

大規模 SMP 並列スーパーコンピューターシステム SR16000 モデル M1 のご紹介

(株) 日立製作所

1. はじめに

2005年3月に納入し、2007年3月に中間レベルアップを行ったベクトル並列型スーパーコンピューターSR11000 モデル J2 (128 ノード) の後継機として、2011年10月に大規模 SMP 並列スーパーコンピューターシステム SR16000 モデル M1 を納入する予定です。SR16000 モデル M1 は、SR11000 の POWER プロセッサ、AIX OS、コンパイラや各種ライブラリーを継承することで、現行システムとの高い互換性を持ち、また、最新の POWER7 プロセッサによる高い計算能力を提供します。ここでは、2011年10月から稼働を開始する大規模 SMP 並列スーパーコンピューターシステム SR16000 モデル M1 (以降、SR16000 と呼びます) について、紹介します。

2. SR11000 と SR16000 の比較

表 1 に、東京大学情報基盤センターの現行システム SR11000 と次期システム SR16000 の比較を示します。

表 1. SR11000 と SR16000 の比較

項目	SR11000 (現行システム)	SR16000 (次期システム)
システム性能	18,841.6GFLOPS	54,906.88GFLOPS
ノード数	128	56
1 ノード当たりの 主記憶容量	128GB	200GB
ノード性能	147.2GFLOPS	980.48GFLOPS
1 コア当たり演算性能	9.2GFLOPS	30.64GFLOPS
1 ノード当たりのコア数	16	32
SMT*機能	なし	あり (最大 4 スレッド/コア)
ノード間ネットワーク	3 次元クロスバー	階層型完全結合
ノード間転送性能	12GB/s (単方向) × 双方向	96GB/s (単方向) × 双方向
磁気ディスク装置	94.2TB	556TB
OS 及び構成	AIX V5.3 クラスター構成	AIX V7.1 クラスター構成

*SMT : Simultaneous Multi-Threading

3. ハードウェアの特長

SR16000 は、科学技術計算システムに要求される「高性能ノード」、「高スケーラビリティ（高いシステム性能）」、「高信頼性」を追求するコンセプトに基づいて開発しています。SR16000 のノード構成を図 1 に示します。

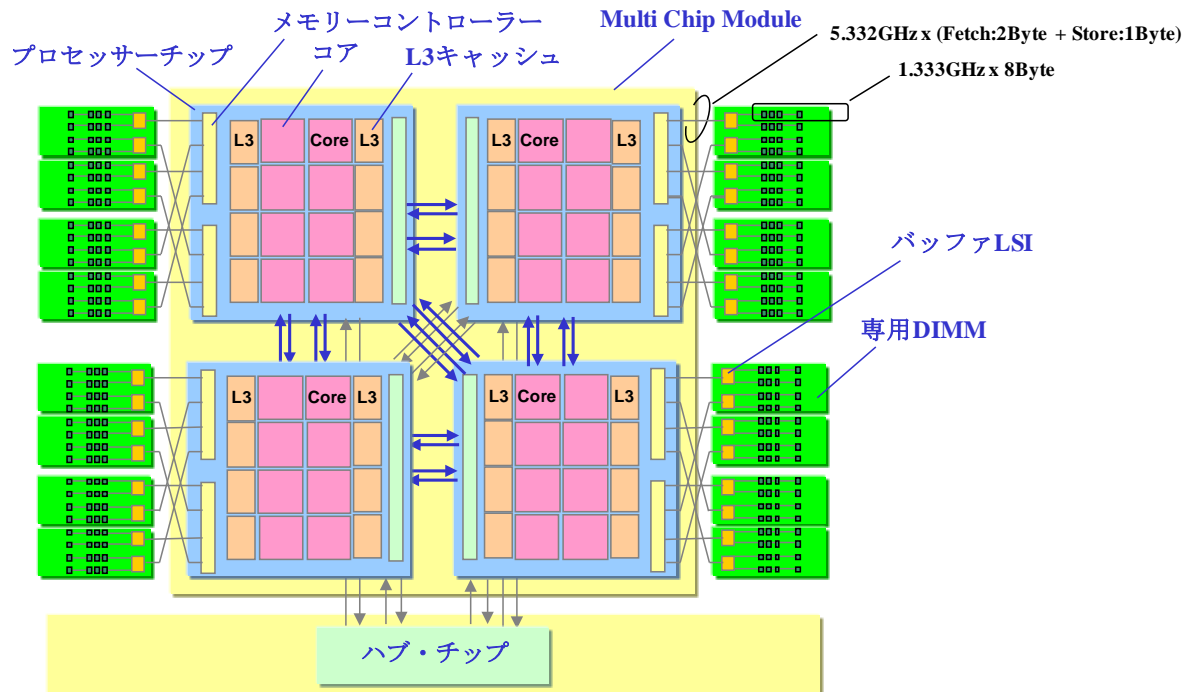


図 1. ノード構成図

SR16000 の 1 ノードは、4 個の POWER7 プロセッサチップ（合計 32 コア）を相互に接続した MCM（Multi Chip Module）と 16 枚の専用 DIMM で構成され、ノード間接続のためのハブ・チップに接続されています。SR16000 の特長は以下の通りです。

