

HPC分野への取り組みと、 APPRO社Xtreme-Xのご紹介

～第八回PCクラスタシンポジウム～

平成20年12月12日
株式会社HPCソリューションズ
根本 雅樹

アジェンダ

- 株式会社HPCソリューションズのご紹介
- 米国APPRO社のご紹介
- スーパーコンピュータXtreme-Xのご紹介
- Appro HyperGreen Clusterご紹介(未発表製品)

会社概要

事業所	本社 東京都中央区日本橋大伝馬町3-2 ダウインチ小伝馬町 1階 TEL:03-5640-7858 FAX:03-5640-2578
	大阪事業所 大阪府大阪市淀川区宮原4-5-36 セントラル新大阪ビル7階 TEL:06-6394-8934 FAX:06-6394-8935
	つくば事業所 茨城県つくば市二宮1-2-2 酒井ビル TEL:029-861-0250
役員構成	代表取締役会長 藤澤 義麿 代表取締役社長 河野 証 専務取締役 東 孝司 取締役(非常勤) 中田 義弘 取締役(非常勤) 伊藤 俊彦 監査役(非常勤) 酒巻 千丈
設立	2006年7月1日
事業内容	PCクラスタシステムの販売およびサポート、 クラスタシステム構築等のコンサルティング、 保守

業務内容

□ 各ベンダーのハードウェア、ソフトウェアの販売

- ◆ ⇒米国APPRO社製品も取り扱っています。

□ マルチベンダーサポート

- ◆ お客様のHPC用途コンピュータシステムの高速化、安定稼働を目指し、あらゆるベンダーのあらゆるコンピュータシステムをワークステーションから大規模クラスタまで一貫してサポート。

- システム設計、クラスタ設計
- システム構築、クラスタ構築
- システム開発
- システムサポート
- サイトプランニング
- HPCアプリケーションベンチマーク
- コンサルテーション

国内初

Windows HPC Server 2008クラスタシステムの 構築・導入

- ❑ OS:Windows HPC Server 2008
- ❑ ノード:PowerEdge1950(CPU:Intel Xeon5406)32台
- ❑ インタコネクト:InfiniBand(4X DDR)(Voltaire)
- ❑ アプリケーション:SCRYU/Tetra(ソフトウェアクレイドル)

http://www.hpc-sol.co.jp/081218_NewsRelease.pdf

Appro社のご紹介

APPRO INTERNATIONAL INC

会社概要

:: HPC専門の会社

- 1991年設立
- 本社:Milpitas - カリフォルニア(シリコンバレー)
事業所:Houston - テキサス
- ハイパフォーマンスコンピューティングに特化したソリューション提供会社
 - 大学官公庁/ 防衛
 - 製造
 - バイオ
 - EDA
 - 金融
 - 石油・ガス



APPRO社のHPC製品

- スーパーコンピュータ
 - Appro Xtreme-X (Nehalem)
- クラスタ
 - AMD & Intel HyperGreen Clusters
- サーバ
 - Intel 1U/2U DP RM Servers (Harpertown)
 - Intel 1U/2U DP RM Servers (Nehalem)
 - AMD 1U/2U DP/QP Servers (Shanghai)
- ワークステーション
 - Entry WH300 Intel & AMD WS



高密度ラックサーバ: Appro 1224X



Appro 1224X

Twin 2way dual/quad-core,
4x hot-swap SATA HDDs

- 1つの1Uシャーシに2台の2プロセッサ用マザーボード搭載
 - 1ノードあたり、0.5U
 - 10Uの高さに20台のノード
- マザーボードあたりの仕様
 - 2つのIntel Xeon Wolfdale/Harpertown プロセッサ
 - Intel 5400 “Seaburg” チップセット - 1600MHz FSB
 - 最大32GB FBDIMM 667/800MHz システムメモリ(8 DIMMスロット)
 - 最大2台の3.5” SATA ホットスワップドライブ
 - 1x LP PCIe Gen2 x16 ライザーズロット
 - IPMI 2.0 リモート管理(オプション)

会社概要

:: HPCでのプレゼンス



	Vendors	Count	Share %	Rmax Sum (GF)	Rpeak Sum (GF)	Processor Sum	Avg. Efficiency
1	Hewlett-Packard	209	41.80%	4193870	7250162	707056	58%
2	IBM	186	37.20%	6511376	10059679	1498276	65%
3	Cray Inc.	22	4.40%	2486731	3226223	376309	77%
4	Dell	19	3.80%	569640	774686	82250	74%
5	SGI	17	3.40%	1218446	1479534	141560	82%
6	Sun Microsystems	7	1.40%	542762	724744	82936	75%
7	Appro International	4	0.80%	158700	191536	30736	83%
8	Fujitsu	3	0.60%	84140	100310	10240	84%
9	Bull SA	3	0.60%	115090	137523	19696	84%
10	Self-made	3	0.60%	52740	82556	8880	64%

