



SNIA Emerald™電力効率測定仕様 ユーザガイド V2.1.1

バージョン 2.0 レビジョン 7



2015年9月30日



SNIA について

Storage Networking Industry Association (SNIA) は、非営利のグローバル組織であり、全世界のストレージ業界の会員企業から構成されています。SNIA の使命は、標準、テクノロジー、教育サービスの開発および推進において、全世界のストレージ業界を牽引し、情報管理面で組織を強化することです。この目的を達成するために、SNIA は、より広範な市場にオープンストレージネットワーキングソリューションを普及させるための標準、教育、およびサービスの提供に取り組んでいます。詳細については、SNIA Web サイト (www.snia.org) を参照してください。また SNIA 日本支部のサイトは <http://snia-j.org/> を参照して下さい。

SNIA のグリーンストレージ分科会について

SNIA のグリーンストレージ分科会 (Green Storage Initiative : 略称=GSI) は、データストレージの利用が環境に及ぼす影響を最小限に抑えることを目的として、ネットワーク化されたすべてのストレージ技術においてエネルギーの効率的な使用と節約を推進するための作業に集中して取り組んでいる。SNIA のグリーンストレージ活動は、SNIA グリーンストレージ技術作業部会 (Technical Working Group : 略称=TWG) とグリーンストレージ分科会の 2 つのグループで行われる。TWG は、エネルギーの消費量と効率を測定可能な試験指標の開発に取り組んでいる。グリーンストレージ分科会は、エネルギー効率の高いストレージネットワーキングを実現するためのベストプラクティスを策定および公開し、IT コミュニティに対する啓発活動を行うとともに、ストレージのフットプリントおよび関連する電力要件を低減するストレージセントリックアプリケーションを推進するための作業に取り組んでいる。

SNIA エメラルドプログラムについて

SNIA Emerald™ プログラムは SNIA のグリーンストレージ分科会が資金を提供し運用している、ストレージ業界とエンドユーザを対象とした、ベンダー中立の誰でも参加可能なサービスである。このプログラムは SNIA Emerald™電力効率測定仕様の普及と発展に努めている。また本仕様に従って測定された製品の試験データを公開し、活用が進むように支援している。SNIA Emerald™電力効率測定仕様は、ストレージの電力効率の測定手順と試験の指標を定めている。本仕様は SNIA グリーンストレージ分科会のガイドの下、グリーンストレージ TWG が開発し、発表し、維持している。

このプログラムはベンダーが測定したストレージシステムのエネルギー使用と効率に関するさまざまな観点からの試験結果を報告する時の標準的な方法を提供している。異なったワークロード下の性能に基づいた試験結果は、他の電力に関連したストレージシステムやその構成部品の情報と組み合わせられる。このプログラムは試験結果を一貫した形式で表示する、誰からもアクセス出来る共通のリポジトリを提供する。このリポジトリはデータストレージの電力使用量や効率とともにそれらを運用するためのコストを決めるのに使用することが出来る。

米国 EPA の ENERGY STAR® データセンター用ストレージプログラムはこの仕様に基づいた方法論を用いているが、本仕様に従って実施した製品の試験結果の公開には別の方法を使っている。



Copyright 2011, 2012, 2013, 2014, 2015 Storage Networking Industry Association

本書に記載されている情報は、予告なく変更される場合がある。本ガイドは、SNIA Emerald™電力効率測定仕様に沿った測定を実施する技術者に対して指針を提供するために、SNIA グリーンストレージ技術作業部会が「ベストエフォート」(出来る限り)で作成したものであり、何時でも内容が更新または置き換えられる場合がある。本書に記載内容の誤りがある場合、SNIA はこれに対して責任を負わないものとする。

本ガイドの改訂についての提案、または SNIA Emerald™電力効率測定仕様の実施に関する質問がある場合は、(電子メールにて) greentwg-chair@snia.org に連絡のこと。



目 次

1	はじめに	5
1.1	対象読者.....	5
1.2	参考資料.....	5
2	適用範囲	6
2.1	概要.....	6
2.2	本ガイドの利用法.....	6
3	分類法に関するコメント	7
4	製品ファミリの定義	7
4.1	概要および目標.....	8
4.2	製品／ファミリの定義.....	8
4.3	変数のレンジに関する検討.....	9
4.4	ベストフィットフォワード（最適値）の試験法.....	10
5	ベストフィットフォワード（最適値）を見つけ出す方法	11
5.1	概要および目標.....	11
5.2	段階的方式.....	11
5.3	推定ツール／シミュレーションツールの検討.....	12
5.4	実施例.....	12
6	測定試験の設定と実施	14
6.1	設定.....	14
6.1.1	試験の構成.....	14
6.1.2	ベンチマークドライバの要件.....	15
6.1.3	電力計の要件.....	15
6.1.4	推奨の電力計.....	16
6.2	試験の手順.....	16
6.2.1	オンラインおよびニアオンライン.....	16
6.2.2	リムーバブルメディアライブラリ.....	17
6.2.3	仮想メディアライブラリ.....	18
6.2.4	データ収集の纏め.....	18
6.3	容量最適化手法（COM）試験.....	18
6.3.1	デルタスナップショット試験.....	19
6.3.2	シンプロビジョニング試験.....	20
6.3.3	パリティ RAID 試験.....	20
6.3.4	データ重複排除と圧縮の試験.....	20
6.4	測定時に発生する可能性のある落とし穴の回避.....	21
6.5	報告する指標.....	21
7	データ提出時の注意事項	22



1 はじめに

本書は、SNIA Emerald™プログラムの一環として作成された SNIA Emerald™電力効率測定仕様 V2.1（本書では単に「測定仕様」と呼ぶ）の参考になる手引きとして使用することを意図している。SNIA Emerald™プログラムは、ストレージシステムベンダが製品の電力効率を実証するための一貫した、信頼できる方法を提供することを目的として策定された。これを促進するために、SNIA のグリーンストレージ分科会（Green Storage Initiative：略称=GSI）とグリーンストレージ技術作業部会（Technical Working Group：略称=GTWG）は、SNIA Emerald™プログラムを通して情報収集および比較のためにアクセス可能な結果をストアするためのメカニズムと併せて、ストレージシステムの電力効率を測定する簡易かつ標準的な方法を開発した。この方法は、顧客先でストレージシステムの真の電力効率を実証することを目的とするものではなく、ストレージシステムの使用中に予測される電力効率に関する全般的な理解を深めることを意図している。

この測定仕様はまた EPA により ENERGY STAR®データセンター用ストレージプログラム及び関連する試験仕様に採用されている。本書に列挙されている手法と注意書はこのプログラムに試験結果を提供する際にも役立つ。もし本書と「測定仕様」に矛盾がある場合は、「測定仕様」が優先される。

1.1 対象読者

本書の対象読者には、「測定仕様」に従って試験を計画し実施する組織と個人が含まれる。試験を実施する組織または個人を試験のスポンサーと呼ぶ。

1.2 参考資料

本ガイドは、以下の資料と併せて参照のこと。これらの資料は SNIA Emerald™の Web サイト (<http://www.sniaemerald.com/download>) から入手できる。

