

日本HP のクラスタビジネス への取り組み

根本 雅樹
テクニカルセールスサポート統括本部
日本ヒューレット・パッカード株式会社
2008.3.7
～ PC クラスタワークショップ in 名古屋 ～

<http://www.hp.com/jp/hpc>



© 2006 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice.

TOP500リストの 4位、5位に HPのブレードクラスタがランクイン！

- ・ 2007年11月発表のTOP500の上位5システム

NOVEMBER 2007

TOP5	MANUFACTURER/COMPUTER	LOCATION	COUNTRY	CORES	R _{MAX} (Mflop/s)
1	IBM eServer Blue Gene Solution	DOE/NNAS/Lawrence Livermore National Lab	USA	212992	478
2	IBM Blue Gene/P Solution	Forschungszentrum Jülich	Germany	65536	167
3	SGI Altix ICE 8200, Xeon quad core 3.0 GHz	New Mexico Computing Applications Center	USA	14336	127
4	HP Cluster Platform 3000 BL460c, Xeon 53xx 3GHz, Infiniband	Computational Research Laboratories, TATA SONS	India	14240	118
5	HP Cluster Platform 3000 BL460c, Xeon 53xx 2.66GHz, Infiniband	Swedish Government Agency	Sweden	13728	103





BL460c

BladeSystem c7000

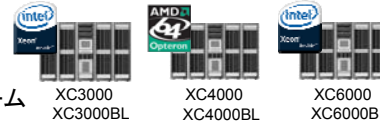
2 2008/4/3



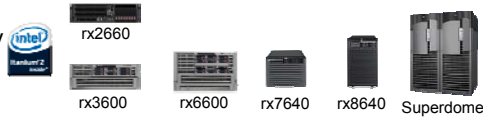
~ HPは標準プロセッサを選択: メニー・コアを牽引

- 豊富な選択肢
- 高性能
- 管理容易

HP クラスタプラットフォーム



HP Integrity
サーバ

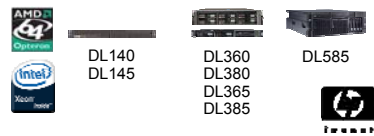


HP BladeSystem

c-Class



HP ProLiant サーバ



HP-UX 11i
Version 2



Microsoft Windows Server 2003
Compute Cluster Edition

3 2008/4/3

(2P、4P) http://h50146.www5.hp.com/products/servers/proliant/dl_lineup.html

1U



2U

2P/2C, 2P/4C, 2C/8C	
オンラインでのご購入	オンラインでのご購入
DL140 G2	DL145 G2
クアドコア または デュアルコア インテル® Xeon®	AMD Opteron™ (デュアルコア)
16GB	

2P/4C, 2P/8C	2P/2C, 2P/4C
オンラインでのご購入	オンラインでのご購入
DL380 G5	DL385 G2
クアドコア または デュアルコア インテル® Xeon®	AMD Opteron™ (デュアルコア)
32GB	

2P/4C, 2P/8C	2P/4C
オンラインでのご購入	オンラインでのご購入
DL360 G5	DL365
クアドコア または デュアルコア インテル® Xeon®	AMD Opteron™ (デュアルコア)
32GB	

4P/4C, 4P/8C		
オンラインでのご購入		
DL580 G4	DL585 G5	DL585 G2
デュアルコア インテル® Xeon®	クアドコア または デュアルコア インテル® Xeon®	AMD Opteron™ (デュアルコア)
64GB	256GB	64GB

1U

4U



4 2008/4/3

HP Integrity サーバ ラインアップ

<http://h50146.www5.hp.com/products/servers/integrity/hardware/>

The diagram shows the HP Integrity server lineup categorized into three main groups:

- ブレード・サーバ (Blade Servers):**
 - BL860c (2P/4C)
 - BL60p (2P/2C)
- エントリーレベル・サーバ (Entry-Level Servers):**
 - rx2660 (2P/4C) - Quick Value
 - rx3600 (2P/4C) - Quick Value
 - rx6600 (4P/8C)
 - rx1620 (2P/2C)
 - rx2620 (2P/4C)
 - rx4640 (4P/8C)
- ハイエンド・サーバ (High-End Servers):**
 - rx7640 (8P/16C)
 - rx7620 (16P/16C)
 - rx8640 (16P/32C)
 - rx8620 (32P/32C)
 - Superdome (64P/128C)
 - Superdome (128P/128C)

