



スケーラブルサーバアプリケーション
ORION SS620Q-G3製品概要
スケーラブルシステムズ株式会社

スケーラブルサーバ アプリケーション



- 最新のプロセッサを搭載したサーバプラットフォーム
 - －サーバ仮想化によるシステム統合
 - －システムはセットアップを完了した状態でご提供
- 基本構成ではバックプレーン用のスイッチなどが不要
 - －各ノード間を直接接続
 - －冗長性の高いバックプレーン構成
- 利用するサーバプラットフォームは、用途や目的に応じて様々な機種から選択可能
 - －高密度実装された4ノードを2Uサイズに搭載したサーバや、2ソケットで1.5TBまでのメモリが搭載可能なサーバなどが利用可能

スケーラブルサーバ ORION SS620Q-G3



ベース構成

- 2Uサイズの8ソケットサーバ
- コンパクトな筐体に多くのコアとメモリを実装し、シングルシステムとして利用可能
- スイッチなし構成でのシステム構築によりラックスペースを無駄にしない構成

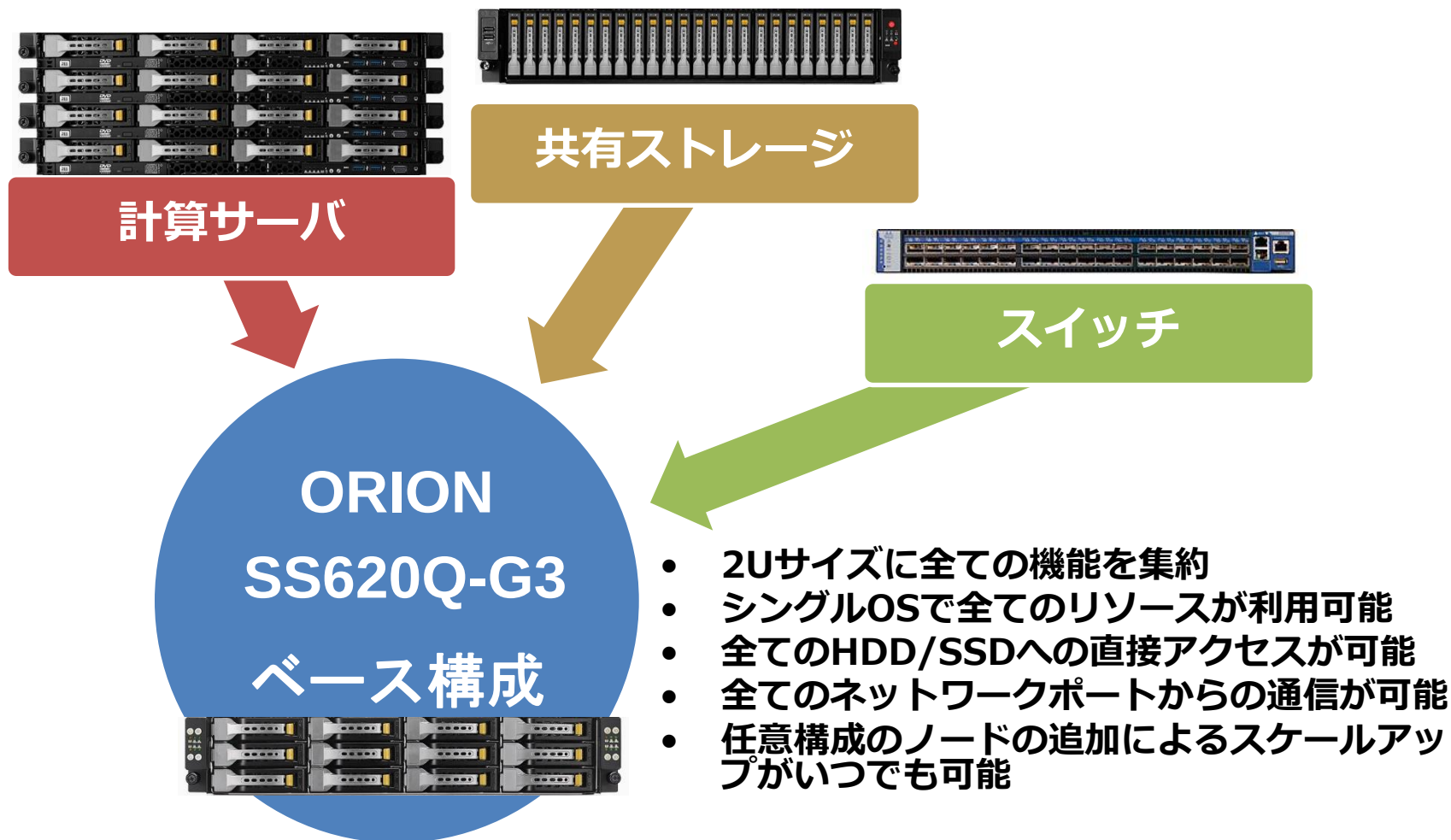
拡張構成

- ベース構成にノードを追加
- ノード追加は任意のサイズと構成で可能

大容量メモリ構成

- 大容量メモリシステムをより廉価に提供
 - 廉価なプロセッサをメモリコントローラとして利用
 - 廉価なメモリモジュールを利用可能
- 詳細なモデル解析やビッグデータの高速処理などに活用
 - インメモリでのデータ処理
 - 大容量メモリをメモリファイルシステムやキャッシュとして利用することでI/O処理を高速化

