

[大規模科学計算システム]

LX 406Re-2 のハードウェア

島本 浩樹 小林 公雄 長沢 富人

日本電気株式会社 IT プラットフォーム事業部 第二サーバ統括部

1. はじめに

LX 406Re-2 は、米インテル社®の Xeon® E5-2600 v2 プロセッサをノードあたり 2 基搭載可能で、4 ノードを 2U の筐体に収容する、高性能、高効率、高密度を実現した PC クラスタ型計算サーバです。LX 406Re-2 は、オンボードの InfiniBand FDR*1 を 1 ポート搭載することにより、データセンターだけでなく HPC(High Performance Computing)からの低 TC0(Total Cost of Ownership)需要に応える製品となっています。本稿では、LX 406Re-2 システムのハードウェアについてご紹介します。

*1 主に HPC やデータセンターで使用されている、計算ノード間を結合するためのインターコネクタ用ネットワーク。ポートあたり 56Gbps の通信帯域があり、低遅延、高信頼性が特長。

2. ハードウェア構成

2.1 特長

表 1 に、LX 406Re-2 の諸元を示します。特長を要約すると以下のようになります。

- (1) 2U 筐体に 4 ノードを搭載でき、ノード単位でのホットプラグを実現
- (2) ノードあたりインテル® Xeon® E5-2600 v2 プロセッサを 2 基搭載可能
- (3) DDR3-1866 対応 DIMM スロットを 8 基有し、16GB DIMM 使用時には最大で 128GB のメモリ容量を実装可能
- (4) IO スロットとしてノードあたり 1 スロットの PCI-Express 3.0 (x16) スロットをサポート
- (5) 高速なインタコネクタでの PC クラスタを構成するために、ノードあたり 1 ポートのオンボード InfiniBand FDR(56Gbps)を実装
- (6) マネージメント LAN として、各ノードにインテル®i350 デュアルポートを搭載し、Gigabit Ethernet をサポート
- (7) HDD はホットスワップに対応し、ノードあたり最大 6 個の 2.5 インチ HDD ドライブが搭載可能
- (8) 電源は、80 PLUS® Platinum レベルの高効率な 1620W 電源を搭載し、冗長構成をサポート
- (9) IPMI v2.0 に対応し、障害の監視、障害情報の収集、障害発生箇所の検出等をサーバ管理ソフトウェア ESMPRO と共に実現

表 1. LX 406Re-2 諸元
筐体諸元

項目	諸元
筐体サイズ	438(幅) x 679(奥行) x 88(高さ) mm
重量	最大 38.6kg(レール含む)
搭載ノード数	4
電源ユニット	1620W 80 PLUS® Platinum、2 個(冗長電源)

ノード諸元

項目	諸元
CPU	インテル® Xeon® プロセッサ E5-2600 v2 、最大 2 台
メモリ	DDR3-1866 対応 8 スロット (最大メモリ容量 128GB[16GB DIMM 使用時])
HDD	2.5” SATA HDD x6
拡張スロット	PCI-E3.0(x16) low-profile x1
LAN	1000BASE-T x2
IPMI	IPMI 専用 LAN ポート(RJ45) x1
InfiniBand	オンボード FDR(56Gbps) x1

写真 1 に、LX 406Re-2 の装置外観(前面/背面)を示します。



前面



背面

写真 1. LX 406Re-2 装置外観

2.2 基本構造

図 1 に、LX 406Re-2 筐体のハードウェア構成を示します。2U シャーシの中に、独立したノードを 4 台(ノード A/B/C/D)搭載しています。各ノードは、電源と FAN を、HDD バックプレーンを介して共有しています。

