

エネルギー効率の良い データ・センター・ストレージ： ストレージ製品の電力効率の評価

2022年7月

編集者：

Donald Goddard NetApp、グリーン・グリッド、SNIA グリーン
ストレージ技術作業部会

寄稿者：

Bill Bailey Seagate Technology、SNIA グリーンストレージ
技術作業部会

Jay Dietrich IBM（退職）

Rona Newmark Dell（退職）

Chuck Paridon SNIA グリーンストレージ技術作業部会

Carlos Pratt IBM、SNIA グリーンストレージ技術作業部会

Patrick Stanko SNIA グリーンストレージ技術作業部会

Herb Tanzer SNIA グリーンストレージ技術作業部会

Dave Thiel SNIA グリーンストレージイニシアティブ

グリーン・グリッド・データ・センター・エネルギー効率技術作業部会（DCEE TWG）は、ストレージネットワークング・インダストリ・アソシエーション（SNIA®）グリーンストレージイニシアティブ（GSI）およびグリーンストレージ技術作業部会（グリーン TWG）との緊密な連携と本ペーパーで提供するデータの収集と分析に対する多大なる貢献に対して感謝の意を表する。

要旨

グリーン・グリッド・アソシエーション（世界中の IT リソースおよびデータ・センター・リソースの効率の改善を目的とした共同事業体）の目的は、ストレージ製品のエネルギー効率を評価するためのツールとしての SNIA Emerald テスト・スイートの有効性を評価し、ストレージ製品の性能／電力効率しきい値の策定および設定におけるその適用および用途を検討することである。このような目的のために、グリーン・グリッドは、本ホワイトペーパーで報告されている SNIA の分析の取り組みを後押しした。

ストレージネットワークング・インダストリー・アソシエーション（SNIA®）は、2020 年 5 月 28 日にリリースされた ENERGY STAR データ・センター・ストレージ 2.0 プログラムの公開を支援して、2020 年 7 月 3 日に SNIA Emerald™ 電力効率測定仕様のバージョン 4.0 をリリースした。SNIA グリーンストレージ技術作業部会（グリーン TWG）は、Energy Star データ・センター・ストレージ・プログラムによって認定された 160 システムに関して EPA から公開された SNIA Emerald 結果だけでなく、追加のアイドル測定データも収集した。SNIA Green TWG は、このデータを分析して、テスト対象の 3 つのワークロード・タイプ（容量、シーケンシャル、およびトランザクション）の SNIA Emerald 性能電力指標に対する構成タイプとコンポーネント選択の影響を把握した。本ホワイトペーパーでは、作業部会の調査結果の詳細を示し、データ・ストレージ製品の電力効率の評価に対する SNIA Emerald 方式の有効性に関する情報をストレージ製品の製造業者、規制機関、およびその他の利害関係者に提供する。

本ホワイトペーパーでは、規制機関に対して、電力効率を評価する場合にどのような製品基準とテスト指標を使用すべきかに関する提言を行う。また、オンライン・データ・ストレージ製品の機能に関する情報と様々な特徴も提供する。この情報には、このような製品の構成方法、組み込まれている容量最適化技術、保存データの整合性とセキュリティの保証方法、およびアイドル状態における動作が含まれる。エンタープライズ・ストレージの新しいトレンドに関する情報も提供される。

目次

1. はじめに	7
2. ストレージ・アーキテクチャの概要	8
ストレージ・アーキテクチャの違い	8
ストレージ・システムのスケール方法：スケールアップとスケールアウト	10
ローカル（DAS）ストレージと共有ストレージの比較	12
クラウドストレージ	13
エッジ・データ・ストレージ	14
データ・セキュリティ	14
データの可用性、持続性、信頼性、および回復力	14
3. ストレージの分類が必要な理由	16
はじめに	16
分類セット	16
分類カテゴリ	17
全般	17
カテゴリ属性	17
分類クラシフィケーション	18
ディスク・セット・オンライン・カテゴリ	20
ニアオンライン・カテゴリ	21
RVMLセット・リムーバブル・メディア・ライブラリ・カテゴリ	22
RVMLセット仮想メディア・ライブラリ・カテゴリ	22
NVSSセット・ディスク・アクセス・カテゴリ	23
NVSSセット・メモリ・アクセス・カテゴリ	24
4. ストレージの物理属性	25
5. ストレージの短期的進化	27
ストレージ・メディア	27
ユース・ケースの比較	27
メディア開発の方向性	27
HDDストレージ・メディア	28
SSDストレージ・メディア	28
まとめ	29
新たなストレージ・トレンド	30
アプリケーションがアプライアンスに取って代わる	30
ハイブリッド・クラウドの導入	30
サービスとしてのストレージ（StaaS）	31

6. 容量最適化手法 (COM)	32
COMのメリットの定量化	33
7. ストレージに関する SNIA EMERALD データが示すもの	35
トランザクション・ワークロード用に最適化されたブロック・ベースのシステムからのデータ	35
シーケンシャル・ワークロード用に最適化されたブロック・ベースのシステムからのデータ	38
所見	40
8. アイドル	41
ディープ・アイドル	41
アイドル状態	41
ソリッド・ステート・ドライブのアイドル状態	41
磁気ディスク・ドライブのアイドル状態	42
9. 日本の標準	44
Japan Energy Efficiency Standard (トップ・ランナー)	44
日本のグリーン購入	46
10. 結論と電力効率認定に関する提案	49
結論	49
提案	49
基準の提案	49
テストの提案	50
まとめ	50
11. グリーン・グリッドと SNIA	51
グリーン・グリッド	51
SNIA	51

Table of Figures

Figure 1: Basic Storage Architectures.....	8
Figure 2: Direct Attach Storage (DAS) Examples	9
Figure 3: Storage Area Network (SAN)	9
Figure 4: Network Attached Storage (NAS).....	10
Figure 5: Scale-up Storage Example	11
Figure 6: Scale-out Storage Example	12
Figure 7: DAS (Local) and Shared Storage	13
Figure 8: Taxonomy Overview	16
Figure 9: Typical Storage Enclosure: Entry Level, Expand Out	25
Figure 10: Hot Band vs. Ready Idle.....	36
Figure 11: Hot Band vs. Ready Idle (Expanded Scale).....	36
Figure 12: Hot Band Metric vs. Drive Speed (HDDs).....	37
Figure 13: Hot Band Metric vs. Drive Speed (HDDs and SSDs)	37
Figure 14: Sequential Read vs. Ready Idle Metrics	38
Figure 15: Sequential Read Metric vs. HDD Speed	39
Figure 16 Sequential Write vs. Ready Idle Metrics	39
Figure 17: Sequential Write Metric vs. HDD Speed	40
Figure 18: 7680 GB SAS SSD.....	42
Figure 19: Example Power Consumption of Idle Magnetic Disk Drive	43
Figure 20: Ready Idle Metric vs. Drive Speed.....	44
Figure 21: Ready Idle Metric vs. Drive Speed for Sequential Optimized Systems	45
Figure 22: Ready Idle Metric vs. Drive Capacity (7200 RPM Drives)	45
Figure 23: Ready Idle Metric vs. Drive Speed (Transactional Systems).....	47
Figure 24: Ready Idle Metric vs. Drive Speed (Sequential Systems)	47
Figure 25: Ready Idle Metric vs. Drive Capacity (All Systems).....	48

Table of Tables

Table 1 – Common Category Attributes	17
Table 2 – Taxonomy Classifications Overview	19
Table 3 – Disk Set Online Classifications	20
Table 4 – Disk Set Near-Online Classifications	21
Table 5 – RVML Set Removable Media Library Classifications	22
Table 6 – RVML Set Virtual Media Library Classifications	22
Table 7 – NVSS Set Disk Access Classifications	23
Table 8 – NVSS Set Memory Access Classifications	24
Table 9 – Active COM Examples (70/30 R/W Random Workload), Based on Physical Capacity Reduction ...	33
Table 10 – Active COM Examples (70/30 R/W, Random Workload) Based on Usable Capacity Optimizations	34
Table 11 – ENERGY STAR configurations of data presented in this paper	35

1. はじめに

SNIA グリーンストレージ技術作業部会は、データ・センター・ストレージ製品に関する ENERGY STAR プログラム要件に基づいて認定したストレージ製品の SNIA Emerald™ 電力効率測定テスト (SNIA Emerald テスト) データを分析した。このデータセットは、オンライン 2、3、および 4 の製品カテゴリに及ぶ Energy Star データ・ストレージ Web サイト上で認定された 160 システムからなる。分析では、ストレージ製品の電力効率を評価するツールとしての SNIA Emerald テスト・スイートの有効性を評価した。そして、ストレージ製品の電力効率しきい値の策定および設定における適用および用途を検討している。この分析に基づいて、データ・センター・ストレージ電力効率特性評価に関する提言がなされる。これには、簡易テストと製品基準が含まれる。

一般的に、データ・センターは、満たすべきサービス・レベル要件を抱えている。ストレージやその他のコンポーネントの選択は、その要件を満たすためのニーズに基づいて行われる。その要件を満たしている製品では、電力効率が検討事項になる。電力効率指標は様々なストレージ製品を比較するための基準となり、この比較によって、より効果的な電力利用が可能になる。電力効率指標は、実際のデータ・センターの電力消費を直接予測できるわけではないが、電力消費の有意義な相対評価を可能にする。つまり、特定の作業負荷では、効率指標の高い製品の方が効率指標の低い製品よりも電力消費が少ないことが期待できる。

2. ストレージ・アーキテクチャの概要

ストレージ・アーキテクチャの違い

従来のストレージ・システム・アーキテクチャには基本的な 3 つのタイプがある。直接接続ストレージ (DAS)、ストレージ・エリア・ネットワーク (SAN)、およびネットワーク・アタッチド・ストレージ (NAS) である (Figure 1)。これらのアーキテクチャには重複した要素が含まれており、それらについて説明する。

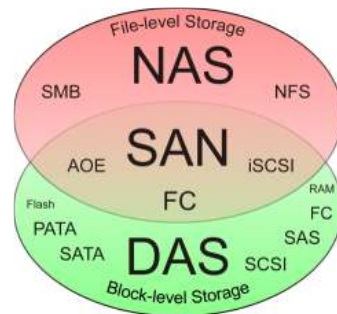


Figure 1: Basic Storage Architectures

DAS は、ブロックレベル・ストレージを提供する最も基本的なストレージ・システムであり、ストレージ・ネットワークを介さずに直接サーバまたはワークステーションに接続される。DAS システムの例を Figure 2 に示す。1 つ目の例では、4 つの Small Computer System Interface (SCSI) ハード・ディスク・ドライブ (HDD) が SCSI ケーブル配線を介してホスト・コンピュータに接続されている。2 つ目の例では、Redundant Array of Independent Disk (RAID) または Just a Bunch Of Disks/Devices (JBOD) ストレージ・サブシステムにホスト・コンピュータを接続するためにファイバ・チャネル・ケーブル配線が使用されている。RAID システムは、コントローラを使用してストレージ・デバイスを管理し、ストレージ・システムの信頼性を向上させるのに対して、JBOD システムは、HDD または SSD (ソリッド・ステート・ドライブ) コントローラを使用してワークステーション/サーバと JBOD システム間のデータを管理する。前者のアプローチだけがデータ・センターの実装に関係する。DAS の主な特徴は、個別のコンピュータ/サーバにストレージ・リソースが結びついていることである。これは、安価なストレージ・ソリューションであるが、専用のストレージ・リソースが何らかの制限になる可能性がある。以下に例を示す。

- バスや筐体の設計によってサポートされる HDD の数によってストレージ容量が制限される。
- コンテンツをサーバ間で共有するのではなく複製する必要があるためストレージ・リソースの使用効率が悪く、あるサーバ上の空きストレージ・リソースをディスク領域が不足している別のサーバで使用することができない。
- 接続されたストレージ・リソースにアクセスできなくなるサーバ障害によって可用性が制限される。
- 個別のサーバの処理速度によって性能が制限される。つまり、並列処理によって複数のサーバにワークロードを分散することができない。

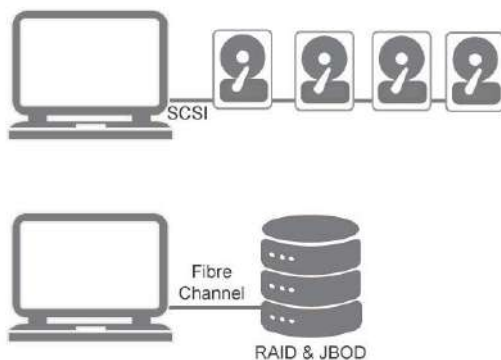


Figure 2: Direct Attach Storage (DAS) Examples

NAS と SAN (Figure 3 と Figure 4 を参照) は、ネットワーク上でストレージを共有するための 2 つの方法であり、DAS に対して本質的メリットを持つ。SAN は、複数のホスト (サーバ) をブロックレベルで単一のストレージ・デバイスに接続することができるため、DAS よりも高度な機能を提供する。通常は、単一ストレージ・ボリュームへの同時接続を許可しないが、あるサーバがボリュームの制御を放棄した場合に別のサーバがそのボリュームを確保することは可能である。

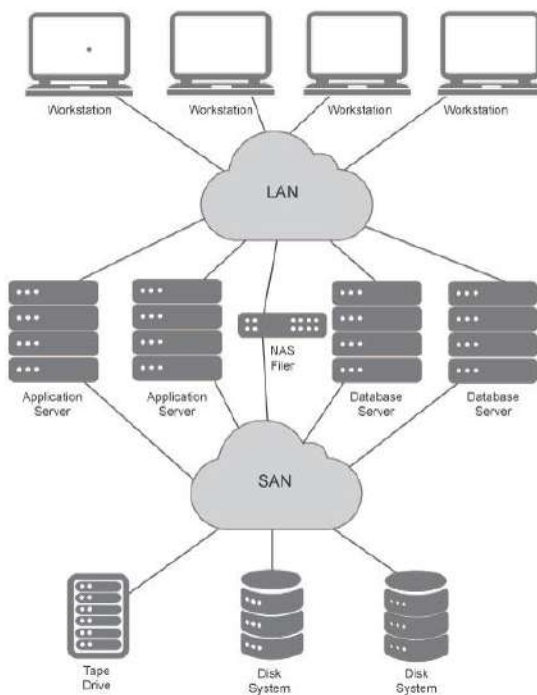


Figure 3: Storage Area Network (SAN)

NASは、基本的にSANまたはDAS技術上に構築された「ファイル・サーバ」であり、NFSやCIFSなどのネットワーク・ファイル・システム・プロトコルを使用してネットワーク（LAN）経由でアクセスされる。NASとSANの違いは、NASが「ファイルレベル I/O」であるのに対して、SANはネットワーク上の「ブロックレベル I/O」である。これは、データがネットワーク経由でブロック（SAN）またはファイル・データ・ストリーム（NAS）として受信者に転送されることを意味する。ファイル・アクセス・モデルは、より高い抽象化レイヤに構築されるため、追加的な処理レイヤがストレージ・サーバにも、NASストレージ・サーバ内でファイル・アクセスとブロック・アクセス間を変換する機能にも必要である。このため、SANブロックレベル・アクセスと比較して、多くのアプリケーションで、処理の遅延が大きくなり、I/Oスループットに影響が出る。NASにおけるより高いレベルの抽象化がもたらすメリットは使いやすさである。共有ストレージは、NASシステムを使い慣れたエンタープライズLAN（イーサネットなど）に接続し、ワークステーションとサーバ上のOSをNASストレージ・サーバにアクセスするように設定することにより、簡単に実装できる。SANやDASと比べたNASのもう1つのメリットは、複数のクライアントで単一のボリュームを共有できることである。SANボリュームとDASボリュームは、同時に単一のクライアントによってしかマウントできない。UNIXやLINUXなどのオペレーティング・システムはNFSプロトコルをサポートし、Windows OSはCIFSプロトコルをサポートしている。

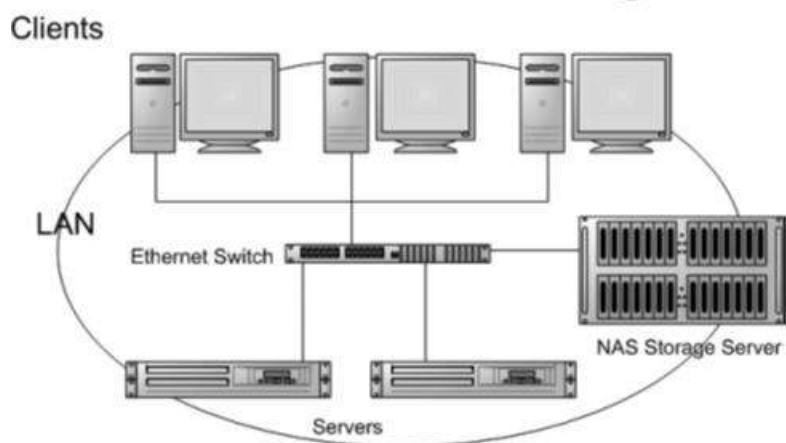


Figure 4: Network Attached Storage (NAS)

ストレージ・システムのスケール方法：スケールアップとスケールアウト

レガシーまたは従来のシステムは、「スケールアップ」ストレージ（Figure 5）を使用して容量を増やしている。長年にわたって、標準的なストレージ・アレイは、2つのストレージ・コントローラ・ヘッド（冗長性のため）とメディア装置を収納する複数のディスク・シェルフで構成されてきた。ストレージ・コントローラはストレージ・エリア・ネットワーク（SAN）に接続して、計算サーバにストレージを提供する。すべてのディスク・シェルフがストレージ・コントローラに接続され、すべての計算サーバが2つのコントローラを介してこれらのストレージ・デバイスにアクセスする。ストレージ・アレイの容量と性能を高めるには、同じコントローラのペアのストレージ・デバイス・シェルフとストレージ・デバイスを増やす。

