

自然科学研究機構 核融合科学研究所様

核融合実験で生成される膨大なデータのアーカイビングにCloudStation dSSを採用

課 題		効 果
逐次、合計で14GBほども生成される100種類ほどの様々なデータを、遅延も手間もなく確実にレプリケーションしたい	➡	データ保存後数秒で自動的に完了するレプリケーション機能により、バックアップリスクと管理者のリソース消費を解消
ハードウェアに依存せず自由に容量追加を行えるクラウド型ストレージを実験データの中長期アーカイビングに導入したい	➡	中長期アーカイブの運用、容量追加に絶大な柔軟性を発揮するKVSを、H/W一体型のアプライアンス形態で余計なコストをかけずに導入
過去のものを含めると膨大なサイズとなる大容量データアーカイブを楽に管理したい	➡	管理GUIのビジュアライザで複数ノードの管理情報をまとめて確認でき、起動停止や書き込み禁止モード切り替えなど設定変更も一括で可能に

CloudStation dSSを導入し、多種のデータで構成される核融合実験の大容量・長期アーカイビングを柔軟で運用に手間のかからないKVSシステムに切り替えることで、ハードウェアに依存しない容量追加やレプリカ遅延の解消といった導入効果を高いコストはかけずに実現しました。

参考：実験ごとに生成される主な計測データ（1GB超のもの）

マイクロ波イメージング反射計(MIR)	2.8GB
燃料ペレット撮像カメラ	2.8GB
遠赤外/赤外線レーザー干渉計	1.9GB
ECEイメージング(ECEI)	1.1GB
マイクロ波反射計	1.1GB
協同トムソン散乱	1.0GB

上記の他、電流・エネルギー、プラズマ位置・形状、赤道面電子温度・密度分布、プラズマ密度分布、イオン温度分布、不純物、周辺プラズマ温度等々、多岐にわたる目的に応じた種々の計測器から、14GB超／回のデータが取得される。各計測器は、直径13.5mにもなる大型ヘリカル装置(LHD)の周囲に設置されている。



自然科学研究機構 核融合科学研究所

世界最大の超伝導プラズマ閉じ込め実験装置である大型ヘリカル装置(LHD)などを擁し、プラズマ・核融合分野の大学共同利用機関として、国内や海外の大学・研究機関とともに双方向の活発な研究協力を進める。優れた人材の育成や社会連携に基づいた基礎的研究・教育の強力な推進も行う。

お問合せ窓口

Tel **03-5213-4370**

〒102-0073 東京都千代田区九段北4-1-3 日本ビルディング九段別館3F
ぷらっとホーム株式会社 営業部
Mail sales@plathome.co.jp / Fax 03-3221-3766

ぷらっとホームおよびPlat'Homeの名称・ロゴは、日本国及びその他の国における、ぷらっとホーム株式会社の登録商標または商標です。
その他、本パンフレットに記載されている会社名および商品・サービス名は、各社の登録商標または商標です。
記載の外観・仕様・価格はすべて2011年12月現在のものであり、予告なく変更される場合があります。

Plat'Home

TECHNOLOGY to serve you.

2011.12

今すぐ導入できる KVS アプライアンス

CloudStation dSS™ R2

巨大なストレージプールをシンプルに管理できる
キーバリューストアを無理なく導入、無駄なく長期運用



容易なノード追加

最大2億ファイルをストア

耐障害性の向上

PB級の単一ストレージを
構成可能

CloudStation dSS™ R2

今すぐ導入できるKVSアプライアンス



概要

Overview

キーバリューストア(KVS)でありながら、既存の環境で今すぐ運用開始できる高い導入性・管理性を備えたアプライアンスが、CloudStation dSS R2です。

KVSによるフラットなファイル管理

KVSであるCloudStation dSS R2なら、一意の文字列をアクセスキーとするため、きわめてフラットなファイル管理が可能です。従来のNASでのディレクトリを必要とする方式と比較し、1000万を超えるような膨大なファイル数のストレージプールをきわめてシンプルかつ効率的に管理できます。またSQLをKVSに置き換えることで、耐障害性の向上やスケールアウト(容量や処理性能の拡張)が非常に容易になります。

