

# Express5800 の並列 PC-Cluster

## Parallel Linux Cluster by Express5800 Series

久保 淳\*

Atsushi Kubo

千田 哲秀\*\*\*

Tetsuhide Senta

岡野 格\*

Tadashi Okano

広瀬 哲也†

Tetsuya Hirose

石井 英志\*\*

Hideshi Ishii

石田 仁††

Hitoshi Ishida

### 要 旨

PC サーバの性能の向上や Linux の機能強化を背景に近年大学や研究機関のみならず、一般企業において、PC クラスターの需要が急速に伸びてきています。このような需要に対し、NEC は新しい製品・サービスを投入・強化して対応しています。

本稿ではその概要について紹介します。

の種類を絞って、比較的小規模な計算を経済的に処理したい場合に適しています。

NEC は前述のように、多種多様なお客様の利用分野や利用形態に合わせた最適な製品、あるいは複数の製品を組み合わせたソリューションを提供できるように、全領域をカバーする豊富な製品ラインナップをそろえ、お客様の多様な要求に対応しています。

### 2. NEC の HPC サーバ製品と PC クラスタ

NEC のサーバ製品ラインナップには、図 1 に示すようにベクトルスーパーコンピュータ SX シリーズをはじめ、米国 Intel 社の Itanium2 プロセッサを搭載した高性能スカラサーバ TX7 シリーズおよびブレードサーバ Express5800/1020Ba, Xeon/Pentium プロセッサを搭載した Express5800/120Ba-

Supported by performance improvement of PC servers and function enhancement of Linux, the demand for PC clusters is rapidly growing in the market of industry as well as universities and research institutes. To meet such a demand, NEC focuses on offering new products and services of PC clusters and continuing their enhancement.

This paper describes the overview of NEC's PC cluster products and services.

### 1. まえがき

HPC (High Performance Computing) 領域におけるサーバの利用分野は多種多様であり、その分野の特性に合った製品を選択することが、利用する側にとって非常に重要なポイントになります。

たとえばベクトル型コンピュータは大容量のデータを使用する大規模計算をできるだけ高速に実行したい場合に、また多数の CPU でメモリを共有する高 SMP (対称型マルチプロセッサ) スカラ型サーバは計算センター業務などのように多種多様なアプリケーションを単位時間にできるだけ多く処理したい場合に、そして同じスカラ型サーバでも、ブレードサーバや PC クラスタは実行するアプリケーション

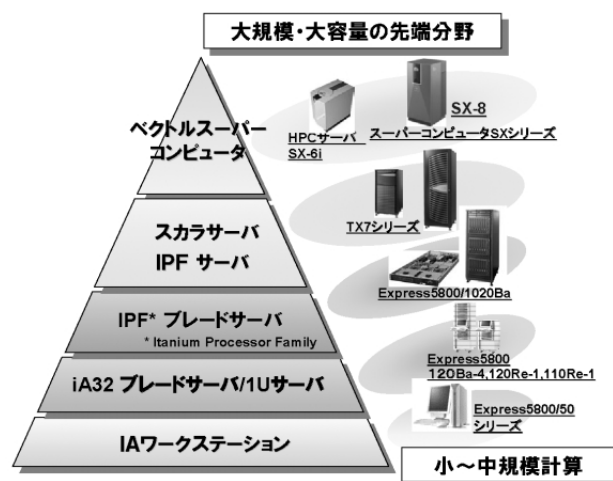


図1 NEC の HPC サーバ製品ラインナップ

Fig.1 NEC HPC Server family lineup.

\* クライアント・サーバ事業部  
Client And Server Division  
\*\* HPC 販売推進本部  
HPC Marketing Promotion Division  
\*\*\* コンピュータ事業部  
Computers Division

† グリッド推進センター  
Grid Promotion Center  
†† NEC フィールディング 東日本サービス営業本部  
NEC Fielding, Ltd.