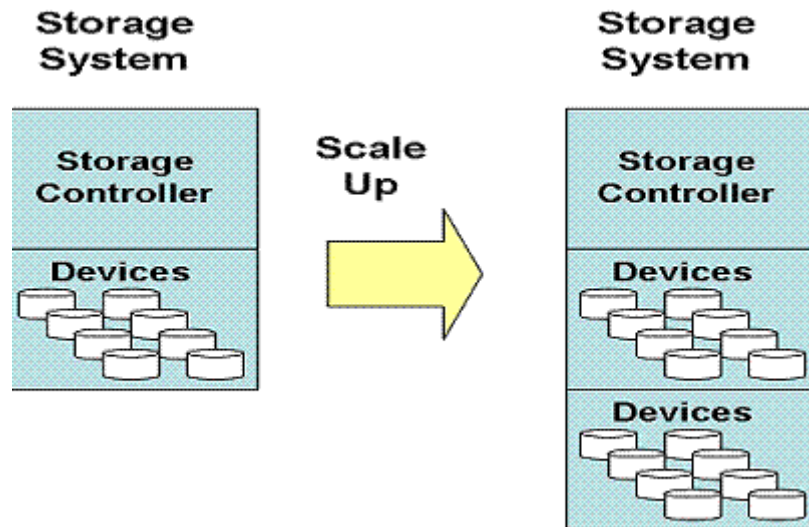


Figure 5: Scale-up Storage Example

スケールアップは、ネットワーク接続などのインフラストラクチャー要素を増やすことなく容量問題を解決できる。ただし、ストレージ・デバイスを追加するには、スペース、電力、および冷却を増やす必要がある。スケールアップでは、コントローラ機能が増えないため、追加のホスト活動を処理できない。既存のストレージ・コントローラを使用してストレージ・デバイスが追加されるだけなので容量をスケールするためのコスト面で有利である。

まず、性能の観点からディスク・スループットが制限要因になることが多いため、ストレージ・デバイスの数を増やせば、通常は、性能が向上する。ただし、ストレージ・アレイに対する負荷が増大し（仮想化によってよく引き起こされる状況）、デバイスの数が増えると、2つのストレージ・コントローラが互いに RAID 計算とその他のコマンド処理をどんどん CPU に要求し始めるため、それら自体がボトルネックになる。最終的に、コントローラが飽和状態になるまでデバイスが追加されると、それ以上性能は上がらない。過負荷状態のコントローラのペアにより速いデバイスを追加してもコントローラ上の過負荷状態が進むだけである。このようになる時点を遅らせる 1 つの方法はストレージの計算能力（コントローラの能力）と帯域幅を予め余分に準備しておく方法である。ただし、コストは高くなる可能性がある。性能ピークに到達した場合の次の選択肢は、ストレージ・コントローラの計算機能と帯域幅機能を交換するか、新しいスタンドアロンのストレージ・システムを購入するかである。これらの選択肢は高価であり、管理の負担になることがある。

容量と性能は「スケールアウト」ストレージ（Figure 6）を使用して拡張できる。通常、スケールアウト・ストレージには、容量と性能の両方を増やすための追加のコントローラとストレージ要素が必要である。スケールアウトと、フロアに追加のストレージ・システムを設置するだけの重要な違いは、スケールアウト・ストレージは単一のシステムとして存在し続けることである。

クラスタ化されたストレージ・システムやグリッド・ストレージ・システムなどのスケールアウトを実現するための方法は複数存在する。Figure 6 に示すスケールアウト・ストレージは、制御機能と容量の両方が追加されているが、単一システムとしてアクセスできる。このスケールアップには、ストレージをコントローラに接続するためのストレージ・スイッチや、クラスタまたはグリッド内のノード間の接続などの追加のインフラストラクチャーが必要な場合がある。スケールアウトでは電力、冷却、およびスペース要件が増え、追加の容量、制御要素、およびインフラストラクチャーのコストが増える。この例のスケールアウト・ソリューションでは、容量が増え、追加のコントローラとともに性能がスケールされる。

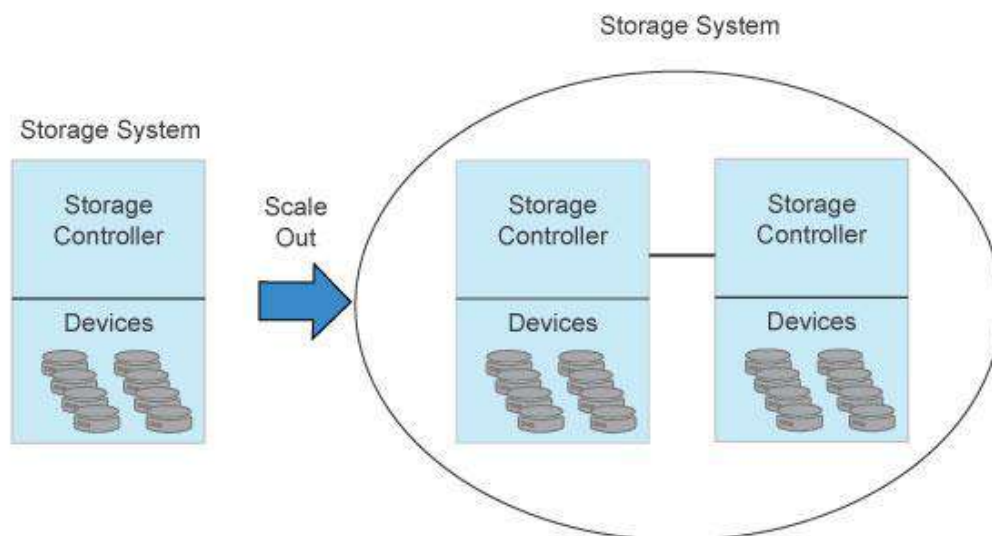


Figure 6: Scale-out Storage Example

スケールアウトは必ずしも新しい概念ではないが、ストレージ・ベンダーが旧式のスケールアップ・アーキテクチャの制限を克服するために最新のサーバの能力を活用して新しい構造を作成したため、新しい命が吹き込まれた。このような新しいスケールアウト・ストレージ・システムは、カスタム・コントローラ・プラットフォームではなく、サーバのグループを使用してストレージ・クラスタを形成する。各サーバにはストレージ・デバイスが搭載され、ネットワーク（通常はイーサネット）を使用して、ストレージ・アレイ内の他のサーバと通信する。サーバのグループは、協力してクラスタ化されたストレージ・アレイを形成し、従来のアレイと全く同様にネットワーク経由で論理ユニット（LUN）またはファイル共有を提供する。スケールアウト・アレイにストレージ・デバイス・ノードを追加するということは、実際には、新たなサーバをクラスタに追加するということである。それと同時に、ネットワーク・ポート、CPU、およびRAMも追加される。スケールアウト・アレイの容量が増えるほど、その性能も増えるため、ノードを追加すると、通常は、性能が線形的に向上する。通常はノードの追加では停止は起きないため、システムを完全に停止することなく、通常のメンテナンス・ウィンドウを使用してアレイの容量と性能を増やすことができる。

ローカル（DAS）ストレージと共有ストレージの比較

ストレージ・システムは単一のサーバ専用の場合（DAS）と複数のサーバ間で共有される場合（SANとNAS）がある。共有ストレージを使用すれば、複数のサーバから同時にまたは別々にストレージにアクセスすることができる。一般的に、共有するように設計されたストレージ・システムは、容量、性能、拡張性、および機能が高く、その代償としてローカル・ストレージ・システムよりコスト高になる。

共有ストレージには、必ず、インターフェイス・カードやネットワーク・スイッチなどのサーバに接続するための追加のハードウェアが必要であり、その分価格が高くなる。共有ストレージは、定期的な調整と追加の管理専門知識が必要なほど複雑である。加えて、サーバがストレージを共有する場合は、共有データが正しく更新されるように各サーバのアクションを調整する必要がある。また、サーバのキャッシュ・データに対応するストレージの部分が別のサーバによって変更された時にキャッシュ・データを確実に無効化することにより、サーバのキャッシュを一貫した状態に保つ必要もある。

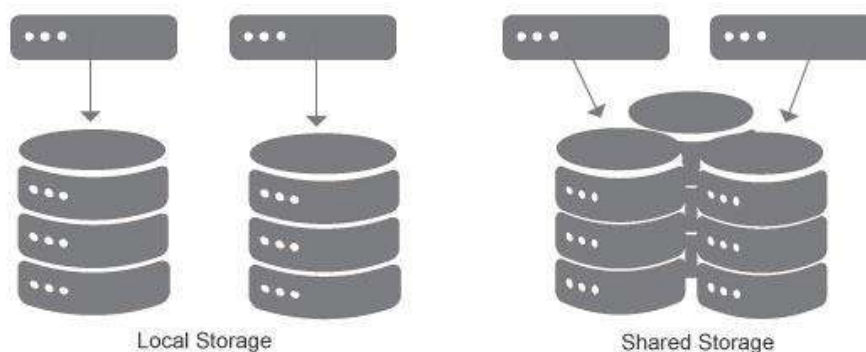


Figure 7: DAS (Local) and Shared Storage

環境やワークロードの要求によっては、ローカル・ストレージで十分な場合と共有ストレージが必要な場合がある。ローカル・ストレージ・デバイスを使用することにより、Input/Output Operations Per second (IOPs) によって測定されるスループットは非常に高くなる可能性があり、ローカル・ストレージのコストは、展開が簡単で管理が単純なため低く抑えられる。ただし、ネットワークに接続された共有ストレージ (SAN/NAS) の拡張性と仮想化の側面は、より高いレベルの冗長性、可用性、および災害復旧性を提供できる。加えて、共有ストレージは、ローカル・ストレージが備えていない様々な追加のソフトウェア機能を提供する。COM¹ (スナップショット、シン・プロビジョニング、圧縮、および重複排除)、自動階層化²、データ保護用の追加のフォーム、データのリモート複製などである。

クラウドストレージ

クラウドストレージは、デジタル・データが「クラウド」と呼ばれる論理的プールに

保存されるコンピュータ・データ・ストレージのモデルである。クラウド・サービスの展開は複数の方法で行うことができる。

パブリック・クラウドは、クラウドのコンピューティングとストレージを展開する際の最も一般的な種類である。クラウド・リソース (サーバやストレージなど) は、サードパーティ・クラウド・サービス・プロバイダによって所有・運用され、インターネット経由で提供される。パブリック・クラウドを使用する場合は、すべてのハードウェア、ソフトウェア、およびその他の補助インフラストラクチャーがクラウド・プロバイダによって所有・管理される。パブリック・クラウドでは、ハードウェア、ストレージ、およびネットワーク・デバイスが他の組織 (クラウド「テナント」) と共有され、サービスはネットワーク経由でアクセスされ、ウェブ・ブラウザを使用して管理される。パブリック・クラウドの展開は、ウェブベースの電子メール、オンライン・オフィス・アプリケーション、ストレージ、テスト環境、および開発環境を提供するためによく使用される。

プライベート・クラウドは、1つの会社または組織によって排他的に使用されるクラウドのコンピューティング・リソースとストレージ・リソースで構成される。プライベート・クラウドは、組織のオンサイト・データ・センターに物理的に配置することも、サードパーティ・サービス・プロバイダによってホストされることもできる。ただし、プライベート・クラウドでは、サービスとインフラスト

¹ COM (Capacity Optimization Methods) は、実際の容量よりも多くのデータを保存できるようにして物理リソースを削減するストレージ・ソフトウェア管理機能である (つまり、使用するディスクと電力を減らす)。

² 自動ストレージ階層化は、スペース、性能、およびコスト要件を満たすために複数のディスク・タイプと RAID レベル間で情報を動的に移動するストレージ・ソフトウェア管理機能である。

ラクチャーが常にプライベート・ネットワーク上で維持され、ハードウェアとソフトウェアは1つの組織専用となる。

ハイブリッド・クラウドは、オンプレミス・インフラストラクチャー、プライベート・クラウド、およびパブリック・クラウドを組み合わせたクラウドの種類である。ハイブリッド・クラウドでは、データとアプリケーションを2つの環境間で移動することができる。

詳細については、<https://www.snia.org/education/what-is-cloud-storage> を参照のこと。

エッジ・データ・ストレージ

エッジ・コンピューティングは、コンピューテーションとデータ・ストレージをそれらが必要とされる場所の近くに配置することによって応答時間を短縮し、帯域幅を節約する分散コンピューティング・パラダイムである。

ガートナーは、エッジ・コンピューティングを「情報処理がエッジ（モノと人がその情報を生成または消費する場所）の近くに配置される分散コンピューティング・トポロジの一部」と定義している。

基本レベルで、エッジ・コンピューティングは、何千マイルも離れた中央の場所に依存するのではなく、コンピューテーションとデータ・ストレージをデータが収集されるデバイスの近くに配置する。これは、データ、特に、リアルタイム・データでアプリケーションの性能に影響を及ぼしかねない遅延問題が発生しないように実施される。加えて、企業は、処理をローカルに行って、中央の場所やクラウド・ベースの場所で処理する必要のあるデータ量を減らすことによって、経費を節約できる。

エッジ・コンピューティングは、クラウドから情報を受信したり、クラウドにデータを戻したりするためにインターネットに接続するIoTデバイスの急増を受けて開発された。また、多くのIoTデバイスが動作中に大量のデータを生成する。

詳細については、<https://www.snia.org/educational-library/storing-edge-architectural-approach-2018> を参照すること。

データ・セキュリティ

データ・セキュリティは、ライフサイクルを通して無許可アクセスや改ざんからデータを保護するプロセスを意味する。データ・セキュリティには、すべてのアプリケーションとプラットフォーム上でデータを保護するデータ暗号化、ハッシュ法、トークン化、および暗号鍵管理の活動が含まれる。

堅牢なデータ・セキュリティ戦略を適切に実装すれば、組織の情報資産がサイバー犯罪者の活動から保護されるだけでなく、今日でもデータ漏洩の主要な原因となっている内部関係者による脅威や人的エラーからも保護される。データ・セキュリティには、重要なデータの保存場所や用途に対する組織の可視性を向上させるツールや技術の展開が含まれる。理想を言えば、これらのツールは、暗号化、データ・マスキング、機密ファイルの伏せ字化などの保護を適用できる必要があり、監査を単純化するためのレポート作成や規制要件の遵守を自動化する必要がある。

詳細については、<https://www.snia.org/educational-library/storage-security-best-practices-2015> を参照すること。

データの可用性、持続性、信頼性、および回復力

エンタープライズ・ストレージ・システムは、データが破損または消失していないことと、データが必要なときに入手可能なことを保証するための多数の複雑なメカニズムを実装している。通常、これらのメカニズムには、ストレージ・システムがアイドル中でもホストからの要求を処理中でも実行されるバックグラウンド活動が含まれる。

システム可用性は、システムが要求に対処できる期間（つまり、システムのアップタイム）の割合である。

データ可用性は、データに素早くアクセスできる度合いである。この用語は、ほとんどの場合、サービス・レベルに関連付けられていて、サービス・レベルは、社内の IT 組織によって設定される場合や、サードパーティのデータ・センターまたはストレージ・プロバイダによって保証される場合がある。

データ持続性／完全性は、保存データがビット腐敗(データ劣化)、ドライブ故障、または何らかの破損形態の影響を受けず、一貫した損傷のない状態を維持する能力である。

システム信頼性は、可用性と密接な関係がある。ただし、システムは「使用できる」場合でも、正しく機能するとは限らない。信頼性は、システムが仕様どおりに機能する確率である。

データ信頼性は、データの正確性と完全性である。

システム回復力は、損傷、故障、過負荷、または攻撃後にシステムが自己回復する能力である。これは、必ずしもイベント中にシステムを連続使用できることを意味しているわけではなく、自己回復するだけであることに注意されたい。

3. ストレージの分類(Taxonomy)が必要な理由

データ・ストレージ・システムの基本的な機能は、プライマリ・データ・ストレージ、ミラー・データ・ストレージ、またはバックアップ・データ・ストレージとしてネットワークにストレージ・リソースを提供することである。様々なストレージ・システムが市販されており、異なる性能と機能で様々な目的に使用されている。ストレージ・システムには、JBOD ディスク・アレイ、RAID ディスク・アレイ、テープ・システム、光ストレージ・システムなどがある。このようなデバイス上で提供されるインターフェイスのタイプには、ファイバ・チャネル、SAS、およびイーサネットが含まれる。

市場では広範囲に及ぶストレージ指向の製品が販売されているため、SNIA Emerald[™]電力効率測定仕様³で分類構造が作成された。分類が存在することで、電力効率を評価する場合に、サイズ、容量、性能、高可用性などを考慮した「同様の」製品同士での比較が可能になる。ストレージ分類によって、異なる電力効率テスト基準が必要なストレージ・システムのクラスが識別される。

はじめに

SNIA Emerald ストレージ分類は、操作プロファイルとサポートされている機能の観点でストレージ製品またはサブシステムを分類する市場分類を定義したものである。この分類は、広範囲に及び、コンシューマ・ソリューションからエンタープライズ・インストールまでの製品の枠組みを定義しているが、その目的はすべてのストレージ・デバイスをカバーすることではない。

