ENERGY STAR®データセンター用ストレージ
初期データ収集方法の草案
2009 年 11 月

概要

ENERGY STAR データセンター用ストレージ基準の策定作業の一環として、EPA は関係者に対して、本書に規定される方法を使用した、データセンター用ストレージに対する一連の試験と性能モデル化の実施を要請する。

この第 1 回データセンター用ストレージ消費電力試験の目的は、稼働およびアイドル状態の両方における、ハードウェアおよびソフトウェアの構成と、システムのエネルギー性能との関係を理解することである。EPA は、単一変数の構成変化による消費電力への影響について感度分析を実施するために、相当数の試験データと模擬（モデル）試験結果を収集したいと考えている。ハードディスクドライブ（HDD）の選択（例：容量 v.s.性能）、信頼性、可用性、保守（RAS）機能（例：制御装置の単一構成 v.s.冗長構成、RAID 水準）、および小型フォームファクタ（SFF：Small Form Factor）および半導体ディスク（SSD：Solid State Disk）技術の使用などの項目が関心事項である。

提出データは、ENERGY STAR データセンター用ストレージ基準バージョン 1.0 の第 1 草案に向けた準備において、分析用に照合および匿名化され、関係者が利用できるように調整される。

定義および頭字語

データに対する最大時間（MaxTTD：Maximum Time to Data）：データオブジェクト全体が、そのストレージ媒体により課された制限範囲内において利用（アクセス）可能になるまでの最大時間。無作為利用型（ランダムアクセス）媒体の場合、一部のバイトが利用される可能性があるとき、データオブジェクトは利用可能である。順次利用型（シーケンシャルアクセス）媒体については、要求されたオブジェクトが既に非稼働状態となっているドライブからの流入を開始したとき、データオブジェクトは利用可能である。

最大持続性能：UUTが規定の作業負荷のもとで提供可能な最大IOPSまたはGB/s。持続性があると見なされるためには、試験段階の間、性能評価基準が 5%の範囲内に維持されていなければならない。

応答時間：UUTがI/Oによる要求を完了するために要する時間。

標準的システム構成（GSC：Generic System Configuration）：検討の基準として、4 つのGSC選択肢が提案される。

- GSC-1：トランザクション指向の重要構成（Transaction-oriented Value Configuration）
- GSC-2：トランザクション指向の高信頼度構成（Transaction-oriented High-reliability Configuration）

GSC-3：アーカイバル／ストリーミング指向の重要構成（Archival/Streaming-oriented Value Configuration）

GSC-4：アーカイバル／ストリーミング指向の高信頼度構成（Archival/Streaming-oriented High-reliability Configuration）

各構成の仕様は、各製造供給事業者（ベンダー）または試験施設に任せられ、製品のオプション、構成要素の有効性、試験機器、およびその他資源に依存していることに留意する。データ収集方法は、消費電力量に関係するハードウェアおよびソフトウェア構成の詳細を捕捉できるほど十分に綿密なものとなる予定である。

多くの分類上の小集団において、GSC 区分の各区分に対し、標準的な量産型の顧客設定を代表するシステムを設定可能であると思われる。例えば、オンライン-3 分類区分において、GSC-1 システムは、最小の RAS ハードウェアおよびソフトウェア機能を設定すると共に、高速・高性能の HDD または SDD を使用すると思われるが、その一方で、GSC-2 システムは、冗長型制御装置、ミラー化、および他の RAS 機能のような追加 RAS 機能を含む可能性がある。一定の GSC 区分が顧客による実際の実施を代表していないために別の分類区分が存在しており（例：取り外し可能なメディアライブラリ分類区分は、トランザクション志向の GSC-1 および GSC-2 アプリケーションには有用ではない）、またこれら非現実的構成の試験は期待されていない。

対象範囲

本書は、図 1 に示されている、the 0.0.18 DRAFT “SNIA Green Storage Power Measurement Specification” のストレージ分類概要を参照している。追加詳細については、http://www.snia.org/green/にて入手可能である。

