

# スーパーコンピュータ SX-7C ハードウェア

百瀬 真太郎<sup>1</sup> 幅田 伸一 萩原 孝

## 1. はじめに

2006年4月より運用を開始した、新並列コンピュータシステムであるベクトル演算サーバ SX-7C のハードウェアについて紹介します。SX-7C は、SX-7 のアーキテクチャを継承しつつ、単一 CPU 性能を約 2 倍に引き上げ、SX-7 向けに開発、または利用可能であるソフトウェア資産の処理が可能なシステムです。ノードあたり 8 台の中央処理装置 (CPU) を有し、ノード間を専用の超高速クロスバネットワークにより接続し、システム全体で 5 ノード、総 CPU 数 40、総合演算性能 640GFLOPS、総メモリ容量 640G バイトを有する、実行演算性能、スケーラビリティ、及び使いやすさに優れたスーパーコンピュータです。特長は以下の通りです。

- ① 90nm 銅配線技術を用いた超高速、高集積 CMOS LSI による高密度実装
- ② ベクトル平方根演算の新規サポートによる性能強化
- ③ スカラユニットにおけるスーパースカラ方式、投機実行、データパイプライン強化、及びクロックサイクル強化による高いスカラ実効性能
- ④ 省電力設計による消費電力・発熱量の削減、及び高密度実装による設置面積の削減

本稿では、上記特性を備えた SX-7C のアーキテクチャ、システム構成、及びテクノロジーの概要について説明します。

## 2. SX-7C のハードウェア

### 2.1. ハードウェア構成

SX-7C は、CPU と主記憶装置 (Main Memory Unit: MMU) を密結合した共有メモリ型のシングルノードシステム 5 台を、超高速なノード間接続装置 (Internode Crossbar Switch: IXS) によりクラスタ接続することにより、分散並列処理を可能としたシステムです。SX-7C のシステム構成を図 1 に示します。

---

<sup>1</sup> 日本電気株式会社 コンピュータ事業部第四技術部

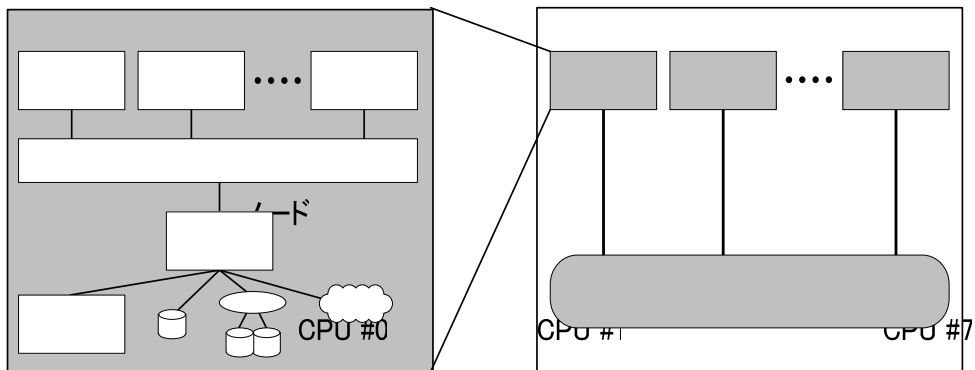


図 1 SX-7C システム構成

シングルノードシステムは、CPU 数 8，総合演算性能 128GFLOPS，主記憶容量 128G バイト，入出力チャンネル数 13 であり，演算性能，メモリスループット性能，入出力性能などのトータルバランスに優れ，高い実効性能を実現します。

### 入出力機構

#### IOF

表 1 SX-7C システム主要諸元

中央処理装置 (CPU)	CPU 数	40
	総合演算性能	SAN 640GFLOPS Network
主記憶装置 (MMU)	主記憶容量	640GB
	データ転送性能	2.5TB/s
入出力機構 (IOF)	チャンネル数	65CH
	転送速度	16GB/s
ノード間接続装置	データ転送性能	80GB/s