この分類は、セット、カテゴリ、およびクラシフィケーションの3つの階層で構成されている。セットは大まかな製品のグループ分けである。セットには、ディスク・セット、RVML（リムーバブルと仮想メディア・ライブラリ）セット、および NVSS（不揮発性ソリッド・ステート）セットが含まれる。カテゴリは、セット内の製品のグループ分けである。カテゴリには、オンライン、ニアオンライン、リムーバブル・メディア・ライブラリ、仮想メディア・ライブラリ、ディスク・アクセス、およびメモリ・アクセスがある。クラシフィケーションは、製品の精巧さ、サイズ、および複雑さのレベルであり、市場での区分に対応する。クラシフィケーションは、カテゴリに付加された数字の接尾辞として識別される。クラシフィケーションによっては、複数のカテゴリで使用される場合がある。

分類セット

分類セットは、同様のシステム特性を共有するストレージ製品の大まかなグループ分けを定義したものである。Figure 8を参照のこと。セットが異なる製品は、一般的に、性能や電力効率特性の点で比較可能でない。

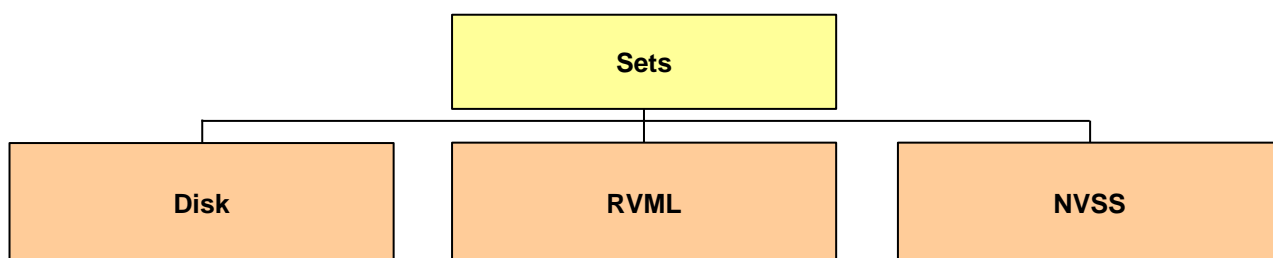


Figure 8: Taxonomy Overview

³SNIA Emerald[™]電力効率測定仕様、<https://www.sniaemerald.com/download>

分類カテゴリ

全般

分類カテゴリは、セット内で大まかに市場セグメントを定義したものであり、共通の機能や性能要件を持つ製品をグループとしてまとめるために使用することができ、その中では意味のある製品の比較が可能である。本書では、6つの大まかな分類カテゴリを定義する（Table 2に要約を示す）。

- ディスク・セット・オンライン
- ディスク・セット・ニアオンライン
- RVML セット・リムーバブル・メディア・ライブラリ
- RVML セット・仮想メディア・ライブラリ
- NVSS セット・ディスク・アクセス
- NVSS セット・メモリ・アクセス

分類カテゴリでは、容量、信頼性、性能、機能などのどの差別化要因に重点が置かれているかに関係なく、製品の1つの特定のモデルまたはリリースが異なる機能セットをサポートする。カテゴリ内の機能の違いは属性によって示される。分類カテゴリごとに、そのカテゴリ内のすべての製品に共通する属性のセットが定義される。

カテゴリ属性

分類カテゴリの中で、属性に対して固有の固定的な設定や範囲が必要な場合があり、その設定を Table 1にまとめる。これは、最初のカテゴリ選択でテスト・スポンサーの役に立つ。

Table 1 - Common Category Attributes

Attribute	Set					
	Disk		RVML		NVSS	
	Category					
	Online	Near-Online	Removable Media Library	Virtual Media Library	Disk Access	Memory Access
Access Pattern	Random/ Sequential	Random/ Sequential	Sequential	Sequential	Random/ Sequential	Random
MaxTTFD	≤ 80 ms	> 80 ms	≤ 5 min	≤ 80 ms	≤ 80 ms	≤ 80 ms
Media Type	Magnetic disk	Magnetic disk	Magnetic tape, optical disk	Magnetic disk, Solid State Storage	Solid State Storage + optional magnetic disk ^a	Solid State Storage
Access Paradigm	Block, File, Object	Block, File, Object	Block	Block	Block, File, Object	Memory
^a Allows a purely Solid State Storage system or a hybrid Solid State Storage and magnetic disk system.						

Table 1 の Attribute 列は次のように定義される。

- Access Pattern は Random と Sequential のどちらかまたはその両方である。

ランダム・アクセス・パターン属性を持つストレージ製品は、任意の保存データに対して大体等しいアクセス時間を提供する。

シーケンシャル・アクセス・パターン属性を持つストレージ製品は、最後にアクセスされたデータの後ろのアドレスにあるデータへのアクセスが他のデータへのアクセスよりも速い可能性がある。

- MaxTTFD (Maximum Time To First Data) は、任意のデータに対する読み取り要求を満たすためにストレージ・システムからのデータ受信を開始するために必要な時間である。
- Access Paradigm は、Block、File、Object、および Memory の 1 つまたは複数である。

ブロック・アクセス・パラダイム属性を持つストレージ製品は、ディスク／テープへの SCSI ブロック・アクセスを介する場合など、データへの固定ブロックまたは CKD アクセスを提供する。

ファイル・アクセス・パラダイム属性を持つストレージ製品は、NFS プロトコルを介する場合など、データへのファイル・アクセスを提供する。

オブジェクト・アクセス・パラダイム属性を持つストレージ製品は、SCSI OSD プロトコル (INCITS 400-2004 と INCITS 458-2011) を介する場合など、データへのオブジェクト・アクセスを提供する。

メモリ・パラダイム属性を持つストレージ製品は、load や store などの CPU メモリ・アクセス命令を介したデータへのアクセスを提供する。

分類クラシフィケーション

クラシフィケーションは、カテゴリ内の属性の設定または値の組み合わせを定義したものである。分類クラシフィケーションは、異なるシステムを区別する。クラシフィケーションの目標は、サイズ、複雑さ、およびターゲット市場が異なるシステムを区別することである。クラシフィケーションの一般的な例を以下に示す。

- コンシューマ／コンポーネント (<クラシフィケーション> 1)
- JBOD (<クラシフィケーション> 1.5)
- ローエンド (<クラシフィケーション> 2)
- ミッドレンジ (<クラシフィケーション> 3) と (<クラシフィケーション> 4)
- ハイエンド (<クラシフィケーション> 5)
- メインフレーム (<クラシフィケーション> 6)

コンシューマ (<クラシフィケーション> 1) システムは、主に、個人、家族、世帯、またはスモールビジネスの目的で購入される多様な量産品である。

コンポーネント (<クラシフィケーション> 1) は、データ・センター・ストレージ製品の一部を構成するコンポーネント (スタンドアロンのテープ・ドライブや個別のハード・ディスク・ドライブなど) である。

JBOD (<クラシフィケーション> 1.5) は、「単純な」JBOD で、ストレージ保護機能が組み込まれていないアーキテクチャである。JBOD 内のストレージ・デバイスは、個別のボリュームとして機能することも、複数個を接続して単一の論理ボリュームを形成することもできる（冗長性を提供しない単一の内蔵コントローラによる）。JBOD は冗長性や回復力を提供しないため、単一のストレージ・デバイスの故障が論理ボリューム全体の故障になる。JBOD はシングル・シェルフである。

ローエンド・システム (<クラシフィケーション> 2) は、一般的に、1 台か 2 台の内蔵コントローラ付きで拡張性が制限されたシングル・シェルフである。

ミッドレンジ (<クラシフィケーション> 3 と <クラシフィケーション> 4) システムは、スモール・ミッドレンジ・システムとラージ・ミッドレンジ・システムに細分される。スモール・ミッドレンジ・システム (<クラシフィケーション> 3) は、一般的に、ラージ・ミッドレンジ・システム (<クラシフィケーション> 4) と比較して拡張性や堅牢性が制限される。

ハイエンド (<クラシフィケーション> 5) およびメインフレーム (<クラシフィケーション> 6) システムは、非常に大きな容量を有し、高度な堅牢性を提供する非常に大規模なシステムである。メインフレームは、カウント・キー・データ（最大データ容量を持つ一定数のトラックによってディスクが構成されていると仮定されたディスク・データ組織モデル）のサポートが要求されるという点で、ハイエンドと区別される。

Table 2に、分類クラシフィケーションの概要を示す。

Table 2 - Taxonomy Classifications Overview

Level	Set					
	Disk		RVML		NVSS	
	Category					
	Online	Near-Online	Removable Media Library	Virtual Media Library	Disk Access	Memory Access
	Classification					
Consumer/Component ^a	Online 1 ^b	Near-Online 1 ^b	Removable 1	Virtual 1	Online 1 ^b	Online 1 ^b
JBOD	Online 1.5	Near-Online 1.5 ^c	Removable 1.5 ^c	Virtual 1.5 ^c	Online 1.5	Online 1.5 ^b
Low-end	Online 2	Near-Online 2	Removable 2	Virtual 2	Online 2	Online 2 ^b
Mid-range	Online 3	Near-Online 3	Removable 3	Virtual 3	Online 3	Online 3 ^b
	Online 4	Near-Online 4 ^c	Removable 4 ^c	Virtual 4 ^c	Online 4	Online 4 ^b
High-end	Online 5	Near-Online 5	Removable 5	Virtual 5	Online 5	Online 5 ^b
Mainframe	Online 6	Near-Online 6	Removable 6	Virtual 6	Online 6	Online 6 ^b
^a Entries in this level of the taxonomy include both consumer products and data-center components, e.g., stand-alone tape drives. ^b No test procedure for this Classification is provided by the Emerald Specification. ^c Classification is not defined; no test procedure is provided by this document.						

ディスク・セット・オンライン・カテゴリ

ディスク・セットは、磁気ディスクを代表とする回転媒体デバイスに基づくストレージ製品で構成される。

ディスク・セット・オンライン・カテゴリは、オンライン、ランダムアクセス、回転媒体ベースのストレージ製品について、特徴と機能を定義したものである。このプロファイル内の製品は、ブロック、ファイル、またはオブジェクト・インターフェースの任意の組み合わせを提供できる。Table 3に、このカテゴリで定義された分類クラシフィケーション（ディスク・セット・オンライン・オンライン1など）に関する要件を示す。

Table 3 - Disk Set Online Classifications

Attribute	Category: Disk Set Online						
	Classification: Online						
	1 ^a	1.5	2	3	4	5	6
Multi-host Shareability	Not Specified	Ability to share with 1 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/Component	Yes	No	No	No	No	No	No
Storage Controller	Optional	Optional	Required	Required	Required	Required	Required
Storage Protection	Optional	No	Required	Required	Required	Required	Required
No SPOF	Optional	Optional	Optional	Optional	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Optional	Required	Required
FBA/CKD Support	Optional	No	Optional	Optional	Optional	Optional	Required
System Capacity (number of drives) ^b	≥ 1	≥ 4	≥ 4	≥ 12	> 100	> 400	> 400
^a No test procedure for this Classification is provided by the Emerald Specification. ^b There is no upper limit on the number of drives in any Classification.							

Table 3のAttribute列は次のように定義される。

- マルチホスト共有可能性は、ストレージ製品を共有可能なホストの数を識別する。
- ストレージ保護属性を持つストレージ製品は、電力喪失やストレージ・デバイスの故障が発生したときにすべての完了したIO操作が保存されることを保証する。この保証は、RAID、NVRAM、ディスク・スペアリング、バックグラウンド・ディスク・スクラビング、バックグラウンド・メディア・スキャンなどのハードウェアやソフトウェアの組み合わせによって提供できる。

- No SPOF 属性を持つストレージ製品には、故障によってシステムが操作不能になったり、データがアクセス不能になったりするコンポーネントやパスがない。
- 無停止サービス性属性を持つストレージ製品は、FRU 交換、コード・パッチ、ソフトウェア／ファームウェア更新、構成変更、データ移行、およびシステム拡張を含むすべてのサービス運用中にデータの連続可用性をサポートする。
- FBA/CKD サポート属性を持つストレージ製品は、IBM Corporation によって定義されたカウント・キー・データを提供する。

ニアオンライン・カテゴリ

ディスク・セットは、磁気ディスクを代表とする回転媒体デバイスに基づくストレージ製品で構成される。

ディスク・セット・ニアオンライン・カテゴリは、ニアオンラインのランダムアクセスの回転媒体ベースのストレージ製品について、特徴と機能を定義したものである。このプロファイル内の製品は、ブロック、ファイル、またはオブジェクト・インターフェースの任意の組み合わせだけでなく、MAID または FCAS アーキテクチャも使用する。Table 4 に、このカテゴリで定義された分類クラシフィケーション（ディスク・セット・ニアオンライン・ニアオンライン1 など）に関する要件を示す。

Table 4 - Disk Set Near-Online Classifications

Attribute	Category: Disk Set Near-Online						
	Classification: Near-Online						
	1 ^a	1.5 ^b	2	3	4 ^b	5	6
Multi-host Shareability	Not Specified		Ability to share with 2 or more hosts	Ability to share with 2 or more hosts		Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/Component	Yes		No	No		No	No
Storage Controller	Optional		Optional	Required		Required	Required
Storage Protection	Optional		Optional	Required		Required	Required
No SPOF	Optional		Optional	Optional		Required	Required
Non-Disruptive Serviceability	Optional		Optional	Optional		Required	Required
FBA/CKD Support	Optional		Optional	Optional		Optional	Required
System Capacity (number of drives) ^c	≥ 1		≥ 4	≥ 12		> 100	> 1 000
<p>^a No test procedure for this Classification is provided by the Emerald Specification.</p> <p>^b Classification is not defined; no test procedure is provided the Emerald Specification.</p> <p>^c There is no upper limit on the number of drives in any Classification.</p>							

The Green Grid

RVML セット・リムーバブル・メディア・ライブラリ・カテゴリ

このカテゴリは、自動または手動メディア・ローダー（テープ・ライブラリや光ライブラリなど）に依存するストレージ製品について、特徴と機能を定義したものである。Table 5 に、このカテゴリで定義された分類クラシフィケーションに関する要件を示す。

Table 5 – RVML Set Removable Media Library Classifications

Attribute	Category: RVML Set Removable Library						
	Classification: Removable						
	1	1.5 ^a	2	3	4 ^a	5	6
Robotics	Prohibited		Required	Required		Required	Required
No SPOF	Optional		Optional	Optional		Optional	Required
Non-disruptive Serviceability	Optional		Optional	Optional		Optional	Required
System Capacity (number of drives)	Not Specified		≤ 4	≥ 5		≥ 25	≥ 25
^a Classification is not defined; no test procedure is provided by the Emerald Specification.							

RVML セット仮想メディア・ライブラリ・カテゴリ

この操作プロファイルは、仮想メディア・ライブラリを提供するために非リムーバブル・ストレージ・メディアに依存するシーケンシャルアクセス・ストレージ製品について、特徴と機能を定義したものである。Table 6 に、このカテゴリで定義された分類クラシフィケーション（RVML セット仮想メディア・ライブラリ仮想 1 など）に関する要件を示す。

Table 6 - RVML Set Virtual Media Library Classifications

Attribute	Category: RVML Set Virtual Media Library						
	Classification: Virtual						
	1	1.5 ^a	2	3	4 ^a	5	6
Storage Protection	Optional		Optional	Required		Required	Required
No SPOF	Optional		Optional	Optional		Optional	Required
Non-Disruptive Serviceability	Optional		Optional	Optional		Optional	Required
System Capacity (number of drives)	≤ 12		> 12 ^b	> 48 ^b		> 96 ^b	> 96 ^b
^a Classification is not defined; no test procedure is provided by the Emerald Specification. ^b There is no upper limit on the number of drives.							

NVSS セット・ディスク・アクセス・カテゴリ

NVSS セットは、不揮発性ソリッド・ステート・ストレージに基づくストレージ製品で構成される。NVSS セット・ディスク・アクセス・カテゴリは、オンラインのランダムアクセスのソリッド・ステート・ストレージ・ベースのディスク・アクセス・ストレージ製品について、特徴と機能を定義したものである。ストレージ製品は、ストレージ・パラダイムを使用したデータ・アクセス（つまり、開く、閉じる、読み取る、および書き込む）を提供している場合にディスク・アクセスを提供する。

このプロファイル内の製品は、ブロック、ファイル、またはオブジェクト・インターフェイスの任意の組み合わせを提供できる。

Table 7 に、このカテゴリで定義された分類クラシフィケーション（NVSS セット・ディスク・アクセ

Attribute	Category: NVSS Set Disk Access Classification: Online						
	1 ^a	1.5	2	3	4	5	6
Multi-host Shareability	Not Specified	Ability to share with 1 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/Component	Yes	No	No	No	No	No	No
Storage Controller	Optional	Optional	Required	Required	Required	Required	Required
Storage Protection	Optional	Not integrated	Required	Required	Required	Required	Required
No SPOF	Optional	Optional	Optional	Optional	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Optional	Required	Required
FBA/CKD Support	Optional	Not integrated	Optional	Optional	Optional	Optional	Required
System Capacity	consumer	JBOD	very small	small	medium	large	mainframe
^a No test procedure for this Classification is provided by the Emerald Specification.							

ス・オンライン1 など) に関する要件を示す。

Attribute	Category: NVSS Set Disk Access Classification: Online						
	1 ^a	1.5	2	3	4	5	6
Multi-host Shareability	Not Specified	Ability to share with 1 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/Component	Yes	No	No	No	No	No	No
Storage Controller	Optional	Optional	Required	Required	Required	Required	Required
Storage Protection	Optional	Not integrated	Required	Required	Required	Required	Required

No SPOF	Optional	Optional	Optional	Optional	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Optional	Required	Required
FBA/CKD Support	Optional	Not integrated	Optional	Optional	Optional	Optional	Required
System Capacity	consumer	JBOD	very small	small	medium	large	mainframe
^b No test procedure for this Classification is provided by the Emerald Specification.							