- ・「測定仕様」
- ・Vdbench ツールの Web サイトへのリンク
- ・関連するワークロード用のスクリプト
- ・COM データ生成ツール
- ・本書
- ・トレーニング資料
- ・試験データレポートのテンプレート

SNIA Emerald™プログラム、および関連する SNIA Emerald™電力効率測定仕様、並びに EPA の ENERGY STAR®データセンター用ストレージプログラムに関する詳細な情報は、以下のウェブサイトから入手できる。

- ・ <http://www.sniaemerald.com>、SNIA Emerald™プログラムウェブサイト
- ・ <http://www.snia.org/forums/green>、SNIA グリーンストレージ分科会ウェブサイト
- ・ <http://www.energystar.gov/products/certified-products/detail/data-center-storage>



2 適用範囲

2.1 概要

ストレージシステムのベンダと顧客が高信頼性かつ一貫した方法で、各種のストレージソリューションのストレージ電力効率を見比べられるように、SNIAは「測定仕様」とSNIA Emerald™プログラムを開発した。「測定仕様」と試験の実施に関する詳細には、これらのシステムの複雑さが反映されている。本書は「測定仕様」の試験方法の理解と効率的な実施の一助となるように意図されている。

「測定仕様」には幾つかの側面がある。まず最初に製品を基本的な機能によって区分する分類法のテーブルがある。しかし実際の試験のためには更に詳細な区分が必要になる。本書はこれに対して製品ファミリーを記述する。また合理的な時間と費用で意味のあるデータを取得するには、試験構成に従って適切な試験のポイントを選択することが大切である。適切な試験のポイントを見つける為の助けとして、所謂ベストフットフォワード（最適値）指標ポイントが定義されている。

「測定仕様」は更に代表的な電力効率指標の値を決める方法を定義し提供する。IO/s/WattとMiB/s/Watt指標はデータをストレージシステムとホストの間で動かす際の電力効率を表現するアクティブな指標である。GB/Wattはストレージシステムにデータを格納しておく際の電力効率を表現するアイドル指標である。

重複排除や圧縮などの容量最適化手法（COMs: Capacity Optimization Methods）の存在と基本的な動作を調べる試験も「測定仕様」に含まれている。これらの機能は複雑であるので、多くの場合アクティブ試験やアイドル試験中に動作させるのは難しい。

「測定仕様」の主要な内容は試験構成、測定機器、ベンチマークドライバの要件とIO試験のプロファイル、試験実行方法と順序、及び指標の計算方法を定義することである。主要な分類のカテゴリ毎に別のセクションが用意されている。

2.2 本ガイドの利用法

本書は、「測定仕様」の補助として、下記の内容を提供する。

- ・ 分類法のカテゴリから製品ファミリーを定義する方法
- ・ 適切な測定構成と測定点を決定する方法
- ・ 測定環境と測定機器の要件を理解する方法
- ・ Vdbanchと関連するワークロードを生成するスクリプトの使用法
- ・ 指標の値の妥当性の検証を含む測定手順の設定と実行の方法
- ・ 問題の回避
- ・ 結果の提出

「測定仕様」と本書では試験中のシステム構成を試験中システム（SUT: System Under Test）と呼ぶ。本書では「測定仕様」の全ての分類カテゴリに関して基本的な情報を提供しているが、主要な内容はオンラインシステムとニアオンラインシステムに関するものである。但し、リムーバブルメディアと仮想メディアライブラリのカテゴリに関して厳選された幾つかのコメントを掲載している。



3 分類法および製品ファミリの構成

ストレージ関連製品は種類が幅広く多岐に及ぶため、分類構造が作成され「測定仕様」に含まれている。この分類法には現在、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、および仮想メディアライブラリのストレージカテゴリがある。この各カテゴリはさらに、種々の特性に基づくクラスに分けられる。これらのカテゴリにはそれぞれ独自の異なる試験基準が設けられているため、有効な測定の為には SUT のカテゴリとクラスを正しく判断することが非常に重要である。

オンラインカテゴリとクラスでは、データブロックの最初のデータを 80ms 以内で読み出せるストレージシステムを扱う。これらのシステムは一般的にディスク又はフラッシュベースである。このカテゴリにはコンシューマ向けの製品/部品から、何百ものストレージデバイスをサポートする大きなシステムまでが含まれる。

あるシステムをオンラインのあるクラスに分類する場合、クラス分類表に載っているストレージデバイス数以上のデバイスをサポートしている必要がある。例えばオンライン 3 クラスのストレージ製品は少なくとも物理的に 12 個のストレージデバイスをサポートしている必要がある。但し、それより少ないストレージデバイスでそのストレージ製品を販売したり試験しても問題はない。

ニアオンラインカテゴリのストレージシステムは、データブロックの最初のデータを 80ms 以内で読み出せる必要は無い。しかし、ランダム IO とシーケンシャル IO の両方をサポートしている必要がある。ニアオンラインはオンラインと同様に、最大サポートデバイス数でクラス分けされる。

オンラインとニアオンラインは似ているので、「測定仕様」では双方に同じ構成と順序を使用しているが、異なる部分に関しては注意書きしている。

リムーバブルメディアライブラリカテゴリにはテープライブラリと光学ジュークボックスが含まれる。これらのシステムは最初のデータの読み出しに 80ms 以上が必要であり、ストリーミング I/O のみをサポートする。SNIA Emerald™プログラムの範囲では、ストレージシステムに限定した電力効率を扱うので、ストレージデバイス（テープ）はライブラリ内にある必要がある（テープをライブラリ外に取り出す必要は無い）。従って、データの読み出しは 5 分以内であればよい。他のカテゴリと同様に、このカテゴリも最大サポートデバイス数でクラス分けされる。全てのカテゴリで番号付けの整合性を保つために、リムーバブル 4 クラスは存在しない。

仮想メディアライブラリのカテゴリは最初のデータを 80ms 以内に読み出せる必要がある。これらは通常、シーケンシャル I/O 向けに設計されたディスクベースのシステムである。このカテゴリも最大サポートデバイス数でクラス分けされる。

4 製品ファミリの定義

分類法はストレージシステムを区別するのに役立つが、ベンダはひとつの分類カテゴリ、クラスの中でも広い範囲の製品を持っている。試験の実施の助けとして、試験のシステム構成を限定し、試験にかかる労力と費用を最小化するために、本書は製品とそのファミリを区分



するための助言を含んでいる。最も小規模なシステムであっても構成オプションは非常に数多くあり、それぞれの構成試験では多大な実施努力が要求される。

4.1 概要および目標

エネルギー効率試験にどのようなシステム構成を選択するかを考慮する際に、いくつかの側面がかかわってくる。特に、顧客は特定の購入製品候補の効率を測定および評価するための明確かつ適度に満足できる方法を要望する。これと同時に、ストレージシステムベンダは場合によっては大規模になる可能性のある一連の製品構成および使用事例からコストをできるだけ低減し、カバレッジ範囲をできるだけ幅広くするために、効率測定試験の種別を最小限に抑えることを望む。