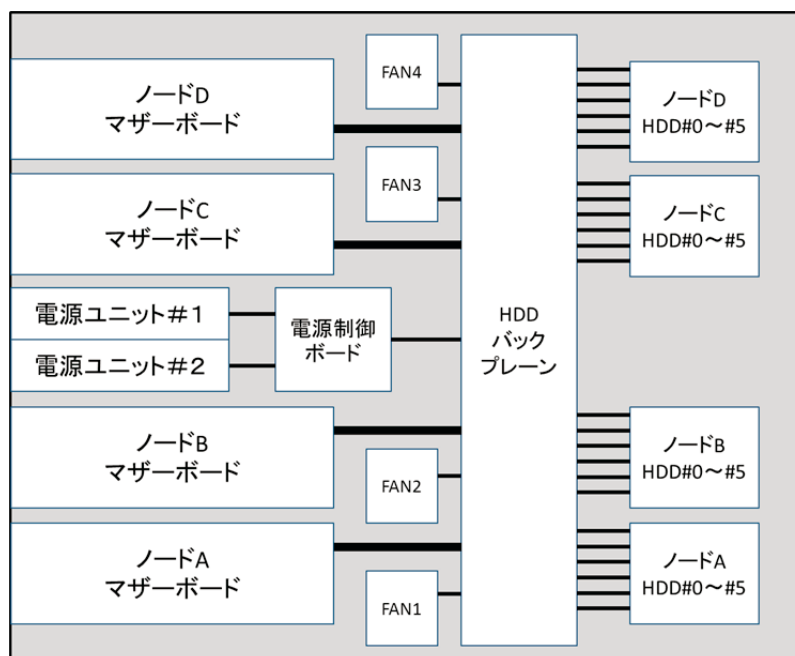


図 1. LX 406Re-2 筐体ハードウェア構成

図 2 にシステムブロックダイアグラム(ノード)を示します。

LX 402Re-2 では、デュアルプロセッサ向け Intel® C602 チップセットを使用しています。プロセッサ間の接続は、2本の QPI (Quick Path Interconnect) を用いています。CPU に統合されたメモリコントローラは、プロセッサあたり 4 チャンネル対応です。プロセッサとチップセット間の接続は DMI を使用しています。SATA ディスクの制御は、チップセットが行い、最大 6 個の SATA ディスクを接続可能です (SATA3 6Gbps x2、SATA2 3Gbps x4)。IPMI 制御用に、Winbond WPCM450 コントローラを搭載しています。OS とは独立して動作しますので、随時、装置内の各種センサー情報を収集・監視することができます。また、IPMI 専用 LAN ポートから WPCM450 へ、直接アクセスすることが可能ですので、CPU やチップセットが正常に機能しない状況でも IPMI による情報収集が可能です。