スケーラブルサーバ ORION SS620Q-G3(ベース構成)



ORION SS620Q-G3(ベース構成)



ベースシステムハードウェア

- 2Uサイズに4ノード搭載
- ノード間InfiniBand直接接続（スイッチなし）
- 冗長バックプレーン構成

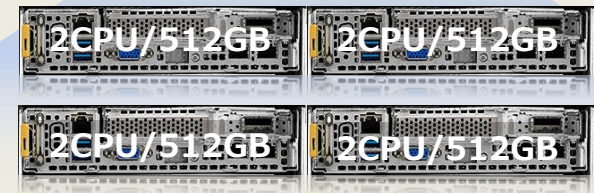
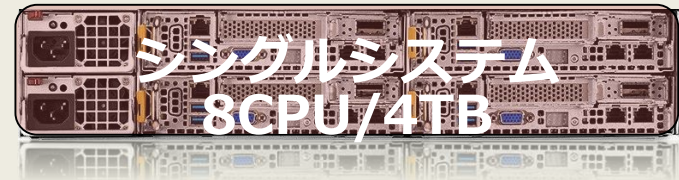
VM構成システム

- シングルシステム（8CPU/4TB）
- 1台のサーバとして利用可能
- クラスタ管理ツールは不用

Linuxオペレーティングシステム

- 標準ディストリビューション
- OSプリインストール
- ストレージ&システムリソース管理ツール
- 初期設定とアプリケーション利用ガイド

ORION SS620Q-G3



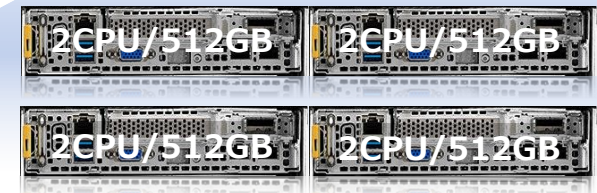
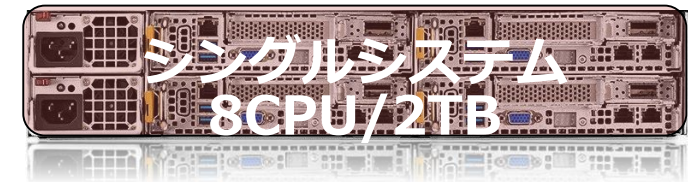
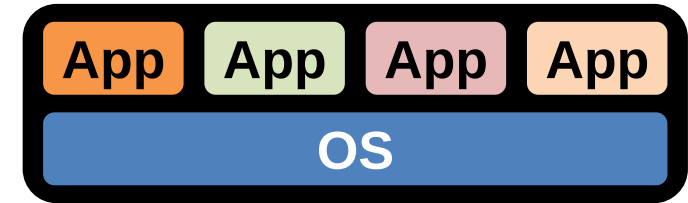
物理サーバリソース

ORION SS620Q-G3(ベース構成)

製品仕様



筐体デザイン	ラックマウント型 (2U)
CPU	インテル XeonプロセッサーE5-2600 v3/v4 製品ファミリー (120W TDPまで)
CPU数	8基 (最大144物理コア、288論理コア)
サポートメモリ	DDR4 2400MHz reg. ECC memory
DIMM数	64 (最大4TB共有メモリ構成)
内蔵ディスク	ホットスワップ可能 12 x 3.5" SATAドライブ (最大72TB)
標準 インタフェース	アナログRGB (ミニ D-sub 15ピン) x 4 GbEx 8 マネージメント用 LANコネクタ x 4