図 1：SNIA GSI ストレージ分類

Storage Taxonomy Summary	Online Storage	Near Online Storage	Removable Media Libraries	Virtual Media Libraries	Infrastructure Appliances	Infrastructure Interconnect
	Prime storage, able to serve random as well as sequential workloads with minimal delay	Intended as second tier storage behind Online Storage. Able to service Random and Sequential workloads, but perhaps with noticeable delay in time to 1st data access.	Archival storage used in a sequential access mode. A Typical example would be Tape based archival, both Stand Along and Robotically assisted libraries.	Storage which simulates removable Media Libraries. Will typically use non tape based storage and as such are able to respond to data requests more quickly	Devices placed in the storage SAN or network adding value through one or more dedicated Storage enhancements. Examples include: SAN Virtualization, Compression, De-duplication, etc.	Devices which enable a SAN or other Storage Network data switching or routing
Maximum Capacity Guidance	Max Storage Devices (Up to 80 Ms MTTD)	Max Storage Devices (Over 80Ms MTTD)	Max Tape Drives		Max Storage Devices Supported*	Max Port Count
Group 1) SoHo & Consumer Storage which is designed primarily for home (consumer) or home / small office usage.	Up to 4 Devices	MTTD = Max Time to Data Maximum time needed to access any data stored in any place on the storage system	Stand Alone Drive		Note: * Infrastructure Appliances by definition have no intrinsic storage, other than what is used for local processing and/or local Caching of data.	
Group 2) Entry, DAS, or JBOD Storage which is dedicated to one or at most a very limited number of servers. Often will not include any integrated controller, but rely on server host for that functionality.	More than 4 Devices	Up to 4 Devices	Up to 4 Drives		Storage Devices Support in this case refers to the number of storage devices controllable down stream of the Appliance	Up to 32
Group 3) Entry / Midrange SAN or NAS connected storage which places a higher emphasis on value than scalability and performance. This is often referred to as 'Entry Level' storage.	More than 20 Devices	More than 4 Devices	More than 4 Drives	Up to 100 Devices	Support for up to 20 Devices	Up to 128
Group 4) Midrange / Enterprise SAN or NAS connected storage which delivers a balance of performance and features. Offers higher level of management as well as scalability and reliability capabilities.	More than 100 Devices	More than 100 Devices	More than 24 Drives	More than 100 Devices	Support for more than 20 Devices	More than 128
Group 5) Enterprise / Mainframe Storage which exhibits large scalability and extreme robustness associated with Mainframe deployments, though are not restricted to Mainframe only deployments.	More than 1000 Devices		More than 11 Drives	More than 100 Devices	Support for more than 100 Devices	© SNIA 2009

第 1 回データ収集の目的として、EPA は、以下の分類区分に関心を持っている。

- オンライン：グループ 2、3、4、および 5
- 近似オンライン：グループ 2、3、4
- 取り外し可能メディアライブラリ：グループ 2、3、4、および 5
- 仮想メディアライブラリ：グループ 3、4、および 5

目的

関係者は、データ収集期間において実際に達成可能な限り多くの製品を、可能な限り標準的システム構成（GSC）にて消費電力のデータを提出することが推奨される。可能である場合において、試験およびモデル化は、ハードウェアまたはソフトウェア構成に対する単一変数の操作を行なった後、類似システムに対して繰り返して実施すること。単一変数の変更例には以下が含まれる。

- ディスクドライブの種類および技術の変更
- ストレージ容量の追加または取り外し
- RAID 構成の変更
- 制御装置に対する単一構成から冗長構成への変更

データを生成するために使用されるすべての模擬モデルは、試験手順の様々な段階に関して消費電力量の変動予測が可能でなければならない。すべてのモデルは、該当する分類区分に対して以下に定義される一連の試験作業負荷を用いて実行すること。

モデルの精度を評価するため、関係者は、結果を容易に比較できるように、試験したすべてのシステム構成をモデル化すること。詳細なモデル化は、ハードウェアおよびソフトウェア構成に関する、様々な追加の単一または複数変数の変化による影響を模擬する際に使用できるようになる。

すなわち、EPA は、優先度順に以下のデータを収集することに関心がある。

1. 可能な限り多くの GSC に対する試験から得られたデータ
2. ハードウェアまたはソフトウェア構成に対する単一変数の変更後の GSC 再試験から得られたデータ
3. 試験されたシステム構成のモデル化から得られたデータ
4. ハードウェアまたはソフトウェア構成に対する追加の単一あるいは複数変数の変更に関するモデル化から得られたデータ

試験設定

試験装置

UUT の入力電圧および電力は、その UUT のすべての構成要素による総消費電力を補足するために、適切な場所において電力計測器を用いて測定される。測定は、PDU において、あるいは他の適切な場