4.2 製品／ファミリの定義

「測定仕様」には、ストレージ製品を比較的粗いカテゴリおよびクラスに分ける分類法が記述されている。この分類法に基づくカテゴリ／クラスに1つの製品を分けた後で、その製品およびこの製品から派生する可能性のあるバリエーションをセクション 4.1 に示した目標に従って実際にどのようにして測定するかという問題が残る。

実際のストレージシステム試験構成の定義を進める作業に役立てられるように、製品および製品ファミリのコンセプトをここでは紹介する。各ベンダがその製品を定義および販売する方法は異なるが、この製品／製品ファミリ方式は一般的に受け入れられるものと確信する。

1つの製品は、見る側によって、異なる側面を持つ。顧客にとって、製品は特定の購入済みおよびインストール済みの構成を表す。ベンダにとって、製品は一連の限られた構成オプションを備えるベース（場合によってはエントリ）ユニットである可能性がある。製品ファミリについても同様に、多くの解釈が適用されることが考えられる。

本書では、製品、および製品ファミリを以下のように定義している。

- ・製品とは、他のすべての関連する可能性のある製品から分離された、基本的な性能能力空間を表すものである。
- ・製品ファミリとは、1つの特定製品について各種の構成編数及びオプションからなるすべてのレンジを表すものである。

このセクションでは「ファミリ」と「レンジ」の用語は同義的に使用され、ストレージデバイスの数とタイプ（スピニングまたはソリッドステートドライブ）、可用性レベルなどの側面が含まれることがある。

図1には、簡略化されているが、考えられる製品／ファミリ（レンジ）差別化記述を表す図を示す。この図は、例えばモノリシック、スケールアップ、またはスケールアウトなど、どのようなストレージシステムアーキテクチャにも適用される点に注目すること（スケールアップとは一般にスケラブルなバックエンドストレージを備える限られた数のコントローラで構成されるシステムを指し、スケールアウトは実または仮想を問わず相互接続されたコンピュータストレージノードで構成されるシステムを指す）。

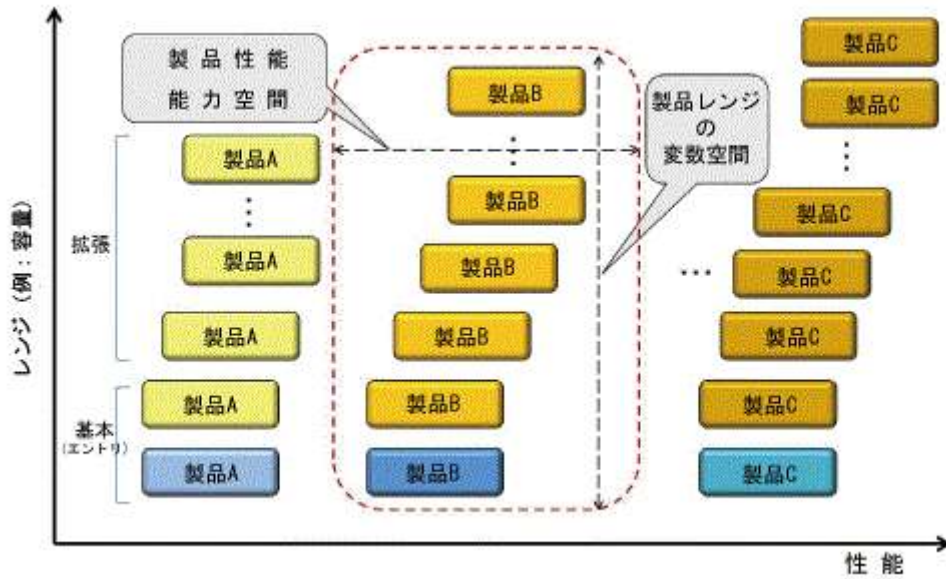


図1 考えられる製品／ファミリ（レンジ）記述

図1に示すレンジの変数空間は容量のみに着目しているが、ストレージデバイスのタイプまたはその他の変数を含むこともできる。製品によっては、容量の増加に伴って性能が向上する製品もあれば、そうでない製品もある。（例えば、特定の容量や変数ポイントを過ぎると性能が低下する製品もある点に注意する。）

4.3 変数のレンジに関する検討

セクション4.2の製品／ファミリの説明で指摘したように、全ファミリレンジは数とタイプ両方の数多くの変数を網羅しており、これらの変数のうち少なくとも25をSNIAは定義している。以下の項目は、エネルギー消費量に対して最も高い影響を及ぼす可能性があると考えられる項目を強調して列挙したものである。

- ・コントローラ、またはこれに関連するコンピュータの要素一般に製品パフォーマンス空間を定義する。
- ・キャッシュ機能一常にコントローラと同調するとは限らず、ユーザアドレッシング可能空間の一部と考えられない。
- ・永続的なストレージデバイスの数とタイプ—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）などで構成されるユーザアドレッシング可能空間を定義する。
- ・RAS項目—信頼性、可用性、およびサービス性の要件を満足するために必要なエネルギー消費機能。
- ・容量最適化—圧縮、複製解除、およびシンプロビジョニングなどの物理的ストレージ空間をより効率的に利用する機能（通常はソフトウェア）。

電源、I/Oポート、冷却部品、相互接続ポートなどのその他の項目は無視されていないが、これらはこのセクションで定義されているパフォーマンスおよび項目と同調およびスケールリングすると考えられている。



変数空間をこのセクションで列挙する5つの項目に低減しても、場合によっては非常に膨大になる可能性のある一連の試験要件および事例をベンダは依然として抱えることになり、それぞれの試験要件および事例には多大な設定および実施時間を必要とする。HDDとSSDの数とタイプが唯一の変数である構成であっても、その実施が極度に困難である場合がある。最大システムサイズ試験は高額のコストを伴い、しかも管理が厄介である。顧客も同様な問題を抱え、大量の試験結果を読み解いて、効果的なベンダ製品比較を実施する必要がある。そこで、この変数セットをもっと低減するように努めるのではなく、別の方法が提案されている。つまり、セクション4.4で定義する「ベストフットフォワード」(「スイートスポット」または「最適点」とも呼ばれる)方式である。

4.4 ベストフットフォワード(最適値)の試験法

ベストフットフォワード(最適値)(Best Foot Forward: 略称=BFF)方式は、ストレージシステム製品を総体的に見る。この方法によって、ストレージベンダは「測定仕様」の測定基準ピーク値に該当するか、またはそれに近いと判断される動作ポイント(すなわち、「スイートスポット」)において、1つまたは複数の特定製品/ファミリの構成を選択および試験することができる。その結果として、製品ファミリ全体を代表する一連の試験結果の数が低減され、ベンダによる試験が容易になるとともに、試験コストが削減され、理解しやすい結果が得られるので、顧客にとっては有用性が高くなる。