5 2008/4/3



BladeSystem c-Classサーバブレード

<http://h50146.www5.hp.com/products/servers/bladeSystem/>



モデル	BL460c	BL465c	BL480c	BL680c	BL685c	BL860c
フォームファクタ	ハーフハイト	ハーフハイト	フルハイト	フルハイト	フルハイト	フルハイト
プロセッサ	Xeon x2	Opteron x2	Xeon x2	Xeon x4	Opteron x4	Itanium 2 x2
メモリ (4GB DIMM 使用時)	FBDIMM 667MHz 8 DIMM/32GB	DDR2 667MHz 8 DIMM/32GB	FBDIMM 667MHz 12 DIMM/48GB	DDR2 667MHz 16 DIMM/64GB	DDR2 667MHz 16 DIMM/64GB	DDR2 533MHz 12 DIMM/48GB
内蔵ディスクベイ	SFF SASベイx2	SFF SASベイx2	SFF SASベイx4	SFF SASベイx2	SFF SASベイx2	SFF SASベイx2
RAID (BBWC装着時)	RAID 0/1	RAID 0/1	RAID 0/1/5/6	RAID 0/1	RAID 0/1	RAID 1
NIC	GbE MF NIC x2	GbE MF NIC x2	GbE MF NIC x2 GbE NIC x2	GbE MF NIC x2 GbE NIC x2	GbE MF NIC x2 GbE NIC x2	GbE NIC x4
メザニンコネクタ	Type I x1 Type IまたはII x1	Type I x1 Type IまたはII x1	Type I x1 Type IまたはII x2	Type I x1 Type IまたはII x2	Type I x1 Type IまたはII x2	Type I x1 Type IまたはII x2

凡例: SFF SAS: Small Form Factor(2.5") SASドライブ、GbE: Gigabit Ethernet、MF: マルチファンクション

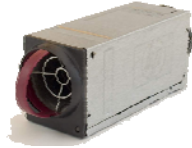
6 2008/4/3



発熱・電力の抑止・コントロール

HP 自身が開発したアクティブ冷却ファン

HP で独自開発。20 個の特許出願中



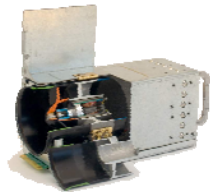
高性能

1 台のファンで 4 台の 1U サーバを冷却できる能力

高寿命

同一冷却性能のファンと比較して

単体寿命が約 4 倍



低消費電力

同一冷却性能のファンと比較して

66%消費電力軽減

c-Class ブレードエンクロージャに
最大10個搭載可能

低騒音

同一冷却性能のファンと比較して

50%騒音抑止



7 2008/4/3

サーバ消費電力の内訳

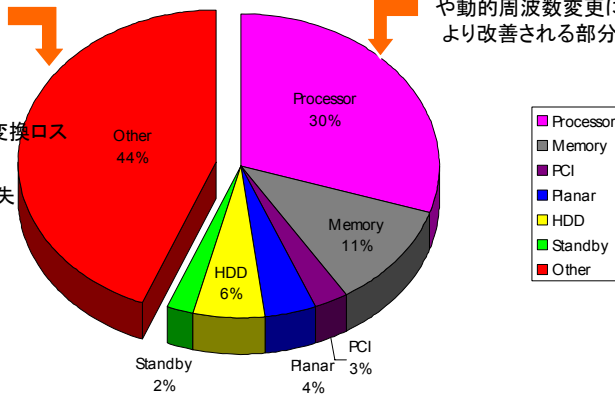
- ・ プロセッサの消費電力が問題視されがちですが、サーバの消費電力を低減させるには、その他の部分を抑える機構が必要です。