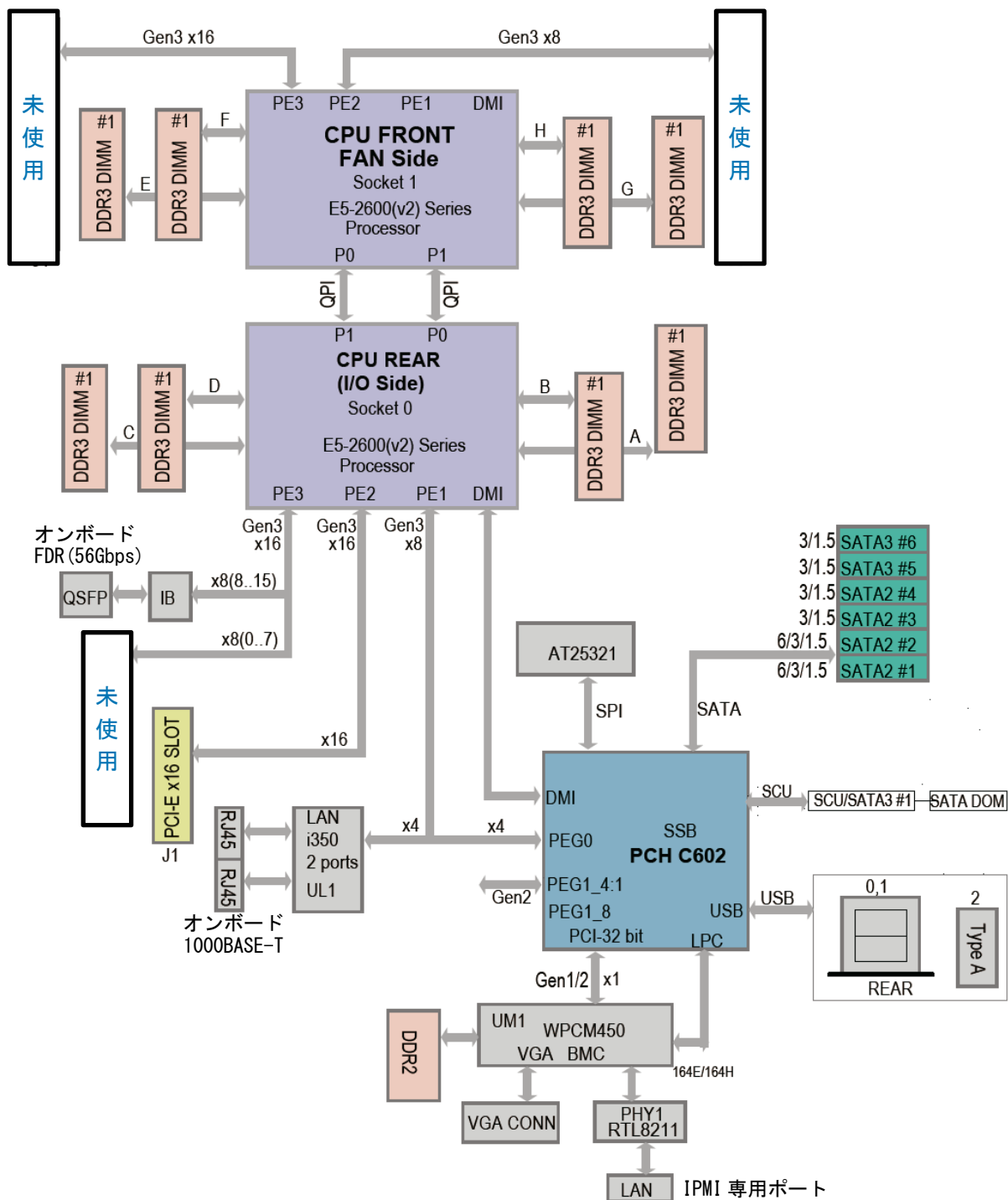


図 2. LX 406Re-2 システムブロックダイアグラム

3. インテル® Xeon® プロセッサ E5-2600 v2 製品ファミリー

3.1 インテル® Xeon® プロセッサ E5-2600 v2 製品ファミリーの特長

インテル® Xeon® プロセッサ E5-2600 v2 製品ファミリーは、前世代製品と比べて最大コア数とキャッシュ容量を 50%拡張し、さらに、メモリの高速化を行うことで、前世代製品と比較し、大幅な性能向上を実現しています。また年々上昇し続けている電力と冷却にかかる費用に対応するため、電力使用の効率化、エネルギー消費に応じた性能の最適化を行い、電力コストの削減に努めています。

具体的な内容を以下に紹介します。

- ・ **大容量キャッシュと高速メモリ**

最大 30MB の L3 キャッシュで頻繁に使用されるデータへの高速アクセスを実現します。また、前世代製品と比べて高速な DDR3-1866 のメモリをサポートします。

- ・ **浮動小数点演算の性能向上を実現する AVX**

インテル®アドバンスド・ベクトル・エクステンション(インテル® AVX)は、256 ビットのベクトルと 32/64 ビット・データ変換の加速をサポートします。ベクトル計算と浮動小数点演算を高速化することで、HPC アプリケーションの性能向上を実現します。

- ・ **ワークロードのピーク時に、より優れた性能向上を実現**

インテル® ターボ・ブースト・テクノロジー 2.0 は、電力と温度のヘッドルームを並列に管理しながらプロセッサの動作周波数を自動的に引き上げます。第二世代となるこのテクノロジーは、必要に応じてよりインテリジェントに動作周波数を引き上げ、性能と電力効率の更なる向上を実現します。

- ・ **インテル® インテグレートッド I/O で帯域幅を拡張**

PCI-Express 3.0 レーンをプロセッサ・ダイに直接統合することで、2-way サーバ構成の場合にノード 1 台当たりの PCI-Express 3.0 のレーン数を最大 80 まで提供します。またこのテクノロジーは、PCIe 2.0 仕様のソリューションと比べて I/O レイテンシーを最大 30%削減し、最大 2 倍の帯域幅の拡張を実現します。さらに、ストレージと科学技術計算向けに最適化された機能として x16 ノン・トランスペアレント・ブリッジ (NTB) を新たにサポートすることで、P2P (ピアツーピア) 間での帯域幅の拡張を実現します。

- ・ **必要なデータを必要な場所に迅速に配置**

インテル®データ・ダイレクト I/O テクノロジーは、ストレージからのデータを直接キャッシュに転送することで、I/O 性能の向上を実現し、メモリアクセスによるオーバーヘッドを軽減します。