Vendors by Performance – November 2008

APPRO社のクラスタシステムはTOP500にも浸透中



HPC 事例

:: APPRO導入事例-1

2006. 9



NOAA Cluster

- 1,424 Cores
- 2.8TB System Memory
- 15 TFlops

2006. 11



LLNL Atlas Cluster

- 9,216 Cores
- 18.4TB System Memory
- 44 TFlops



2007. 6



LLNL Minos Cluster

- 6,912 Cores
- 13.8TB System Memory
- 33 TFlops



2008. 2



DE Shaw Research Cluster

- 4,608 Cores
- 9.2 TB System Memory
- 49 TFlops



HPC 事例

:: APPRO導入事例-2

2008. 4



TLCC Cluster

- 72,462 Cores
- LLNL, LANL, SNL
- 638 TFlops



2008. 6



Tsukuba University Cluster

- 10,784 Cores
- Quad-rail IB
- 95 TFlops



University of Tsukuba

筑波大学



2008. 7



Renault F1 CFD Cluster

- 4,000 Cores
- Dual-rail IB
- 38 TFlops

2008. 8



LLNL Hera Cluster

- 13,824 Cores
- 120 TFlops



HPC Cluster Solutions

SC08会場のAPPRO社のブース



新型Xtreme-Xを展示



***Appro Xtreme-X* スーパーコンピュータご紹介**

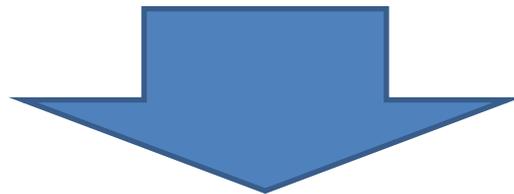
APPRO INTERNATIONAL INC

クラスタの課題

:: Xtreme-X によって克服

◆ クラスタはまだ利用や管理が難しい

- 電源、冷却、設置面積への対応が重要。
- RAS機能の充実が急務。
- ストレージやデータ管理がボトルネックになりつつある。
- インタコネクットの性能が重要。