科学研究など、膨大なデータの長期アーカイビング用途に

効率的、柔軟、かつかんたんなアーカイブ長期運用ができるCloudStation dSS R2は、逐次発生する膨大なデータの長期アーカイビングを必要とする科学研究分野などで、きわめて大きな導入効果を発揮します。

導入を飛躍的に容易に

これまでのKVSでは、使用中のアプリケーションを一挙に対応させる手間が必要で、導入のハードルを高くしていました。これを克服し、KVSデータに通常のNAS方式でもアクセス可能なのがCloudStation dSS R2です。これにより、無理なくスムーズなKVS導入が可能となりました。

アーカイブを破棄せず更新し長期運用

ストレージプールに自由にノード追加・容量追加ができ、サービスを止めずに機器を新陳代謝することが可能です。膨大なタイムコスト・人的コストを節約しつつ、巨大なアーカイブをきわめて柔軟に長期運用できます。

キーバリューストアとは - シンプルかつフラットなファイル管理

従来のストレージとは一線を画す、シンプルかつフラットなファイル管理の可能なデータストア方式が、KVSです。従来のNAS方式ストレージでは、ハードウェアに依存したファイル位置情報(ディレクトリ情報)とファイル名がデータへのアクセスキーとなりますが、このようなシステムでは、大量のデータを扱う場合などで管理が非常に複雑になります。

しかし、複数のストレージハードウェアを一つの抽象的なストレージプールとして扱い、そこに保存されたデータ(バリュー)を一意のファイル名/文字列(キー)だけ扱うことができるKVSなら、ユーザーはストレージプールの物理的な組成・構成を知る必要がなく、非常にフラットかつシンプルなデータ管理が実現します。

キーバリューストアとそれ以外のスケールアウトストレージの比較

	キーバリューストア	KVS 以外のスケールアウトストレージ
性質	様々な機能を削ぎ落としたシンプルなストレージ。オーバーヘッドの少ない効率的なデータ IO を実現。データベースとして活用すれば、複雑でスケールさせにくくメンテナンス性の悪い SQL などを使わなくてよい。	従来からある HDD などの基本的なストレージ概念を増強・拡張したもの。NAS などとして働き、従来のアプリケーションから通常のファイル IO で利用可能。
共通の機能	耐障害性、スケールアウト、負荷分散	
異なる機能	キーバリューアクセス	ファイル IO アクセス (POSIX)
主な導入目的	データベースやファイルアーカイブの 高効率化、管理の簡略化	基幹システムにおけるストレージの 管理の簡略化・既存ストレージ強化
IO	○ シンプルで高効率なデータ IO	△ ファイルシステムに依存した IO
メリット・デメリット	△ アプリケーションの改修が必要 (→CloudStation dSS R2 なら段階的な移行が可能!)	○ アプリケーションをそのまま使える

CloudStation dSS R2 のアドバンテージ

Advantages

通常のストレージとしても KVS としてもアクセス可能 - 段階的な移行を実現

KVSの機能を保ったまま、通常のNASの方式と同じようにCloudStation dSS R2をマウントできるFUSE (Filesystem in Userspace) クライアントを標準搭載。

いずれのストレージ方式でもアクセス可能

従来のNAS方式ストレージに対応のアプリケーションを、GETとPUTの二つのコマンドで操作できるKVS方式ストレージ上で動かすには、アプリケーションの改修が必要でした。

そこで、CloudStation dSS R2には、KVSであるCloudStation dSS R2を仮想HDD化し、従来のNAS方式でもアクセスできるようにする専用のFUSEクライアントが付属しています。これにより、例えば3ノードのCloudStation dSS R2が、84TBのHDDとして接続されているように見え、通常のHDDのようにディレクトリ作成やファイル読み書きができるため、従来のアプリケーションも改修なしで動かすことができます。

FUSEクライアントからPUT(書き込み)したファイルは、フルパスをキーとし、KVS方式でもGET(読み出し)することができます。またそれとは逆に、KVS方式でフルパスをキーとしてPUTしたファイルを、FUSEクライアントから読み出すことも可能です。

アプリケーションを徐々に改修しながらの導入が可能

NAS方式のアクセスが可能で、アプリケーションを一気にKVS対応させる必要がないため、アプリケーションを徐々に改修しながらの段階的なKVS移行が可能です。また、性能をフルに引き出したい特定のアプリケーションのみをキーバリューに対応させ、それ以外のユーティリティソフトはそのまま、といった使い分けも可能です。