この方式は、「測定仕様」のアクティブ指標には、より小さい、つまりより容易に測定可能な製品/ファミリの構成範囲内に「ピーク」値ポイントが存在するという考え方に基づいている。ベンダは1つまたは複数の適切とされる代表的な構成を選択し、上記の「測定仕様」指標のピークポイントを見つけ出す。この方法で重要な点は、構成が最大になる試験、並びに外挿法や補間法などのその他の複雑な方法を回避することである(小規模システムの一部のケースでは、最大構成が事実上BFFである場合がある点に留意すること)。

性能が容量にスケールし、途中から性能が低下する仮想的なストレージシステムの例を図2に示す。

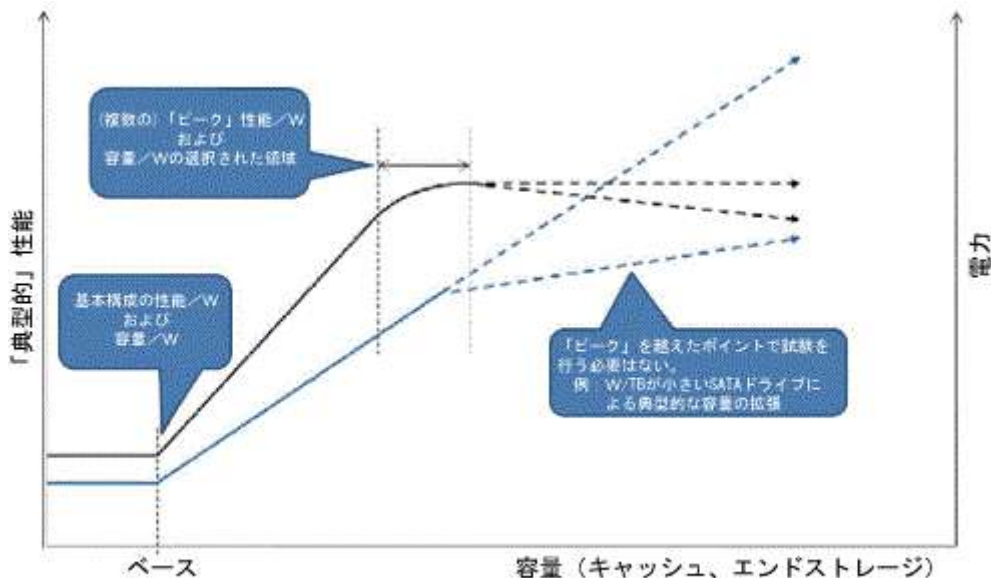


図2 仮想的なストレージシステムの性能/電力例



図2に描く線は非常に簡略化して表示した近似線であり、実際のシステムではこれらと異なることがある（スケールアウトシステムでは、性能の線が広範囲に低下しないことがある）。それにもかかわらず、この例ではベンダが決定したピークの「測定仕様」測定指標値ポイントで比較的小さな代表的システムを選択および試験できる方法を示している。基本構成（エントリポイント）で試験を行うことも可能である。ただし、ピークポイントを越えた領域で試験を実施しなければならない要件はない。実際のところ、多くのシステムではすでに高いGB/ワットと大容量のHDDを使用して容量を拡張している。これと同様に、スケールアウトシステムはBFF方式に従って試験されるデバイスと同じデバイスのステップ&リピートインスタント化によって性能と容量をスケーリングすることが可能である。

5 ベストフットフォワード（最適値）を見つけ出す方法

5.1 概要および目標

電力効率指標のピーク値における製品／ファミリの構成の試験法として、ベストフットフォワード（最適値）（「スイートスポット」とも呼ばれる）方式をセクション4.4で紹介した。この方式の利点は、製品ファミリ全体を代表する試験結果を少なくするとともに、試験可能な一連の変数を大きな変数範囲から少ない数（場合によっては1つのみ）に低減することである。

このセクションでは、予測ツールを用いてベストフットフォワード（最適値）の構成を見つけ出す1つの方法について説明する。さらに、この方式の特性についても記述する。説明するツールを使用することによって、幅広い範囲の構成変数の評価が可能になり、予測されるスイートスポットに比較的迅速に到達できる。

5.2 段階的方式

ベンダは以下のステップに従って、ベストフットフォワード（最適値）を求めることができる。

1. 分類法の定義の範囲内に適合する製品から作業を開始する。製品をいくつかの分類に適合するように構成できる場合、ベンダはそれぞれ適用されるカテゴリおよびクラス別にデータ提示を別途行うことを考慮する必要がある。
2. すべての見込まれる（および有効な、つまり販売可能な）製品のSKU（ストック管理ユニット：Stock Keeping Unit）を考慮に入れた上で、電力効率指標のピーク値が得られる最適化された構成を特定する。

6つの異なるSNIA Emerald™プログラム試験プロファイル（5つのアクティブと1つのアイドル）があるため、指標のピーク値に達する構成として以下のように最大で6つまでの異なる最適化（調整された）構成が存在する可能性がある。

- ・1×ホットバンド [IO/s/ワット]、2×ランダム [IO/s/ワット]、
2×シーケンシャル [MiB/s/ワット]、および1×アイドル [物理（Raw）容量、GB/ワット]



3. 推定ツールを用いて、指標のピーク値を予測する。これに代わる方法として、経験に裏付けられた推測があるが、結果としてこれには多大な労力を伴い、資源集約的試験になる可能性がある。シミュレートされた結果が適度に正確なものである限り、ピーク値を（測定によって）特定するために選択された物理的構成は妥当なサイズと範囲になりうる。
4. 最初のスイートスポットについて、セットアップ、試験、および測定指標のピーク値の測定を実施する。
 - ・ SNIA エメラルド™試験プロファイルの全シーケンスを網羅するように試験を行う。
 - ・ 試験、妥当性確認、および予測された結果とのデータ相関を行う。
5. 対象となる追加のスイートスポットについてそれぞれ再構成と再試験を行う。

各スイートスポットに、特定の試験プロファイルに関する指標のピーク値を発生する調整された構成が存在する。しかし、1つの調整された構成から実際には関連するワークロード（すなわち、ランダムまたは順次）について複数の測定基準ピーク値が発生することがある。スイートスポットデータを提示する際には、特定の試験プロファイル「X」において最高の性能を発揮するように最適化されたシステムとしてSUTを特定することが有利であると考えられる。

5.3 推定ツール／シミュレーションツールの検討

全製品／ファミリー範囲の構成について指標のピーク値を見つけ出す作業に直面する際に、推定ツールが非常に有用な手段になることがある。ストレージベンダは、そのストレージ製品向けとして各種の電力計算機や性能推定ツールを整備している場合がある。限定された一連の電力効率指標を予測出来るツールを備えているベンダもある。これらのツールは複雑なシミュレーション方法に基づき、補間と外挿によって求められるいくつかのデータポイントを根拠とするか、またはこれらのいずれかを適用することが可能である。予測の精度は常に疑わしいため、予測された結果についてそれが正確であることを主張する前に、完了したデータ相関を常に明確化する必要がある。試験データレポート（TDR: Test Data Report）には、顧客がこれらの計算を実施することも可能にする仕様データシートが記載されている場合がある。

