic Disk Pools テクノロジーの導入により、管理者は従来のボリューム・グループよりも多くのメリットを得ることができます。このテクノロジーによるメリットは以下です。

- 再構築時間を大幅に短縮し (最大 8 倍)、重要なセグメントの再構築を優先順位づけすることにより、追加のドライブ障害のリスクを軽減
- 再構築時のパフォーマンスへの影響を軽減
- 一度に最大 60 台のドライブをオンラインで追加または削除可能
- 専用のアイドル状態のホット・スペアを排除

ETERNUS AB/HB series ストレージシステムは、DDP と従来のボリューム・グループの両方をサポートこれらを 1 つのシステム内で組み合わせることで、ワークロード要件を満たす優れた柔軟な構成を実現できます。

付録 A
用語集

次の表では、本書で使用されている用語のうち、読者にはなじみのない用語について説明します。

用語	説明
クリティカルセグメント再構成	複数のドライブ障害の影響を受ける2つのDピースがDストライプにある場合、システムはそれらをプール内のほかのドライブに再作成する必要があります。これらは、影響を受けるDピース内の各セグメントのRAID 6再構築によって再作成されます。これらのセグメントはクリティカルセグメントと呼ばれます。
Dピース	1つのドライブに含まれるDストライプの部分。
Dピースサイズ	Dピースのサイズはプラットフォームによって異なります。ETERNUS AB6100を除くすべてのストレージシステムでは、Dピースの総容量は512 MiBです。ETERNUS AB6100の場合、Dピースの総容量は1 GiBです。
ドロワー損失保護(DLP)	プールにDLPが搭載されている場合、障害やドライブ・シェルフの1つのドロワー全体の取り外しにも、可用性を損なうことなく対応できます。
Dストライプ	ボリュームを含むDDP要素。各Dストライプは、プールのサイズに関係なく、プール内の正確に10ドライブに存在します。Dストライプは、インテリジェントなアルゴリズムによってプール全体に分散され、パフォーマンスと障害許容度を最適化します。
Dストライプサイズ	Dストライプのサイズはプラットフォームによって異なります。ETERNUS AB6100以外のすべてのストレージシステムでは、Dストライプに4 GiBのユーザー・データと1 GiBのRAID 6パリティが含まれます。ETERNUS AB6100の場合、Dストライプには8 GiBのユーザー・データと2 GiBのRAID 6パリティが含まれます。
GB、MB、TB、PB	これらの単位は、10進数で容量を参照するときに使用されます。I/Oスループット値と未フォーマット時のドライブ容量は、10進数で表されます。
GiB、MiB、TiB、PiB	これらの単位は、2進数で容量を参照するときに使用されます。プール、ボリューム・グループ、ボリューム容量は、すべて2進数で表されます。
論理ユニット番号(LUN)	ボリュームがホストに割り当てられると、ホストがそのボリュームを認識するための参照番号が取得されます。
プール	ボリュームを格納するDDPテクノロジーのコンテナ。
RAID 6ストライプ	8つのデータ・セグメントと2つのパリティ・セグメントを含むDストライプのサブセット。PおよびQパリティ・セグメントは、Dストライプに含まれるドライブ間に分散されます。ETERNUS AB6100のDストライプには8,192個のRAID 6ストライプがあり、他のすべてのストレージシステムのDストライプには4,096個のRAID 6ストライプがあります。
RAIDグループ	ボリュームを格納し、従来のRAIDを使用するコンテナを指す業界用語。ETERNUS AB/HB seriesではボリューム・グループと呼びます。

用語	説明
セグメント	1つのドライブ上に存在するRAID 6ストライプの部分。セグメントのサイズは128 KiBで、ETERNUS AB6100のDピースには8,192個のセグメントがあり、他のすべてのストレージシステムのDピースには4,096個のセグメントがあります。
ボリューム	アプリケーション・データの格納に使用される容量の単位。ボリュームはボリューム・グループまたはプールに存在し、ホストに関連づけられるとLUNに割り当てられます。
ボリューム・グループ	RAIDグループを指す。ETERNUS AB/HB seriesでの呼び名。

付録 B
シンプロビジョニング

ETERNUS AB/HB series では、DDP 機能と組み合わせて使用する場合にシンプロビジョニングをサポートしますが、ほとんどのワークロードでは、シックボリュームの方が適しています。

注 意

ETERNUS AB6100 では、RAID 10、RAID 5、RAID 6 のシン・プロビジョニングがサポートされていません。

シンプロビジョニング機能はオーバーコミット容量を提供しますが、シック・ボリュームと同レベルの IOPS またはスループット・パフォーマンスは提供しません。そのため、シン・ボリュームは通常、ETERNUS AB/HB series やトランザクション・ワークロードには推奨されません。シンプロビジョニングされるボリュームの最大サイズは 256 TiB です。

最初にシンボリュームを構成し、シックボリュームに変更する場合は、既存のデータを新しいシックボリュームにコピーする必要があります。したがって、データをコピーする時間、変換能力に関連するコスト、およびアプリケーションの切り替えロジスティックスを慎重に計画する必要があります。

FUJITSU Storage
ETERNUS AB series オールフラッシュアレイ ,
ETERNUS HB series ハイブリッドアレイ
SANtricity OS Dynamic Disk Pool 機能の説明とベストプラクティス

P3AG-5672-01Z0

発行年月　2020 年 11 月
発行責任　富士通株式会社

● 本書の内容は、改善のため事前連絡なしに変更することがあります。
● 本書の内容は、細心の注意を払って制作致しましたが、本書中の誤字、情報の抜け、本書情報の使用に起因する運用結果に関しましては、責任を負いかねますので予めご了承願います。
● 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。

FUJITSU