- ・ **業界最高水準のエネルギー効率を実現**

インテルの 22nm 3-D トライゲート・トランジスタは、同じ性能レベルで動作する前世代のトランジスタと比べて、消費電力を大幅に削減します。

3.2 インテル® Xeon® プロセッサ E5-2695 v2 諸元

表 2 に、インテル® Xeon® プロセッサ E5-2600 v2 製品ファミリーのなかで、コア数と L3 キャッシュ容量が最大の E5-2695 v2 に関する諸元を示します。

表 2. インテル® Xeon® プロセッサ E5-2695 v2 諸元

項目	仕様
コア数	12
スレッド数	24
動作周波数	2.4 GHz
ターボ・ブースト利用時の最大周波数	3.2 GHz
インテル® スマート・キャッシュ	30 MB
インテル® QPI 速度	8 GT/s
QPI リンク数	2
命令セット	64-bit
命令セット拡張	AVX
リソグラフィー	22 nm
最大 TDP	115W
メモリの種類	DDR3-800/1066/1333/1600/1866
メモリチャネル数	4
最大メモリ帯域幅	59.7 GB/s
物理アドレス拡張	46-bit

3.3 インテル® Xeon® プロセッサ E5-2695 v2 のキャッシュメモリ仕様

表 3 に、インテル® Xeon® プロセッサ E5-2695 v2 のキャッシュメモリ仕様を示します。

表 3. インテル® Xeon® プロセッサ E5-2695 v2 キャッシュメモリ仕様

キャッシュレベル	形式	容量
L1 (命令)	コア毎	32KB
L1 (データ)	コア毎	32KB
L2	コア毎	256KB
L3	全コア共有	30MB

4. ハードウェア性能

インテル® Xeon® プロセッサ E5-2600 v2 製品ファミリーでは、拡張命令セット AVX を新たに採用し、さらに性能向上を実現したマイクロアーキテクチャと、単一 CPU ソケット内に最大 12 プロセッサコアを内蔵したマルチコア CPU 構造を採用しています。図 3 に、プロセッサコアの内部構造を示します。Port0 の AVX が 256-bit SIMD 演算(乗算)、Port1 の AVX が 256-bit SIMD 演算(加算)を行うことで、1 クロックサイクルあたり最大 8 回の倍精度演算を行います。インテル® Xeon® プロセッサ E5-2695 v2 の場合、CPU ソケットあたりの最大計算能力は 230.4 GFLOPS となります。

LX 406Re-2 は、2CPU/8DIMM で構成するノードを基本単位とし、各ノードを高速な InfiniBand ネットワークで結合したクラスタを構築することで、大規模並列計算を行います。CPU の計算能力向上にあわせて、ノード内の各種インターフェイスも高速化しています。

たとえばメモリについては、1866MHz で動作する DDR3 を新たにサポートし、プロセッサあたり、59.7GB/S のメモリ転送性能を実現しました。図 4 に、ノード内の主要コンポーネントのインターフェイス性能を示します。