HP BladeSystem導入により
改善される部分

OTHERとは？

- ・ パワーサプライによる変換ロス
- ・ 冷却ファンの効率性
- ・ 電力供給ラインでの損失

低電圧プロセッサ導入
や動的周波数変更により
改善される部分



8 2008/4/3

c7000とc3000

c7000

- サーバ台数が多い(8台以上～数千台)
- I/O本数が沢山必要
- ネットワークなどの高い冗長化を求める
- データセンター、マシンルームに設置



c3000

- サーバ台数が少ない(8台以下)
- AC100Vで利用をしたい
- 手軽にブレードをはじめたい
- I/O本数があまり必要ない

9 2008/4/3



「コンパクト第3世代ブレード」 HP BladeSystem c3000 前面

- ・ 8個のブレードデバイスベイ
- ・ サーバブレードだけでなく、テープブレード・ストレージブレードも挿入可能



**工事不要
壁コンセント対応**
通常のAC100V
壁コンセントが利用可能



※電源容量はご注意ください



Onboard
Administrator
(管理モジュール)

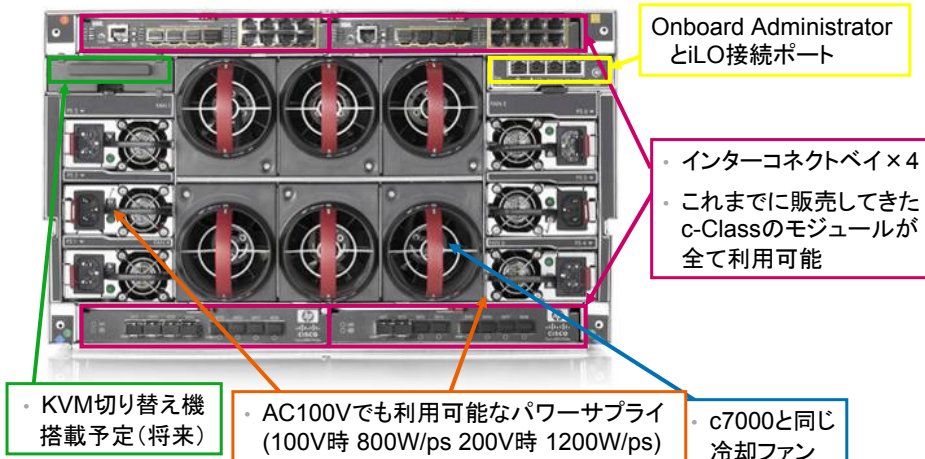
サーバの稼働状況・修理方法
まで教えてくれる
インサイトディスプレイ

複数サーバで同時に利用できる
DVDドライブ

10 2008/4/3



「コンパクト第3世代ブレード」 HP BladeSystem c3000 背面



11 2008/4/3

HP Confidential



Blade-Everything



12 2008/4/3



国内シェアNo.1*記念
<BLADE 3.0>
ありがとうキャンペーン

日本国内で最も導入されているHP BladeSystem c-Classを
 お得に導入できるチャンス！



※2007年第3四半期国内x86ブレードサーバ出荷台数・金額(出典:IDC
 Worldwide Quarterly Server Tracker Q3 2007)

- ・ **キャンペーン期間:2008年3月28日(金)まで**
- ・ サーバブレードを2枚以上ご購入いただくと.....
 - － 特典1:エンクロージャが210円(c3000)、315円(c7000)
 - － 特典2:サーバブレードが30%OFF
 - － 特典3:オプション製品が30%OFF
 - － 特典4:バーチャルコネクが21万円
 - － 特典5:標準保証アップグレードが最大50%OFF

キャンペーン詳細は:

http://h50146.www5.hp.com/products/servers/bladeSystem/campaign/blade_no1.html



13 2008/4/3

HPC クラスタ・ソリューション

日本HPがHPC用途クラスシステムに最適なコンポーネントを選定

★高性能、高機能、高信頼性、低価格★

最適な サーバ、インタコネク、関連ソフトウェアの選択、提案

構築

設置

稼動

基本サービス

HP工場(国内)

- ・ハード組立・試験
- ・ソフト導入
- ・CMU導入・基本設定
- ・動作試験
- ・出荷



お客様設置場所

- ・システム再設置
- ・調整
- ・引渡し

システム稼動

- ・保守サービス



HPC SIサービス

14 2008/4/3

HP logo.