SX-7C システム全体は，ノード間を超高速に接続する専用の IXS により，5 ノードをクラスタ接続したシステムであり，表 1 に示すように総 CPU 数 40，総合主記憶容量 640G バイト，総入出力チャンネル数 65 の構成であり，総合演算性能 640GFLOPS という高い演算性能を実現します。

設置性と運用面に関しては，SX-7C の 1 ノードにおける保守エリアを含む設置面積，及び消費電力はそれぞれ 2.55m<sup>2</sup>，及び 7.27KVA であり，1 チップベクトルプロセッサ，及びメモリモジュールを最新の高密度実装技術により装置パッケージングすることにより，設置環境，及び消費電力を大幅に改善しています。次節以降，SX-7C システムのハードウェアについて説明します。

## 2.2. CPU

SX-7C の CPU は従来の SX アーキテクチャを継承しつつ、さらなる機能・性能の強化を図っています。図 2 に CPU の構成を示します。CPU は、スカラユニット部、及びベクトルユニット部により構成され、プロセッサ／メモリ間ネットワークを介して MMU と接続されます。スカラユニットは命令の解読、ベクトルユニットへのベクトル命令の供給・起動、及びスカラ命令の実行を行います。一方、ベクトルユニットは 4 セットのベクトルパイプラインを備え、2GHz で動作します。クロックサイクルあたり 4 個の加算と乗算を同時に処理することが可能であり、最大 16GFLOPS のベクトル演算性能を実現します。

ベクトルユニットはベクトル演算部、及びベクトル制御部から構成されます。ベクトル演算部は、論理演算、乗算、加算／シフト演算、及び除算／平方根演算の 4 種のそれぞれ独立に動作可能なパイプライン、マスク演算パイプライン、ロード／ストアパイプライン、マスクレジスタ、及びベクトルレジスタにより構成されています。ベクトル演算部全体は、以上のユニット 4 セットからなる多重パイプライン構成であり、強力にベクトル演算処理を実行します。特に、SX-7C は除算パイプラインに平方根演算機能を追加することにより、倍精度浮動小数点平方根演算において 1GFLOPS という高い演算性能を実現しています。

スカラユニットは 64 ビット RISC アーキテクチャであり、L1 キャッシュ、命令同時デコード数 4、命令同時発行数 4 であるスーパースカラアーキテクチャ、アウトオブオーダー実行、及び命令投機実行の採用により短ベクトル処理性能、及びスカラ性能を大幅に向上させ、2GFLOPS の最大演算性能を実現しています。

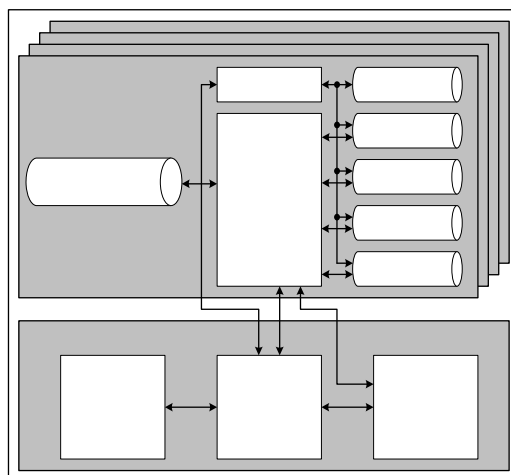


図 2 CPU 構成

## 2.3. 主記憶装置(MMU)

スーパーコンピュータにおいて高い実効性能を実現するためには、高い演算性能に見合うだけの高いデータ供給性能が必要となります。SX-7C の MMU は、CPU とリモートアクセス制御装置(Remote access Control Unit:RCU)に接続され、高速、かつ均一にメモリアクセス可能な共有メモリ方式を採用しています。メモリを 4,096 個のバンクに分割し、それぞれ独立に制御するインターリーブ方式、及びメモリバンクキャッシュの採用によりバンク競合を最小限に抑えつつ、高いメモリスループット性能を実現しています。