5.4 実施例

代表的なオンライン3アレイに関して電力計算機および性能推定ツールを使用して、SNIA Emerald™プログラム試験プロファイルについて性能、電力、および電力効率指標の特性プロットをいくつか作成した。アレイコントローラの性能オプションをハイレベルに固定し、ドライブタイプとドライブ数の構成変数を評価した（注：この一連の実施では、SSDは評価しなかった）。このアレイの最大構成サイズは、240 LFF（Large Form Factor）ドライブおよび／または450 SFF（Small Form Factor）ドライブである。予測実施の目的は、各試験プロファイルについて電力効率指標のピーク値を見つけ出すことである。いくつかのプロット図を図3、図4、および図5に示す。

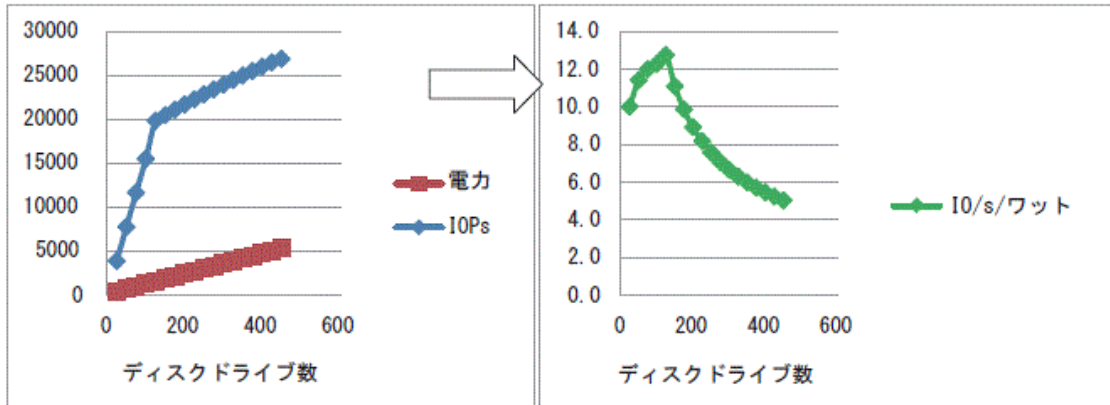


図3 性能、電力、および電力効率指標値 対 ドライブ数
[SFFのランダムワークロード、15K rpm SAS ドライブ]

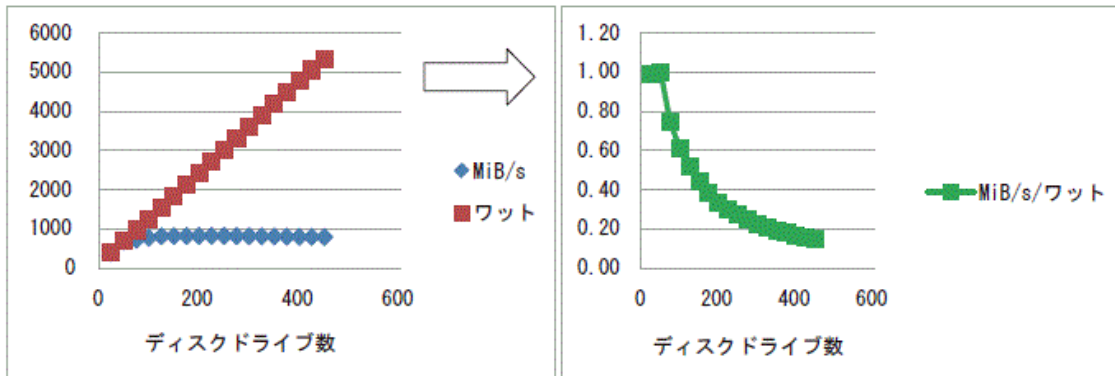


図4 性能、電力、および電力効率の測定指標値 対 ドライブ数
[SFFのシーケンシャルワークロード、15K rpm SAS ドライブ]

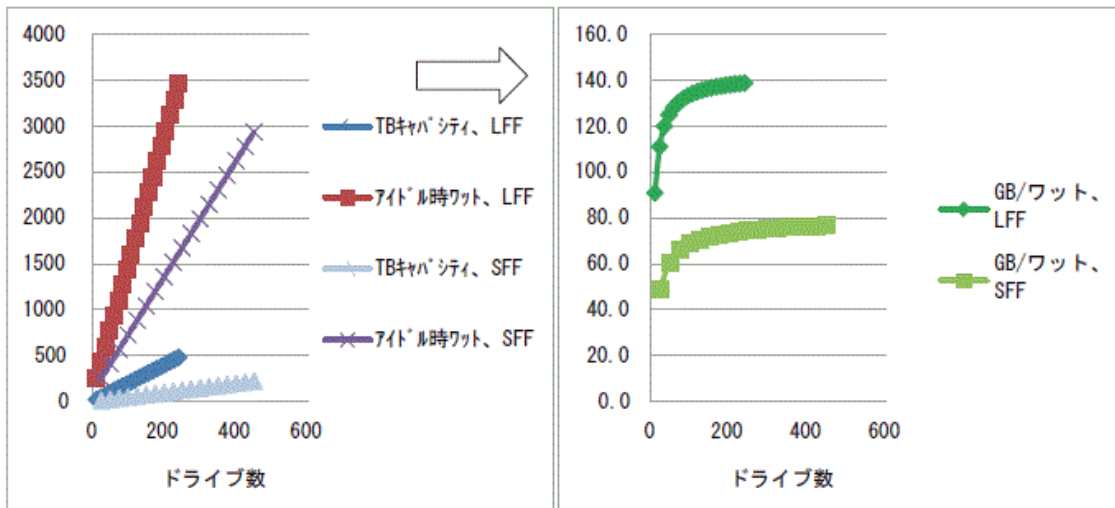


図4 性能、電力、および電力効率の測定指標値 対 ドライブ数
[SFFのシーケンシャルワークロード、15K rpm SAS ドライブ]

明確なことであるが、コントローラの性能、帯域幅、およびハードウェア効率に基づき、上記曲線の勾配と形状は変化する。しかし、以下の例からこれらの観察を行うことができる。



- ・すべてのケースで、構成サイズが大きくなるに従って電力は絶えず、かつ一様に増加する。
- ・すべてのアクティブなケースで、性能は最大のドライブ数よりも大幅に小さい構成時にピークに達し、その後で横ばいになるか、またはわずかに低下する。
- ・すべてのアクティブなケースで、比較的少ないドライブ個数の構成時に指標のピーク値〔性能／電力〕にも達する。
- ・ランダムおよびシーケンシャルワークロードの場合、SFF、15K rpm のスピニングドライブのときに、測定基準のピーク値に達した。
- ・アイドルのケースでは、(コントローラの電子回路の電力はドライブ数の増加に従って償却されるため) 測定指標のピーク値はドライブ数の増加に伴って上昇し続ける。