HP CMU (クラスタ管理ユーティリティ)

SCoReにCMUの管理機能を追加

HPCクラスタシステムの構築・管理を容易にするツール

- HPCクラスタにおいて実績のある管理ツールを提供します
 - ・ 全世界で150システム以上の納入実績

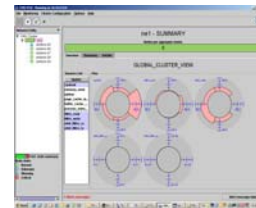
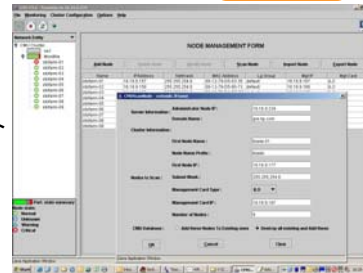
主な機能

演算ノードの一括制御
電源ON/OFF、ブート、シャットダウン、リブート

演算ノードのシステム管理をGUIで一元化

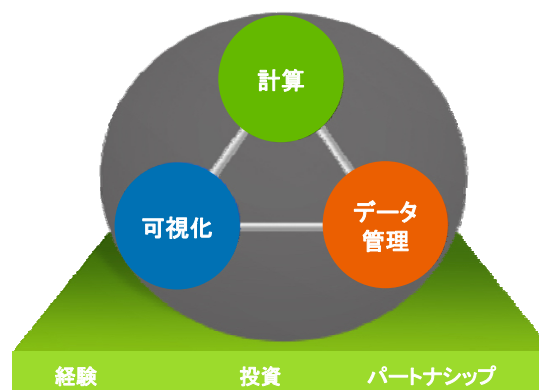
- 1. 論理グループによる管理
- 2. ノード状態の表示
- 3. コンソールブロードキャスト機能
 - 全演算ノードへのコマンド一括入力
- 4. ノード監視とイベント処理
- ・ クローン機能によるシステムディスクの一括複製
- ・ CMUモニタ: 演算ノードの負荷状況等をグラフ表示
- ・ UPS/自動シャットダウン(オプション)

CMU
CLUSTER
MANAGEMENT
UTILITY



完全なHPCソリューションの提供

- ・ 標準技術をベースにした革新性
- ・ HPCソリューションにフォーカスした広範な選択肢
- ・ 手頃で、利用しやすいスーパーコンピューティングクラスの性能



実運用が容易な、HPC用Linuxクラスタソリューション HP XCクラスタ



5ノードから構成可能
大規模クラスタに対応可能

パッケージ製品なので、
HPC用Linuxクラスタ運用の手間が
大幅に削減！

- 業務へすばやい展開が可能
- 不具合時に迅速な対応可能
- 計算負荷の変化に伴うシステム構成の変更が容易
- ソフトウェア設定が容易
- ブート時間を短縮
- システムソフトウェア更新が容易