4, 120Re-1, 110Re-1などが用意されています。

ここでは、Express5800/1020Ba, 120Ba-4, 120Re-1を用いたPCクラスタについて説明します。

#### (1) Itanium2搭載のブレードクラスタソリューション

高い演算性能能力を持つItanium2プロセッサを搭載したブレード型サーバExpress5800/1020Baは、ラックサーバ並みの諸元を保ちつつ、より高密度実装を実現しています。

**写真1左**にブレードの外観を、**写真1右**にラックに搭載した際の外観を示します。なお、具体的な特徴は下記のとおりです。

1) Itanium2プロセッサを搭載することで高い演算性能を実現。また、4CPUシステム向け大容量キャッシュItanium2プロセッサに比較して、安価な2CPUシステム向けのItanium2プロセッサも選択可能。

2) メモリ容量は最大24GBまで拡張可能で、大規模データを扱う科学技術計算分野においても、高い演算性能を実現。

3) 大容量データを高速に処理できるUltra320 SCSI HDDをブレードサーバ当たり最大2個内蔵。ファイルサーバとのデータ通信オーバーヘッドを減らし、高速性能を実現。

4) PCI-X拡張スロットを1個装備。I/Oインタフェースに柔軟性を持ち、高速なInfinibandのHCAカードを搭載することで、分散並列処理において、高い処理性能を実現。

これらに加えて、多数のサーバを効率よく運用、監視するために下記の機能も搭載しています。

5) OSをネットワーク経由で効率よくインストールするDeploymentManager機能、サーバの状態を監視するESMPPro/ServerManager, ServerAgent機能を標準で添付。

これらの特長を備えるItanium2ブレードサーバにより、Itanium2のハイパフォーマンスを、効率の良い運用で利用可能なPCクラスタシステムが構築可能です。

#### (2) EM64T機能搭載のブレードクラスタソリューション

EM64T対応のXeonプロセッサを搭載したブレード型サ

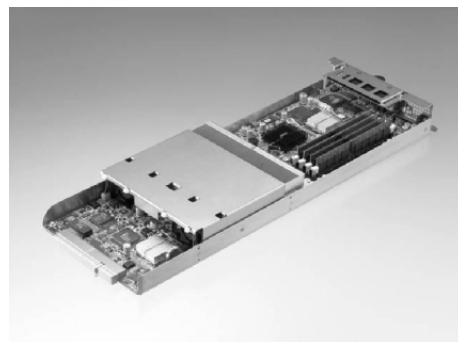
ーバExpress5800/120Ba-4は、従来製品化していたExpress5800/420Maの後継機に当たり、互換性を維持しつつ機能および性能を大幅に向上しています。

**写真2**にブレードの外観、**写真3**に収納ユニットのイメージを示します。ブレードサーバは1Uサーバに対し高密度な設置が可能となっています。なお、具体的な強化ポイントは下記のとおりです。

1) EM64T対応の高性能Xeonプロセッサをブレードサーバとしていち早く採用（最大2CPU）。またメモリ容量を最大8GBまで拡張し、より少ないノード数で大規模な演算問題に対応。

2) FSB（フロントサイドバス；CPUバス）=800MHzのCPU採用に加え、DDR2-400 SDRAM DIMM（ECC付き）を採用。科学技術計算ではキャッシュの効果を期待しにくい大規模データを扱うケースが多く、メモリ性能の向上により、演算性能の改善を期待することが可能。

また、従来機からの特長を強化し、以下の特長も備えています。



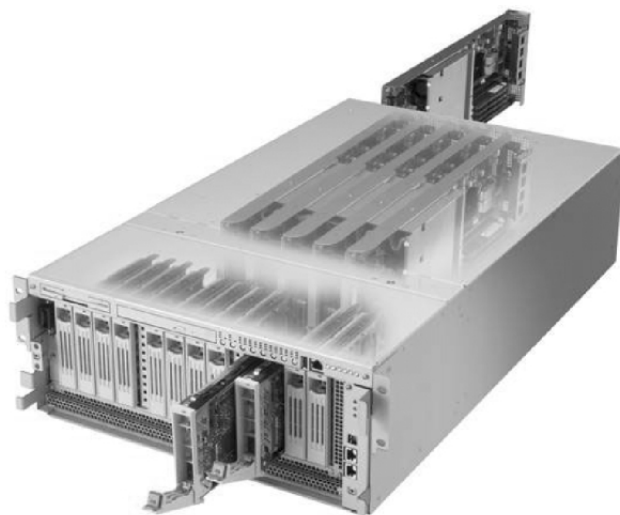
**写真2** Express5800/120Ba-4

Photo 2 External view of Express5800/120Ba-4.