6 測定試験の設定と実施

「測定仕様」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、及び仮想メディアライブラリのカテゴリの、ストレージ電力効率指標を導き出す手順を記述している。全ての手順は基本的な同一のフローに従っているが、カテゴリ毎の本質的な性質により、それぞれ差異がある。

このセクションではオンラインとニアオンラインのカテゴリに注力する。これらふたつの試験手順はほぼ同じなので、「測定仕様」ではひとつの手順を用い、違いがある部分だけを補足している。リムーバブルメディアライブラリと仮想メディアライブラリについても記述している。

セクション 6.1 から 6.5 では試験構成、ベンチマークドライバ及び電力計の要件、並びに試験手順と指標の計算方法を詳細に記述している。

6.1 設定

6.1.1 試験の構成

ストレージシステムの電力効率指標測定は、データセンター環境で実施することが想定されている。ストレージシステムの入力電源は、「測定仕様」に記載された電圧要件を満足する必要がある。温度と湿度などの環境面は、「ASHREA 2012 サーマル指針クラス A1 仕様」に準拠するものとする。

一般的な設定を図 6 に示す。

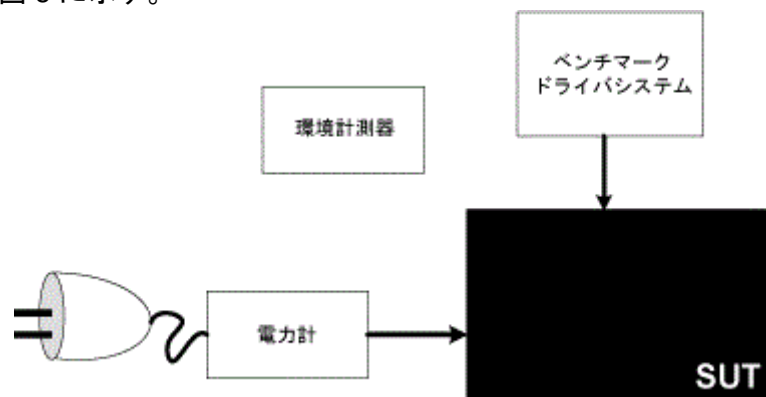


図 6 基本的な電力測定設定



6.1.2 ベンチマークドライバの要件

Vdbench ドライバはワークロードの生成とデータ収集の目的で使用される。このドライバと関連するスクリプトが、「測定仕様」に挙げられている必要なワークロードを生成する。指定されたバージョンの Vdbench とスクリプトの組み合わせを試験中ずっと使う必要がある。スクリプトはまた試験のスポンサーが特定の試験構成で最適な結果（BFF）を出すためのパラメータを設定可能である。

「測定仕様」は下記の 5 つのアクティブ試験のワークロードを定義しており、その実際の IO がスクリプトで実現されている。

- ・ホットバンド
- ・ランダムリード
- ・ランダムライト
- ・シーケンシャルリード
- ・シーケンシャルライト

ホットバンド試験は、リードキャッシュの効果を見るための、13 種類のランダムとシーケンシャルなワークロードの組み合わせである。残りの 4 つの試験は業界で 4-コーナーと呼ばれる有名な試験である。オンラインとニアオンラインの試験手順はこれらの 5 つのワークロードを全て用いる。リムーバブルメディアライブラリと仮想メディアライブラリの試験手順はシーケンシャルリードとシーケンシャルライトのワークロードだけを用いる。

Vdbench スクリプトを実行する前の推奨事項をいくつか以下に記載する。

- ・ Vdbench に精通していない場合には、Vdbench パッケージに含まれているユーザガイドを読むことを特に推奨する。
- ・ 各自のストレージニーズに基づいたエネルギーの最適利用のためにシステムを構成する作業に役立つ利用可能ツールを使用する。
- ・ 可能な限りホストを 1 台だけ使用するように努める。これを順守すれば、測定の実施が大幅に容易になる。

6.1.3 電力計の要件

電力計はベンチマークドライバと同期して、試験中に SUT の正確な電力消費量を測定する。電力計の要件を表 1 に記載する。

表 1 電力計の要件

消費電力(p)	最低精度
$p \leq 10 \text{ W}$	$\pm 0.01 \text{ W}$
$10 < p \leq 100 \text{ W}$	$\pm 0.1 \text{ W}$
$p > 100 \text{ W}$	$\pm 1.0 \text{ W}$



6.1.4 推奨の電力計

推奨される電力計のリストは、「測定仕様」の付属書 A に掲載されている。

6.2 試験の手順

6.2.1 オンラインおよびニアオンライン

「測定仕様」ではオンライン試験とニアオンライン試験の手順を下記のように定めている。

1. ランダムなデータで SUT を満杯にする事前ロード試験
2. SUT を既知の状態にする SUT コンディショニング試験
3. アクティブ指標のためのデータを収集し、収集したデータの妥当性を確認するアクティブ試験フェーズ
4. 容量指標のためのデータを収集するアイドル試験
5. ストレージシステムが容量最適化手法を持つかどうかを見るための COM 試験

これらの試験は、COM 試験を除いて、中断のないシーケンスで実施する必要がある。アクティブ試験やアイドル試験の途中で COM の機能を動作させるかどうかはベンダ試験のスポンサーの判断に任されている。

Vdbench ワークロードジェネレータは事前ロード試験、SUT コンディショニング試験、及びアクティブ試験の際に、SUT にワークロードを提供するために使用される。アイドル試験は外部 IO (ホストからの IO) を必要としないが、ネットワークに接続され、IO 要求を処理する準備が完了している必要がある。

COM 試験はそれぞれ専用の実行方法があり、SNIA Emerald™ の Web サイトにあるプログラムで生成する特別なデータを使用する場合がある。6.3 容量最適化手法 (COM) 試験を参照。

SUT の電力効率を適正に表すには、SUT に事前に適切にデータをロードして満杯にし、事前調整しなければならない。これが事前ロード試験と SUT コンディショニング試験の目的であり、この試験はアクティブ測定が開始される前にシステムを既知の状態にさせるように設計されている。事前ロードフェーズでは、2 対 1 に圧縮可能なデータによってストレージシステムに必要なだけのデータを満たす。どれだけの量のデータを満たす必要があるかは「測定仕様」で規定されている。

ホットバンド IO プロファイルで構成されている SUT コンディショニング試験は、SUT が I/O 要求を満足する能力の実証、システムストレージデバイスが完全に動作することの確認、動作温度の達成、およびストレージシステムを既知の状態に保つことを目的としている。最小の SUT コンディショニング試験持続時間は、「測定仕様」で規定されるが、必要に応じてこの時間を長くすることができる。オンラインシステムは、SUT コンディショニング試験の最後の 4 時間にわたり平均応答時間が 20ms 未満である必要がある。ニアオンラインシステムには平均応答時間の要件がない。

それぞれのアクティブ試験のフェーズでは連続した 30 分 (1800 秒) の測定間隔が必要である。この間に収集されたデータは主要な指標 (「測定仕様」のセクション 8) の計算に使用される。この 30 分間では応答時間や安定性指標などの有効性基準を満足しなければならない。