CloudStation dSS R2 ならではの自由度

他のKVSでもFUSEによるアクセスが可能なのはありますが、FUSEクライアント経由で読み書きしたファイルをKVSでもさらに読み書きできるアプライアンスは、CloudStation dSS R2だけです。

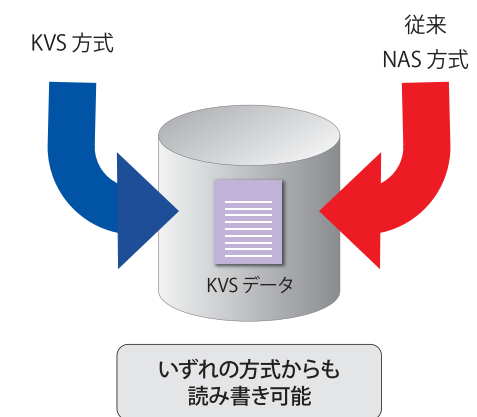
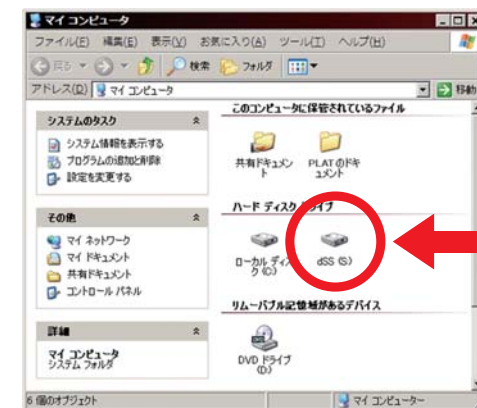
効率的な分散アクセス

クライアントコンピュータにインストールされるFUSEクライアントは、NASプロトコルとは異なり、常にクラスタ内の全ノードに対して分散して、また調子の悪いノードは回避しながら、効率よくデータにアクセスします。そのため、ちょうど接続しているノードが故障しディスクが見えなくなった、といったトラブルも回避できます。

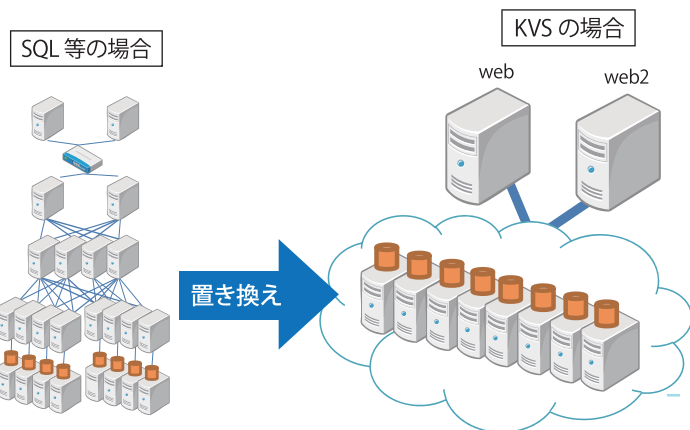
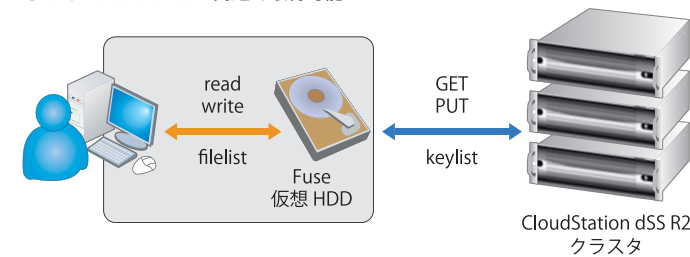
SQL からの置き換えに

主キーのみの簡単な検索にSQLを使用してしまうと、システム管理やスケールアウト、あるいはレプリケーションなどの障害対策に非常に手間がかかってしまいます。

こうした検索にKVSを導入することにより、シンプルでフラットなファイル管理、ノードを物理的に接続・切断だけでよい容易なスケールアウト、そもそも持っているレプリケーション機能や耐障害性など、非常に高い導入効果を発揮できます。