(1) 高性能プロセッサと大容量キャッシュによる高性能ノード

銅配線技術や SOI（Silicon on Insulator）技術による先進の 45nm CMOS テクノロジーによって実現した POWER7 プロセッサは、3.83GHz の高速クロックレートで動作し、8 台のコアを実装するプロセッサチップ当たり 245.12GFLOPS の高い性能を有します。さらに、32MB の L3 キャッシュをプロセッサチップ上に搭載することで、L3 キャッシュに対するレイテンシーを大幅に短縮し、実効性能を向上させています。表 2 に SR16000 が搭載する POWER7 プロセッサの仕様を示します。POWER7 プロセッサでは、1 台のコア内で複数のスレッドを同時実行できる SMT（Simultaneous Multi-Threading）機能を有しており、SR16000 では 1 台のコアを仮想的に 2 または 4 台のコアとして扱うこともできます。複数のスレッドを同時実行することでプロセッサの利用効率が向上できます。

また、ノードを構成する 4 台のプロセッサを一斉に、しかも高速に同期させる「協調型マイクロプロセッサ機構」を採用し、効率のよい並列処理が実行できます。

表 2. POWER7 プロセッサ仕様

項目	仕様
クロックレート	3.83GHz
コア当たりの理論演算性能 (倍精度浮動小数点)	30.64GFLOPS (乗算 2FLOP+加算 2FLOP) × 2 演算器 × 3.83GHz
L1 キャッシュ	命令 32kB/コア+データ 32kB/コア
L2 キャッシュ	256kB/コア
L3 キャッシュ	32MB/チップ

(2) メモリー先読み機構とソフトウェアパイプライン

メモリー先読み機構は、ハードウェアが自動的に主記憶データアクセスのパターンを検出し、先読みデータをキャッシュに取り込むことで、キャッシュヒット率を向上し、実効メモリーレイテンシーの削減と実効メモリースループットの向上を図る機能です。

ソフトウェアパイプライン技術によりプロセッサの演算器をパイプライン的に動かすことによって、高速処理を実現します。

(3) 高速ノード間ネットワーク

SR16000 では、専用ノード間ネットワークによる階層型完全結合の高速ネットワークを採用しました。各ノードに専用のノード間ネットワーク制御 LSI (図 1 におけるハブ・チップ) を実装し、プロセッサとハブ・チップ間の転送性能は、96GB/秒 (単方向) × 双方向と高速です。図 2 にノード間ネットワークの構成図を示します。