Table 7 - NVSS Set Disk Access Classifications

NVSS セット・メモリ・アクセス・カテゴリ

NVSS セットは、不揮発性ソリッド・ステート・ストレージに基づくストレージ製品で構成される。NVSS セット・メモリ・アクセス・カテゴリは、オンラインのランダムアクセスのソリッド・ステート・ストレージ・ベースのメモリ・アクセス・ストレージ製品について、特徴と機能を定義したものである。ストレージ製品は、メモリ・プリミティブを使用したストレージへのホスト・アクセス（読み込みや保存など）を提供している場合にメモリ・アクセスを提供する。

このプロファイル内の製品は、メモリ・インターフェイスを提供する。Table 8 に、このカテゴリ内の分類クラシフィケーションに関して可能な要件を示す。

Table 8 - NVSS Set Memory Access Classifications

Attribute	Category: NVSS Set Memory Access						
	Classification: Online						
	1 ^a	1.5 ^a	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a
Consumer/Component	Yes	No	No	No	No	No	No
Storage Protection	Optional	TBD	Optional	Required	Required	Required	Required
No SPOF	Optional	TBD	Optional	Optional	Required	Required	Required
Non-Disruptive Serviceability	Optional	TBD	Optional	Optional	Optional	Required	Required
System Capacity	consumer	JBOD	very small	small	medium	large	mainframe
^a No test procedure for this Classification is provided by the Emerald Specification.							

加えて、ストレージ業界は、新しいタイプのストレージ・アーキテクチャと製品を急速に進化させ、市場に投入している。対象となる製品領域には、オブジェクト・ストア、ビッグ・エンタープライズ仮想化ストレージ・サーバ、およびコンバージド、ハイパーコンバージド、コンポーザブル、およびソフトウェア定義ストレージが含まれる。ストレージ製品とサーバ製品との境界の定義を正確に記述するのが困難になっており、ソフトウェア定義の実装が注目されていることから、これらを解決するためには、業界の IT パートナー全体での共同の分類の取り組みが必要である。

4. ストレージの物理属性

ストレージ・システムは数種類のメディア、コントローラ、および筐体で構成される。ストレージ・デバイスには、ハード・ディスク・ドライブ（HDD）、ソリッド・ステート・ドライブ（SSD）（PCIe や NVMe のバス方式も含む）、非ドライブ形式のソリッド・ステート・メディア、テープ、および光メディアなどがある。光メディアとテープは、主に、バックアップやアーカイブに使用される。ストレージ・コントローラまたはプロセッサは、物理ストレージ・デバイスを管理し、ハードウェア RAID を実装し、フロントエンドのコンピュータ HBA（ホスト・バス・アダプタ）とバックエンドで制御されるデバイスにインターフェイスを提供する。ドライブ、コントローラ、I/O モジュール、ミッドプレーン、電源、および冷却ファンが物理筐体に内蔵されている。このような筐体は、単一の統合筐体にすることも、コントローラ筐体とドライブ筐体に分けることもできる。単一障害点（SPOF）がないという要件を満たすために、このような筐体は冗長なハードウェア・モジュールを備えることになる。ディスクが内蔵された一般的なストレージ筐体を Figure 9 に示す。

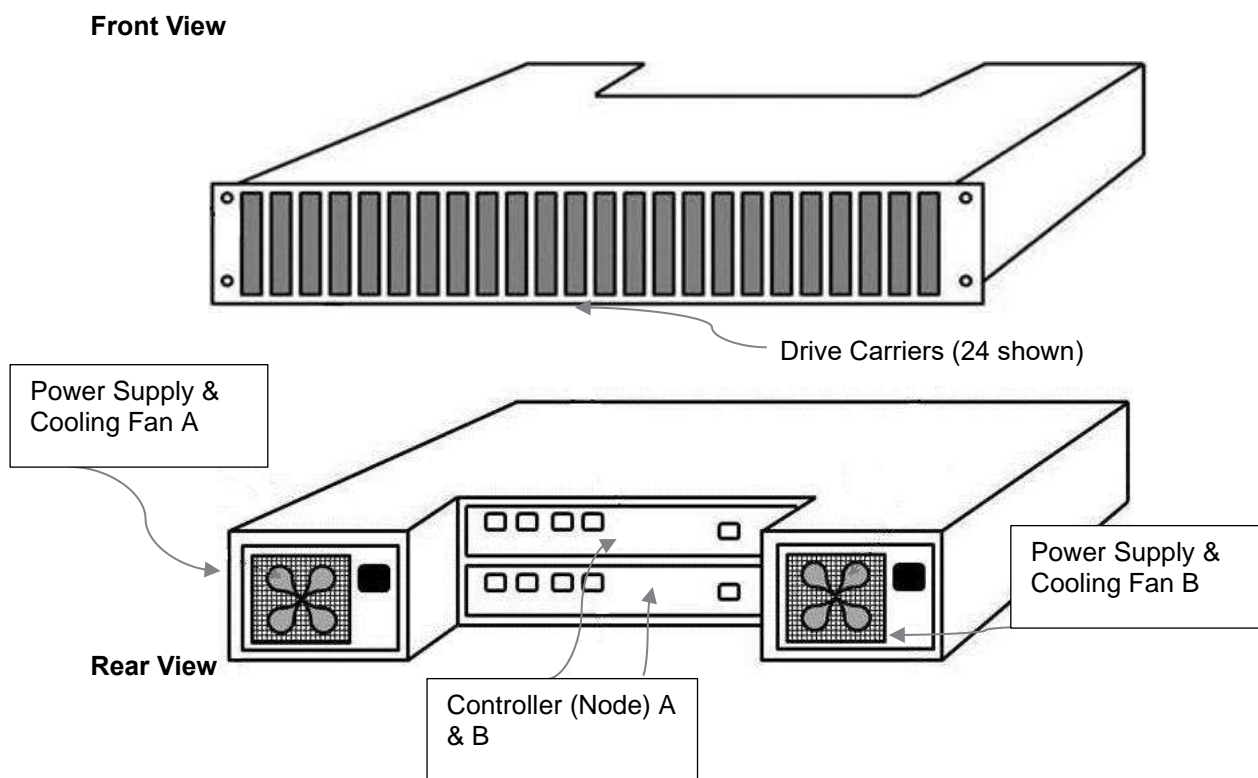


Figure 9: Typical Storage Enclosure: Entry Level, Expand Out

ストレージ・システムのエネルギー消費は、大まかに、ストレージ・メディアの特性（タイプ（SSD または HDD）、ストレージ容量、個別のデバイスの電力消費、およびデバイス数）によって決定される。その他の要因は、RAID レベルの選択、RAS（信頼性、可用性、および保守容易性）機能、使用中の容量最適化手法（COM）、およびストレージ・コントローラの性能である。一般的に、ストレージ・システムの能力が高いほど（ローエンドではなくハイエンド）、追加の機能をサポートするために必要な電力が高くなる。

エンタープライズ・ストレージ製品に使用されているハード・ドライブの中では、2.5" ドライブの平均電力需要は約 6W（7.2K RPM）/ドライブで、10K RPM と 15K RPM ではそれぞれ 7.5W と 8.3W と電力需要が少し高くなる。3.5" ドライブの平均電力需要は 7.2K RPM ドライブで約 10W である。

SATA ソリッド・ステート・ドライブと SAS ソリッド・ステート・ドライブの平均電力需要は 7～15W である。NVMe ソリッド・ステート・ドライブの電力要件ははるかに高い。電力は、容量に応じて変化し、25W を超える可能性がある。

多くのストレージ・ベンダーから机上電力シミュレーターを入手できる。これを使用すれば、各ストレージ筐体の電力寄与を決定し、様々な構成の合計を計算することができる。

コントローラ・モジュールは性能によって一般的に 50W～1000W を消費し、ドライブ筐体 I/O モジュールは 20W～50W を消費する。24 台の 2.5"スピニング・ドライブを収容する標準的な 2U ドライブ筐体は 200W～400W を消費する可能性がある。加えて、4U あたり 100 台を超える 3.5"ドライブまたは 3U あたり 120 台を超える 2.5"ドライブを収容可能な高密度筐体がある。これらの筐体 1 つの電力は 1500W を超える可能性がある。ドライブ筐体が 3～4 台を超える大規模なストレージ構成は、ドライブの電力が大部分を占める。大規模なストレージ構成におけるドライブの電力消費は、通常は合計電力消費の 60%以上になる。

5. ストレージの短期的進化

ストレージ・プロバイダは、特定の製品の設置面積に対して保存可能なデータの量を増やし、ストレージ製品のワットあたりの性能を向上させ、データ・センター全体の電力消費を減らすために、ハードウェア・ベースとソフトウェア・ベースのイノベーションの研究、開発、および提供を継続している。ストレージ・デバイス（メディア）とストレージ・システム（特にデータ・センターに実装されるもの）の両方のレベルで改善が進められている。

ストレージ・メディア

ストレージ・メディアは、ハード・ディスク・ドライブ（HDD）とソリッド・ステート・ドライブ（SSD）の両方の継続的技術開発を反映し続けている。この両方のタイプのメディアは、今後少なくとも5~10年間はデータ・センター内に展開され続ける。一部のデータ・センター・ストレージ製品ベンダーは1種類のメディアしか使用しないが、他の多くのベンダーは運用要件、容量要件、およびビジネス要件をより良く満たすためのメディア・タイプの組み合わせを顧客に提供ようになる。このセクションでは、主な違いと一部の類似点を中心にして、両方のメディア・タイプの特徴について説明する。その違いにより、場合によってはストレージにSSDと同様にHDDを使用し続ける必要があるかもしれない。

ユース・ケースの比較

多くのデータ・センターのユース・ケースでは、SSDの採用が今後も増え続ける。従来のデータ・センター内のアプリケーションの多くが高い性能を要求する。アプリケーションによっては、一部の機能で高い性能が要求される。このような性能制約にさらされているユーザはHDDの代わりにSSDを採用する可能性がある。この状況では、SSDの使用は、大規模システム内の1つのストレージ階層または専用のストレージ・アプライアンスの使用に限定される。このクラスのアプリケーションでは、生産性/経費の観点からSSDへの投資はビジネスに有益である。

一部のクラウド・データ・センターは、データがインターネット経由で配信され、そこで遅延が生じて性能全体が制限されるため、性能は重視されない。

他のデータ・センターは、大量のデータをオンラインで維持しなければならないという別の基本的問題に直面している。このような状況では、適正な性能レベルでの容量が性能そのものよりも重要なため、高容量HDDの展開が続く可能性がある。

HDDに依存し続ける可能性が高いSME（中小企業）とエンタープライズ・データ・センターのもう1つのユース・ケースは、信頼性を確保するためのデータのバックアップである。ほとんどの企業が、複製から高度なバックアップ・アプライアンスまで、何らかの形態のデジタル・バックアップを使用している。これは、データが事故や装置の故障で破損しないことを保証するためである。これは、大容量が必要な別の状況である。一般的にバックアップ・ソフトウェアは変化した項目のみを保存するように管理するため、性能はそれほど大きな問題にはならないが、バックアップは安価に抑える必要があるオーバーヘッドのコストであるため、非常に大きな問題になる。

メディア開発の方向性

前述したように、SSDとHDDのどちらも今後展開されるユース・ケースは十分にある。そのため、両方の分野で研究と開発が継続されている。本書の後半のセクションで、両方の分野の可能性の高い開発について見ていく。

ディスク・ドライブ・サプライヤは、主要顧客である、ドライブをサーバやストレージ製品に統合するシステム・ベンダーによって定義された設計パラメータ内に留まるために各ドライブによって消費される電力を維持または削減する努力を強いられている。この制約の中で、サプライヤは次の点に注目している。

- 容量を増やす。限られた電力予算の中でこれを実現するという事は、プラッタをドライブに追加できないことを意味する。この領域の研究は、主に、受け入れ可能な信頼性のレベルで記録密度を上げるための新しいまたは改良された記録技術に焦点が当てられる。
- ドライブの機械部品の変更を試すことによってドライブのエネルギー効率を高める。密封され、ヘリウムで満たされたドライブがこの進行方向の1つの例である。

HDDストレージ・メディア

エンタープライズ・システムのHDD（ハード・ディスク・ドライブ）は高性能と容量の2つのタイプに分けられる。高性能ディスク・ドライブは、SASとファイバ・チャネルのどちらかの高性能データ・インターフェイスを備えたディスク・ドライブのことである。このドライブの回転速度は10K RPMまたは15K RPMと高い（この市場は大幅に縮小している）。また、比較的密度（データ・ストレージ容量）は低い。容量ディスク・ドライブは、高性能ディスク・ドライブより高いデータ容量を備えたディスク・ドライブのことである。このドライブは、性能の低いデータ・インターフェイス（通常はSATA）を備え、回転速度も遅い（通常は7.2K RPM以下）。

多くの企業が従来のディスク・ドライブ技術の物理的な制限を克服してストレージ容量を増やす技術の開発を進めている。

ヘリウム充填ディスク・ドライブ

ヘリウム充填ディスク・ドライブでは、プラッタが回転してヘッドが移動する環境としてヘリウムが使用される。この低密度・低粘度のガスを使用することで、プラッタをより密接に配置して特定のフォーム・ファクターのドライブの容量を増やすことができる。また、プラッタを回転させるために必要な電力が減るため、使用するモーターを小型化し、ドライブを操作するために必要な電力を減らすこともできる。

熱アシスト磁気記録ドライブ

熱アシスト磁気記録では、読み書きヘッド内で生成された光エネルギーまたはマイクロ波エネルギーを使用して、書き込まれるディスク領域の周辺を温め、その領域の磁気特性を一時的に変化させることで、書き込みやすくしている。この技術では、使用する磁気材料をより高いデータ密度をサポートするように改良することができる。

瓦磁気記録ドライブ

瓦磁気記録（SMR）は、ディスク・プラッタ上の隣接するトラックを部分的にオーバーレイさせてトラック密度を増やすことで、ディスク・プラッタとディスク・ドライブの容量を増やす磁気ディスクの記録技術である。この記録技術の名称は、瓦が屋根の上に並べられるようにトラックをオーバーラップさせることに由来する。隣接するトラック上のデータを単独で書き込むことができないため、書き込みプロセスは複雑になる。様々な管理技術によって、SMRディスクにデータを書き込む際にかかる余分な時間が短縮されている。

デュアルアクチュエータ・ドライブ

一部のドライブ・メーカーは、性能を向上させるために単一のドライブ内に2個のアクチュエータを備えたディスク・ドライブを導入している。これは、特に、高容量ドライブでGBあたりの性能を維持する場合にメリットがある。

SSDストレージ・メディア

ソリッド・ステート・ドライブは、ソリッド・ステート・メモリ技術に基づくストレージ・デバイスである。

SSD技術は、次のような様々なフォーム・ファクター、インターフェイス、およびプロトコルで利用できる。

The Green Grid

- HDDのプラグイン交換（通常はインターフェイスとしてSASまたはSATAをサポート）
- HDDのフォーム・ファクターを有するが、インターフェイスとしてNVMeをサポートするデバイス
- SSDに最適化されたフォーム・ファクター（U2やM2など）を有し、インターフェイスとしてNVMeをサポートすることが多いデバイス

NVMe（Non-Volatile Memory Express）は、高い性能を発揮するPCIeベースのSSDのコマンド・セットと機能セットを定義する標準的なインターフェイスである。M2やU2などの一部のフォーム・ファクターがNVMe SSD用に展開されている。このようなフォーム・ファクターの一部は小型で、そのすべてで従来のSATAデバイスやSASデバイスとは異なるコネクタが必要である。

SSDは、同等のコストで優れた性能を発揮するため、10K RPMと15K RPMの磁気ディスク・ドライブを置き換えている。

メモリ・アクセスSSD技術が現在開発中である。

SSDは、シリコン技術の改良から生まれたオンライン・ストレージ・メディア（便利なデータ移動手段や不揮発性キャッシュではなく）として登場した。SSDは、次のような特徴の結果として、HDDと比べて様々なメリットを提供する。

- **絶対性能：**SSDストレージはすべてのHDDより性能が優れており、大幅に上昇した小規模I/O操作数/秒と高速化したストリーミング・データ・レートの両方を提供する。
- **消費電力：**SASまたはSATAインターフェイスをサポートする中小容量のSSDは、必要な電力が高性能HDDよりはるかに少なく、通常は、7.2K RPM HDDの消費電力のオーダーである。これは、特定の容量の高性能HDDストレージを同等の容量のSSDストレージに置き換えることにより、エネルギー消費が下がることを意味する。
- **振動：**SSDは、可動部分がないため、ストレージ・システムに機械的振動を伝達しない。
- **フォーム・ファクター：**磁気ディスク・ドライブの円形プラッタの制約を受けないフォーム・ファクターによって、エンタープライズ・システムとクライアント・システムの両方でより小さいフォーム・ファクターが可能になる。
- **重量：**中小容量のSSDは、同等の容量のHDDより軽量である。このメリットは、多くの場合、非常に大きい容量では維持されないが、家電製品からラップトップまでのポータブル・デバイスに信頼できる軽量ストレージを提供する点において非常に重要であることが証明されている。

SSDの半導体メモリ・セルには、各セルで処理する書き込み操作の回数に寿命による制限がある。そのため、SSDには通常、デバイスの早期老化を避けるための「ウェアレベリング」アルゴリズムが組み込まれている。

ソリッド・ステート技術の進歩は、以下の点に主に集中している。

- 性能の向上
- パッケージ・サイズ（ボリューム）あたりの容量の増加。これにより、複数の「ディスク相当物」またはシステム・スロットを使用せずにシステム全体の容量が増えるため、より狭い設置面積でより大きなストレージ・システムをサポートできる。
- GBあたりのコストの削減