**写真1** Express5800/1020Ba

Photo 1 External view of Express5800/1020Ba.



**写真3** Express5800/120Ba-4 ブレード収納ユニット

Photo 3 External view of Express5800/120Ba-4 Blade Server assemble unit.

3) 高速インターコネクトカードに威力を発揮するPCI-X インタフェースを標準装備。多数台ノードでシステムを構成する場合、演算性能全体に影響を及ぼすノード間通信遅延時間を、最大8.5Gbpsの帯域により軽減。

4) 大容量データを高速に処理できるU320 SCSI HDDをサーバ当たり最大2個内蔵。300GBのHDDを2個搭載した場合、ノード当たり600GBのデータをローカルに配置することができ、ファイルサーバとのデータ通信オーバーヘッドを削減。

これら性能面での強化に加え、多数のサーバを効率よく運用管理するための下記の機能も新たに搭載しています。

5) KVMスイッチおよびローカルデバイスを内蔵した収納ユニットを新たに製品化、煩雑なケーブル処理を効率化。さらにフルワイヤの性能を持つGiga bitインテリジェントL2スイッチ内蔵モデルも選択可能で、ノード間ネットワークの効率的な構成が可能。

6) サーバをネットワーク経由で遠隔操作する管理ソフトSystemGlobe/DianaScope、OSをネットワーク経由でインストールするSystemGlobe/DeploymentManager Liteを標準で添付、装置の初期設定、障害時の復旧、日常の運用管理を効率よく行うことが可能。

これらの特長を備えるブレードサーバにより、PCクラスタシステムをより高密度・高性能かつ効率良く運用可能です。

(3) EM64T機能搭載の1Uサーバクラスタソリューション  
EM64T対応のXeonプロセッサを搭載した1U型サーバExpress5800/120Re-1は、従来製品化していたExpress5800/120Rd-1の後継機として、機能および性能を向上しています。

**写真4**に外観を示します。この1Uサーバを使用することで、より安価なシステムが構築可能となっています。なお、具体的な強化ポイントは下記のとおりです。

1) EM64T対応の高性能Xeonをいち早く採用（最大2CPU）。

2) DDR-333/266 SDRAM DIMM（ECC付き）を採用し、メモリ容量も最大12GB（DDR266時）まで搭載可能。より安価に大規模なメモリシステムを構築可能。

3) Ultra 320 SCSI HDD（最大3個）に加え、S-ATA HDD（最大2個）が選択可能（工場出荷時選択）。300GBのHDDを3個搭載した場合、ノード当たりの900GBの大容量

システムが構築可能。また、S-ATA HDDを搭載した場合は、より安価なシステム構築が可能。

4) 1Uのサイズながら冗長FANや冗長電源をサポート（SCSIモデル）。可用性への要求にも対応。

5) PCI-X インタフェースを標準装備。オプションで新しいPCI-Express インタフェースにも対応。より高速なインターコネクトカードに威力を発揮。

これらの特長を備える1Uサーバにより、PCクラスタシステムをより高性能かつ安価に構築可能です。

### 3. グリッド推進センター

グリッド推進センターは、PCクラスタの技術センターとして、2001年に設立された「HPCエンジニアリングセンター」を統合し、PCクラスタ構築のための技術支援、保守支援をはじめ、PCクラスタ上で動作する並列アプリケーションのポーティングや動作検証、および並列化支援、性能チューニングに関するコンサルティングなどを行っています。グリッド推進センターでは、PCクラスタをより効率的に使用するために、HyperClusterEngineを製品化するとともに、PCクラスタコンソーシアムへメンバー参画し「SCoreクラスタシステムソフトウェア」（以下SCore；エスコア）発展へも寄与しています。またグリッド推進センターにはベンチマークや、動作検証を行うために64ノード/128CPUのPCクラスタおよび最新のCPUを搭載した16ノード/32CPUのPCクラスタを設置しています。これらシステムは登録すればインターネット経由でも利用可能となり、ISVやお客様からも利用されています。