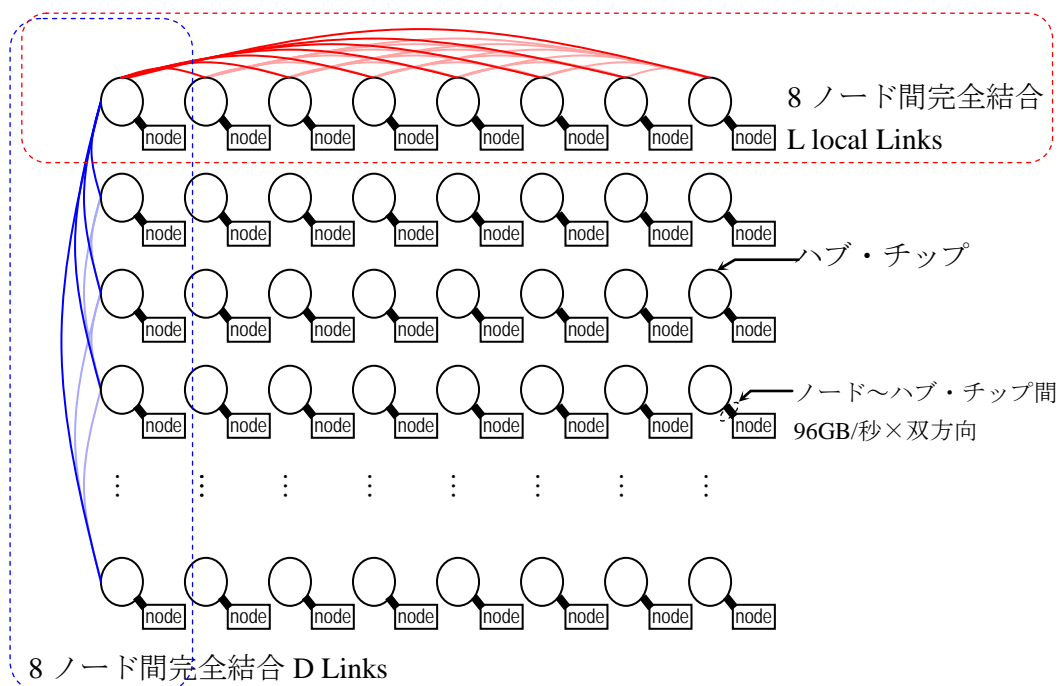


図 2. ノード間ネットワーク

(4) 高信頼性

多彩な高信頼化、高可用化機能により、高速処理の安定化を図っています。

- ・プロセッサには、データ／アドレス系に ECC (Error Correction Code) またはパリティチェッカーを備えます。
- ・L1 キャッシュには、パリティチェックによる誤り検出を備えます。L2/L3 キャッシュには、1 ビット誤り訂正、2 ビット誤り検出可能な ECC を備えます。
- ・主記憶にはメモリー素子の 1 チップ誤り訂正、2 チップ誤り検出可能な ECC を備えます。メモリーラインの 1 ビットエラーが閾値に達した場合に、ラインを予備メモリーチップに交代するメモリービット交代機能を備えます。
- ・ハードウェア障害情報の自動通報機能を備えます。

(5) 環境への配慮

プロセッサの発熱を冷却水により効率的に冷却します。この冷却水は、ラック内に装備される冗長熱交換ユニットにおいて、設備冷却水と熱交換を行い、排熱されます。さらに排気熱の一部は高効率熱交換水冷リアドアによって吸収されます。これらの水冷技術により、総発熱量の 100%を排熱します。

4. ソフトウェアの特長

SR16000 の OS (オペレーティングシステム) には、米国 IBM 社の AIX の最新版であるバージョン 7.1 (AIX V7.1) を採用しています。SR11000 でご利用いただいている AIX V5.3 の後継バージョンであり、SR11000 と同様のコマンドインターフェースを提供します。

また、コンパイラや数値計算ライブラリーも、弊社の最適化 FORTRAN や MATRIX/MPP といった製品をご提供し、SR11000 上でご利用いただいているプログラムを変更することなく、ソースプログラムを SR16000 上でリコンパイルするだけで移行できます。

4.1 最適化 FORTRAN90

最適化 FORTRAN90 は、ISO 国際規格 ISO1539;1991、米国標準規格 ANSI X3.198-1992 及び、JIS X 3001-1994 (Fortran90) /JIS X3001-1:1998 (Fortran95) 規格に準拠した Fortran コンパイラです。また、OpenMP 2.5 仕様をサポートします。

最適化 FORTRAN90 は以下の特長を有します。

・POWER7 新命令への対応

POWER7 アーキテクチャー単体プロセッサ性能を引き出すために、POWER7 で新規追加された Vector Scalar eXtension 命令 (VSX 命令) を出力することが可能です。

・性能チューニング機能

プログラム実行に要した CPU 時間、総実効命令数、浮動小数点演算数をプロセス単位、関数単位、SMP 並列単位に取得することが可能です。

- ・プログラム高速化を支援する最適化診断

SMP 並列化および最適化についての診断メッセージをソースリスト上に出力できます。この診断メッセージにより、ユーザーが容易に、並列化および最適化の状態を把握することが可能です。更に、並列化不可部分については、メッセージに基づきユーザーが依存性を解消する指示文を挿入することにより、更なるプログラムの並列化を行うことが可能です。

4.2 最適化 C

最適化 C は、国際標準規格 ISO/IEC 9899:1990 及び米国標準規格 ANSI X3.159-1989 に準拠する C コンパイラーです。また、コンパイルオプションの指定により旧言語仕様 (K&R 仕様) に対応した互換仕様を利用できます。OpenMP2.5 仕様をサポートします。

4.3 最適化標準 C++

最適化標準 C++ は、国際標準規格 ISO/IEC 14882:1998 をサポートし、高度な最適化機能を持つ C++ 言語のコンパイラーです。また、コンパイルオプションの指定により、旧言語仕様 (ARM 仕様) に対応した互換仕様を利用できます。OpenMP2.5 仕様をサポートします。

4.4 IBM XL C/C++

IBM XL C/C++ は、国際 C++ 標準 ISO/IEC 14882:1998、2003 年改訂国際 C++ 標準 ISO/IEC 14882:2003(E)、および C (ISO/IEC 9899:1999) に加えて、OpenMP 3.0 仕様をサポートし、Unicode のサポートに関する最新の C 標準委員会の勧告を取り入れています。VSX 命令のサポートや POWER7 アーキテクチャー向けに調整されたベクトル MASS ライブラリーを提供します。

以上