- FUSE クライアント (標準添付) にて 1 台の巨大な仮想 HDD として使用可能
- ディレクトリリストも高速で取得可能



高可用性とボトルネック排除 – マスターノードレス方式

CloudStation dSS R2は、従来の分散ストレージが抱えるSPOF (Single Point of Failure、単一障害点) を排除した「メタデータ同期式マスターノードレスアーキテクチャ」を採用、高い可用性とボトルネック排除を実現しました。

それぞれに特性の異なるデータ分散保存の方法

分散ストレージのアーキテクチャは、それぞれに性能や特性の異なるデータ分散保存の方法によって区別されます。一般的な分散ストレージであれば、ファイルデータはファイル本体とメタデータに分けられ、メタデータ側にファイル本体の保存場所、つまり「ファイル本体が分散保存されているのはどのノードとどのノードか」が記述されています。ファイルの出し入れやレプリカ管理はこのメタデータをもとに行われます。

旧来の分散ストレージでは、このメタデータは、マスターと呼ばれるシステム全体を司るノードに保存されていました。しかしこれではマスターノードがSPOFとなり、故障するとストレージシステム全体が停止してしまいます。

一方、最近主流になりつつあるのが、ファイル本体とメタデータとを両方分散保存する方法です。これならノード追加による性能強化は容易になりますが、ファイルリスト取得に時間がかかる、書き込んだファイルの最新版読み出し可能時期がわからない、などの特性を考慮する必要が出てきます。

メタデータ同期マスターノードレス方式、高可用性とボトルネック排除

CloudStation dSS R2が採用するのは、メタデータを「あえて」全ノードで同期する方式（メタデータ同期マスターノードレス方式）です。すべてのノードがメタデータの完全なコピーを保持し、常に同期処理を行っているため、ファイルアクセスへの高いレスポンスが可能です。もちろん、マスターノードがないため、前述のSPOFの回避、またボトルネックの排除も実現します。さらには、FUSEクライアントによるKVS方式とNAS方式の間でのデータ共有も可能です。

メタデータ同期の場合、ノード追加時の入力IOPS増加は望めませんが、出力IOPSとスループットはメタデータ分散方式と同等もしくはより大きく増加するため、出力の多いアーカイブ用途で高い導入効果を発揮します。

※1 メタデータ：ファイル名、サイズ、所有者など、ファイルに対する付随的(=メタ)データ。

アーカイブを破棄せず更新、長期運用

CloudStation dSS R2の各ノードは、ネットワーク上で自分の状態を常に確認しあい、自律的にクラスタの管理を行っています。このことにより、様々な機能を実現し、容易なシステム構築と、ストレージプールを破棄せず更新し続けるアーカイブなどの長期運用が実現します。

自由にノード追加・取り外し

ノード追加時のクラスタ再構成や容量・出力性能の向上、あるいはノード故障および取り外し時の該当ノード検索・複製データの補完が、すべてサービス無停止で自動的に行われます。

サービスは止めずに機器を新陳代謝

導入後CloudStation dSS R2の新世代モデルが発売されても、既存モデルで構成されたストレージプールにそのまま新世代モデルを追加可能。ノードを一台ずつ新世代モデルに交換すれば、サービスは無停止でストレージシステムの更新ができます。

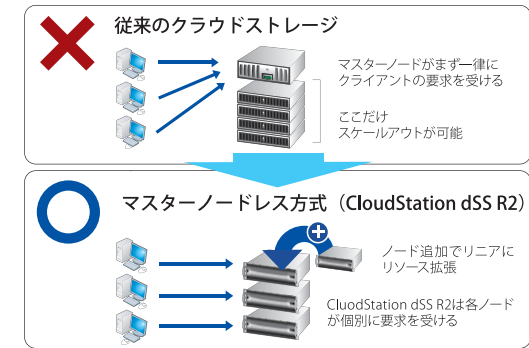
各ノードの性能・容量を考慮

ノードの状態のみならず、その性能や容量も考慮した自動管理が行われます。例えば、容量の小さなノードと大きなノードがクラスタに併存する場合、各ノードの容量は自動的に同じ割合で消費され、全体のパフォーマンス向上のための様々な調整は必要ありません。このため導入するノードの性能や容量をそろえる必要もなく、比較的小容量のノードで構成されたクラスタに、新しく発売された大容量ノードをさらに追加、といったことが可能です。