まとめ

このセクションで説明した特徴を組み合わせると、SSDとHDDの両方が今後何年間も多くのデータ・センター環境で相補的役割を担って共存し続けるという結論になる。

新たなストレージ・トレンド

アプリケーションがアプライアンスに取って代わる

ここ数年の間に、データ・センター業界は新しい「装置」のトレンドを経験した。高度なアプリケーション・ソフトウェア製品は、ストレージ・システムやネットワーク・スイッチなどの従来のデータ・センター向け専用装置をエミュレートするまでに進化した。文献や論文では、このような製品カテゴリは、ソフトウェア定義ネットワーク（SDN）、ソフトウェア定義ストレージ（SDS）、コンバージド、ハイパーコンバージド、ストレージ・サーバなどという名前で呼ばれている。ここでは、単純にSDNやSDSと呼ぶことにする。このようなアプリケーション・カテゴリの両方が初期採用者の間で大幅な市場シェアの成長を経験しており、両方がメインストリーム・データ・センターでの存在感を強めている。

このようなカテゴリの1つであるSDSを詳しく調べると、運用ソフトウェア製品はインストール先のストレージとサーバのブランドに依存しないように設計されていることが分かった。代わりに、このような製品では、他のアプリケーションと同じように、適切な実行環境についてのガイドラインの類が存在する。つまり、命令セット、オペレーティング・システム、使用可能なメモリ、ソリューションのサポートに必要なストレージ領域などの要件が列挙される。このようなアプリケーションは、それが使用するドライブとともに、どのようなサーバにも自由にインストールすることができ、そのサーバは、既に別のアプリケーションを実行しているものでもよい。事実、SDSまたはSDN製品を使用する主な目的は専用装置の購入を避けることである。SDSアプリケーションやSDNアプリケーションは、通常、SDSを使用してデータを保存する同じアプリケーションとサーバを共有する仮想化されたサーバ環境で動作する。

SDSまたはSDNのエネルギー効率と性能は、ソフトウェアの動作とハードウェアの動作という2つの独立した変数によって決められる。このことはストレージ専用装置またはネットワーク専用装置に表面的には当てはまるが、基本的な違いがある。例としてストレージを使用すると、特定のストレージ専用装置は実際には特定のコントローラ設計だったり、専用のI/Oインフラストラクチャーだったり、特定のストレージ・アプリケーションからなる。これをテストした場合は、同じ動作が繰り返し行われることが期待できる。SDS環境では、エネルギー効率と性能がストレージ・アプリケーション、特定のサーバ構成（すべてのアプリケーションに必要なメモリやI/Oコントローラなどを含む）、およびサーバ上で使用中の他のアプリケーションに左右される。

2つ目の結果として、このようなSDSまたはSDNアプリケーションのどちらかを実行中のサーバには、データ・センター内の専用の計算サーバとは大きく異なる構成が必要になる。少なくとも、特定の専用装置の代用として機能する適切な数の高性能I/Oデバイスをシステムに組み込む必要がある。最後に、SDSアプリケーションには、追加ストレージ容量とその追加ストレージ容量をサポートするための追加のコンポーネントが必要になる。

ハイブリッド・クラウドの導入

クラウドストレージのセクションで既に説明したように、クラウドストレージは、パブリック、プライベート、またはハイブリッド展開が可能である。全員が全データをパブリック・クラウドに移動させる準備ができていくわけではなく、ハイブリッド・クラウド（マルチクラウド）戦略を導入する企業が増えている。これは、組織がコスト効率性を獲得し、データの完全性を損なうことなく、パブリック・クラウドとプライベート・クラウド間のデータ移動性を向上させるためのデータ・ストレージ戦略である。ハイブリッド・クラウド戦略を使用している組織は、オンプレミス上でもオフプレミス上でも、データを柔軟に収集、分離、および保存することができる。これにより、市場投入までの時間が短縮され、環境の展開と管理が容易になる。ハイブリッド・クラウドストレージ・ソリューションは、一般的に、ユーザに様々なストレージ・オプションを利用可能にしている。これには、ファイル・ストレージとファイル共有、オブジェクト・ストレージ、ブロック・ストレージ、およびバックアップ・システムが含まれる。ほとんどの企業は、自社の作業のうち、セキュリティやアクセ

ス時間が重要でないいくつかの側面にパブリック・クラウドを使用できる。ただし、このような企業も、より機密性の高いデータやより速いアクセス時間が必要なデータを保存するためにプライベート・クラウド・サービスが必要である。つまり、ハイブリッド・クラウド・ソリューションが必要である。

サービスとしてのストレージ (SaaS)

SaaSは、プロバイダがサブスクリプションを通してストレージ・リソースを顧客に貸し出すデータ・ストレージ・ビジネス・モデルである。SaaSは、運営費用 (OpEx) の機動性によってコストを削減するものであり、利用者は、必要な時に必要なストレージに対してのみ料金を支払う。SaaSモデルでは、長期間に及ぶバルク・データ・ストレージの複雑な側面 (ハードウェア・コスト、セキュリティ、およびデータの完全性) のほとんどをストレージ・プロバイダが処理する。サービスは、単一顧客専用のインフラストラクチャーからオンプレミスで提供されることも、サブスクリプションによって購入され、1つ以上の利用状況の指標に基づいて料金が請求される共有サービスとしてパブリック・クラウドから提供されることもある。SaaSを使用する組織は、オンプレミスでデータを保存する代わりに、一般的にストレージ・ニーズのためにパブリック・クラウドを利用する。パブリック・クラウドストレージは、SaaSに様々なストレージ方式を使用することもできる。SaaSの顧客は、標準的なシステム・インターフェイス・プロトコルやアプリケーション・プログラム・インターフェイス (API) を通して個別のストレージ・サービスにアクセスする。このようなストレージ方式には、バックアップとリストア、災害復旧、ブロック・ストレージ、ファイル・ストレージ、オブジェクト・ストレージ、およびバルク・データ転送が含まれる。

6. 容量最適化手法 (COM)

容量最適化は、特定のストレージ・デバイスの設置面積に対してより大きなデータ・ストレージ容量を可能にするためにストレージ製品内で使用される一連のソフトウェア・テクニックを意味する一般用語である。保存データ量を減らすことで、必要なドライブ容量を減らすことができ、結果的に特定の操作のエネルギー消費が減る。このようなテクニックのそれぞれが容量最適化手法 (COM) と呼ばれている。一般的な COM の説明を以下に示す。

- **高度なデータ保護：** (パリティ RAID とイレイジャーエンコーディングを含む)
これらは、フル・コピーを使用しないデータ保護の手法である。ミラーリングされたコピーまたはフル・コピーを高度なデータ保護に置き換えた場合の予想スペース節約は 25~50% である。
- **シン・プロビジョニング：**
プロビジョニング時にすべての物理容量を割り当てるのではなく、アプリケーションがデータを書き込む時に物理容量を割り当てる技術。
- **デルタ・スナップショット：**
複数のフル・コピーを使用せずに、ファイルの複数のバージョンを再構築するタイプのポイントインタイム・コピー。変更が小規模であれば、大きなスペース節約が可能になる。
- **圧縮：**
サイズを削減するためのデータ・エンコーディングのプロセス。通常は、2 : 1 の圧縮率が達成される。
- **データ重複排除：**
様々な粒度レベルで、データの重複を共有コピーへの参照に置き換えること。これは大きなスペース節約が得られる。

概して、COM は、概ね相互に独立しているが、完全に独立しているのではない。COM は、任意の組み合わせでメリットを提供するが、複合効果は個別のメリットの合計と一致せず、単一のストレージ製品に 3 つ以上の COM を展開した場合は非常に効率が悪くなる。

ENERGY STAR データ・センター・ストレージ V2.1 は、SNIA Emerald テスト中に適格なオンライン製品カテゴリで高度なデータ保護を有効にするよう求めている。ディスク・セット・アクセス・ストレージ製品カテゴリと NVSS ディスク・セット・アクセス・ストレージ製品カテゴリでは、高度なデータ保護が必要である。加えて、選択可能な機能として使用可能にする必要がある COM (高度なデータ保護以外に上で示したもの) の数を以下に示す。

- **オンライン 2：** 使用可能にする必要がある COM の数は 1 である。
- **オンライン 3：** 使用可能にする必要がある COM の数は 2 である。
- **オンライン 4：** 使用可能にする必要がある COM の数は 3 である。

テスト検証では、特定の COM の存在と有効性を確認する検出手法が適用される。テスターは、すべてのアクティブ測定テストやアイドル測定テスト中に無効にすることが可能なすべてのオプション COM を無効にする必要もある。

COMのメリットの定量化

COMの簡易存在証明テストを使用するよりも、COMを有効にしたアイドル指標とアクティブ指標に対する影響を把握するための定量化可能な手法を見出す方が有益かもしれない。アクティブCOMのメリットの見積もりはいくつか存在するが、アクティブ操作中の効率を正確に考慮する有効な手法を確立するにはもっと多くの特性評価テストが必要である。テストによってアクティブCOMの影響を比較する方法の例を以下にいくつか示す。Table 9は物理容量削減（つまり、論理容量は一定に保つが、エネルギー消費を削減するために物理ドライブを取り外す）に基づいており、Table 10は使用可能容量最適化（つまり、物理ストレージの実効仮想容量の増加を定量化する）に基づいている。「Delta (GB/W)」行と「Delta (IO/W)」行は、アイドル状態指標とアクティブ（この例ではランダム・ワークロード）指標の正味変化（プラスまたはマイナス）を示している。すべてのケースにおいて、特に、使用可能容量手法の場合に、アイドル指標の値に正味での向上が見られることに注意されたい。アクティブ指標の結果にはばらつきが多い。物理容量削減（実際のドライブ数が削減されるため）の場合は、アクティブ性能および指標が低下（2ケース）または向上（1ケース）している。すべての使用可能容量ケースにおいて、アクティブ性能指標は同じか向上している。

Table 9 - Active COM Examples (70/30 R/W Random Workload), Based on Physical Capacity Reduction⁴

COM Type	Thin Provisioning	TP + Data Dedup	Data Compression
Baseline Config	Raid 5 (3+1 Data + Parity)	Full Provisioned No Dedup	Uncompressed
Baseline Capacity (GB)	54,000 (160 – 450 GB disks)	54,000 (160 – 450 GB disks)	36,000 (160 – 450 GB disks)
Baseline Power (W)	2360(I) 3134(B)	2360(I) 3134(B)	2360(I) 3134(B)
Baseline (GB/W)	22.88	22.88	15.25
Baseline Perf. (IOPs)	25,250	25,250	36,010
Baseline (IO/W)	8.06	8.06	11.49
Optimized Config	Thin Provisioning (50% written)	TP+Deduped Patterns (2:1)	Compressed (2:1)
Optimized Capacity (GB)	54,000 (50% written)	54,000 (50% written-dedup 2:1)	36,000 (80 – 450 GB disks)
Opt. Power (W)	1464(I) 2030(B) (80 disks)	1016(I) 1478(B) (40 disks)	1464(I) 2030(B)
Optimized (GB/W)	36.88	53.15	24.59
Delta (GB/W)	+61%	+132%	+61%
Opt. Perf (IOPs)	13,560	8,580 (2:1 Dedup Ratio)	38,280
Optimized (IO/W)	6.68	5.81	18.86
Delta (IO/W)	(17)%	(28%)	+64%

⁴ 出典：『Software Defined Storage Energy Efficiency Features』、The Green Grid Forum 2016

Table 10 - Active COM Examples (70/30 R/W, Random Workload) Based on Usable Capacity Optimizations⁵

COM Type	Thin Provisioning	TP + Data Dedup	Data Compression
Baseline Config	Raid 5 (3+1 Data + Parity)	Full Provisioned No Dedup	Uncompressed
Baseline Capacity (GB)	54,000 (160 – 450 GB disks)	54,000 (160 – 450 GB disks)	36,000 (160 – 450 GB disks)
Baseline Power (W)	2360(I) 3134(B)	2360(I) 3134(B)	2360(I) 3134(B)
Baseline (GB/W)	22.88	22.88	15.25
Baseline Perf. (IOPs)	25,250	25,250	36,010
Baseline (IO/W)	8.06	8.06	11.49
Optimized Config	Thin Provisioning (50% written)	TP+Deduped Patterns (2:1)	Compressed (2:1)
Optimized Capacity (GB)	108,000	216,000 (50% written, 2:1)	72,000 (2:1 Comp Ratio)
Opt. Power (W)	2360(I) 3134(B)	2360(I) 3134(B)	2360(I) 3134(B)
Optimized (GB/W)	45.76	91.52	30.50
Delta (GB/W)	+100%	+300%	+100%
Opt. Perf (IOPs)	25,250	32,825 (2:1 Dedup Ratio)	72,020 (2:1 Comp Ratio)
Optimized (IO/W)	8.06	10.47	22.98
Delta (IO/W)	0%	+30%	+100%

物理容量削減手法を使用した場合は、ドライブ数が減ることによる電力節約を容易に特定できる。ただし、存在する物理ドライブが減るため、性能指標（IOPs など）とアクティブ指標が下がる可能性がある。加えて、実際のストレージ・システムでは、運用中ではなく、初期購入時またはセットアップ中のみ物理ドライブ削減が行われる。使用可能（または仮想）容量最適化手法を使用した場合は、実際の物理容量が事実上増加する。ただし、SNIA Emerald™仕様では、使用可能容量の改善ではなく、アイドル状態指標の計算時の分子として物理容量が使用されるため、このCOMのメリットは反映されない。アイドル状態指標とアクティブ指標の両方がCOMのメリットを評価するための手法の選択の影響を受けることに注意されたい。アイドル状態指標は、使用可能または仮想容量手法によって大幅に向上する。ただし、物理ドライブ数の削減が優先された場合は、アクティブ性能指標が低下する可能性がある。COMのメリットを比較するための推奨される定量化可能な手法を導出するにはさらなる調査が必要である。

複数のCOMの同時運用の実現可能性とメリットは、実装と構成に依存する。COM機能の運用を可能にするアレイ・リソースによっては、最小限の相互作用が生じる場合がある。例えば、コントローラ経由の重複排除、IOスタックの「バックエンド」上にあるSSD内で提供される圧縮機能、およびシン・プロビジョニング機能を有効にする製品を考へてみる。これらのそれぞれが、ストレージ・サブシステムによって提供される全体的な使用可能（論理）容量に貢献できる。理想的なケースでは、圧縮率とシン・オーバープロビジョニング率の両方が2:1であれば、アプリケーションは論理空間全体を（あたかも物理的に存在しているかのように）利用することができる。重複排除は、圧縮では不可能なより大きなデータの「チャンク」を重複排除することにより圧縮率を高めることができる。最後に、提出者が有効にするCOMの数と選択は提出者次第である。それらのいずれかが性能上の問題や干渉を引き起こす場合は、それらを評価に含めない選択がなされるかもしれない。

連続運転（24時間無休）が不可欠のエンタープライズ・データ・センター環境では、アクティブ指標がアイドル電力指標よりはるかに重視され、COMは最適なメリットを提供するように選択的に適用する必要がある。

⁵ 出典：『Software Defined Storage Energy Efficiency Features』、The Green Grid Forum 2016

7. ストレージに関する SNIA EMERALD データが示すもの

データ分析に使用された SNIA Emerald テスト・データは 2022 年 3 月にまとめられた。200 製品以上のデータがあるが、データが同一になる OEM システムのデータを削除した後の 160 製品のブロック I/O システムからなる。ファイル・アクセス・システムに使用可能なデータは非常に限られていたため、そのデータはこの分析に含まれていない。36 個のオンライン 2 システム、55 個のオンライン 3 システム、および 69 個のオンライン 4 システムが存在する。データはさらに、それぞれの構成が最適化されたワークロード・タイプ（トランザクションまたはシーケンシャル）に区分できる（Table 11）。

Table 11 - ENERGY STAR configurations of data presented in this paper

System type	Number of Systems					
	Total	7.2K RPM	10K RPM	15K RPM	Mixed RPM	SSD
Online 2 Transactional	15	1				14
Online 3 Transactional	33	4	18	10		1
Online 4 Transactional	41	4	8	9	2	18
Online 2 Sequential	21	10	10	1		
Online 3 Sequential	22	7	14	1		
Online 4 Sequential	28	4	12	12		

ストレージ・システムの性能は、コントローラ、コントローラ・キャッシュ、フロントエンドとバックエンドの相互接続、ストレージ・メディアの容量、およびストレージ・メディアのドライブ速度（該当する場合）に依存する。以降のセクションでは、ワットあたりのトランザクション性能、ワットあたりのシーケンシャル性能、および分類、ドライブ・タイプ、およびドライブ速度が値に与える影響について見ていく。

トランザクション・ワークロード用に最適化されたブロック・ベースのシステムからのデータ

ブロック・ストレージ・システムのトランザクション・ワークロードは、SNIA Emerald のホット・バンド指標が特性値となる。Figure 10 と Figure 11 は、ホット・バンド性能とアイドル電力指標に相関関係がないことを示している。Spec 1 と Spec 2 という用語は、Energy Star for Data Storage 仕様の 1.0 と 2.0 を意味する。