物理サーバリソース



システム管理・動作モニター

```
root@vsmp1:~  
File Edit View Search Terminal Help  
1 [|||||] 100.0% 21 [|||||] 100.0% 41 [ 0.0%] 61 [ 0.0%]  
2 [|||||] 100.0% 22 [ 0.0%] 42 [ 0.0%] 62 [ 0.0%]  
3 [|||||] 100.0% 23 [ | 0.4%] 43 [ 0.0%] 63 [ 0.0%]  
4 [|||||] 100.0% 24 [ 0.0%] 44 [ 0.0%] 64 [ 0.0%]  
5 [|||||] 100.0% 25 [ 0.0%] 45 [ 0.0%] 65 [ 0.0%]  
6 [|||||] 100.0% 26 [ 0.0%] 46 [ 0.0%] 66 [ 0.0%]  
7 [|||||] 100.0% 27 [ 0.0%] 47 [ 0.0%] 67 [ 0.0%]  
8 [|||||] 100.0% 28 [ 0.0%] 48 [ 0.0%] 68 [ 0.0%]  
9 [|||||] 100.0% 29 [ 0.0%] 49 [ 0.0%] 69 [ 0.0%]  
10 [|||||] 100.0% 30 [ 0.0%] 50 [ 0.0%] 70 [ 0.0%]  
11 [|||||] 100.0% 31 [ 0.0%] 51 [|||||] 100.0% 71 [ 0.0%]  
12 [|||||] 100.0% 32 [ 0.0%] 52 [ 0.0%] 72 [ 0.0%]  
13 [|||||] 100.0% 33 [ 0.0%] 53 [ 0.0%] 73 [ 0.0%]  
14 [|||||] 100.0% 34 [ 0.0%] 54 [ 0.0%] 74 [ 0.0%]  
15 [|||||] 100.0% 35 [ 0.0%] 55 [ 0.0%] 75 [ 0.0%]  
16 [|||||] 100.0% 36 [ ||| 19.1%] 56 [ 0.0%] 76 [ 0.0%]  
17 [|||||] 100.0% 37 [ 0.0%] 57 [ 0.0%] 77 [ 0.0%]  
18 [|||||] 100.0% 38 [ 0.0%] 58 [ 0.0%] 78 [ 0.0%]  
19 [|||||] 100.0% 39 [ 0.0%] 59 [ 0.0%] 79 [ 0.0%]  
20 [|||||] 100.0% 40 [ 0.0%] 60 [ 0.0%] 80 [ 0.0%]  
Mem[|||||] 27497/459775MB Tasks: 129, 247 thr; 22 running  
Swp[ 0/0MB] Load average: 28.49 31.66 33.14  
Uptime: 5 days, 03:18:40
```

```
root@vsmp1:~  
File Edit View Search Terminal Help  
vSMP Foundation: 5.5.155.13 (Apr 28 2014 21:11:29)  
System configuration:  
Boards: 4  
Processors: 8, Cores: 80  
Intel(R) Xeon(R) CPU E5-2690 v2 @ 3.00GHz Stepping 04  
Memory (MB): 466461 (out of 524108), Cache: 39919, Private: 17728  
Link Rate: 54.3Gb/s  
Boot device:  
Serial number:  
System key:  
Supported until:  
[root@vsmp1 ~]# numactl --hardware  
available: 8 nodes (0-7)  
node 0 cpus: 0 1 2 3 4 5 6 7 8 9  
node 0 size: 56415 MB  
node 0 free: 8471 MB  
node 1 cpus: 10 11 12 13 14 15 16 17 18 19  
node 1 size: 60672 MB  
node 1 free: 4460 MB  
node 2 cpus: 20 21 22 23 24 25 26 27 28 29  
node 2 size: 56320 MB  
node 2 free: 47653 MB  
node 3 cpus: 30 31 32 33 34 35 36 37 38 39  
node 3 size: 60416 MB  
node 3 free: 55774 MB  
node 4 cpus: 40 41 42 43 44 45 46 47 48 49  
node 4 size: 56064 MB  
node 4 free: 97 MB  
node 5 cpus: 50 51 52 53 54 55 56 57 58 59  
node 5 size: 60416 MB  
node 5 free: 35 MB  
node 6 cpus: 60 61 62 63 64 65 66 67 68 69  
node 6 size: 56064 MB  
node 6 free: 34795 MB  
node 7 cpus: 70 71 72 73 74 75 76 77 78 79  
node 7 size: 60061 MB  
node 7 free: 42989 MB  
node distances:  
node 0 1 2 3 4 5 6 7  
0: 10 20 254 254 254 254 254 254  
1: 20 10 254 254 254 254 254 254  
2: 254 254 10 20 254 254 254 254  
3: 254 254 20 10 254 254 254 254  
4: 254 254 254 254 10 20 254 254  
5: 254 254 254 254 20 10 254 254  
6: 254 254 254 254 254 254 10 20  
7: 254 254 254 254 254 254 20 10  
[root@vsmp1 ~]#
```