17 2008/4/3

HPがHPC向けの‘ターンキー’Linuxクラスタを提供

HPCでクラスタシステムを利用するために
必要な機能はすべてまとめて提供・サポート

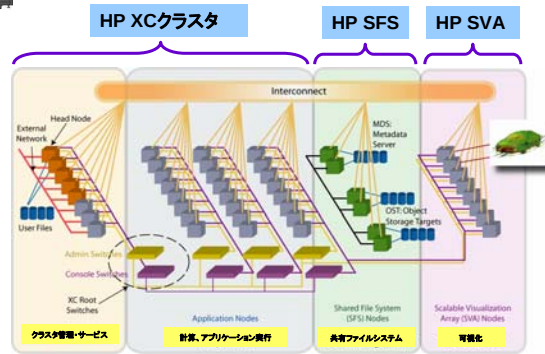
ノード: Opteron, Xeon, Itanium2

インターコネクト: InfiniBand, GigE, Quadrics, Myrinet

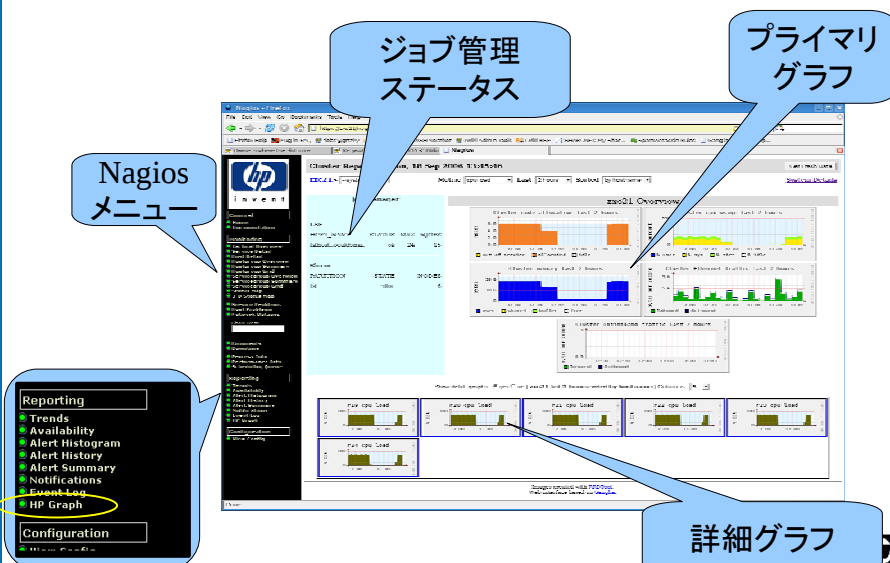
OS/管理ツール: XCシステムソフトウェア

- RedHat EL互換
- LVS, LSF, SLURM, NAT
- Kickstart, XC Config, SystemImager, Nagios, SuperMon, pdsh, MySQL
- HP-MPI

⇒HPが標準製品としてのサポートを提供



グラフウィンドウのレイアウト

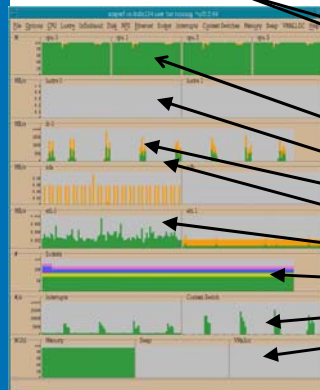


18 2008/4/3

Xtoolsの階層型GUI

「システムレベル」表示

1. Interconnect (in/out)
2. CPU utilization
3. Memory utilization
4. Ethernet utilization
5. Disk (in/out)



「ノードレベル」(アイコンをクリック)

- CPU Utilization
- Luster Traffic
- Infiniband Traffic
- Disk Accesses
- GigE Traffic
- Number and Type of Open Sockets
- Interrupts and Contact Switches
- Memory Utilization



HP XCクラス

5ノードから
構成可能



システム拡張



32 ノード構成

128 ノード構成 (さらに大規模構成へ)

HP XC3000

Xeonベース

- DL140 (2CPU)
- InfiniBand, GigE, Myrinet

HP XC4000

Opteronベース

- DL145 (2CPU)
- DL585 (4CPU)
- InfiniBand, GigE, Quadrics, Myrinet

HP XC6000

Itanium2ベース

- rx2660 (2CPU)
- rx4640 (4CPU)
- Quadrics, InfiniBand, GigE

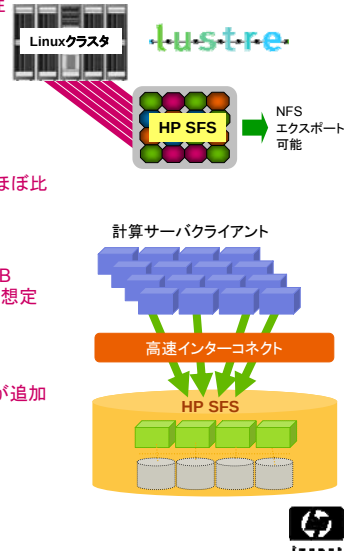
c-Classブレードにも対応



HP StorageWorks Scalable File Share(HP SFS)