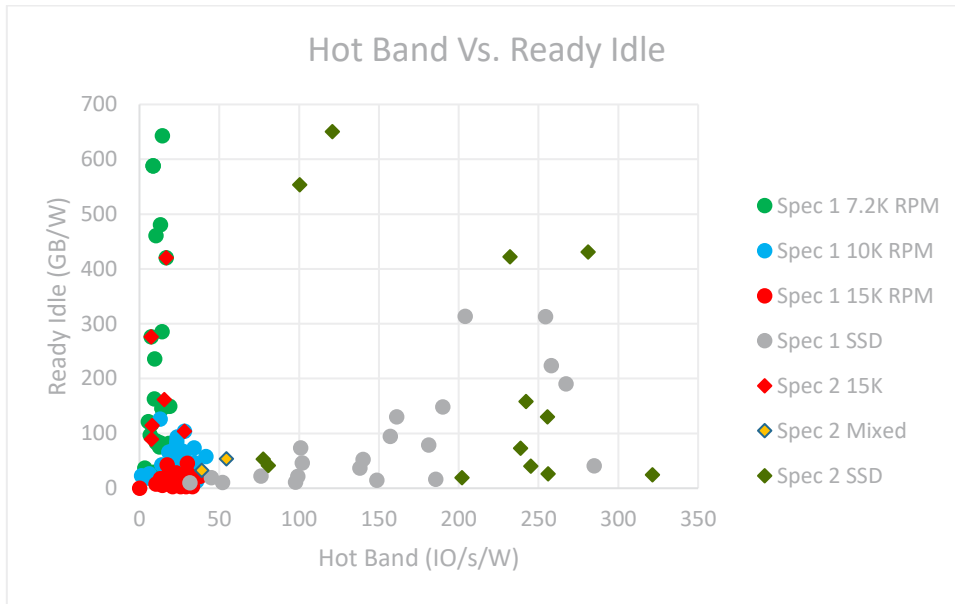


Figure 10: Hot Band vs. Ready Idle

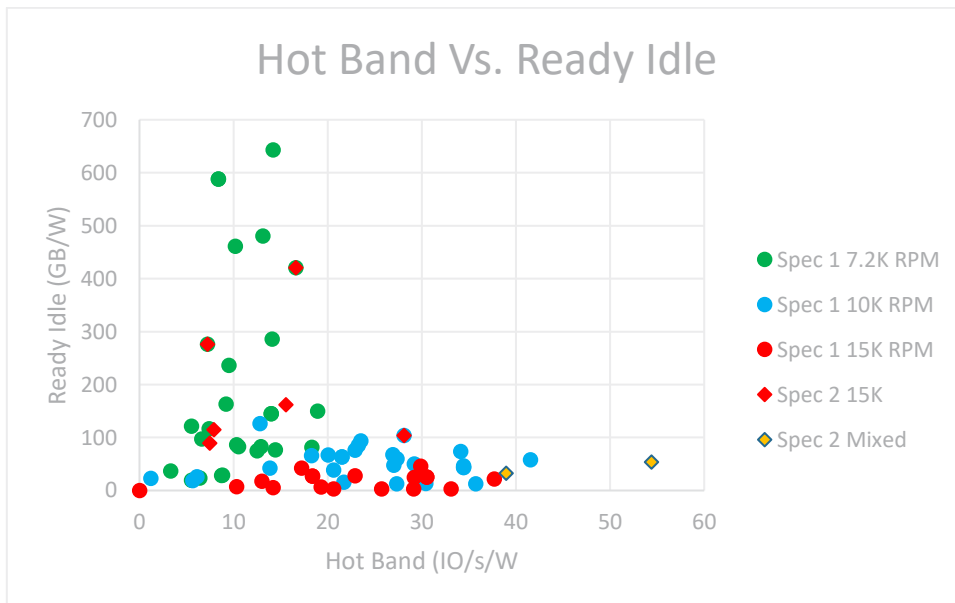


Figure 11: Hot Band vs. Ready Idle (Expanded Scale)

Figure 12 と Figure 13 は、ドライブ速度と SNIA Emerald のホット・バンド指標の関係を示している。これらの図には、参考として、最新の Energy Star for Data Center Storage バージョン 2.1 の限界値⁶が含まれている。この要件について、この限界値を超える 7.2K RPM ドライブは存在しない。

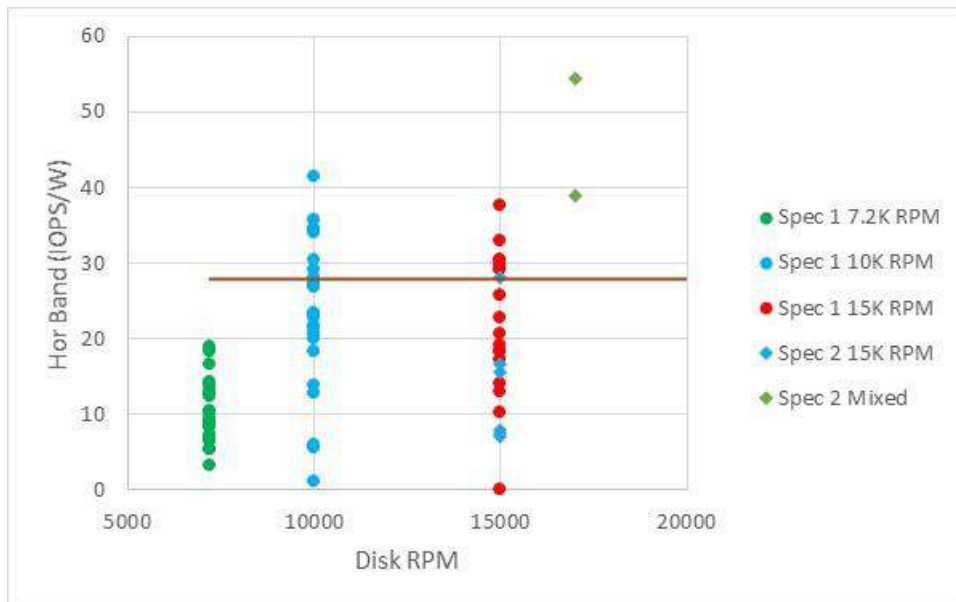


Figure 12: Got Band Metric vs. Drive Speed (HDDs)

Figure 13 は、SSD ドライブ・データを追加したものである。これは 2 つの事実を示している。ソリッド・ステート・ドライブ・システムはスピニング・メディア・ドライブより性能が大幅に高い。グラフは、すべての SSD ドライブ・システムが Energy Star Data Center Storage V2.1 の限界値⁶を超えていることも示している。

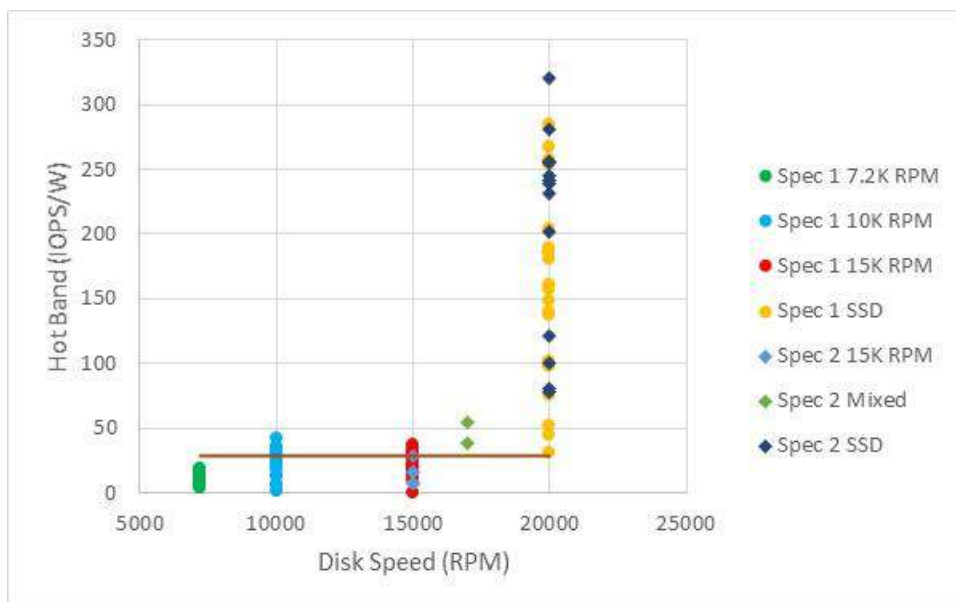


Figure 13: Hot Band Metric vs. Drive Speed (HDDs and SSDs)

⁶ ENERGY STAR Data Center Storageバージョン2.1最終仕様、第3.3.1項。

シーケンシャル・ワークロード用に最適化されたブロック・ベースのシステムからのデータ

シーケンシャル・データ用に最適化されたブロック・ベースのシステムは、SNIA Emerald のシーケンシャル読み取りワークロードとシーケンシャル書き込みワークロードが特性値となる。

Figure 14: Sequential Read vs. Ready Idle Metrics は、シーケンシャル読み取り電力指標とアイドル状態電力指標に相関関係がないことを示している。

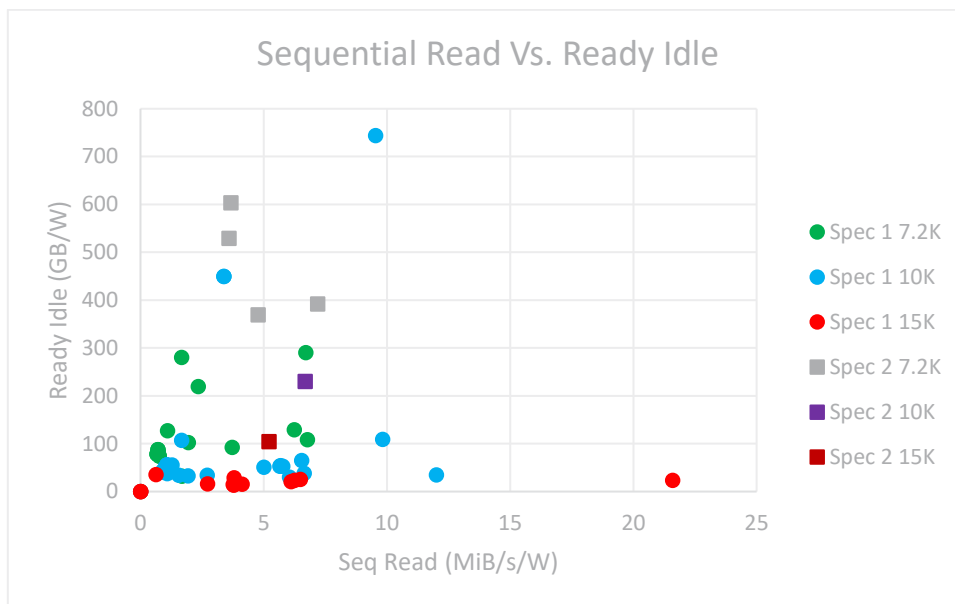


Figure 14: Sequential Read vs. Ready Idle Metrics

Figure 15 は、ドライブ速度とシーケンシャル読み取り指標の関係を示している。Energy Star for Data Center Storage バージョン 2.1 の限界値⁷が参考としてグラフに含まれている。

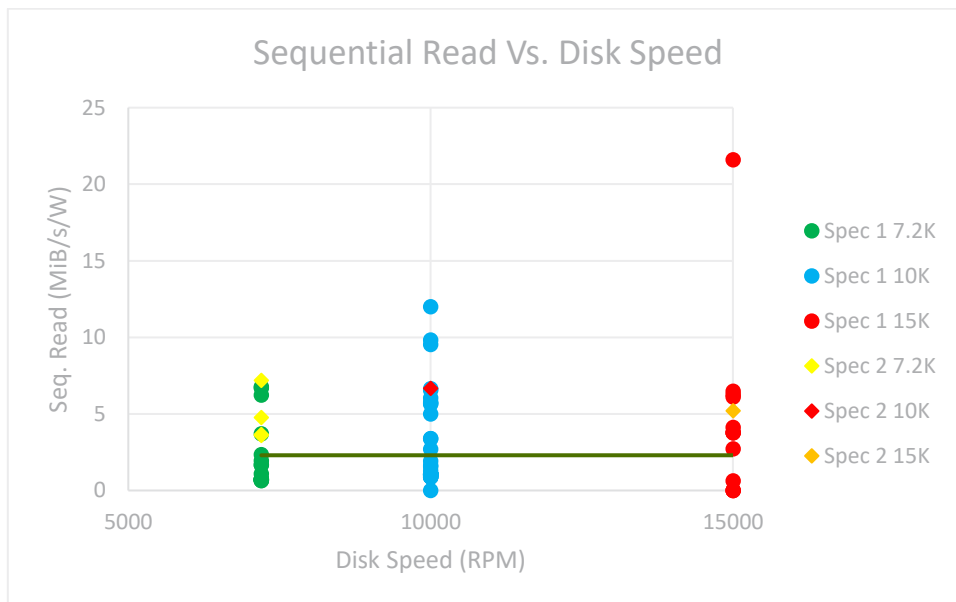


Figure 15: Sequential Read Metric vs. HDD Speed

Figure 16 は、シーケンシャル書き込み電力効率指標とアイドル電力指標に相関関係がないことを示している。

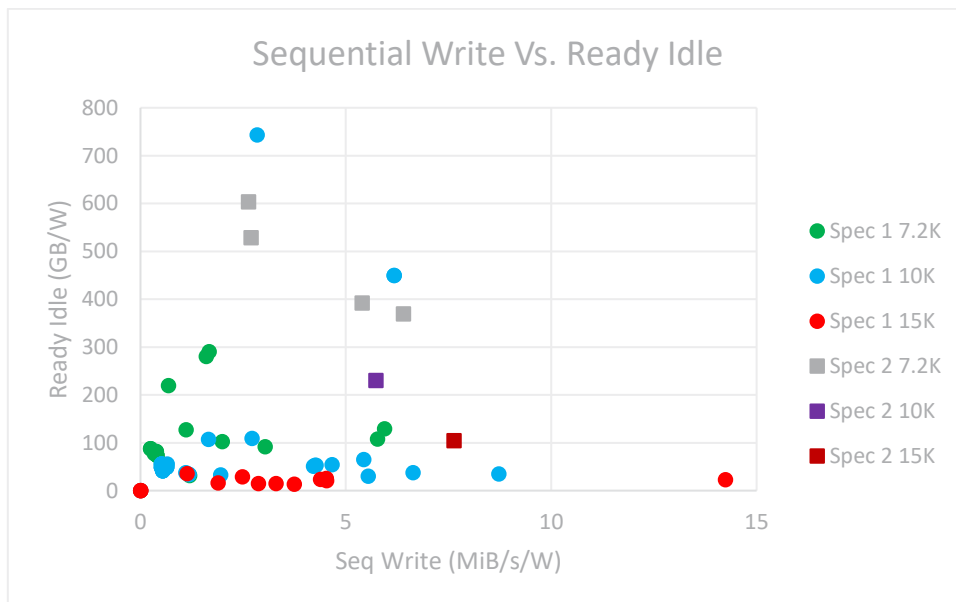


Figure 16 Sequential Write vs. Ready Idle Metrics

⁷ ENERGY STAR Data Center Storageバージョン2.1最終仕様、第3.3.1項。

Figure 17は、ドライブ速度とシーケンシャル書き込み指標の関係を示している。Energy Star for Storage バージョン 2.0 の限界値が参考として含まれている。

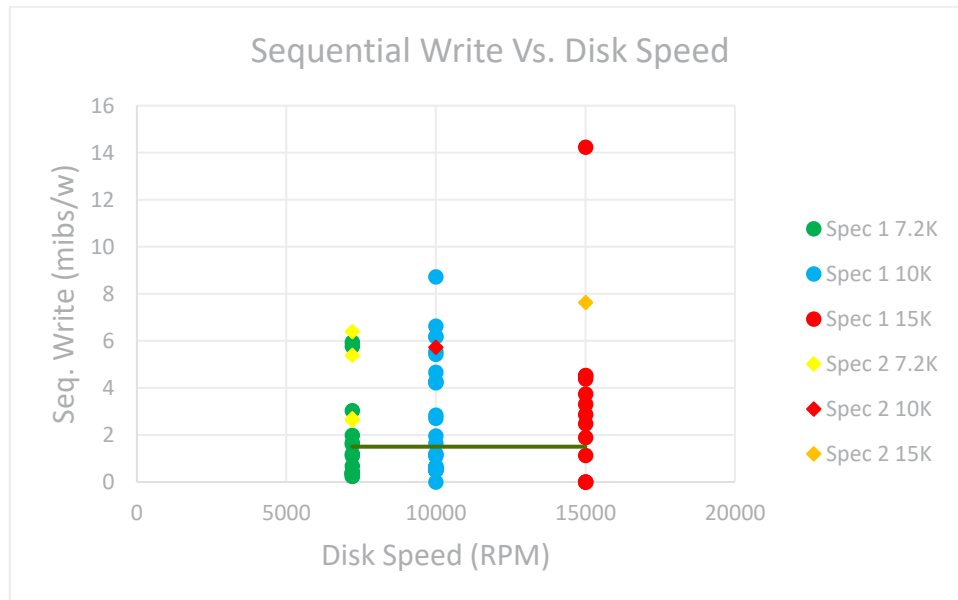


Figure 17: Sequential Write Metric vs. HDD Speed

所見

ブロック I/O システムに関するデータからの所見：

- アイドル状態は、他の性能指標との相関関係がなく、性能が高いシステムほど悪化する傾向がある。
- 7.2K RPM ドライブ・システムは、ホット・バンド基準での成績があまり良くない。
- 7.2K RPM ドライブ・システムは、シーケンシャル・テストで相応に良い成績を出す。
- SSD システムは、ホット・バンド基準でとても良い成績を出す。
- 10K および 15K RPM ドライブ・システムは、ホット・バンド基準とシーケンシャル基準で相応に良い成績を出す。
- Energy Star Data Center Storage バージョン 2.1 の限界値は、より性能の高いシステムに報いるという目標を満たしている。
- トランザクションまたはシーケンシャル用に最適化されたブロック・システムは、分類カテゴリにあまり依存していない。ドライブ速度とドライブ・タイプには、大きく依存している。