これにより SX-7C は表 2 に示すように、シングルノードにおいて、128GB のメモリ容量、及び 512G バイト/秒のメモリスループット性能、システム全体では 640GB の総メモリ容量、及び 2.5T バイト/秒の総合メモリスループット性能を実現します。また、MMU は通信レジスタ(Communication Register:CR)を備えており、自動並列処理や、Open MP 指示行による共有メモリ内並列処理時の CPU 間同期制御を高速に実行することが可能です。

SX-7C は、SX-4 以降採用している 3 次元実装構造の MMU カード実装を引き続き採用しています。これにより、RAM とメモリ制御部間の物理的距離を短くし、RAM の高速動作、及び RAM と CPU 間的高速信号伝送を実現しています。一方、メモリ信頼性確保のための ECC(Error Check and Correct)符号、タイミング、パリティ、2 重化回路などの採用による高いメモリ故障検出率の実現、擬似障害によるチェック回路の診断機能、エラー内容から即座にエラー箇所を指摘するビルトイン機能などにより、RAS(Reliability, Availability, Serviceability)機能の充実を図り信頼性を高めています。

表 2 主記憶装置諸元

	シングルノード	システム全体
主記憶装置容量	128GB	640GB
インターリーブ数	4,096	20,480
CPU あたりのデータ供給能力	64GB/秒	64GB/秒
最大データ供給能力	512GB/秒	2.5TB/秒

## 2.4. 入出力装置

入出力機構(Input Output Features:IOF)はシステムのスループットを高く保つために、SX-7C の高いプロセッサ性能、及びメモリ性能に見合った高速なデータ転送性能を備えており、ノードあたり 13 チャンネル、3.2G バイト/秒、SX-7C システム全体では 65 チャンネル、16GB/秒の性能を有します。入出力インタフェースは、ファイバチャンネル、Ultra320 SCSI チャンネルなどの汎用インタフェースをサポートしており、様々な周辺機器を接続することが可能です。また、ネットワークインタフェースとして、ジャンボフレームに対応したギガビット・イーサネットなどが利用可能で

す。

I/O 動作において、CPU は全ての I/O 装置に対して対等にアクセスすることが可能であり、実行負荷の低いCPUをI/O制御に割り当てるなど、CPUの効率的利用を可能としています。また、I/O 制御方式として Direct I/O 方式を採用し、低レイテンシ、かつ高スループットな I/O 制御を実現しています。

## 2.5. ノード間接続装置

SX-7C は図 3 に示すように、共有メモリ型のシングルノード 5 台を超高速専用クロスバススイッチである IXS を介して結合することにより、分散並列処理を可能としています。各シングルノードシステムは、RCU を介して IXS とケーブル接続され、分散並列処理において低通信レイテンシ、及び高通信スループットを実現します。

RCU はノードあたり 2 台実装され、各 RCU のデータ受信部、及び送信部はそれぞれ独立に動作可能であり、ノードあたり 16G バイト/秒×2(双方向)の通信バンド幅を実現します。また、RCU は CPU とは独立に動作するデータムーバを持つことにより、異なるノードのメモリ間でデータ転送を行なうリモートメモリアccessを、CPU 動作とは完全に独立して行なうことが可能です。IXS は回線交換型のデータ交換を行なうクロスバススイッチです。各ノードの RCU とのインタフェースを持ち、IXS 全体ではポートあたり 16GB/秒(双方向)のスループット性能を有しています。