APPRO社 Xtreme-X で解決！

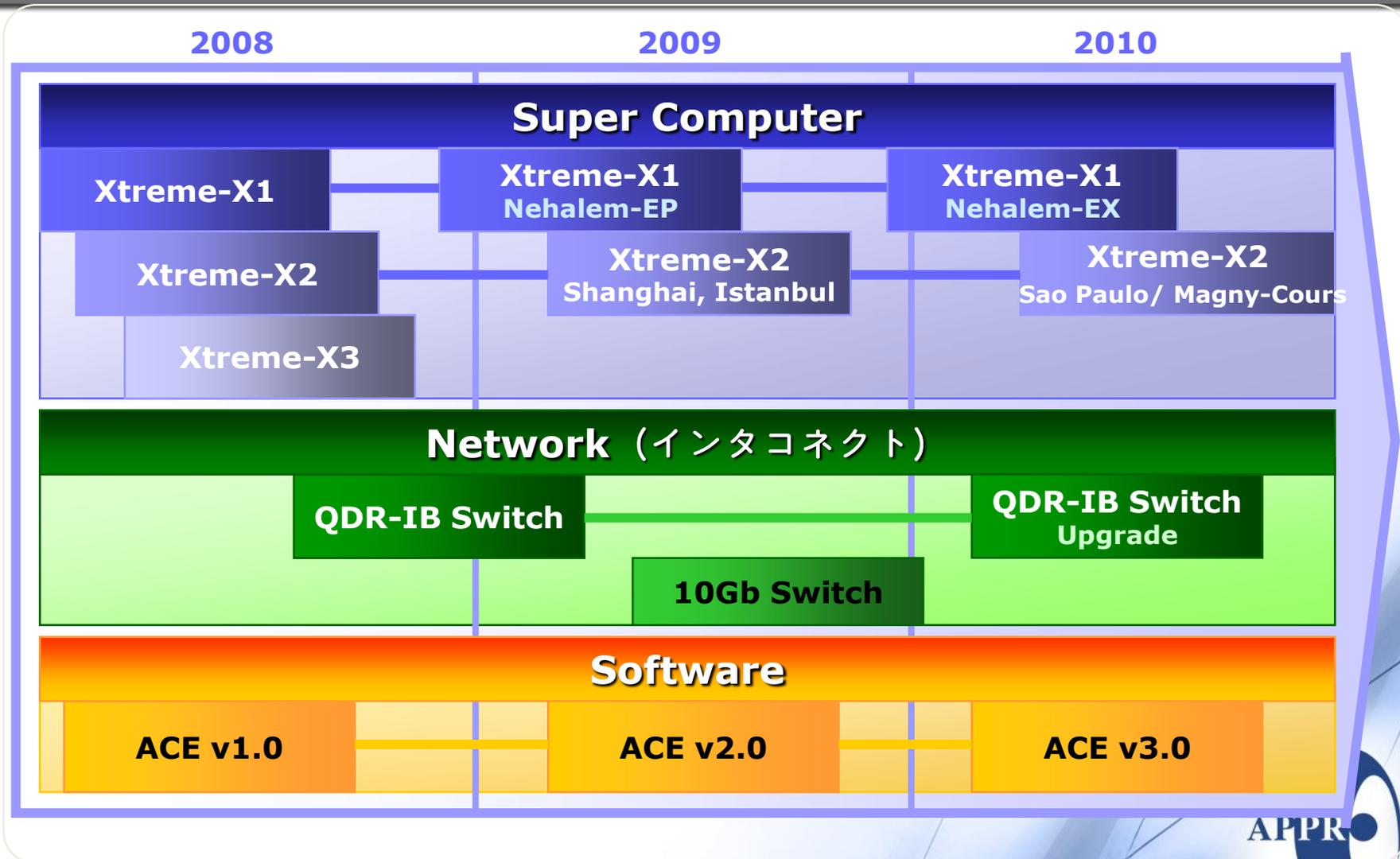
Xtreme-X スーパーコンピュータ

:: 概要



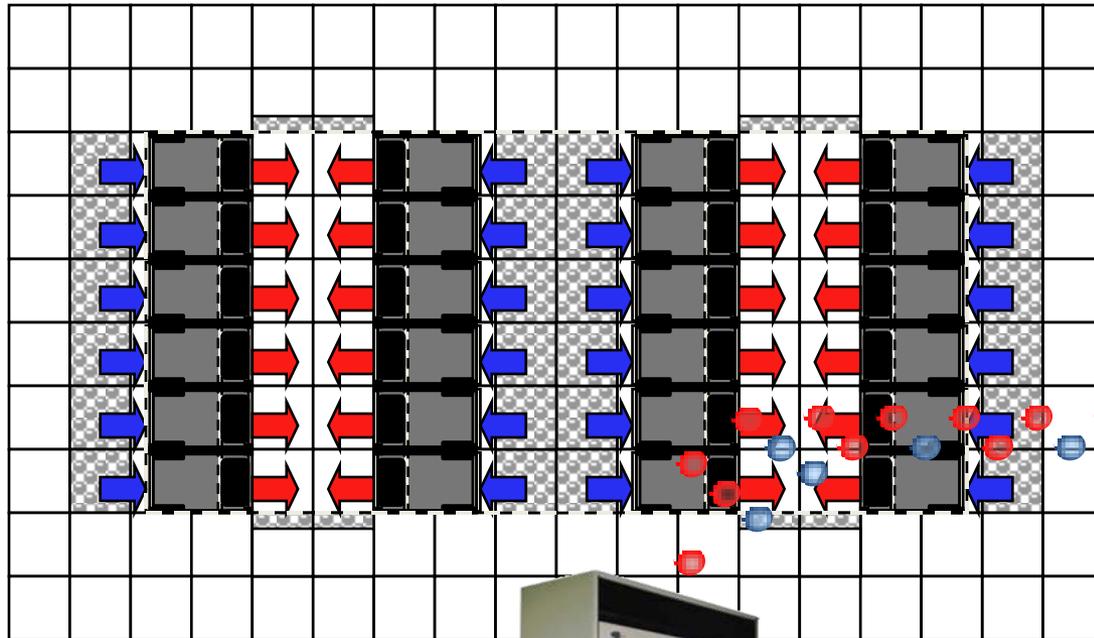
- 完全に統合した、プリパッケージソリューション
→それぞれのラックをスケラブルユニット(SU)として構成
 - TCOの削減
 - クラスタを容易に構成、増設可能
 - クラスタの信頼性、利用可能性、保守性の改善
- **1000TF以上へのスケラビリティ**
- 業界標準製品の活用により、アプリケーションソフトウェアの移植が容易
 - Xeon、Opteron
 - インターコネク
 - ネットワーク