#### (1) LSFとHyperClusterEngine

グリッド推進センターではPlatform社とジョブスケジューラLSFについての包括的OEM契約を締結し、HyperClusterEngineの名称で製品化しています。HyperClusterEngineは、LSFに対してNECでFailover機能を独自強化しています。この機能により、ノードの異常を検出した際に実行中のジョブをキャンセルし、自動的に予備ノードを計算ノードとして組み込み、実行していたジョブを自動的にキューに再投入し、システムの運用を継続することが可能になります。さらに、このHyperClusterEngineを核として、数千台のCAD/CAEワークステーション群とPCクラスタを統合し、WebベースのポータルUIによるジョブ実行操作の簡素化、システム管理の省力化を実現する、統合運用ソリューションの提供も行っています。

#### (2) SCoreクラスタシステムソフトウェア

PCクラスタを構築する手法としては、経済産業省の新情報処理開発機構（RWCP）によって開発されたSCoreを使用したSCore型クラスタと、標準のLinuxディストリビューションパッケージを組み合わせたBeowulf型クラスタとが主流になっています。

官公庁系のお客様に主に利用されているSCoreは、**図2**に示すように多数の計算サーバから成るPCクラスタをあた



**写真4** Express5800/120Re-1

Photo 4 External view of Express5800/120Re-1.

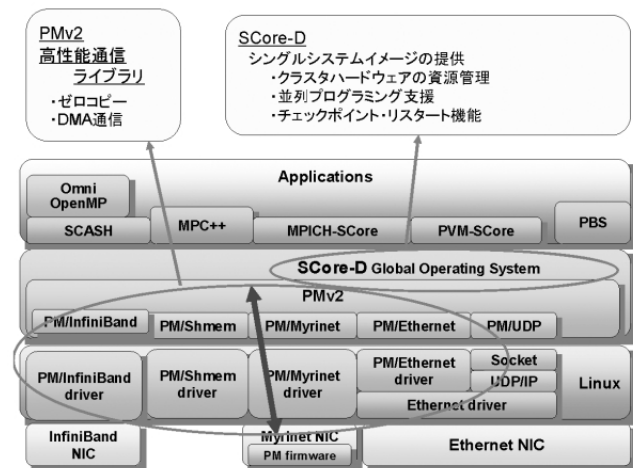


図2 SCore  
Fig.2 SCore.

かも1つのシステムとして見せるシングルシステムイメージの実現や、信頼性の高い並列処理を実現するためのチェックポイントリスタート機能を実現したSCore-D、および計算ノード間的高速通信を実現する通信モジュール (PM) のサポートなどの特長を有しています。

なおSCoreは、現在は2001年に設立された「PCクラスタコンソーシアム」に移管され、オープンソースのソフトウェアとして強化、発展が行われています。

### (3) 並列型アプリケーションの品揃え

グリッド推進センターでは、今後ともPCクラスタで利用できる、並列型アプリケーションの品揃えを推進していく予定です。

## 4. クラスタシステムインテグレーションセンター

NECの保守サポートを担当するNECフィールドイングでは、PCクラスタが普及する以前から、設置環境調査、環境工事、コンピュータやネットワーク機器のハードウェア環境構築と設定、保守サポートなど、PCクラスタに関するサポートサービスを行ってきましたが、近年のPCクラスタの普及に伴い多様化する要件に対し、より効率的に柔軟に対応するため、2003年1月に専門的な組織として「クラスタシステムインテグレーションセンター：CSIC」を設立しています。