アーカイブを破棄せず長期運用

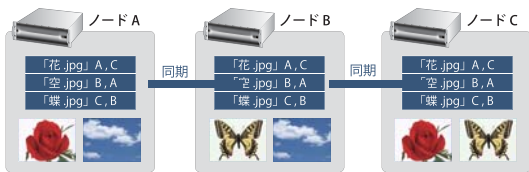
ストレージプールの初期導入時や拡張時、従来は多大な導入計画やコスト資産、余裕を持ったシステム設計が必要でした。容量追加やノード入れ替えが自由に行えるCloudStation dSS R2なら、膨大なタイムコスト・人的コストを節約しつつ、ストレージプール破棄は不要で長期にわたるアーカイブ運用を実現します。

マスターノードを排除

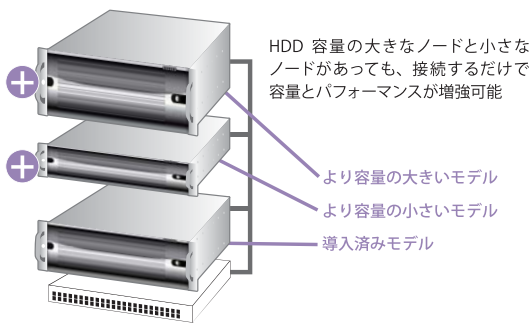
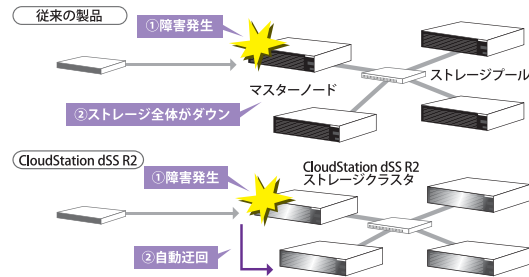


メタデータ同期方式 (CloudStation dSS R2 で採用)

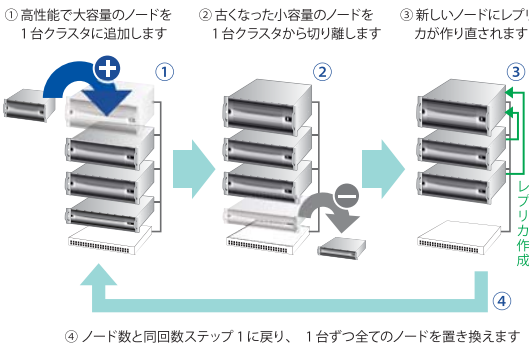
ファイル保存時に各ノードのメタデータを同期する。



単一障害点を排除



サービスを止めずに機器のみ新陳代謝



アプライアンスでスムーズに KVS 導入

CloudStation dSS R2なら、ハード・ソフト・サポートが一体となったアプライアンス製品として、KVSを導入できます。

現在ほとんどのKVSは、Hadoop、Apache Cassandraなどオープンソースプラットフォームを、別のベンダによるハードウェアにインストールする形で運用されます。この場合、運用開始までの手間はもちろん、ソフトウェアと

