

スーパーコンピュータ「京」の CPU SPARC64 VIIIfx

SPARC64 VIIIfx: CPU for the K computer

● 吉田利雄 ● 本藤幹雄 ● 菅 竜二 ● 杉崎 剛

あらまし

スーパーコンピュータ「京」のプロセッサとして開発されたSPARC64 VIIIfxは、富士通セミコンダクターの45 nm CMOSプロセスを採用し、八つのコアと6 Mバイトの共有2次キャッシュ、メモリコントローラから構成される。動作周波数2 GHzでピーク性能128 GFLOPSを58 Wという低消費電力で実現した。電力あたりの性能は当社前機種のSPARCプロセッサに対し、6倍以上にまで達する。

著者らは、この電力あたりの性能を実現するためSPARC-V9アーキテクチャを拡張し、科学技術計算に最適な命令セットHPC-ACE(High Performance Computing-Arithmetic Computational Extensions)を開発した。また水冷によるリーク電力削減、クロックゲーティングによるダイナミック電力削減により低消費電力化を行った。さらに8万個以上のプロセッサを接続するシステムの安定稼働のため、メインフレームやUNIXサーバの高信頼性技術を備える。

本稿では、SPARC64 VIIIfxの高性能、低消費電力、高信頼性を実現する技術について概説する。

Abstract

SPARC64 VIIIfx, which was developed as a processor for the K computer, uses Fujitsu Semiconductor Ltd.'s 45-nm CMOS process for semiconductors and is composed of eight cores, a 6 MB shared level 2 cache, and memory controllers. Peak performance of 128 GFLOPS at an operating frequency of 2 GHz is achieved with power consumption as low as 58 W. The performance per unit of power is more than six times that of the SPARC processor, our previous model. To achieve this performance per unit of power, we extended the SPARC-V9 architecture to develop high performance computing-arithmetic computational extensions (HPC-ACE), the optimum instruction set for scientific computations. In addition, we successfully reduced the leakage power by water cooling and dynamic power by clock gating to achieve a lower power consumption. Furthermore, high-reliability technology for mainframes and UNIX servers is used to ensure stable operation of a system connecting more than 80 000 processors. This paper outlines the technologies used to achieve the high performance, low power consumption and high reliability of SPARC64 VIIIfx.

まえがき

SPARC64 VIIIfx⁽¹⁾(**図-1**)はスーパーコンピュータ「京」^(注)のプロセッサとして開発された。「京」は10 PFLOPSを超える性能を実現するために8万个以上のプロセッサがシステムに搭載される。そのプロセッサには、高性能、低消費電力、高信頼性がいずれも高いレベルで求められる。

本稿ではこれらを実現するための技術について概説する。

SPARC64 VIIIfxの開発のねらい

SPARC64 VIIIfx開発のねらいについて述べる。

(1) 高性能

SPARC64 VIIIfxは八つのコアと共有2次キャッシュ、メモリコントローラ (MAC), そして高速シリアル伝送のシステムインタフェースを内蔵するマルチコアプロセッサである。

各コアが実アプリケーションで高い実行性能を発揮するために、著者らはSPARC-V9アーキテクチャ⁽²⁾⁻⁽⁴⁾の拡張を行い、科学技術計算を効率良く実行可能な命令セットHPC-ACE (High Performance Computing-Arithmetic Computational Extensions)⁽⁵⁾を開発した。

(注) 理化学研究所が2010年7月に決定したスーパーコンピュータの愛称。

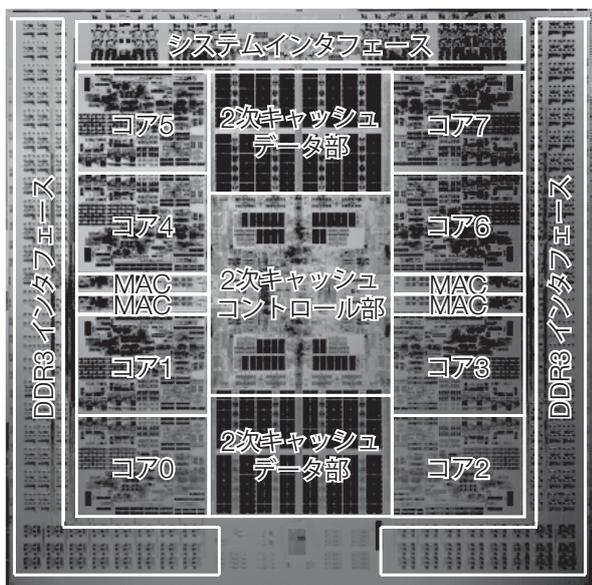


図-1 SPARC64 VIIIfx チップ

またチップ上の八つのコアによる並列処理を高速化するため、全てのコアで2次キャッシュを共有し、さらにコア間の同期処理をハードウェアで行う機能を備える。これに富士通の自動並列コンパイラを組み合わせることで、ユーザはプログラミングの際に複数コアであることを特に意識せず、複数コアをあたかも高速な一つのCPUとして扱うことが可能となる。これを富士通はVISIMPACT (Virtual Single Processor by Integrated Multi-core Parallel Architecture) と呼んでいる。