標準のLinuxのツールやオープン
ソースの管理ツールやシステム
モニターが利用可能

ORION SS620Q-G3 用途と利点



システムへの要望	用途・アプリケーション	スケーラブルSMPの利点
複数ジョブの同時実行	MCAEシュミレーション 数値解析シュミレーション ・ MATLABなど	柔軟なワークロード処理 対話処理・バッチ処理の双方可能 OSによるリソースの一元管理
大規模共有メモリ	MCAEシュミレーション ・ 陰解法CSMソルバー ・ プリ処理(モデル作成、メッシュ分割) ライフサイエンス	スケーラブルなメモリ構成 ・ 大きなメモリ空間 ・ 高いメモリバンド幅 ・ I/Oボトルネックの解消 ・ メモリファイル&キャッシュ
並列アプリケーション 性能	MCAEシュミレーション ・ 陽解法CSMソルバー ・ CFDアプリケーション	アプリケーション実行性能 ・ 最適化MPIライブラリ ・ 並列処理支援ツール ・ Linux最適化カーネル
スレッドAPI (Java、OpenMPなど) アプリケーション	数値解析シュミレーション アプリケーション開発 ライフサイエンス	スケーラブルな性能 ・ より多くのスレッド利用が可能 ・ 並列処理支援ツール ・ Linux最適化カーネル

MCAE:メカニカル・コンピュータ支援エンジニアリング CFD:計算流体力学 CSM:計算構造力学

仮想化によるシステム統合

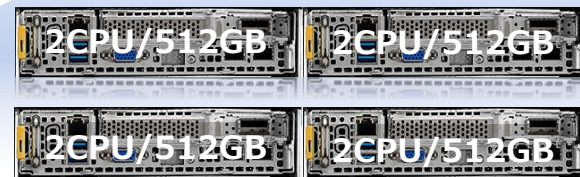
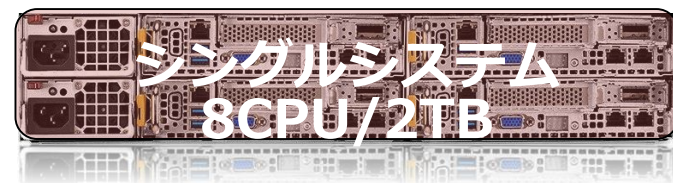
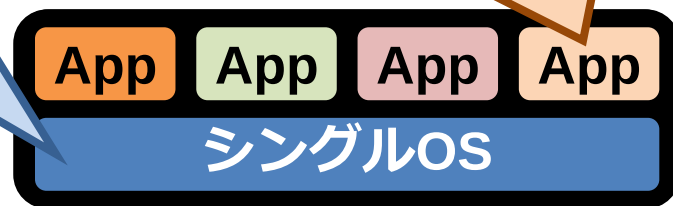


シングルメモリアドレス
全てのリソースへのアクセスが可能

透過的に全てのCPU、メモリ、
I/Oリソースを利用可能



仮想化



物理サーバリソース

管理

- 標準のLinuxオペレーティングシステム
- クラスタ管理ソフトウェア不要

利用

- ワークステーションやPCサーバと同様に利用可能
- スレッドやOpenMPなどでの並列化やMPIなどでのパラレル処理も可能

増設

- ノードを容易に追加可能
- ソフトウェア環境の変更なし



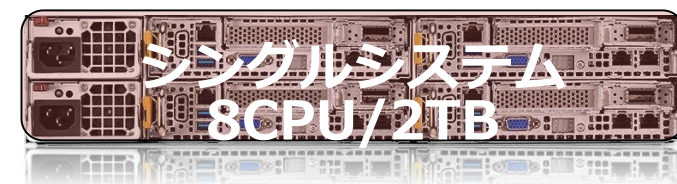
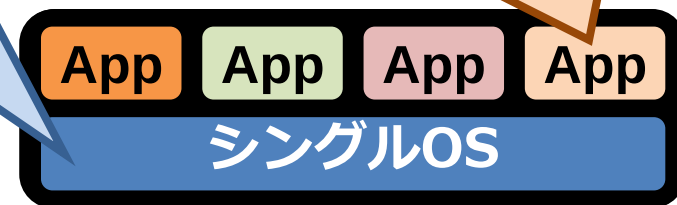
システムの運用管理

シングルOSで複数サーバ分のリソースの管理が可能です。

アプリケーション実行での高い互換性を提供します。

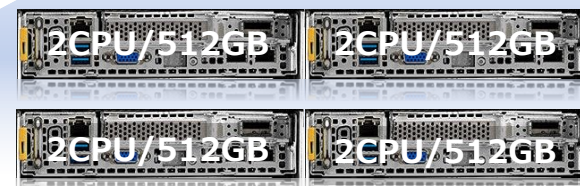


仮想化



質問 複数のシステムを通常のOSだけで容易に管理することは出来ないだろうか？

回答 複数サーバを仮想化することで通常OSだけで複数のシステムを統合し、1台のサーバとして利用可能です。クラスタなどの運用に時間とリソースを掛けることなく、システムを利用することが可能となります。