Appro Xtreme-X ロードマップ

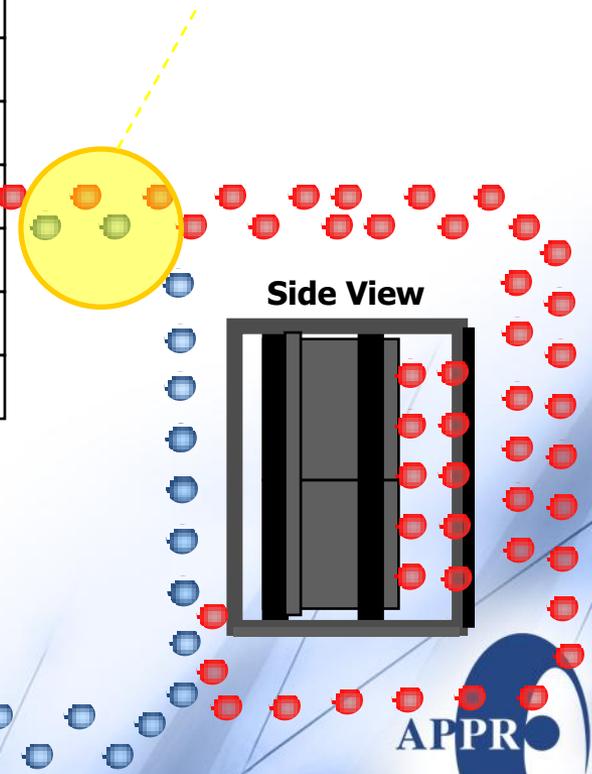


電力と冷却

::従来の冷却方法



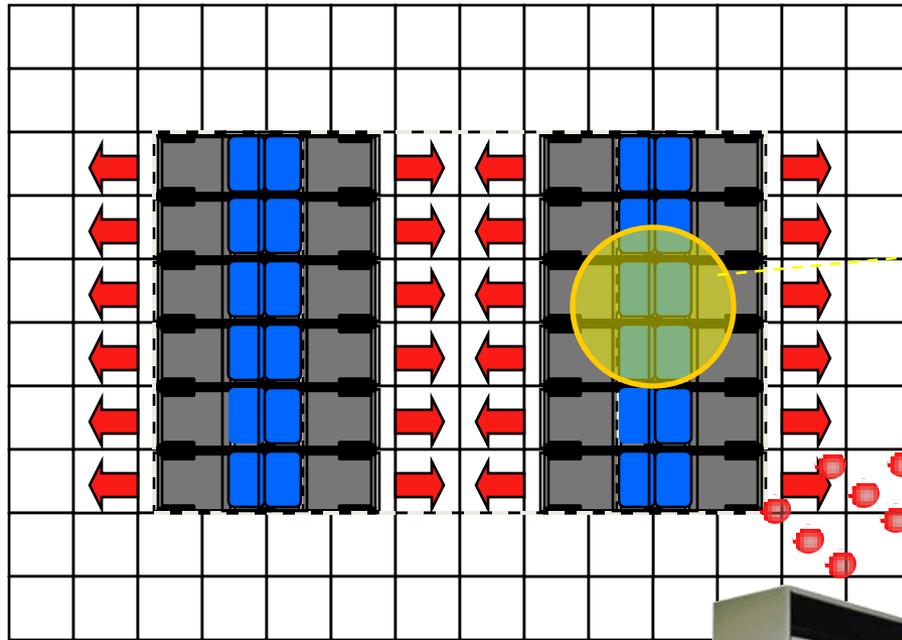
熱風と冷風が
混合してしまうために
冷却効率損失が発生



Side View

電力と冷却

：Xtreme-Xの新冷却方法 Directed Airflow™ Cooling

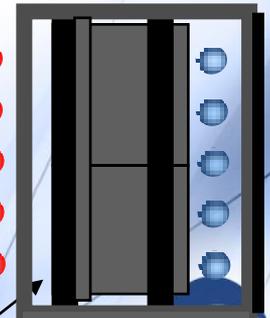


前面アクセス方式の採用で
設置面積を30%削減でき、
さらに
熱風と冷風分離により
冷却効率が改善



Cable area

Side view



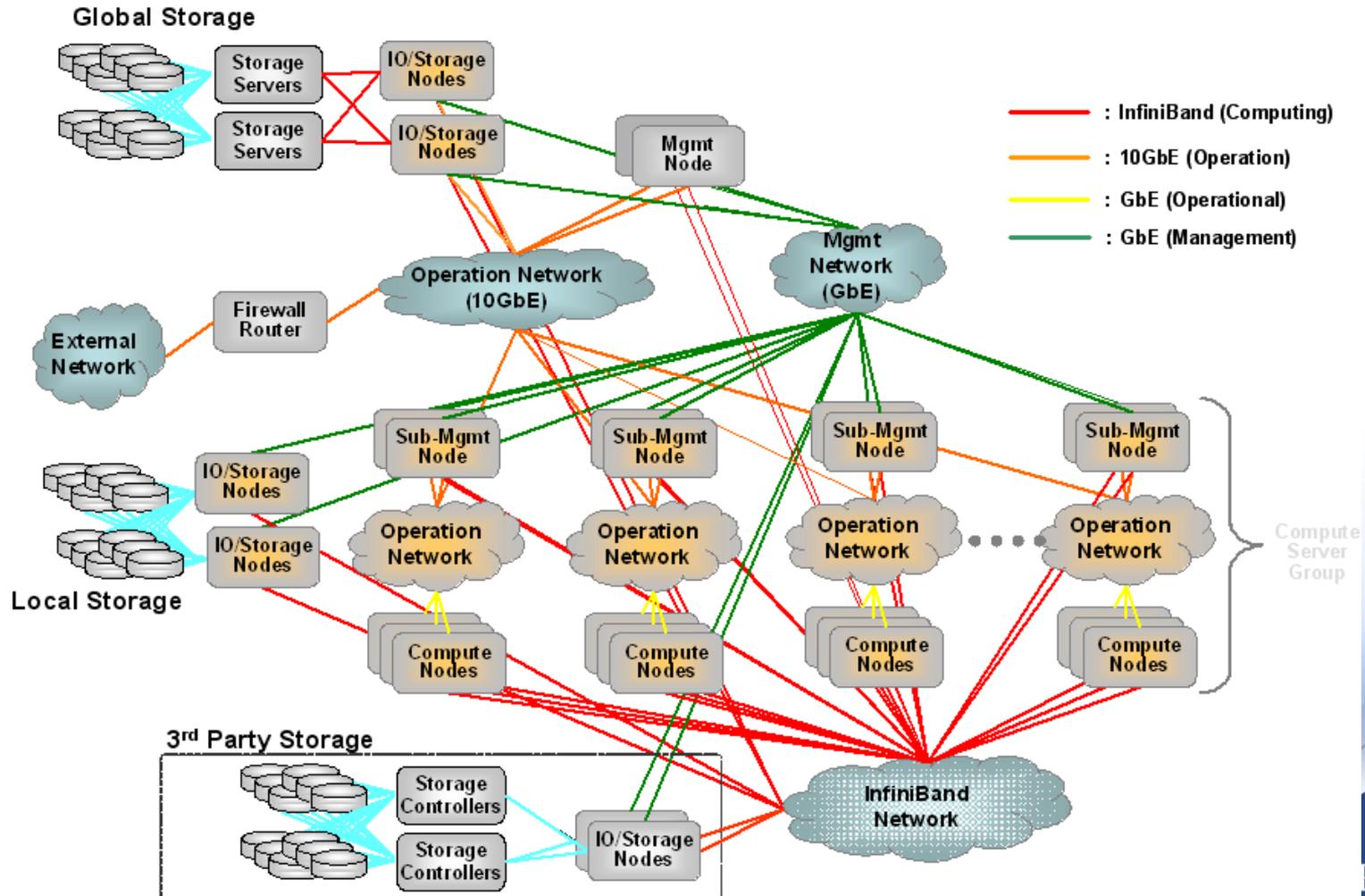
クラスタの信頼性への対策

:: Xtreme-Xの信頼性向上対策

- 不具合発生個所に対する信頼性向上の対策
 - ディスクドライブ
 - ディスクレスアーキテクチャ(計算ノード)
 - ファン
 - **N+1 冗長構成**
 - パワーサプライ
 - サブラックにはホットスワップ可能な**N+1** パワーサプライを配備
 - ネットワーク
 - **InfiniBand**とイーサネットについて冗長構成採用の**Xtreme-X** アーキテクチャ
 - サーバ
 - **RAID**ディスクアレイを管理ノードに採用の**Xtreme-X**アーキテクチャ
 - 冗長構成(アクティブ/スタンバイ)の管理ノード
 - 冗長構成(アクティブ/アクティブ)のサブ管理ノード
 - 計算ノードはホットスワップ可能、および、他のノードへの影響なしに簡単に取り外し可能