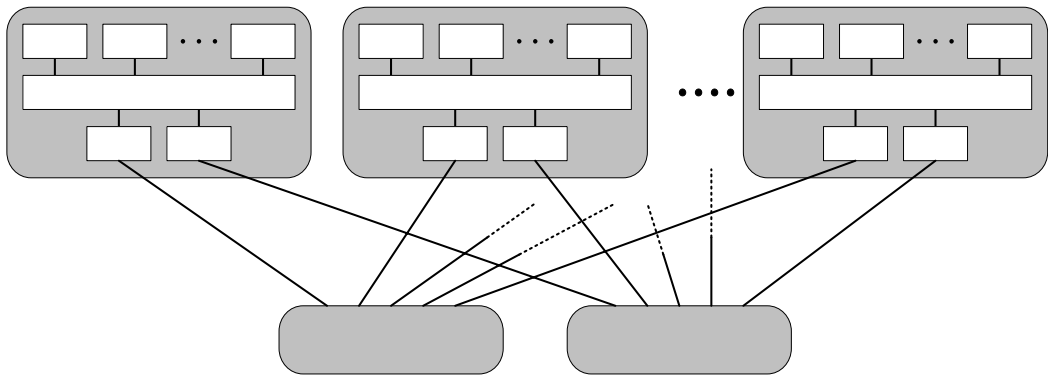


図 3 ノード間接続

### 3. SX-7C のテクノロジー

#### 3.1. LSI 技術

SX-4以降、SXシリーズではCMOSテクノロジーによる高集積化、及びプロセッサの平行化により高性能化を実現しつつ、コストパフォーマンスを向上させてきました。

SX-7C では、さらに高い性能を実現するために LSI 技術、及び回路技術を発展させています。

システムの性能向上のためには、LSI間信号伝送の高速化も必要になります。SX-7Cでは、低レイテンシのマルチチャネル・シリアル・インタフェースを開発することにより、LSI間の高速データ転送を実現しています。また、インタフェース回路を低消費電力化、小面積化することによりLSIへの多チャネルの搭載を実現しています。図4、及び表3にSX-7CのCPUチップ外観、及び諸元をそれぞれ示します。

表3 CPUチップ諸元

テクノロジーロード	90nm
搭載トランジスタ数	8,800 万トランジスタ
電源電圧	1.0V
ピン数(内信号ピン)	8,210(1,923)
配線層構成	銅9層
I/O インタフェース	1.0V シリアル
実装形態	ベアチップ実装

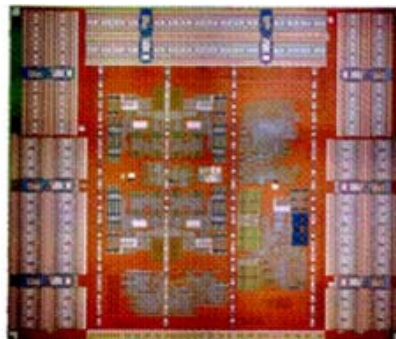


図4 CPUチップ外観

SX-7CはLSI共通仕様として、90nm CMOSプロセス、9層銅配線、及び低誘電率層間絶縁膜などの採用による配線遅延の改善、ラジカル窒素プロセスの開発によるゲート酸化膜の薄膜化による高性能な低電圧電源の実現、新規低誘電率層間膜材料(K=2.9)の採用による配線ピッチ微細化に伴う配線寄生容量増加の低減などを行なっています。さらに、チップあたり世界最高の実効データ伝送レートを実現するマルチチャネル・シリアル・インタフェースの開発により、CPU/MMU間の低転送レイテンシを実現、同時にLSI上の電源分離、アナログ回路の削減、制御信号のデジタル化などによりノイズ耐力、及びエラーレートの格段の向上を実現しています。

LSI内部のRAM回路は、専用設計によりデバイス性能を最大限引き出すことが可能です。一方、LSIの低消費電力化のために、従来のワンショット動作と非選択回路の停止に加え、読み出し回路の制御信号簡略化によるダイナミックパワーの削減、及びリーク電流の少ないトランジスタの利用によるスタティックパワーの削減を実現しています。また、高速なクロック動作を実現するために、LSI外部からクロックを逡倍するAPLL(Analog Phase-Locked Loop)回路を採用しています。

高速システムにおける処理能力の向上には、LSI 内信号伝送の高速化とともに、LSI 間信号伝送の高速化が必要となります。同様に、信号伝送の高速化を妨げる要因となる電源ノイズ対策も重要になります。SX-7C では高速、かつ安定した信号伝送を実現するために、信号伝送時の減衰が小さい低損失材料を使用した基板、伝送信号の波形を改善するプリエンファシス機能を備えた回路、及び波形ひずみの少ないソケットやコネクタなどを採用しています。また、トランジスタが高速化し、電源電流の時間変化が大きくなることにより電源ノイズが増加するため、デカップリング用コンデンサの搭載数最適化などにより電源ノイズの低減を実現しています。