アクティブ試験中の全ての有効性基準を満足する、任意の連続した 30 分間に収集したデータが主要な指標の計算に使われる。

アクティブ試験フェーズのデータは 1 分間 (60 秒) の間隔で収集される。「測定仕様」に詳細があるが、性能データは 1 分間隔で収集され平均値が計算される[O (60)]。電力計は同じ 1 分間のうちの 5 秒毎の電力使用量を収集し平均する[PA (60)]。これら性能の平均値と電力の平均値からワット当りの性能値が計算される[EPP (60)]。

連続した 30 分の測定が有効であるためには下記が必要である。

- ・ 1 分毎の平均応答時間[RTA (60)]がアクティブ試験で定められている 80ms の基準を満たす必要がある。これはホットバンド、ランダムリード、およびランダムライトにのみ適用される。
- ・ 30 分間全体の平均応答時間[RTA (1800)]がアクティブ試験で定められている 20ms の基準を満たす必要がある。これはホットバンド、ランダムリード、およびランダムライトにのみ適用される。
- ・ 30 個の連続した EPP (60) の値が安定である。

安定の定義は 30 個の連続した 1 分間隔の間、EPP (60) の値の変動が少ないことである。変動が少ないかどうかを決めるのに下記の二つの方法が使われる。安定であると言えるにはこの両方の方法が必要である。

- (a) 候補となる 30 個の EPP (60) の線形近似の傾きの最大許容値
- (b) 30 個の候補となる EPP (60) を移動平均で平滑化した値と基準との比較

(a) において、安定かどうかは 30 個の候補の値を最小二乗法により線形近似して決めることが出来る。線形近似の結果得られた傾きの絶対値が「測定仕様」で規定された値より大きくないことが条件である。

(b) においては、30 個の候補の値に指定された平滑化関数を適用する。結果として得られた全ての値と定義された基準値との偏差が、「測定仕様」で規定された有効範囲内であることが条件である。

アイドル試験は少なくとも 2 時間継続する必要がある。最後の 2 時間の値が指標に使用される。

6.2.2 リムーバブルメディアライブラリ

リムーバブルメディアライブラリの試験手順には事前ロードや COM 試験がない。またコンディショニング試験やアクティブ試験には応答時間の要件がない。アクティブ試験ではシーケンシャルリードとシーケンシャルライトのワークロードだけを用いる。

「測定仕様」では、これらのシステムのシーケンシャル I/O 特性により、その電力効率測定値が公称データスループットの 80%以内に維持されることを要求している。システムの公称スループットは、各種のメディアデバイスのスループット平均値を求めることによって計算しなければならない場合がある。



6.2.3 仮想メディアライブラリ

仮想メディアライブラリの試験手順には事前ロードや COM 試験がない。またコンディショニング試験やアクティブ試験には応答時間の要件がない。アクティブ試験ではシーケンシャルリードとシーケンシャルライトのワークロードだけを用いる。

「測定仕様」では、これらのシステムのシーケンシャル IO 特性により、その電力効率測定値が公称データスループットの 90%以内に維持されることを要求している。

6.2.4 データ収集の纏め

「測定仕様」から直接引用したカテゴリ毎のデータ収集要件を表 2 に纏める。

表 2 データ収集の要約

試験	収集間隔 (秒)		ベンチマークドライバが収集するデータ		最小試験持続時間 (分)	
	電力計	温度計	オンライン/ ニアオンライン	リムーバブル/ 仮想	オンライン/ ニアオンライン	リムーバブル/ 仮想
コンディショニング	5	60	応答時間 (1 分間隔)	スループット (MB/s)	720	7
アクティブ	5	60	応答時間 (1 分間隔)	スループット (MB/s)	30	30
アイドル	5	60	該当なし	該当なし	120	120

6.3 容量最適化手法 (COM) 試験

COM 試験は実質的には存在試験であり、オンラインとニアオンラインカテゴリのシステムのみ適用される。

COM はある種の (潜在的には相当な効果がある) ストレージ効率化機能であり、アクティブ試験実施中にその存在に気付くのは難しい。信頼度の高い手法を提供するために「測定仕様」では COM の存在と動作状態を突き止めるための、一連の発見的な手法に基づく試験を用いる。これらの試験の目的は、試験中システムが指定された COM の機能を本当に持っているかを第三者が検証する為の方法を提供することである。

COM には、以下の項目が含まれる (しかし将来においてはこれらに限定される訳ではない)。

- 1) デルタスナップショット (読み出しと書き込み)
- 2) シンプロビジョニング
- 3) データ重複排除
- 4) 圧縮
- 5) パリティ RAID

発見的な手法の試験の意味は COM があるか無いかを調べるだけで、効率を測定するのではない、ということである。従って、それぞれの試験は単純な yes/no 試験である。試験のスポンサーはどの試験を実行し存在証明を得るかを選択する。全ての試験は比較的分かりやすい



がCOMの種類により種類がある。実際にサポートしているCOMの機能によって、実行するテストが決まる。

ベンダは所定のSUT上で存在を証明したいCOMの所定のステップに従わなければならない。試験の途中でメディアの追加または削除、もしくはメディアの状態変更（オンラインにする、オフラインにする、スペア品として使用するか、あるいは組み込むなど）をすることは出来ない。またRAIDグループを変更することはできない。試験の途中でディスクに不良が発生し自動的にRAIDのリビルドが発生した場合には、リビルドが完了し、インストールした動作中のシステムのメーカーの指針に従って不良ディスクを交換した後で、試験をやり直す必要がある。

存在を示すためにある種のデータセットを必要とするCOM試験がある。これらのデータセットはCOM試験データセット生成Cプログラムが生成する。本プログラムはwww.sniaemerald.com/downloadからダウンロード出来る。このプログラムは試験の前にコンパイルして試験のホストにロードされる。プログラムの使い方はREADMEファイルに書かれている。

このプログラムは三つの約2GBのデータセットを生成する。

- －削減不可能なデータセット：圧縮と重複排除どちらを使ってもサイズをあまり減らせないデータセット
- －重複排除可能、圧縮不可データセット：重複排除ではサイズを相当減らせるが、圧縮では減らせないデータセット
- －圧縮可能、重複排除不可データセット：圧縮ではサイズを相当減らせるが、重複排除では減らせないデータセット

これらのデータセットにより複数のCOMが動作しているシステム、言い換えると個別のCOMの動作を止めることが出来ないSUTの試験が可能になる。

6.3.1 デルタスナップショット試験

事前と事後の空き容量によりデルタスナップショットの有無を調べる。割り当てられた空間と試験データセットから成るコンテナを定義する。

コンテナの空き容量は試験の開始時に決定される。COM試験が実施された後に再度空き容量を調べる。ふたつの空き容量の値から存在の閾値を指定する。このプロセスは下記のステップから成っている。

- a) SUTにコンテナを作成し、その空き容量を調べる。
- b) 2GBの削減不可能なデータセットをコンテナに書き込み、読み出し専用スナップショットは書き込み可能スナップショットを作成する。
- c) 作成したスナップショットのタイプにより、幾つかのデータを読み出すか書き込む。
- d) 大量のストレージ空間が使用されているかどうかを確認するために、コンテナの空き容量の確認を行う。