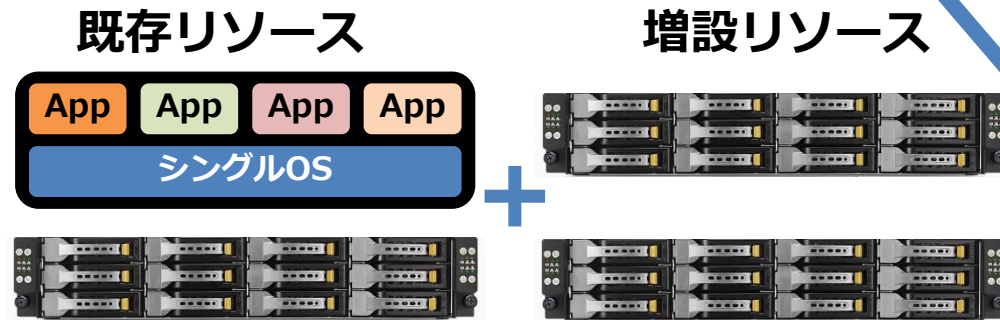
物理サーバリソース

増設・アップグレード



増設・アップグレードする前
と同じOS環境が利用可能

アプリケーション実行は、従来
と同じです。



質問) 計算リソースの増設時に、従来の環境のまま
計算能力の向上を図ることが出来ないか？

回答) ハードウェアを追加しても、vSMPでの仮想化
によって、シングルOSで、全ノードを管理し、アプ
リケーションもそのまま利用可能

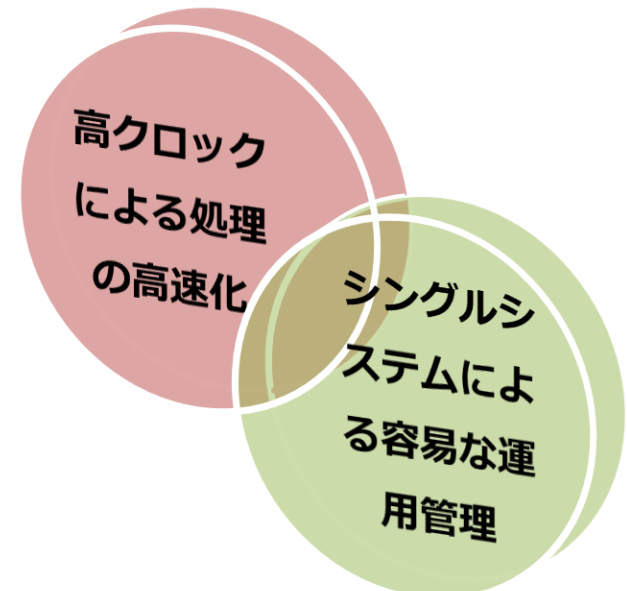


高クロックで多くのコアの搭載

- 最新プロセッサの傾向
 - 高クロック・少ないコア数 .vs. 低いクロック・多くのコア数
- スケーラブルサーバの利点
 - 高クロックで多くのプロセッサコアの利用が可能
 - 廉価で高クロックなプロセッサの選択が可能

プロセッサ仕様比較（＜120W TDP）

	E5-2623 v3 SSMPサーバ	E5-2695 v3 DP サーバ
コア数	4 (32)	14 (28)
スレッド数	8 (64)	28 (56)
プロセッサ・ベース 動作周波数	3.0 GHz	2.3 GHz
ターボ・ブースト利用 時の最大周波数	3.5 GHz	3.3 GHz
TDP	105W	120W





スケーラブルサーバ製品比較

- 2Uサイズに8ソケットを搭載し、大容量のメモリとストレージを提供
- 144物理コア、2TBメモリ、72TB HDDまでの拡張性
- ハイエンドアクセレータ搭載も可能
- 標準Linuxと最適化ツールパッケージ

	DIMMs	サイズ	搭載可能ドライブ	アクセレータ カードサポート
ORION SS	64	2U	12 HDD/SSD	YES (1)
HP DL560	48	2U	5 SAS/SSD	NO
IBM x3750 M5	48	2U	16 HDD/SSD	NO
Dell R820	48	2U	16 SAS/SATA/SSD	NO