Xtreme-X バランスのとれたソリューション

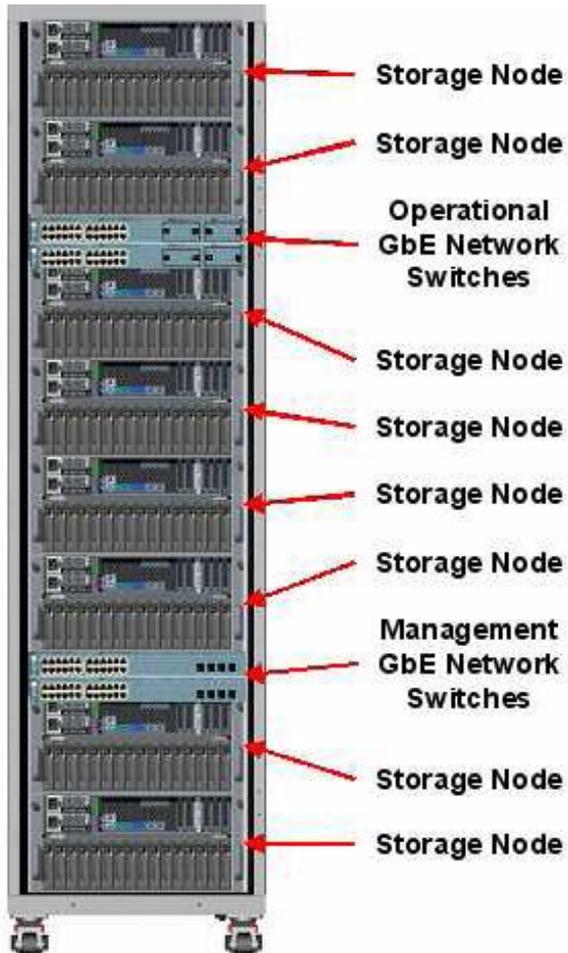
:: 高性能かつ安定性を追求した設計



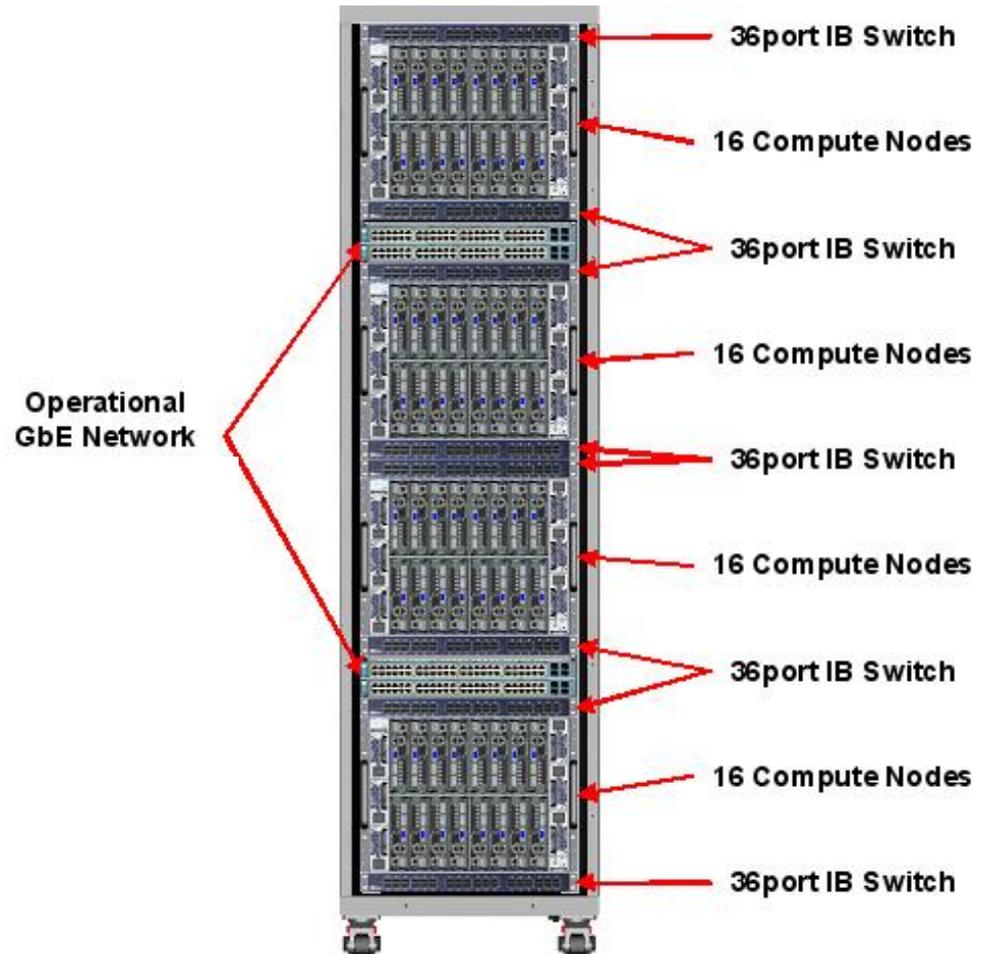
Xtreme-Xスーパーコンピュータ

::スケラブルな設計

ストレージユニット

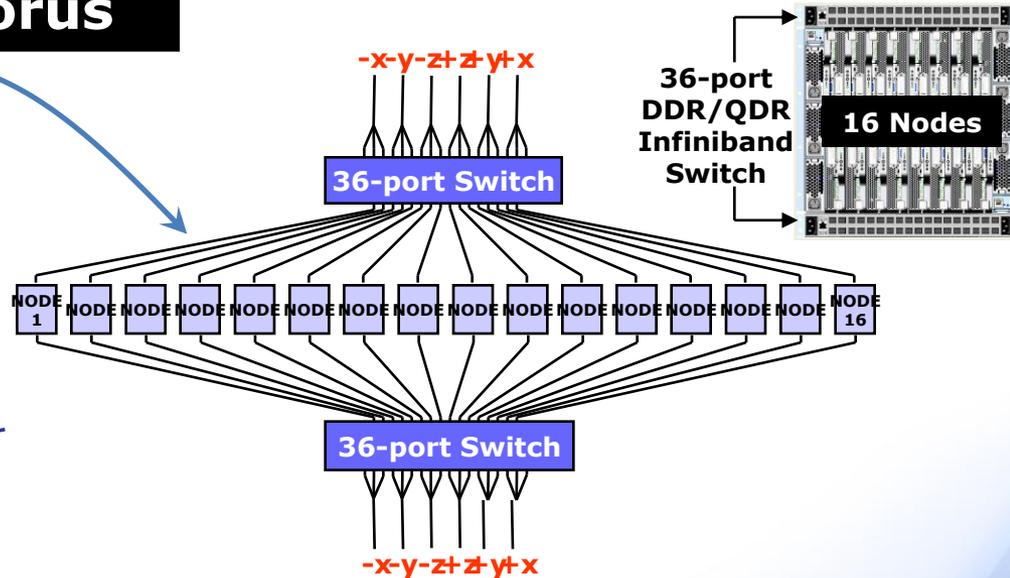
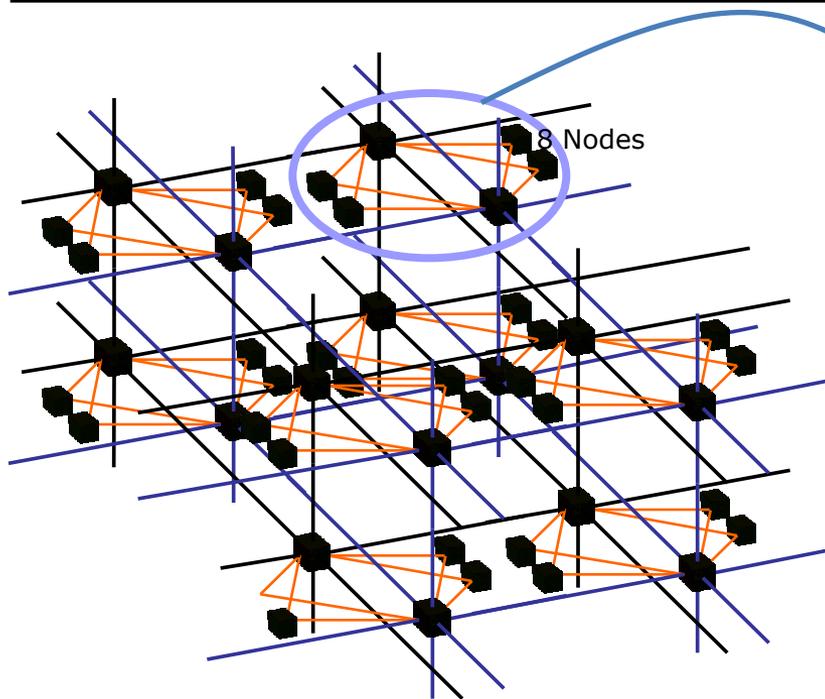


計算ユニット



インタコネクト トポロジ

Dual-Rail DDR(QDR) 3D Torus



IB インタコネクトとサブラックの構成

- 2つの独立したInfiniBandサブネット
- MVAPICH2についてポートを統合可能
 - フェールオーバ及びロードバランス
- フォールトトレランスの提供
- サブラック(16ノード)に対して2倍のバンド幅を提供

3D Infiniband 接続イメージ図

- **Dual-Rail DDR, 4X links**
- **スイッチ各軸間バンド幅: $2 \times 12\text{GB/sec} = 24\text{GB/sec}$**
(QDRでも構成可能。その場合は、**48GB/sec**になります。)

ペタフロップスへの展開

:::スケラブル クラスタ管理ソフトウェア ACE

3D トーラス ネットワークトポロジ

ジョブスケジューリング

統合オペレーション

デュアルレール ネットワーク

BIOS 表示

Sun Grid Engine との連携