CSICは、営業担当をはじめ、電源・空調・耐震などの設置環境調査・設計、特殊ラック設計、ラック内ケーブル配線設計、コンピュータ搭載設計を主に担当するファシリティエンジニア、ネットワーク設計・構築の専門知識を持つネットワークエンジニア、Unix/Linux技術とデータベース設計・構築の専門技術を持つシステムエンジニア、そしてコンピュータハードウェア保守のトレーニングを積んだカスタマエンジニアで編成されており、それぞれの専門分野の知識と豊富な経験を持つエンジニアがチームワークを活かして、対応に当たっています。

またCSICは、環境整備された専用コンピューター室を有し、小規模から中規模システムはもちろん、大規模システムまで19インチラックへのコンピュータ搭載から環境設定、電源、ネットワークケーブルのラック配線、ネットワーク機器との接続にいたるまで、PCクラスタがシステムとして機能するまでの基本構成を組み上げ、連続稼働試験、負荷試験を通して品質確認を行ったうえでお客様へシステムをお渡しすることを可能にしています。

特にPCクラスタのような疎結合型コンピュータシステムでは、薄型ラックマウントモデル、ブレード型モデルなどの機種別、そして構成するコンピュータの台数やネットワーク機器、電源系統など様々な要素によって、個別に排熱効率を考慮することが重要になります。CSICでは、ラックの構成・実装設計、コンピュータの実装方法、そしてラック内のケーブル配線方法など、過去に多数設置したシステム構築経験からそのノウハウを活かし、排熱効率を高めた高信頼性、高品質なシステム造りを実現し続けています。

システム納入後のアフターサポートにおいても、同社の社内教育カリキュラムによりExpress5800シリーズのトレーニングを受けた全国の保守サポート要員が、整備された保守マニュアルに則って対応しており、PCクラスタに特化した事象については、CSICが支援する体制を整えています。

## 5. むすび

官公庁系の研究機関で利用され始めたPCクラスタも現在は、構造解析や衝突解析をはじめ幅広く実業務のなかで活用されています。図3にすでに提供済み、または評価済みのアプリケーションを示します。

今後もNECは広がるPCクラスタの活用分野にExpress5800を核として、グリッド推進センター・CSICと一体となった、製品・サービスを提供してこれら活用分野に対応していく予定です。

バイオ	流体解析
BLAST :相同性検索	STAR-CD
AMBER:分子動力学シミュレーション	VECTIS
Gaussian	FLUENT
気象	SCRYU
WRF (評価予定)	STREAM
SWAN (評価予定)	Icepak
衝突解析/構造解析	PHOENICS
LS-DYNA	PowerFLOW
PAM-CRASH	PAM-flow
PAM-stamp	音響解析
RADIOSS	SYSNOISE
ADventureCluster	
GT-Power	

図3 提供・評価済みアプリケーション一覧

Fig.3 List of already worked/evaluated applications.

- \* Inanium, Xeon, Pentiumは、米国またはその他の国におけるインテルコーポレーションおよび子会社の登録商標または商標です。
- \* UNIXは、The Open Groupの登録商標です。
- \* LSFは、Platform Computing Corporationの米国およびその他の国・地域における登録商標です。

## 筆者紹介



Atsushi Kubo

くぼ あつし  
**久保 淳**

1990年、NEC入社。現在、第二コンピュータ事業本部クライアント・サーバ事業部第一製品技術部技術エキスパート。



Tadashi Okano

おか の ただし  
**岡野 格**

1984年、NEC入社。現在、第二コンピュータ事業本部クライアント・サーバ事業部第一技術部技術エキスパート。



Hideshi Ishii

いし い ひでし  
**石井 英志**

1982年、NEC入社。現在、第一コンピュータ事業本部HPC販売推進本部HPC製品計画マネージャー。



Tetsuhide Senta

せん た てつひで  
**千田 哲秀**

1985年、NEC入社。現在、第一コンピュータ事業本部コンピュータ事業部第三技術部長。



Tetsuya Hirose

ひろ せ てつ や  
**広瀬 哲也**

1992年、NEC入社。現在、システムソフトウェア事業本部グリッド推進センター主任。



Hitoshi Ishida

いし だ ひとし  
**石田 仁**

1982年、NECフィールディング入社。現在、東日本サービス営業本部ソリューション営業部クラスタシステムインテグレーションセンター長。