(1) アクセラレータカード利用時は、アクセラレータカード用ノードサーバを追加



vSMP Foundation

アーキテクチャとシステム概要



スケーラブルサーバ構成

標準Linuxだけで、全リソースの
管理が可能
全ノードのメモリが利用可能
全てのリソースへのアクセスが可能

アプリケーションは通常Linuxで
利用出来るものならバイナリ互換
透過的に全てのメモリ、CPU、I/O
リソースを利用可能

アプリケーション

オペレーティングシステム

vSMP Foundation

Memory

CPU

CPU

I/O

Memory

CPU

CPU

I/O

Memory

CPU

CPU

I/O

Memory

CPU

CPU

I/O

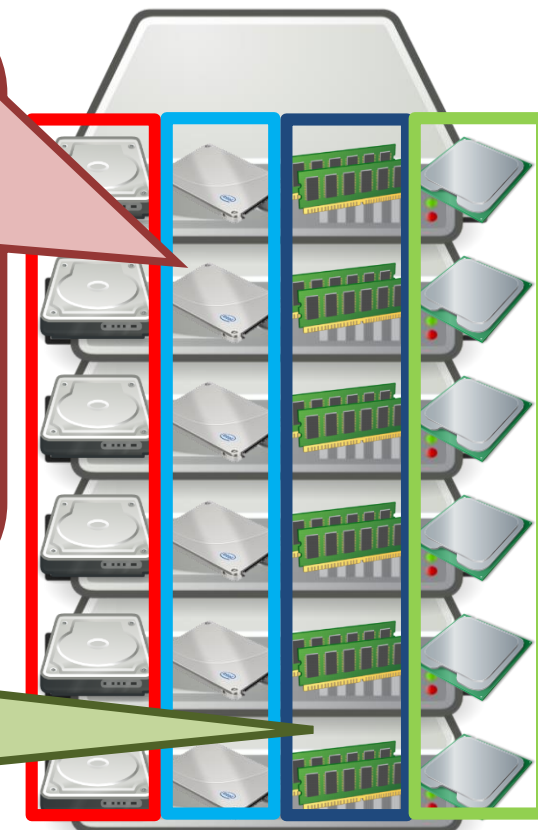


シングルシステム

ソフトウェアRAIDなどの構成によって、同時に複数のHDDドライブへのアクセスが可能

また、各ノードにSSDなどの高速デバイスを搭載して、それらを共有して、大きなリソースプールとすることも可能

tmpfs/ramfsなどのメモリファイルシステムを利用することで、より高速なIO処理が可能



ディスク SSD メモリ CPU

大容量の共有リソース

オペレーティングシステム

APP

APP

APP

APP

APP

全てのリソースへのアクセスが可能
シングルメモリアドレス
全プロセッサコアのスレッドでのアクセス

InfiniBand Fabric

透過的に全てのメモリ、CPU、I/Oリソースを利用可能



ワークロード処理の効率化

大規模メモリを必要とするワークロード

- 大きなメモリ空間を利用
- プログラムの変更なしでデータセットのサイズに柔軟に対応

CPU負荷の大きなアプリケーション

- スレッドアプリケーション（OpenMPやPthreadsなど）
- Java

スケーラブルサーバ



システム運用とアプリケーションの実行をよりシンプルに....

- リソースを統合して運用管理を容易に
- アプリケーションの実行手順をより簡便に

I/O負荷の大きなワークロード

- 大容量ファイル処理
- メモリをI/Oのバッファとして利用
- メモリファイルシステムの活用

スケーラブルサーバの利点



管理が容易

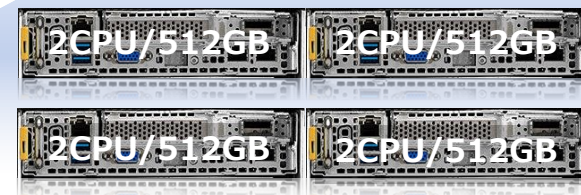
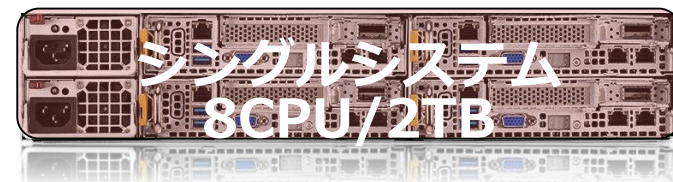
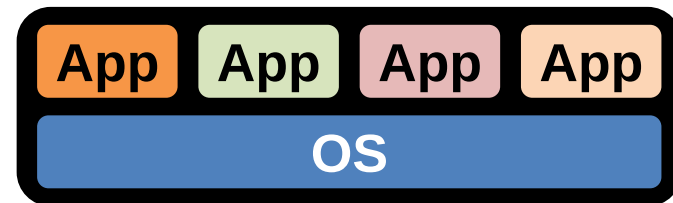
- 標準のLinuxオペレーティングシステムだけ利用可能
- 特別なクラスタ管理ソフトウェアなどを必要としない