データが不十分なために、ファイル I/O システムに関する所見はない。

8. アイドル

ディープ・アイドル

ストレージ・デバイスは、電力が完全に切り離されている場合やデバイスがスピン・ダウン状態にある場合（HDDの場合で、電子回路はオンまたはオフ）に、ディープ・アイドル状態になる。電子回路がオフになっている場合は、操作をして再有効化するまでデバイスはIO要求に応答できない。エンタープライズ（データ・センター）用途の配備では、設置済みのストレージ・デバイスがディープ・アイドル状態になることは非常に稀である。このことは、ネットワーク接続された外部デバイスだけでなく、直接接続されたデバイス（計算リソースと同じ筐体内に存在することがある）にも当てはまる。

これ以外の状態では、何らかのレベルのバックグラウンドの動作が存在し、これは常にではないが多くの場合、データ完全性、容量最適化、またはフラッシュ・ベースのデバイスの場合にはウェアレベリングが目的である。

レディアアイドル状態

レディアアイドル状態は、I/O要求は存在しないが、システムはI/O要求を受け入れて指定された分類カテゴリのファースト・データまでの平均時間（MTTFD）以内に応答する準備ができていない状態である。分類クラシフィケーションでのオンライン・システムでは、MTTFDを80ミリ秒未満にする必要がある。

磁気ディスク・デバイスのレディアアイドル状態電力消費は、ドライブ速度と容量に左右される。一定のドライブ速度とフォーム・ファクターにおいて、ドライブ容量を増やすことによって、レディアアイドル状態指標を改善することができる。例えば、容量が1TBの7200RPM 3.5"ドライブは、容量が10TBの同様のドライブと比べて電力消費が約半分であるため、10TBのドライブのレディアアイドル状態指標は1TBドライブの約5倍大きい。

レディアアイドル状態は、ストレージ・デバイスがホストIO要求を実行していない場合に移行する最も一般的な状態である。この状態は、ネットワーク接続されたデバイスと直接接続されたデバイスに該当する。前述したバックグラウンドの動作またはプロセスの例を以下に示す。これらは通常、応答時間を最適化するためにホストIO要求が少ないときに実行される。

レディアアイドル状態の間に、ストレージ・システムとストレージ・デバイスの多くが「スマート」データ完全性チェックを実施する。これには、データの読み取りとエラーのチェックが含まれる。これにより、読み取り再試行回数がしきい値を超えると、HDDセクターまたはSSDページが予備のセクターまたはページに置き換えられる可能性がある。また、読み取り再試行回数がしきい値を超えると、ストレージ・システムがストレージ・デバイス全体を置き換える可能性がある。

レディアアイドル状態の間に、ストレージ・システム／デバイスがディープ・データ重複排除を実施する可能性がある。これには、同じデータの複数のコピーの検索とそれらの単一のデータのコピーへの参照との置き換えが含まれる。データ重複排除は、ファイルのコレクション、複数のLUN、複数のストレージ・デバイスの全体、および／またはデータ・センター全体に対して行われる可能性がある。

上記の例のことを考えれば、ホストI/O要求が存在しない間にはこのようなプロセスが抑制なしで実行されるため、明らかに、アイドル状態のデバイスはまったくアイドルではない。この状態は、ストレージ・デバイスの電力消費が最大になる状態であることが多い。

ソリッド・ステート・ドライブのアイドル状態

ディープ・アイドル時間中の活動は、SSD内で動作するメーカー独自アルゴリズムによって大きく異なる。この活動には、以下の挙動の一部または全部が共通に含まれる。

- デバイスがエネルギー節約モードに移行する。
- デバイスがメモリ・バンクを切り離す。
- 場合によっては、外部アルゴリズムが、エネルギー節約モードに基づいて、特定の時間が経過後に、メモリをシャットダウンする。
- 大半の最新の実装では、ディープ・アイドル状態に移行したデバイスを再アクティブ化するために100ミリ秒以上かかる。
- 様々な形態の「ウェイクアップ」を除いて、ディープ・アイドル（ディープ・スリープ）モード中には何も起きない。
- ウェイクアップ期間にデバイスがメモリの完全スキャンを実施する場合があります、その場合、さらに時間がかかる。

したがって、オンライン・ストレージ・システムにとってディープ・アイドルは受け入れ可能なモードではない。オンライン・ストレージ・システムではファースト・データまでの最大時間が80ミリ秒であることが必要なためである。

複数のSSDドライブの電力消費の経時変化が、システムからのI/Oがないアイドル状態で測定された。

この測定では、可能であれば、Idle_A電力状態だけを有効にしたため、Active状態とIdle_A状態のみが使用されている。

Figure 18は、代表的な結果である。

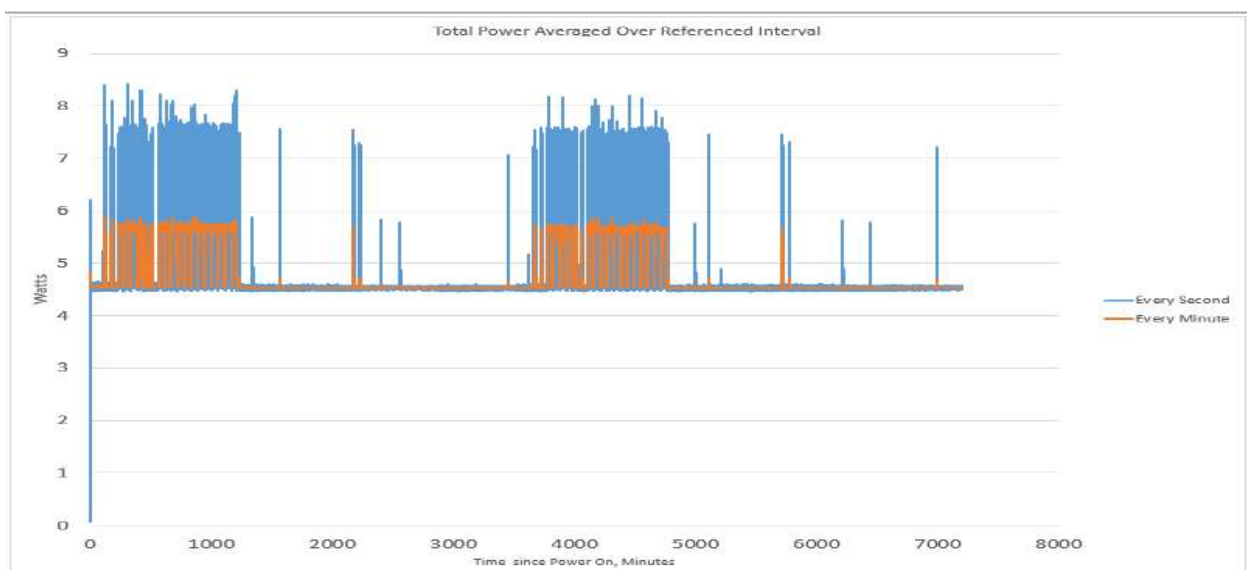


Figure 18: 7680 GB SAS SSD

Figure 18の2つの色つきの線は、a) 1秒間隔とb) 1分間隔の平均電力を示している。

図から分かるように、アイドル状態では、多くのバックグラウンド活動が存在する。

磁気ディスク・ドライブのアイドル状態

下のグラフは、アイドル状態の磁気ディスク・ドライブの電力消費の経時変化を示している。

アイドル・モードは、INCITS T10 および T13 の仕様で規定されている。以下の詳細は、特定のメーカーの実装に関するものだが、他のメーカーとの差はあまりないと思われる。

Idle_A は、オンライン・ストレージの 80 ミリ秒のファースト・データまでの平均時間要件をサポートするアイドル・モードである。Idle_A 状態では、ドライブのサーボ・システムのほとんどが無効になるため、ドライブのプロセッサとチャネルの電力消費が抑えられる。Idle_A 状態では、ディスクが最大定格回転数で回転する。

Figure 19 は、磁気ディスク・ドライブの一般的な挙動である。

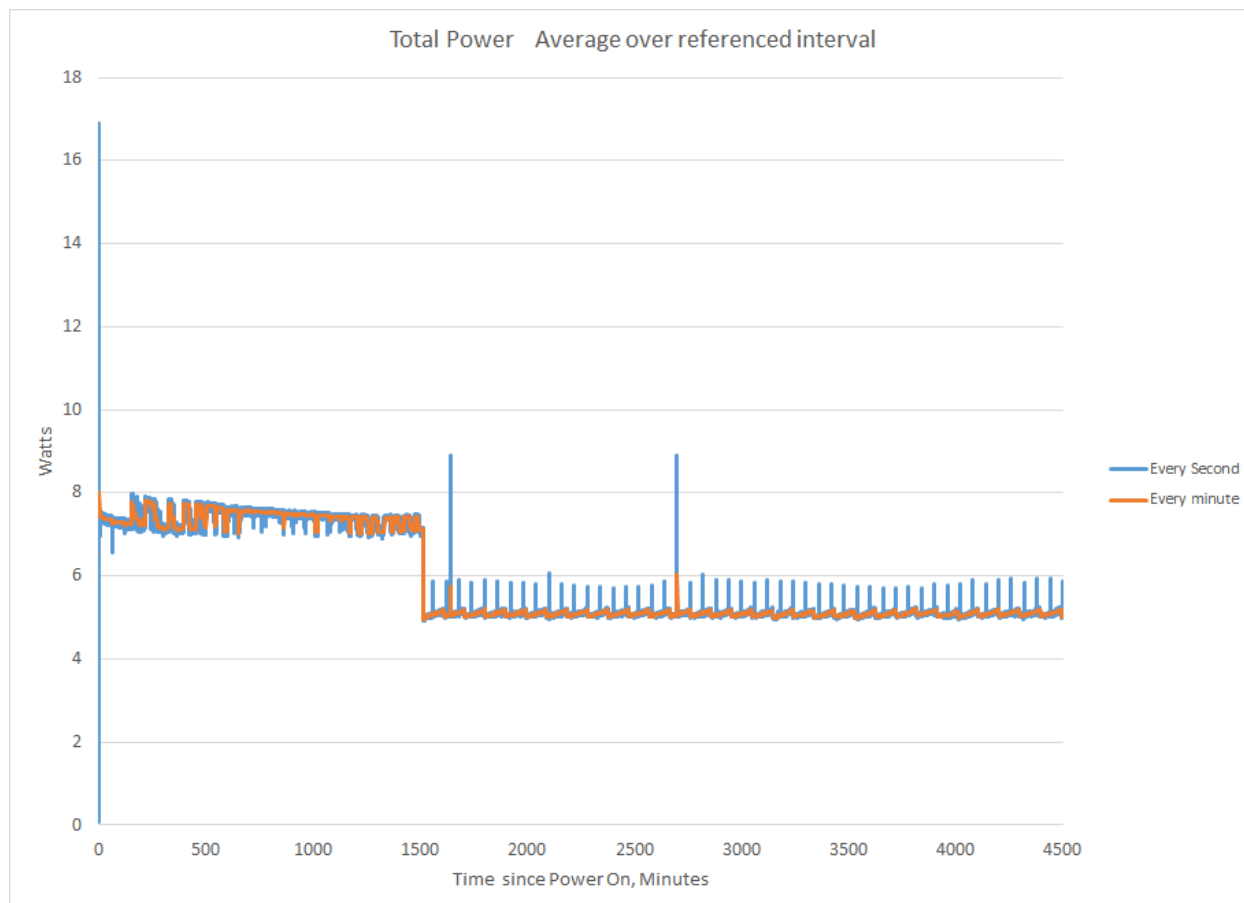


Figure 19: Example Power Consumption of Idle Magnetic Disk Drive

Figure 19 の 2 つの色つきの線は、a) 1 秒間隔と b) 1 分間隔の平均電力を示している。

Figure 19 の活動は、データ・スクラブを実行（一般的に週 1 回行われる）してから idle_A 状態に移行するドライブを示している。図から分かるように、ディスク・ドライブは「アイドル」期間にかなりの活動を行う。

9. 日本の標準

日本省エネ法（トップ・ランナー）

日本は、磁気ディスク記憶装置に関する新しい基準を採用した。経済産業省（METI）の省令・告示（2021年4月19日）新しい2023年度を目標年度とする磁気ディスク装置の基準エネルギー消費効率（トップ・ランナー）は以下で入手できる。

- <https://www.meti.go.jp/press/2021/04/20210419002/20210419002.html>
- https://www.enecho.meti.go.jp/category/saving_and_new/saving/enterprise/equipment/toprunner/07_jikidisk.html
- エネルギーの使用の合理化等に関する法律（昭和五十四年法律第四十九号）は以下で入手できる。
- <https://www.japaneselawtranslation.go.jp/en/laws/view/3959>

Figure 20 は、トランザクション最適化システムの様々なドライブ速度に関する SNIA Emerald アイドル状態指標（GB/W）のプロットである。20000 RPM 列は、実際は、ソリッド・ステート・デバイスである。各線は、日本省エネ法で規定されている限界値である。

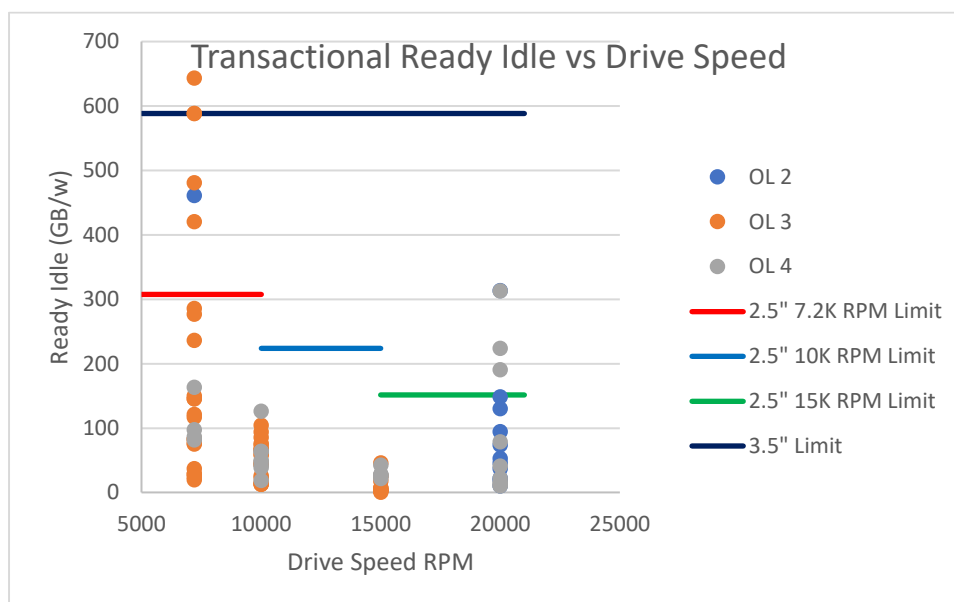


Figure 20: Ready Idle Metric vs. Drive Speed

図から分かるように、10000 RPM と 15000 RPM のドライブ・システムはすべて日本省エネ法基準を満足しない。

Figure 21 は、シーケンシャル最適化システムの様々なドライブ速度に関する SNIA Emerald アイドル状態指標 (GB/W) のプロットである。

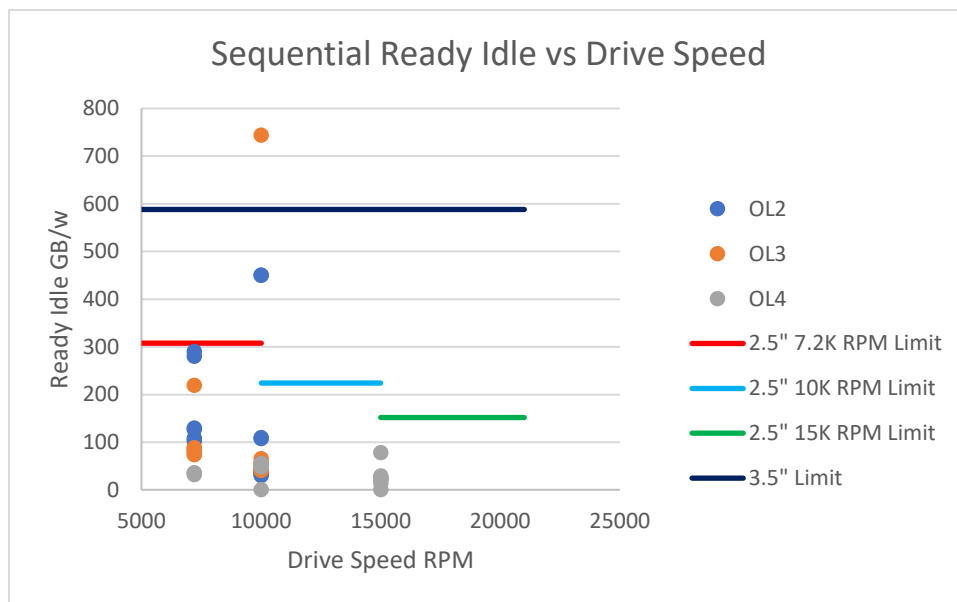


Figure 21: Ready Idle Metric vs. Drive Speed for Sequential Optimized Systems

繰り返しになるが、10000 RPM と 15000 RPM のドライブ・システムはすべて日本省エネ法の基準を満足しない。

Figure 22 は、すべてのシステムの 3.5" 7200 RPM ドライブのアイドル状態指標とドライブ容量の関係を示している。

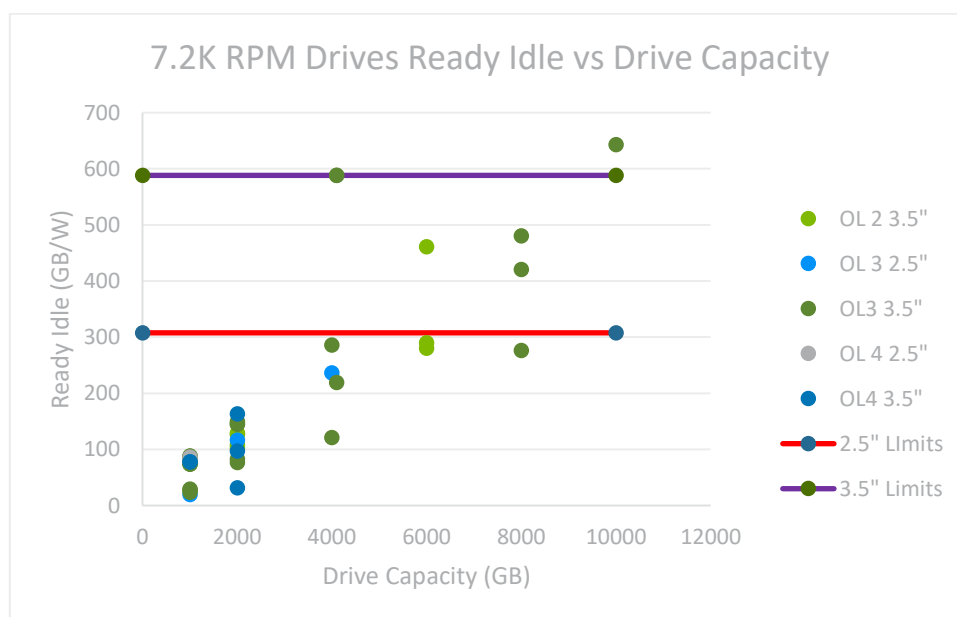


Figure 22: Ready Idle Metric vs. Drive Capacity (7200 RPM Drives)