読み込み専用および書き込み可能なデルタスナップショットは別個に扱われるため、読み出し専用スナップショットのみに対応するシステムはこれらに対する存在証明を取得することができる。

6.3.2 シンプロビジョニング試験

シンプロビジョニング試験は簡単である。

- ・もしシンプロビジョニング機能が無ければ、SUT に有効容量以上の容量を割り当てられない。
- ・もしシンプロビジョニング機能があれば、SUT に有効容量以上の容量を割り当てられる。

6.3.3 パリティ RAID 試験

パリティ RAID の存在は RAID1 との比較による簡単な計算式で確認できる。RAID グループのサイズとパリティの要件は単純で広く周知されているからである。

6.3.4 データ重複排除と圧縮の試験

データ重複排除と圧縮では基本的な COM 機能を調べるために事前と事後の空き容量を使用する。割り当てられた空間と試験データセットから成るコンテナを定義する。コンテナの空き容量は試験の開始時に決定される。COM 試験が実施された後に再度空き容量を調べる。ふたつの空き容量の値から存在の閾値を指定する。

これらの試験では全ての3つの生成されたデータセットを用いる。従ってこれらのデータセットが重複排除や圧縮の別の実装にも有効なことが重要である。

SUT の中には重複排除や圧縮機能が実行される為の最小サイズをもっているものがあるかもしれない。この場合 2GB 以上のデータセットが必要になることがあるが、定義された 2GB のデータセットを組み合わせてより大きなデータセットを作ることが出来る。例として、少なくとも 8GB のデータセットサイズが無いと重複排除機能が働かないストレージを考えてみる。

必要なデータは、定義された重複排除可能な 2GB のデータセットと、少なくとも3つの削減不可能な 2GB のデータセットを連結して作ることが出来る。3つの削減不可能なデータセットはそれぞれの間でも重複排除不可能でなければならない。このためそれぞれの削減不可能なデータセットは異なるベースとなる値から作られる。これにより削減不可能なデータセットの組み合わせも全体として削減不可能となる。このためにはベースの値として素数を用いることが提案されている。圧縮のデータセットにも同様の作り方が適用可能である。

重複排除可能なデータセットは、ブロック境界にアラインされた多くの重複したブロックからなる、さまざまなサイズの多くの重複したファイルを含んでいる。それはまたブロック境界にアラインされていない、様々な長さの重複したブロックを含んでいる。これにより実際に使用された時にブロック方式、可変長方式、SIS 方式などの方式を検地できる。重複排除に関してより深く理解するには DDSR SIG による "Understanding Data Deduplication ratios" を参照のこと。この文書は下記の Web サイトに掲載されている。



http://www.snia.org/sites/default/files/Understanding_Data_Deduplication_Ratios-20080718.pdf

圧縮機能の存在試験に関する他の懸念として、データセットが様々な圧縮方法に対応しているか、というものがある。存在試験は如何に圧縮するか、或いはどんなタイプのデータが圧縮されるかに関しては判断しない。この目的のためにはデータセット生成ツールはビットレベルとパターン指向の両方の方法を提供する。

6.4 測定時に発生する可能性のある落とし穴の回避

ストレージシステムは非常に複雑であるため、測定時にシステム上で発生する可能性のあるすべての落とし穴に「測定仕様」が対処できるとは限らない。このセクションでは、個々の測定の実施時に試験のスポンサーが対処しなければならないと考えられるある種の問題点を列挙する。これはすべての包括的リストを意図するものではなく、測定を実施する前に考慮および／または対処する必要のある全般的項目のリストである。

- ・安定したシーケンシャルリード試験のためには十分なライトデータが必要なため、シーケンシャルライト試験時に十分な書き込みが含まれていることを確認する。
- ・測定時には、すべての開示された RAS 機能をオンに設定する必要がある。バッテリー充電などの作業は通常の日常的作業ではないため、測定を開始する前に完了しておく。
- ・ワークロードジェネレータと電力計／温度計との間のタイミングを一致させる必要がある。これにオフセットが生じると、測定指標の生成がオフになることが考えられる。時間設定は互いに 1 秒以内とする。
- ・ストレージシステムにワークロードを提供するホストがボトルネックになってはならない。「測定仕様」では、Vdbench とそのスクリプトにユーザ設定変数が無い限り、クライアント側の設定を必要としない。Vdbench はその仕様通りに使用すること。良い結果を出すには、クライアントの適切なサイジングが必要である。

SNIA はストレージシステムの電力消費に大きな影響を持っている

Reliability/Availability/Serviceability (RAS) 機能を特定し、「測定仕様」に列挙している。これらの RAS 機能は、最近の高い可用性とサービス性をもつストレージシステムの要件である。これらの機能に関する問題は、これらが存在することにより消費電力は増えるが、性能には直接的なプラスの影響が無いので、「測定仕様」の指標には有害なことである。

6.5 報告する指標

「測定仕様」の 8 節では分類カテゴリ毎に、最終的な電力効率の計算方法と COM の指標に関して詳細に記述している。アクティブ試験の値は 30 分（1800 秒）の測定間隔の間の有効なデータから計算される。アイドル試験は 2 時間（7200 秒）の測定間隔で計算される。最終的な COM の仕様は COM 試験が関連する発見的手法を満足したかどうかで、単純に真（1）または偽（0）で表現される。



7 データ提出時の注意事項

測定結果の SUT 指標と構成情報が組み合わされて SNIA Emerald™プログラム及び EPA ENERGY STAR®データセンター用ストレージプログラムに提出される。

SNIA Emerald™プログラムはデータ提出のための SNIA 向け試験データレポート (Test Data Report: TDR) を用意している (1.2 節の参考資料を参照)。このレポートには、基本システム情報、試験設定情報、および分類カテゴリのために定義された試験段階それぞれの結果を入力するためのエントリが用意されている。データ提出の際に、平均応答時間を TDR に記入する必要があるが、この時間が公開されるかどうかは TDR の平均応答時間を公開するかどうかと言う質問への yes か no の回答による。各ベンダは、レポートを作成・確認し、SNIA Emerald™ プログラムの Web サイト経由で提出できる。EPA は ENERGY STAR®データセンター用ストレージプログラムへのデータ提出の為に、別の、しかし良く似た方法を用いている。詳細は EPA の Web サイトで入手できる (1.2 節の参考資料を参照)。

第 4 章「製品ファミリの定義」と第 5 章「ベストフットフォワード (最適値) を見つけ出す方法」で指摘したように、容量、性能、および電力の間にはトレードオフの関係がある。ストレージシステムのベンダは製品をプロモートする市場ごとに、これらのトレードオフを評価しなければならない。すべての可能なシステムの設定について、1 つの代表的なストレージ電力効率指標を定義することはほとんど不可能である。したがって、ストレージ全体の電力効率を実証するために、複数のシステム構成を SNIA Emerald™プログラムに適宜提出することがベンダにとって得策である。