大規模Linuxクラスタのための共有ファイルシステムLustreの実装製品

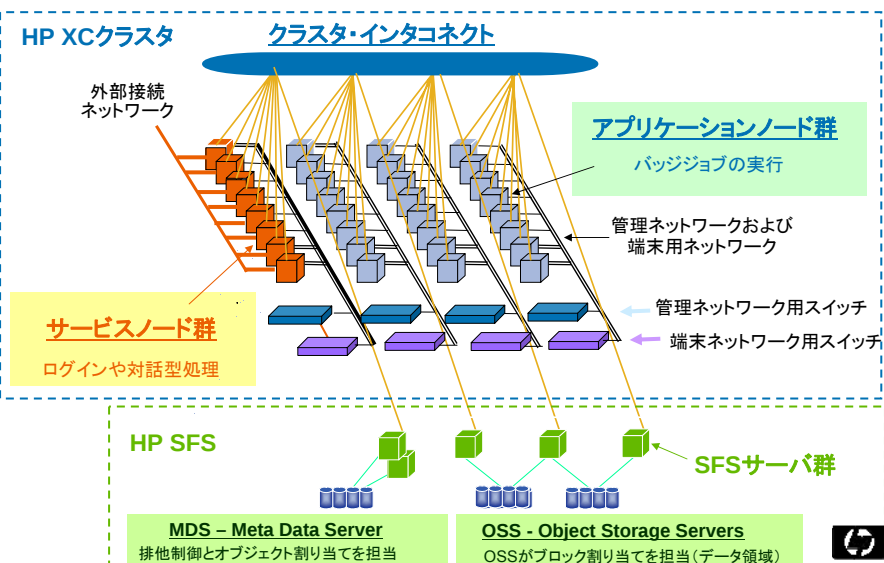
- 全ノードから一貫性のあるファイルシステムアクセス
 - Lustreファイルシステム: POSIX準拠のデータの一貫性、整合性
 - すべてのノードから同じファイルを同じパス名で直接参照可能
 - ファイル・システムを単一イメージで提供
 - MPIジョブからの並列I/Oに伴う特殊なプログラム変更が不要
- スケーラブルなI/Oバンド幅
 - 200MB/s ~ 35GB/s さらに拡大可能
 - 総合I/Oスループット(実際に利用者が享受する性能)
 - I/O用ノードやストレージ装置の増設によるスケーラビリティ
 - 合計性能だけでなく、単一ファイルの処理性能もサーバー数にほぼ比例して増加
 - 高速インターコネクト経由のデータ転送
- スケーラブルな容量: 最大数百TB
 - I/O用ノード、ストレージ装置の増設でSFS一式あたり最大512TB
 - スケーラブルなアクセス: 数千クライアントからの同時アクセスを想定
- サービスの継続性
 - 障害に強い構成、縮退運転機能(フェールオーバー機能)
 - RAID5、RAID5+1、RAID6、RAID6+1サポート
 - 自動アラート機能、診断機能など安定運用のための機能をHPが追加
- 用途に合わせたストレージ構成
 - 価格性能重視: StorageWorks SFS20
 - 高信頼性重視: StorageWorks EVA
- HP製品としての保守体制
 - オープンソースLustreファイルシステムを製品としてHPが保守



21 2008/4/3

XCクラスタとSFS

XCクラスタの共有ストレージには、高性能・大容量・高信頼性のSFSファイルサーバが最適！

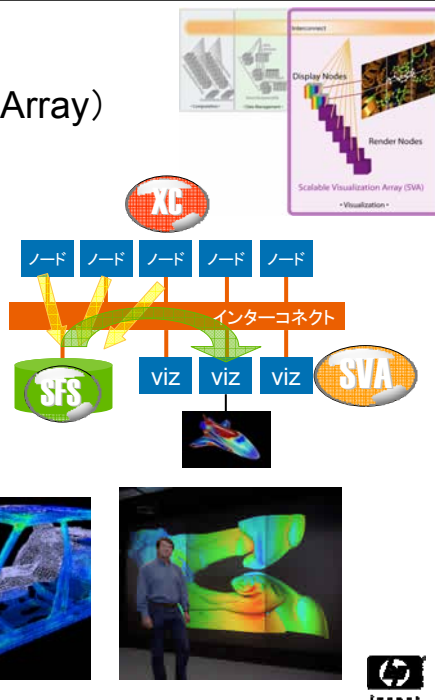


22 2008/4/3

HP SVA

(HP Scalable Visualization Array)