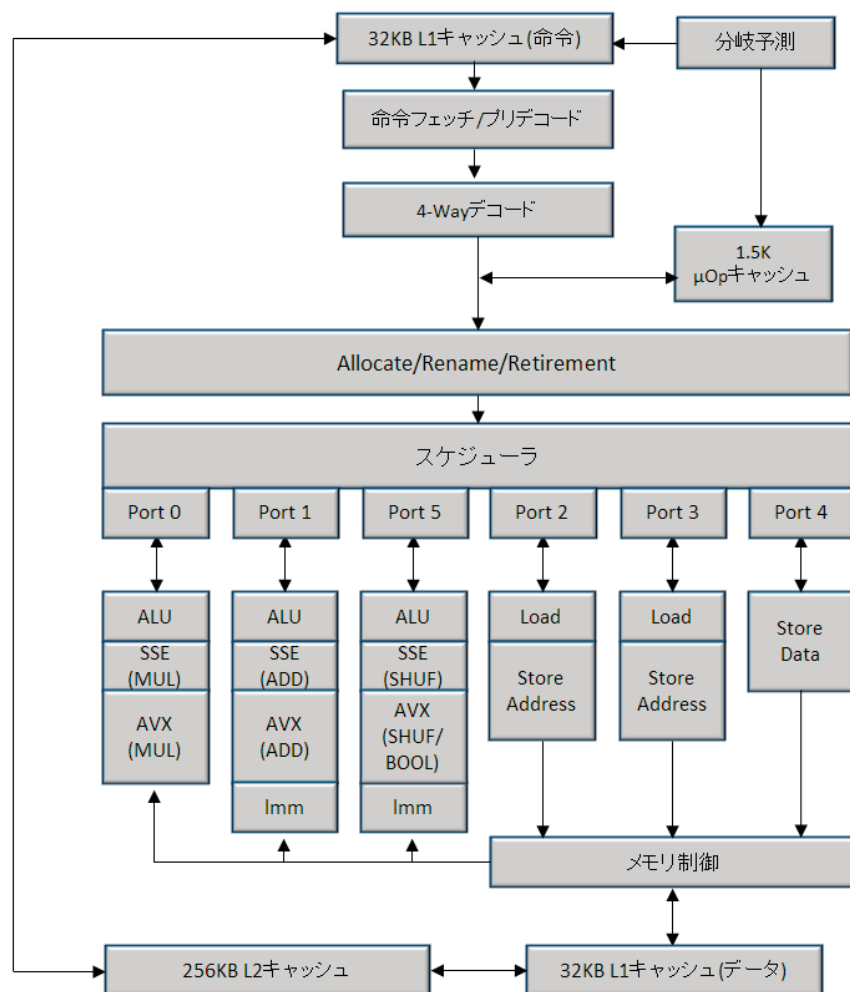


図 3. インテル® Xeon® プロセッサ E5-2695 v2 プロセッサコア内部構造

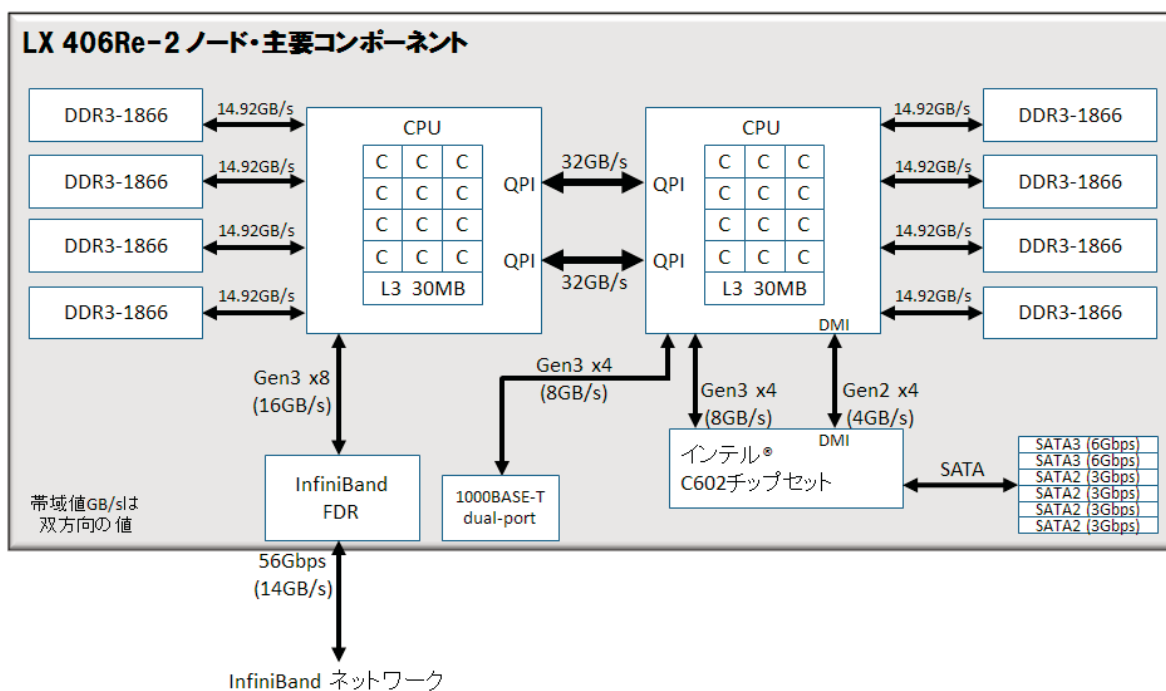


図 4. LX 406Re-2 ノード主要インターフェイス性能

5. むすび

以上、PC クラスタ型計算サーバ LX 406Re-2 について述べました。科学技術計算の重要性と需要が年々高まっているのと同時に、消費電力の削減や冷却コストの削減が大きな課題として重要視されています。NEC は、最新テクノロジーの採用と、HPC 分野で培ったノウハウをもとに、高性能・省スペース・省電力の製品を継続して開発し、低 TCO を実現する最適なソリューションを提供してまいります。