主な機能

Features

アクセスインタフェース

C言語用API

CloudStation dSS R2へ最も高速でアクセス可能なC言語用ライブラリを提供しています。

GET/PUT

クライアントマシンへクライアントソフトのインストールを行うことにより、GET/PUT操作を基本とするコマンドでの操作が可能です。

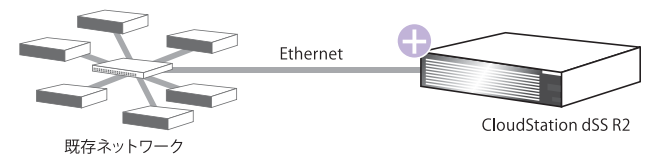
Fuseによるシステムボリュームへのマウント

一般のファイルシステムのように利用可能になります。

容易なシステム構築

ハードウェア性能やネットワーク性能、各ノードの接続状況を認識し自己調整を行うため、システム構築時はIPアドレスとライセンス登録のみで即運用が可能です。

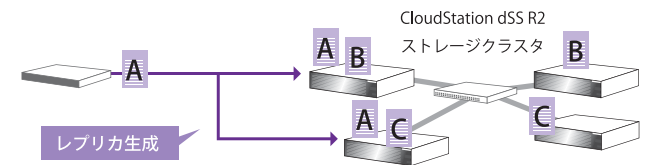
ネットワーク接続した瞬間から運用開始



レプリケーション機能

データファイルは、ファイル単位でクラスタ内の空き領域の大きいノードに優先して保存されます。またクライアントからの保存命令の終了と同時に他ノードへレプリケーションを開始します。レプリケーション数の設定は任意であり、例えば3と設定されれば、オリジナルを含めてクラスタ内のノードの3つへ保存されます。このレプリケーション数は常時監視されており、もし保存先のノードのどれかに障害があった場合は、不足した分のレプリケーションを正常な他のノードに作成します。

レプリケーション機能（自動レプリカ）



人気コンテンツのキャッシュ処理

リクエストを受け付けたノードは、自分が該当のデータファイルを持っていない場合、別ノードにある該当データを中継して返しつつ、そのファイルを自らもキャッシュします。結果、アクセスの多いデータファイルは、最終的に全てのノードがその供給元となります。なお、元データファイルに更新があれば、該当のキャッシュは削除されます。

ハードウェアでサポートの利用が分断あるいは制限され、KVS導入のハードルをいっそう高くすることにつながっていました。

CloudStation dSS R2は、ぶらっとホーム開発のKVSプラットフォームdSS Engine™を自社ハードウェアにインストール済みのアプライアンス製品のため、容易な導入と一括したサポートサービスを実現しています。

スケールアウト

ノードを追加接続するだけで、簡単にスケールアウト（運用開始後の容量拡張・性能拡張）が可能です。

処理を行うノードの数が増えることで、読み出し性能も比例して強化されます。ノード追加は動作中でもサービスを止めずに行えます。

スモールスタート、柔軟に拡張

いつでも柔軟にスケールアウトが行えるため、システム設計時に将来を予測した大きなストレージシステムを組む必要はありません。現況に合わせた規模で初期導入を行い、運用中のデータ増加に伴い随時ノードを追加するなど、スモールスタートからスケールアウトしてゆくシステム管理が可能です。

フェイルオーバーとロードバランス

専用クライアントソフトは、問合せ先のノードに異常がある時、クラスタ内の正常なノードへ処理をフェイルオーバーします。問い合わせ先はランダムに選ばれ、いずれか一つのノードにリクエストが集中することはありません。

グラフィカルな監視・管理

クラスタの監視・管理をPCのWEBブラウザから行うことができます。ストレージ容量、保存バランス、負荷状況などを一目で把握できる、わかりやすい設計が採用されています。障害発生時のエラーログチェックや、手動による障害復旧オペレーションなども一元化できます。



管理画面例

自動ファイルメンテナンス

未アクセスで長期保管されるデータには読み出し不良などの障害が発生しがちですが、CloudStation dSS R2なら、保存ファイルを定期的に読み込み、エラー検出時は正常なレプリケーション先からデータファイルを上書きしリフレッシュすることで、自動ファイルメンテナンスを行います。

仕様と価格

Specifications & Prices

dSS 各モデル比較表（特記なき場合はノードあたり）

システム性能

モデル		2U 8ベイモデル	3U 16ベイモデル	4U 24ベイモデル
製品写真				
初期構成	ノード数	3ノード		
	参考価格(税込・最小構成時)	4,980,000円	5,980,000円	6,980,000円
	型番	PSDS2-2T08-N3	PSDS2-2T16-N3	PSDS2-2T24-N3
追加ノード	ノード数	1ノード		
	参考価格(税込・最小構成時)	1,660,000円	1,990,000円	2,330,000円
	型番	PSDS2-2T08	PSDS2-2T16	PSDS2-2T24
最大データ容量				
ノードあたりデータ容量 (OS領域含む、RAID6、スベアなし)		12TB	28TB	44TB
同一ノード構成時の最大クラスタ容量 (36ノード時)		432TB	1008TB	1584TB
最大ファイル数(クラスタ全体)		2億		

ハードウェア仕様(ノード毎)