インスタントSW プロビジョニング

Appro Cluster Engine™

仮想クラスタ管理

フェイルオーバ & リカバリ

IB-サブネットマネージャ

RedHat EL5サポート

リモート管理

“Appro Cluster Engine™ software turns a cluster of Servers into a,
functional, usable, reliable and available computing system”

Jim Ballew, CTO Appro



ACE(Appro Cluster Engine)概要 (1)

1. ネットワーク管理	<ul style="list-style-type: none">a. インターコネクットのハードウェア構成自動検出b. ネットワーク冗長構成c. ロードバランスとフェイルオーバーd. ネットワークステータスの管理システムへの統合
2. サーバ管理	<ul style="list-style-type: none">a. ノードのハードウェア構成自動検出b. 遠隔サーバ管理(パワーオン、パワーオフ、パワーサイクル)c. 遠隔サーバ初期化 (リセット、リブート、シャットダウン)d. 大規模クラスタでのスケーラブルな高速ディスクレスブートe. サーバ冗長化、フェールオーバーf. サーバステータスの管理システムへの統合
3. クラスタ管理	<ul style="list-style-type: none">a. クラスタをパーティショニングし、複数の論理コンピュータとして運用可能b. 複数OSのサポートc. 論理コンピュータの管理、監視d. クラスタステータスの収集