所で実施される可能性がある。電力測定には、UUT の一体化と動作を提供するために必要なすべての要素が含まれていること。この要素には、制御装置、トレー、ロボット組み立て部、配電／PDU、さらに UUT 内部で使用されるデータネットワーク（例：一体型 SAN スイッチ）が含まれる。

電力計測器には、以下の能力があること。

- 1%精度および 5 秒以下の測定周波数により UUT の入力電圧を測定し記録する。
- ±1.0 ワット精度および 5 秒以下の測定周波数により UUT の瞬間消費電力を測定し記録する。

入力電圧

UUT に供給される電力は、以下に示される選択肢の 1 つと一致していること。

表 1:入力電力要件

入力電圧範囲	相	交流入力周波数範囲
100-120 VAC RMS	1	47-63 Hz
180-240 VAC RMS	3	47-63 Hz
200-240 VAC RMS	1	47-63 Hz
380-508 VAC RMS	3	47-63 Hz

試験手順

オンライン＆近似オンライン区分

以下の試験手順は、オンラインおよび近似オンライン分類区分のシステムに適用される。

表 2:オンラインおよび近似オンラインに対する試験手順

段階	作業負荷	最大持続可能性能の割合	ブロックサイズ	持続時間[2]
未設定	無作為 70%読み込み 30%書き込み	100%	8K	10 分
稼働「A」	無作為 読み込み	100%	8K	10 分
稼働「B」	無作為 書き込み	100%	8K	10 分
稼働「C」	順次 読み込み	100%	256K	10 分
稼働「D」	順次 書き込み	100%	256K	10 分

段階	作業負荷	最大持続可能性能の割合	ブロックサイズ	持続時間[2]
稼働「E」	無作為 70%読み込み 30%書き込み	25%	8K	10 分
稼働「F」	無作為 70%読み込み 30%書き込み	80%	8K	10 分
稼働「G」	無作為 70%読み込み 30%書き込み	100%	8K	10 分
稼働準備（レディ）アイドル	適用なし	0%	適用なし	30 分
ディープアイドル（任意）[3]	適用なし	0%	適用なし	10 分

注記：

1. 稼働試験段階の間、オンラインおよび近似オンラインのシステムの応答時間は、30ms を超えてはならない。
2. 各試験段階は、ストレージシステム全体が動作し、安定したシステム性能が達成されるように、十分な時間でなければならない。安定性が達成された後、測定時間が開始され、表に規定される「持続時間」にわたり継続される。
3. ディープアイドル試験段階には、1 つまたは複数の記憶装置を電力低減状態にすることにより、ストレージシステムが高度な省電力機能を実行できるようにすることが含まれている。ディープアイドルは、システム運用者により有効化および設定することができる、ディープアイドル機能を提供するシステムに対してのみ実行される。電力低減状態になるため意図的に記憶装置の一部を必要とし、その結果 30msの応答時間要件を満たすことができないストレージシステムは、近似オンライン区分に入ると見なされる。

取り外し可能＆仮想メディアライブラリ区分

以下の試験手順は、取り外し可能メディアライブラリおよび仮想メディアライブラリ分類区分のシステムに適用される。

表 3:取り外し可能＆仮想メディアライブラリに対する試験手順

段階	作業負荷	最大持続可能性能の割合	ブロックサイズ	持続時間[5]
未設定	順次 書き込み→ 巻き戻し→ 読み込み	100%	128K	10 分

段階	作業負荷	持続可能な最大性能の割合	ブロックサイズ	持続時間[5]
稼働「A」	順次書き込み	100%	128K	10 分
稼働「B」	順次読み込み	100%	128K	10 分
稼働準備（レディ）アイドル	適用無し	0%	適用無し	30 分
ディープアイドル（任意）[6]	適用無し	0%	適用無し	10 分

注記：

4. 稼働時試験段階における仮想および取り外し可能メディアライブラリシステムの応答時間は、30ms を超えてはならない。
5. 各試験段階は、ストレージシステム全体が動作し、安定したシステム性能が達成されるように、十分な時間でなければならない。安定性が達成された後、測定時間が開始され、表に規定される「持続時間」にわたり継続される。
6. ディープアイドル試験段階には、1 つまたは複数の記憶装置を電力低減状態にすることにより、ストレージシステムが高度な省電力機能を実行できるようにすることが含まれている。ディープアイドルは、システム運用者により有効化および設定することができる、ディープアイドル機能を提供するシステムに対してのみ実行されること。