キャッシュメモリ	8GB		
SSD搭載	標準搭載		
外部接続	1000Base-T 10GbE(オプション)		
電源	AC100V(47～62Hz)		
消費電力			
アイドル時 ※1	156.98W(160.09VA)	244.10W(251.18VA)	315.27W(318.12VA)
負荷時 ※2	196.93W(200.33VA)	290.95W(298.95VA)	341.87W(347.04VA)
最大時 ※3	249.69W(254.03VA)	338.00W(344.00VA)	389.03W(394.42VA)
エネルギー消費効率 ※4	0.144232689		
省エネ法に基づく表示 ※4	i		
寸法 (mm)	W437×H89×D705	W437×H132×D648	W437×H178×D660
形状 (19インチラックマウント)	2U	3U	4U
重量	21.6Kg	33.7Kg	43.5Kg

- ※1. 消費電力アイドル時
- 標準構成のCPU(Xeon E3-1220)1個、標準構成のメモリ(4GB ECC)2枚、最大搭載本数のHDD(容量問わず)にCloudStation dSS R2 OSを搭載したシステムでdSSクラスタがI/O要求を受けられる状態で5分間動作させた場合の平均
- ※2. 消費電力負荷時(CloudStation dSS R2運用時、負荷が掛かった際の消費電力となります。)
- 標準構成のCPU(Xeon E3-1220)1個、標準構成のメモリ(4GB ECC)2枚、最大搭載本数のHDD(容量問わず)にCloudStation dSS R2 OSを搭載したdSSクラスタ全ノードに対し常にI/O要求がある状態で5分間動作させた場合の最大値
- ※3. 消費電力最大時(ハードウェアの最大消費電力となります。)
- ※4. エネルギー消費効率
- ストレージはサブシステム扱いとなり、エネルギー消費効率は搭載HDD回転数(7,200rpm)によって求められ、区分は「i」となります。
- ◇ 計算式
- エネルギー消費効率E =exp(2.0*ln(N)-19.7) N=回転数(7,200rpm)

オプション

Options

オプション機器

10GbE ネットワーク (SFP+)



IBM BNT RackSwitch G8124

参考価格：1,639,000 円 (税込) ～



Chelsio N320E

参考価格：81,900 円 (税込) ～

IBM BNT RackSwitch G8124 シリーズを使用した SFP+ 接続により、低消費電力かつコストパフォーマンスに優れた 10G システムを実現します。

インターフェイス：10G SFP+ ポート 24 個
消費電力：168W
スループット：480Gbps
形状：19 インチラックマウント 1U
外形寸法：幅 43.9cm、奥行 38.1cm
重量：6.40kg
※ 背面吸気、前面吸気をご指定下さい
専用 SFP モジュール (ポート毎)
参考価格：110,669 円 (税込) ～

インターフェイス：10G SFP+ ポート数：2 ポート
※ ノード毎に増設が必要です
専用 SFP モジュール (ポート毎)
参考価格：27,000 円 (税込) ～

オプションサービス

保証期間延長

ご購入時標準の製品保証に加え、別途年単位でお申込みいただける保証延長の価格を明確化、無理のない長期運用が可能です。

	4U 24ベイモデル	3U 16ベイモデル	2U 8ベイモデル
型番	PSDS2-2T24/1Y	PSDS2-2T16/1Y	PSDS2-2T08/1Y
参考価格 (年額)	438,000円(税込)～	408,000円(税込)～	378,000円(税込)～



設置・設定サービス

機器納品時の設置支援のほか、初期設定・初期構築支援も行います。
ラックマウントからケーブル接続、システム接続まで、システムが安全に稼働するまで万全のサポートを行います。



オンサイトメンテナンス

現地での障害調査、設定のバックアップやバージョンアップ、ソフトウェアのアップデートなどを行います。
専門技術者の派遣により、メンテナンスをより安心に行うことができます。

参考価格 (年額)	初期構成3ノード	240,000円(税込)～
	ノード単位	96,000円(税込)～

Cold-Standby の提供

障害時の予備系統となるCold-Standby機で多重化を行い、信頼性を向上します。



その他の追加サービス

先出しセンドバックサービス、講習・簡易トレーニングなど、
その他の追加サービスも豊富にご用意いたします。