利用が容易

- ワークステーションやPCサーバと同様に利用可能
- スレッドやOpenMPなどでの並列化やMPIなどでのパラレル処理も可能

増設が容易

- ノードを容易に追加可能
- ソフトウェア環境の変更を必要としない



物理サーバリソース

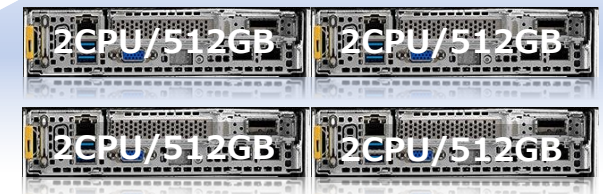


システムの運用管理

課題) 複数のシステムを通常のOSだけで容易に管理することは出来ないだろうか？



スケーラブルサーバ) 複数サーバを仮想化することで通常OSだけで複数のシステムを統合し、1台のサーバとして利用可能です。クラスタなどの運用に時間とリソースを掛けることなく、システムを利用することが可能となります。

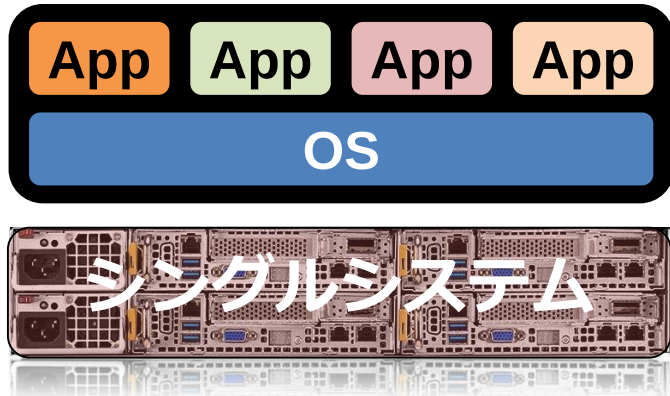


物理サーバリソース

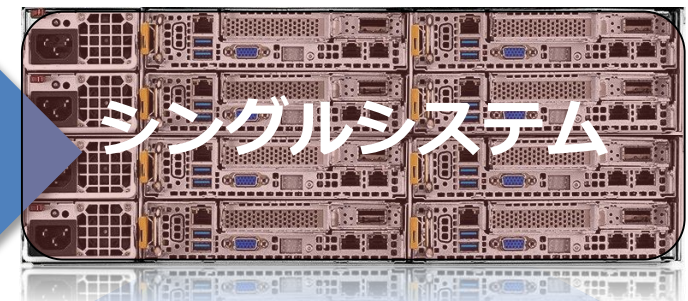
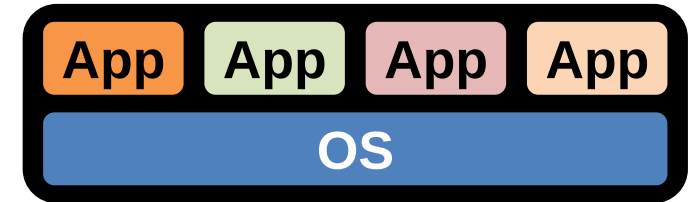
増設・アップグレード



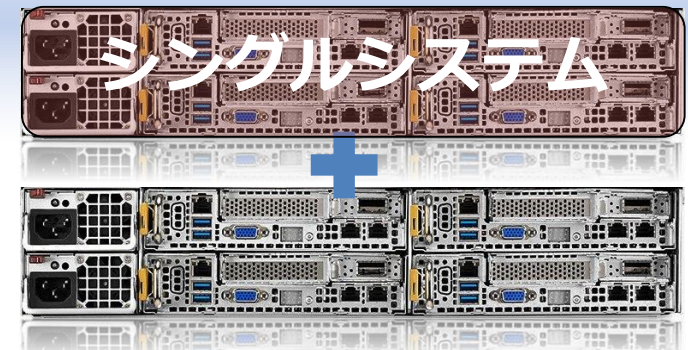
課題) 計算リソースの増設時に、従来の環境のまま計算能力の向上を図ることが出来ないだろうか？