Figure 22 のデータによれば、小容量 7.2K RPM ドライブに基づくシステムの日本のアイドル限界は問題となるが、新しい 7.2K RPM ドライブは容量が 12,000 GB と大きくなっている一方で、フォーム・ファクターと回転速度が同じ小容量ドライブと消費電力がほぼ同じである。すると、アイドル状態値は 1,000 GB ドライブの値の約 12 倍である 12,000 GB ドライブに相当する値になり、上記システムの一部を合格値まで高めるはずである。

Figure 20 と Figure 21 でプロットされたデータによれば、10K RPM または 15K RPM ドライブを搭載した 2.5" ドライブ・システムはどれもトップ・ランナー限界内に収まらない。15K RPM ドライブは、10K RPM ドライブとソリッド・ステート・ドライブのどちらかに置き換えられつつある。現在、10K RPM ドライブの新しい開発は行われていないため、10K RPM ドライブを使用しているシステムのアイドル性能の向上は期待できない。ドライブ容量が増加している領域は、7.2K RPM 3.5" ドライブである。上記データは、大容量 7.2K RPM ドライブを使用しているシステムが限界内に収まる可能性が高いことを示している。

日本のグリーン購入法

「グリーン購入法」は、「エネルギーの使用の合理化等に関する法律」（トップ・ランナー）とは異なる別の規制である。

「グリーン購入法」は、トップ・ランナー規制プログラムを管轄する METI（経済産業省）ではなく、環境省の管轄である。

「グリーン購入法」がシステムを購入するユーザについて各システムの基準を規定しているのに対して、トップ・ランナー規制は、システムを製造・供給しているベンダーについて合計システム出荷の基準を規定している。

「グリーン購入法」の基準は、基本的に、トップ・ランナーの基準に従って規定されている。12 台以上の HDD が搭載されたストレージ・システムのカテゴリ V および VI の基準は、トップ・ランナー基準より緩和されている。

- カテゴリ V（3.5" HDD が搭載されたストレージ・システム）
 - トップ・ランナー： $E=0.00170$
 - グリーン購入： $E=0.00213 (= 0.00170/0.8)$
- カテゴリ VI（2.5" HDD が搭載されたストレージ・システム）
 - トップ・ランナー： $E=\exp(0.952 \times \ln(N)-14.2)$
 - グリーン購入： $E=\exp(0.952 \times \ln(N)-14.2)/0.5$

「グリーン購入法」は以下で入手できる。

- https://www.env.go.jp/policy/hozen/green/attach/gpp%20pamphlet_jpn.pdf

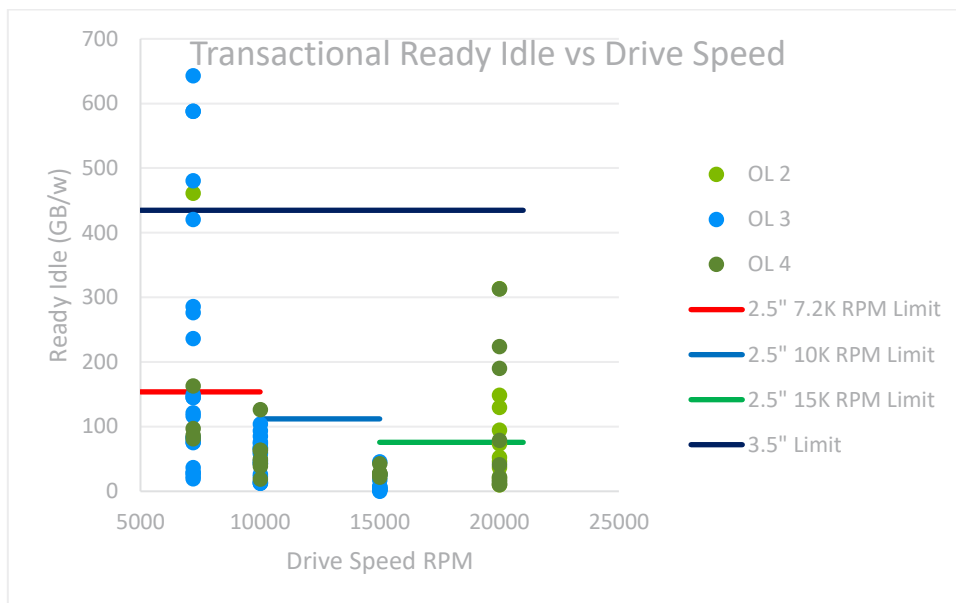


Figure 23: Ready Idle Metric vs. Drive Speed (Transactional Systems)

Figure 23 から分かるように、10000 RPM と 15000 RPM のドライブ・システムは、まだほとんどが日本省エネ法の基準を満足しない。

Figure 24 は、シーケンシャル最適化システムのような様々なドライブ速度に関する SNIA Emerald アイドル状態指標 (GB/W) のプロットである。

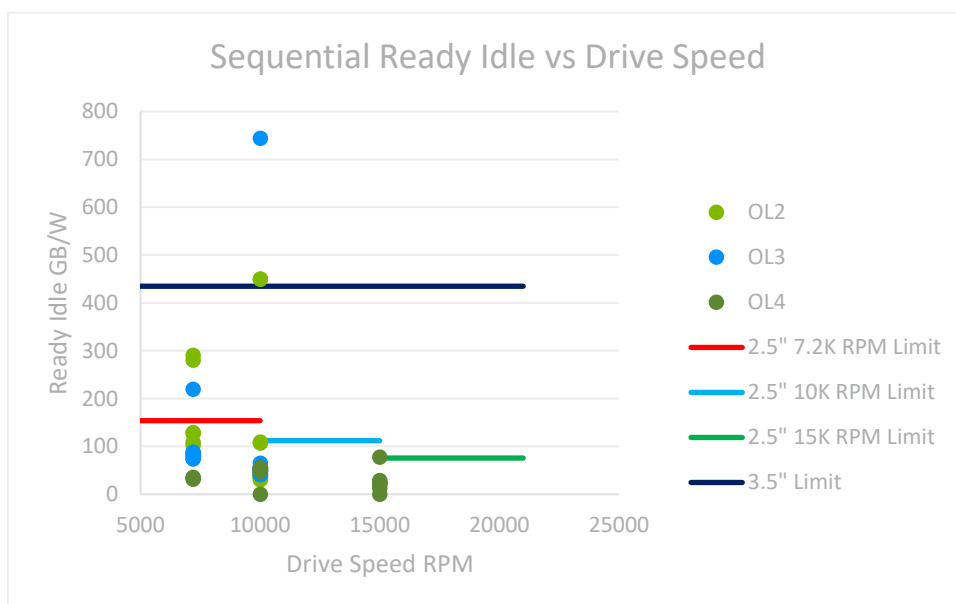
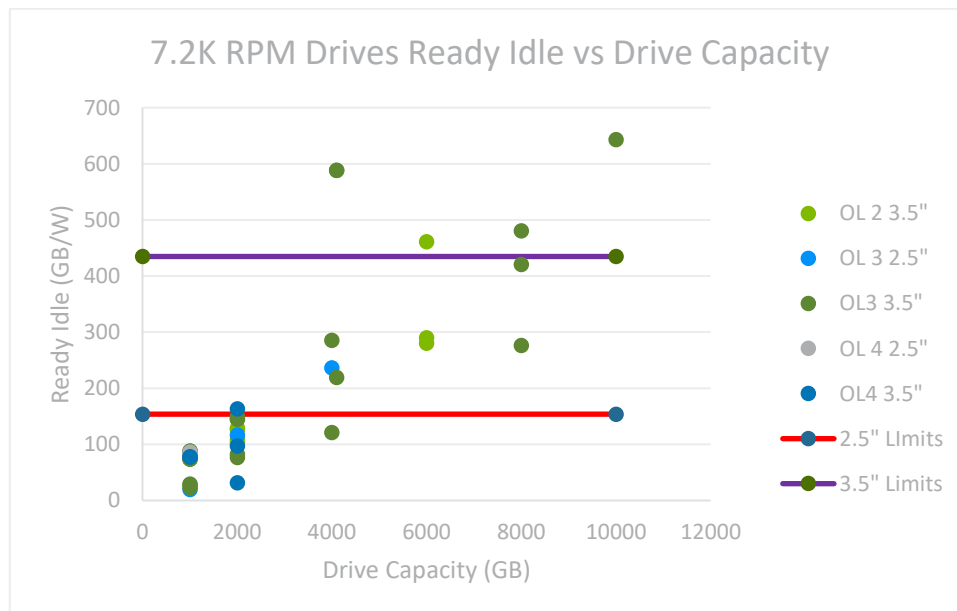


Figure 24: Ready Idle Metric vs. Drive Speed (Sequential Systems)

Figure 24 でも、10000 RPM と 15000 RPM のドライブ・システムはすべて日本省エネ法で規定された限界の範囲外である。

Figure 25 は、すべてのシステムの 3.5" 7200 RPM ドライブのアイドル状態指標とドライブ容量の関係を示している。



10. 結論と電力効率認定に関する提案

結論

ENERGY STAR データ・センター・ストレージ・プログラムに提出されたデータの分析から興味深い所見が得られた。

アイドル状態指標に関する所見：

- アイドル状態指標は、トランザクション指標とストリーミング指標のどちらとも相関関係がない。特に、アイドル状態指標はトランザクション指標と逆相関を示すことが多い。
- 特定の速度の磁気ディスクに基づくストレージ・システムに関するアイドル状態指標はドライブ容量に直接関係する。アイドル状態指標を高めるには、同じ速度で容量の大きいドライブに変更する必要があるだけである。ドライブ容量は定期的増加するため、次の容量がリリースされるまで待てば容量を増やすことができる。特定のドライブ・ファミリーでは、電力増加が容量増加に対して無視できるほどである。
- アイドル状態電力は、「アイドル」中にシステムによって実行される活動（COM、デバイスレベルおよびシステムレベルのハウスキーピングなど）の量に左右される。

最小、最適、および最大ドライブ数に関するファミリー・データのばらつきは、トランザクション指標の10%未満、シーケンシャル指標の20%未満である。これに基づく、最小/最大テストは追加的価値をもたらさないという結論になる。

データ・センター・ストレージは、電力効率用に特徴付けるには難しいカテゴリである。規模の範囲は十数台のデバイスから数百台のデバイスまでに及ぶ。デバイス自体にはフォーム・ファクター、性能レベル、インターフェイス、および容量の様々な組み合わせがあり、購入者のワークロードと性能期待に応じて高度に個別化した構成に展開できる。技術的変化が急速な進化を続けており、運用環境と実際のドライブの両方に付加価値機能をもたらしている。同時に、コンポーネント製造業者の数が極端に多くないため、複数のストレージ・システムの製品ラインで多くのハードウェア・コンポーネントが共通している。

この複雑さは、電力効率のための幅広い構成の特性評価に多くの労力が要求されることを意味する。また、テスト用の設備にも多額の投資が必要になることを意味する。リソース投資は、企業が電力効率テストを実施しないと決断する可能性があるほど大規模になる。

提案

電力効率テストの目的は、最適に動作する製品を認識して、電力効率の向上を動機付けることである。このような目的を踏まえて、次のような提案をする。

1. 製品認定のための一連の基準を定義する。
2. ベンダーによる幅広い参加を促すために簡略化されたアクティブ・テスト計画を定義する。

基準の提案

1. 製品は最低でも ASHRAE A2 レベルの運用に適合する必要がある。
2. PSU は、複数出力電源付きの製品の場合は 80 PLUS ゴールド以上に、単一出力電源付きの製品の場合は 80 PLUS プラチナ以上に必要がある。

3. COMの最小数は、SNIA Emerald 電力効率測定仕様で規定された分類によって与えられるクラシフィケーションに基づいてサポートする必要がある。

オンライン 1.5 :	なし
オンライン 2 :	1
オンライン 3 :	2
オンライン 4、5、6 :	3

4. 吸入温度報告を提供する必要がある。
5. 入力電力報告を提供する必要がある。

テストの提案

システムの特性評価プロセスの実行は複雑で膨大なリソースを必要とする。効果的でリソース要件が限定的なテスト・アプローチを実現するために、製品ファミリーに対してベンダーが選択した単一の「最適な」構成を望ましいデバイス・タイプでテストする SNIA Emerald 電力効率測定仕様に基づくテスト・アプローチを推奨する。このアプローチは、データ・センター・オペレータにストレージ製品のアクティブ性能／電力指標の評価を提供する。製品ファミリーのすべての構成の特性評価が単一のテストによって行われる。

特に、アイドル状態指標などの「アイドル」電力がストレージ製品の運用電力効率の評価に使用されないことが重要である。アイドル状態指標は、一般的なエンタープライズ・ストレージ・システムが動作する条件を代表していない。アイドル状態指標は、運用電力効率との相関関係が全くなく、エンタープライズ環境で作業を最も効率的に実行するシステムを特定できない。

まとめ

これらの提案は、市場参入または調達目的で検討されているすべてのストレージ製品の電力効率要件またはプログラムに関係する。

11. グリーン・グリッドと SNIA

グリーン・グリッド

グリーン・グリッド (TGG) は、情報技術産業評議会 (ITI) の関連会員組織である。TGG は、世界中で、ツールを作成したり、技術的な専門知識を提供したり、低炭素経済を可能にするデータ・センター・エコシステムのエネルギーおよび資源効率の最適化を提唱したりしている。また、TGG は、データ・センター・オペレータ、クラウド・プロバイダ、技術や装置のサプライヤ、ファシリティ・アーキテクト、およびエンドユーザの開かれた産業コンソーシアムである。TGG は、データ・センター業界や政府がデータ・センターの効率性や持続可能性に関する知見や助言を求めて参考にする技術的専門知識を提供している。詳細については、<https://www.thegreengrid.org/en> を参照されたい。

SNIA

ストレージネットワーキング・インダストリ・アソシエーション (SNIA) は、情報技術全体にわたる会員企業で構成された非営利組織である。世界的に認知および信頼された機関である SNIA の使命は、ストレージ業界による情報の効率的な管理、移動、およびセキュリティを促進するベンダー中立のアーキテクチャ、標準、および教育サービスの開発と普及をリードすることである。

SNIA グリーンストレージイニシアティブ (GSI) は、すべてのネットワーク接続ストレージ技術におけるエネルギー効率および保存の発展とデータ・ストレージ運用の環境への影響の最小化を専門とする。GSI の使命は次のとおりである。

- SNIA Emerald™ エネルギー効率測定のための SNIA Emerald™ プログラムを確立して維持し、SNIA Emerald™ テスターと業界関係者のトレーニングを実施する
- IT 業界、ベンダー・コミュニティ、および規制当局に対して、エンタープライズ・ストレージ環境向けのエネルギー節約技法について教育する
- SNIA グリーンストレージ技術作業部会 (TWG) の技術的作業の支持と支援を外部から提供する
- グリーンストレージの測定仕様、指標、および標準に関する要件に対する意見を SNIA グリーンストレージ TWG に提出する
- データ・センターのエネルギー効率に関連したプログラム、テスト方法と測定方法、および標準を調整して発展させるための業界間の関係と提携を確立して維持する

グリーンストレージ TWG は、ストレージ・ネットワーキング製品のエネルギーと冷却に関連した技術的事項を特定して、充実化し、調整するための SNIA の主たる技術的組織として機能している。有効性を持つために、標準指標は効率性の現実的な反映（投入エネルギーあたりの達成作業量）でなければならない。これにより、明確な運用コストの決定が可能になる。SNIA は、様々な種類のストレージ作業を検討しなければならず、複数のはっきりと異なる「性能値」が必要になると考えている。また、SNIA は、適切な参加者で構成された公平なフォーラムとして、ユーザにとって有益な指標を作成するのに比類なく適任であると信じている。

SNIA Emerald™ 電力測定仕様は、新製品の開発、エンドユーザ顧客の評価、および規制標準の策定の目的でストレージ・ネットワーキング製品の電力消費と電力効率を測定可能な指標を特定し、手法を定義している。この仕様は、SNIA 標準であり、バージョン 3.0.3 が国際標準 ISO/IEC 24091:2019 として採用されている。