ACE(Appro Cluster Engine)概要 (2)

4. ノード管理	<ul style="list-style-type: none">a. ジョブスケジューリング及び管理b. ノードの監視、管理c. ノードステータスの管理システムへの統合
5. ストレージ管理	<ul style="list-style-type: none">a. ディスクレスノードへのスケーラブルなrootファイルシステムのサポートb. 複数のグローバルストレージ構成サポートc. サーバステータスの管理システムへの統合
6. システム管理	<ul style="list-style-type: none">a. システム構成の管理b. 管理サーバの冗長化及び自動フェイルオーバー

Appro Cluster Engine による管理

Appro Cluster Engine Xtreme-X1 (on norbert)

Monitor View Actions Help

Switches Servers Clusters

Name	Type	L1	L2	L3	N	Grp	Rack	Slot/U#	Host	Net1	Net2	State
server-0001	management	1	1	1	1	1	1	1		●	●	● active
server-0002	management	1	1	1	2	1	1	2		●	●	● unknown
server-0003	group	1	1	1	3	1	1	3	sysgrp-0001	●	●	● active
server-0004	group	1	1	1	4	1	1	4	sysgrp-0002	●	●	● active
server-0005	compute	1	1	1	5	1	1	5	prod-0001	●	●	● active
server-0006	compute	1	1	1	6	1	1	6	prod-0002	●	●	● active
server-0007	compute	1	1	1	7	1	1	7	prod-0003	●	●	● fail on
server-0008	group	1	1	1	8	2	1	8	sysgrp-0003	●	●	● active
server-0009	group	1	1	1	9	2	1	9	sysgrp-0004	●	●	● active
server-0010	compute	1	1	1	10	2	1	10	prod-0004	●	●	● active
server-0011	compute	1	1	1	11	2	1	11	prod-0005	●	●	● active
server-0012	compute	1	1	1	12	2	1	12	prod-0006	●	●	● fail on

Console Progress

ACE Log Console

```
2008-05-08 15:54:36.599 INF Host prod-0005 started on server-0011
2008-05-08 15:54:36.599 INF Host prod-0006 started on server-0012
```

● connected

NEC様との欧州でのパートナーシップ

- EMEA(欧州、中東、アフリカ地域)において、APPRO社とNEC様は、NEC様のHPCソリューションのひとつに、APPRO社のXtreme-X及びACEを加えるという戦略的パートナーシップを発表しました。



High-Performance and Enterprise Computing

Cor

Home | Company | Products & Solutions | Support

Press Room

Press Releases

News Coverage

Success Stories

Event Schedule

Awards & Honors

Press Release

SUBJECT  [9/16/08 - Appro and NEC Partner to Strengthen Supercomputer Sales](#)

Strategic Partnership to Capitalize on Growing Market for Large HPC Cluster Installations

Milpitas, CA -- 9/16/2008 -- Appro (<http://www.appro.com>), a leading provider of high-performance enterprise computing systems, and NEC Corporation, one of the world's leading providers of Internet, broadband network and enterprise business solutions, announce a strategic partnership today that will see Appro Xtreme-X™ Supercomputer and Appro Cluster Engine™ Software Management added and branded as part of NEC's HPC solution offering.

The partnership will enable Appro and NEC to work together towards a common goal, focus on reducing complexity of technology integration when deploying and managing integrated solutions, while lowering customer total cost of ownership (TCO). Commencing September 2008, Appro Supercomputer products will be added to NEC's HPC offering as a first step in this partnership.

<http://www.appro.com/press/view.asp?Page=1&Num=160>



***Appro HyperGreen Cluster*ご紹介** **(未発表製品)**

APPRO INTERNATIONAL INC

Appro GreenBlade

10 Blade Nodes



Nehalem,
Shanghai
対応予定

iSCB

4 Pluggable PSU (1625W/PS)

(3+1 or 2+1 Redundant configuration)

19" Standard Rack

- Weight for 1 Enclosure with full node
: about 65.25kg per subrack including 10 nodes and 3 PSUs.



GreenBladeの利点

- Greenソリューション
 - 共有冷却ファンで、ノードあたりの消費電力を低減
 - 代表的な1U サーバの冷却ファンの消費電力(18w x 6 =**108W**/ノード)
 - 代表的な1U twinサーバの冷却ファンの消費電力(18w x 3 =**54W**/ノード)
 - Appro GreenBladeの共有冷却ファン(20w X 6 =120W/10ノード=**12W**/ノード)
 - 共有パワーサプライで、ノードあたりの消費電力を低減
- 信頼性向上
 - 冗長/ホットスワップ可能な共有冷却ファン(3+3 構成)
 - 冗長/ホットスワップ可能な共有パワーサプライ. (2+1または3+1 構成)
- 保守容易
 - ホットプラグ可能な計算ノード
- より高密度
 - 標準19インチラックに、80台の計算ノードを搭載可能