増設・アップグレード



スケーラブルサーバ) ハードウェアを追加しても、vSMPでの仮想化によって、シングルOSで、全ノードを管理し、アプリケーションもそのまま利用可能となります。



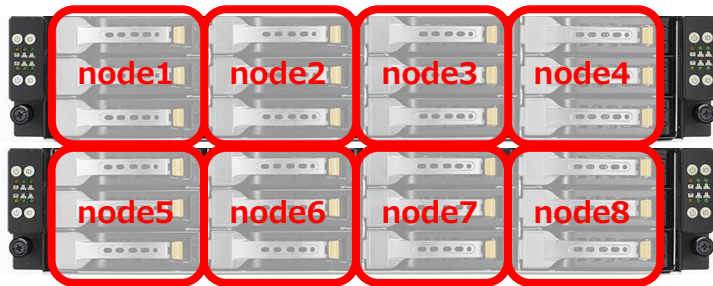
物理サーバリソース増設

ストレージ共有



- 大容量の共有ストレージ（ローカルストレージ）
- パラレルI/O（HW+SW RAID）による高速I/O
 - SSD/HDDを目的に応じて選択

通常のクラスタ構成



ストレージは個々のノードに分散

データ共有：共有ファイルシステムが必要
アクセス速度：ローカルディスク（DAS）
管理運用：個々のディスクを別々に管理

スケーラブルサーバ



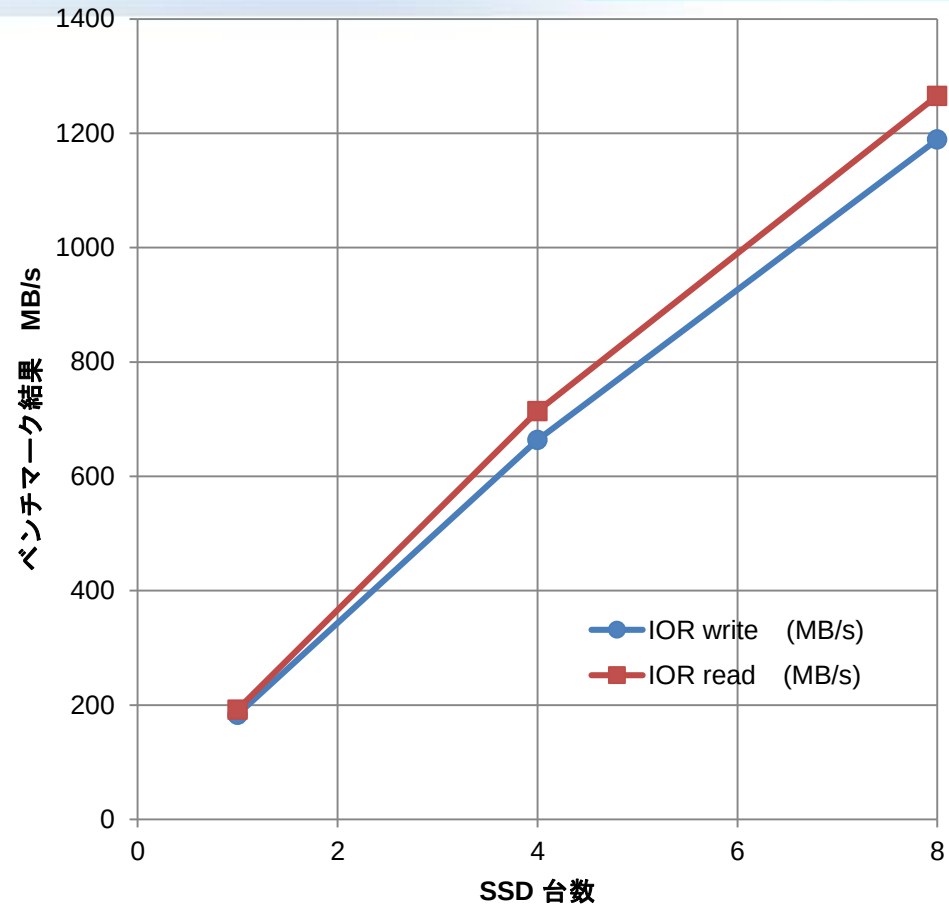
シングルシステム+ローカルストレージ

データ共有：大容量共有ストレージ
アクセス速度：RAIDによる分散処理
管理運用：シングルOSによる統合管理

パラレルI/O性能



- システム構成
 - 各ノードに2台のSSDを搭載
 - 4ノード構成
 - 合計8台のSSDをLinuxのソフトウェアRAIDでRAID0構成
- 評価プログラム
 - IORベンチマークで評価
 - SSDの台数によるベンチマーク性能を評価
- 結果
 - 台数に比例した性能向上
 - より性能の高いSSD利用による高い性能の実現も可能



IOR ベンチマーク

http://sourceforge.jp/projects/sfnet_ior-sio/releases/

OCZ Vertex 3

http://www.ocztechnology.com/res/manuals/OCZ_VerTEX3_Product_sheet.pdf

ホームページとお問い合わせ先



Facebookでも最新情報やニュースを適時紹介しています。

お問合せとお見積りはオンライン問合せをご利用ください。

ホームページの更新と資料アップデートはこちらに掲載されます。

電話でのお問い合わせ

0120-090715 

携帯電話・PHSからは(有料)

03-5875-4718

9:00-18:00 (土日・祝日を除く)

WEBでのお問い合わせ

www.sstc.co.jp/contact

スケーラブルシステムズ



この資料の無断での引用、転載を禁じます。

社名、製品名などは、一般に各社の商標または登録商標です。なお、本文中では、特に®、TMマークは明記しておりません。

In general, the name of the company and the product name, etc. are the trademarks or, registered trademarks of each company.

Copyright Scalable Systems Co., Ltd., 2005-2016.
Unauthorized use is strictly forbidden.

新製品情報とニュースを紹介しています。

各製品ラインの概要説明と各製品へのリンクになります。