- ✓ クラスタでの高速・大規模・並列可視化が容易
- ✓ HP SFSによる高速・大容量データ共有
 - ✓ シミュレーション計算 ⇒ 可視化
 - ✓ レンダリングノード ⇒ ディスプレイノード
- ✓ 主要な可視化アプリケーションに対応
 - ✓ EnSight
 - ✓ AVS 等
 - ✓ 可視化アプリケーションのランチャー提供
- ✓ HP XCクラスタによる運用管理
 - ✓ Nagiosによる稼働状況監視
 - ✓ SLURMによる資源管理
 - ✓ LSFジョブスケジューラ 等
- ✓ ベンダ製品としての保守体制



23 2008/4/3

HP-MPI: 性能、移植互換性に優れ、多くのISV様がサポート表明しています

- ・MPI 2.0準拠
- ・HPが研究開発、販売、保守
- ・ASCI-Q, PNNL 等の世界最大級のクラスタ計算機環境での稼働実績
- ・TCP/IP サポート
- ・PGI, Intel コンパイラ サポート
- ・豊富なプロファイリング機能
- ・TotalViewとの連動
- ・HP-UX、Linuxサポート
- ・Windowsサポート
 - CCS, XP

XCクラスタに
標準装備



24 2008/4/3

HP-MPIは主要なISVのアプリケーションソフトウェアに採用されています



Note: end users of these ISV applications on industry-standard clusters may receive HP-MPI transparently from these ISVs as part of the ISV's application distribution. "Transparent" means no cost and no FlexLM licensing requirement. The first point of contact for support is the ISV.
All ISVs listed may distribute HP-MPI on Linux & HP-UX. Those with * have also signed up for Windows CCS.

Example ISV web sites listing platforms supported with HP-MPI:
www.abaqus.com/support/v66/v66_mpi.html
www.ansys.com/services/ss-interconnects.htm



HPCソリューションセンター

- ・ HPCスペシャリストがお客様の検証作業をご支援致します。
 - 新製品、新技術の検証やベンチマーク
 - ポーティング・チューニング
 - お客様の課題や問題の解決

弊社HPCポータルサイトからご予約いただけます。

<http://www.hp.com/jp/hpc/>

その他、お問い合わせは

Mailto: hp-scs-collabo-center@hp.com



設備:

HP XCクラスタ、HP SFS
HP BladeSystem
HP ProLiant、HP Integrity
各種ストレージ
ワークステーション 等々



まとめ

- ・ HPCソルバーのプラットフォームとして、クラスタが主流
 - － HPC用途クラスタは性能のみならず、安定稼動が重要
- ・ 業界標準技術の活用
 - － IntelやAMDの性能ロードマップ及びHPCシステムアーキテクチャ
 - － ISVアプリケーションと業界標準技術のマッチング
- ・ HP製品で稼動するCAE・EDAアプリケーションの完全なポートフォリオ
 - － 多くのCAE・EDA ISV様と技術やビジネスについて深いパートナーシップ
 - － 業界標準クラスタについてISV様のアプリにHP-MPIを提供し、最適化
- ・ HPCに対する統合ソリューション提供
 - － 計算、IO、可視化
 - － XCクラスタや、統合環境、パートナー様のソリューションを提供
 - － ベンチマーク、導入、保守サポートなどを提供
- ・ HPCへのブレードの適用
 - － 性能、設置場所、電力、冷却
 - － XCクラスタとしても提供可能
- ・ HPCにおけるお客様のROIを最適化
- ・ HPCに関する情報
 - － www.hp.com/jp/HPC



27 2008/4/3