### 3.2. 実装技術

SX-7C は、世界最高性能、高コストパフォーマンス、及び優れた設置性を実現するために、高密度 LSI 実装技術、高密度接続技術、高効率冷却技術、及び高性能電源モジュール技術により、従来のワンチッププロセッサをさらに進化させました。

CPU、及び MMU モジュールの外観を図 5、及び図 6 に、CPU、及び MMU モジュールの諸元を表 4 にそれぞれ示します。超高速動作が要求される CPU/MMU モジュールは、高密度実装により大型で多ピンの LSI を搭載可能としています。CPU モジュールは、CPU LSI をビルドアップ基板表面にベアチップ実装し、裏面には入出力用のパッドを並べた LGA (Land Grid Array) が形成されています。MMU モジュールは、メモリ制御用の MMC LSI とメモリキャリア用コネクタをビルドアップ基板表面に実装しています。一方、SX-7C は環境配慮型製品として、ベアチップ接続部を含む全てのはんだ付け部に Sn-Ag-Cu 系の鉛フリーはんだを採用しています。

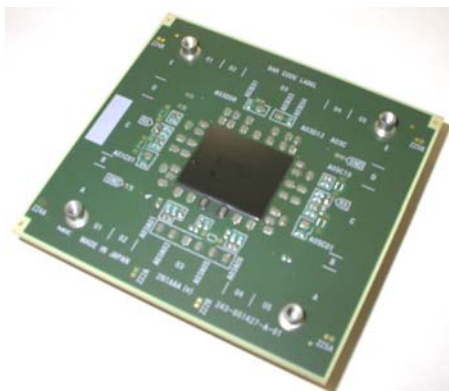


図 5 CPU モジュール外観



図 6 メモリモジュール外観

表 4 CPU/MMU モジュール諸元

	CPU モジュール	MMU モジュール	
基板層数	5ビルド-8コア-5ビルド		
配線密度( $\mu\text{m}$ )	配線幅/間隙=20/20		
via 仕様( $\mu\text{m}$ )	via 径/ランド径=50/80		
モジュール	入力端子数	5,329	824
	冷却	空冷	
	消費電力(W)	75	95

次にシステム実装技術について述べます。SX-7C はノードボードと呼ばれるメインボードを介して、CPU モジュールと MMU モジュールを相互接続するケーブルレス構造を採用しています。これにより、1 ノードの実装密度の改善、ケーブルレスによる製造の効率化、及び部品点数の削減によるコスト低減を実現しています。ノードユニットは図 7 に示すように、2 枚のノードボードを向かい合わせ、その間に MMU モジュールを 32 枚挟み込み、両面にそれぞれ CPU モジュール 4 個、及び RCU モジュール 1 個を MMU モジュールとは逆面実装する構造となっています。これにより、コンパクトなノードユニットを実現しています。

SX-7C はこのようなノードユニット構造を採用することにより実装密度を高めていると同時に、保守を考慮に入れた設置性の高い筐体構造となっています。図 8 にノードユニットの構造を示します。図 8(a) に示すように、ノードユニットは筐体から引き出すことにより表面に実装された CPU モジュールの保守を可能としています。さらに図 8(b) に示すように、ノードユニットを回転させることにより、ノードボード間の MMU モジュールの保守を可能にしています。このため、1 方向からのアクセスのみによりノードユニットの 4 方向へのアクセスを可能にし、保守性、及び設置性を大幅に改善しています。



(a)引き出し状態 (b)回転状態

図 7 ノードモジュール

図 8 ノードユニット実装構造

#### 4. おわりに

本稿では SX-7C のハードウェア概要について説明しました。SX-7C は、スーパーコンピュータの要件である CPU の高い演算性能と、それに見合う主記憶からのデータ供給能力のバランスを重視し、ユーザに使いやすい大規模な分散共有メモリ型スーパーコンピュータとして開発しました。NEC は、今後も様々な研究分野の発展を支える強力なツールであるスーパーコンピュータを開発していきます。