(2) 低消費電力

システム全体の電力制限からプロセッサの消費電力を58 W以下にする必要があった。そのため低リークのトランジスタの使用や、水冷による冷却方式によりジャンクション温度を30℃まで低下させてリーク電力を抑えている。さらに、クロックゲーティングなどを徹底的に行い、ダイナミック電力を低減する。

(3) 高信頼性

システムの安定稼働を実現するために、メインフレーム、UNIXサーバ^{(6), (7)}で用いる高信頼性技術を備える。

SPARC64 VIIIfxのマイクロアーキテクチャ

SPARC64 VIIIfxのパイプラインを**図-2**に、諸元を表-1に示す。

コアは、命令制御部、演算処理部、1次キャッシュ部からなる。命令制御部は命令フェッチ、命令デコード、命令のアウトオブオーダー処理制御、そして命令完了の制御を行う。

演算部は、二つの固定小数点演算器 (EXA/B), 二つのロード、ストアのアドレス計算を行う演算器 (EAGA/B), および浮動小数点積和演算器 (FMA: Floating-point Multiply-and-Add) を四つ (FLA/B/C/D) 備える。FMA演算器はSIMD (Single Instruction Multiple Data) 構成を採り、一つの命令で二つの演算を並列に行う。一つのFMA演算器は毎サイクル浮動小数点の乗算と加算を実行することが可能であり、各コアで毎サイクル8個、チップでは64個の倍精度浮動小数点演算が実行可能である。動作周波数は2 GHzであり、ピーク性能は128 GFLOPSとなる。レジスタは固定小数点系で192本、浮動小数点系では256本である。

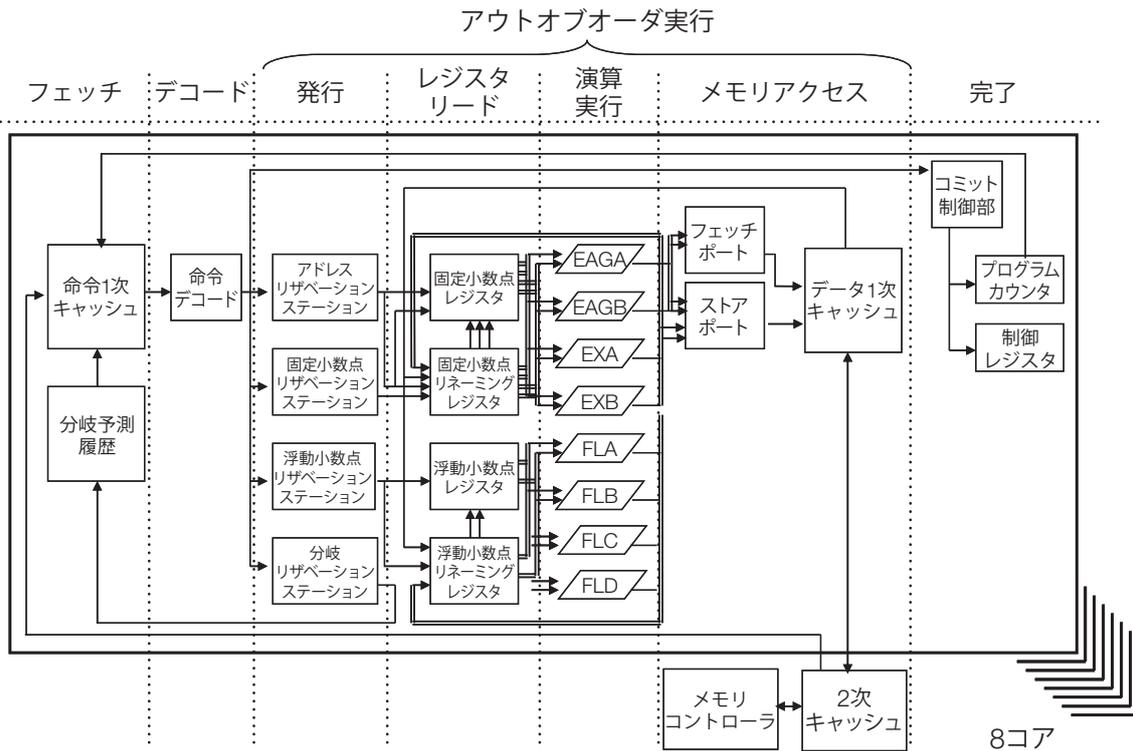


図-2 SPARC64 VIIIfx パイプライン

表-1 SPARC64 VIIIfx 諸元

項目	諸元
コア数	8
2次キャッシュ	6 Mバイト
動作周波数	2 GHz
プロセステクノロジー	FSL 45 nm CMOS
ダイサイズ	22.7 mm×22.6 mm
トランジスタ数	約7億6000万個
ピーク性能	128 GFLOPS
メモリ帯域	64 Gバイト/秒 (理論ピーク値)
消費電力	58 W (プロセス条件TYP)

FSL：富士通セミコンダクター

1次キャッシュ部は、ロード、ストア命令を処理する。コアごとに32 Kバイト2ウェイの命令キャッシュとデータキャッシュをそれぞれ有する。データキャッシュは、二つ同時にロードアクセスが可能なデュアルポート構成であり、16バイトのSIMDロードを二つ、または16バイトのSIMDストアを一つ実行する。

2次キャッシュ部は八つのコアで共有され、各コアを含めたキャッシュコヒーレンスを保証する。またコア間的高速同期処理のため、コア間のハー

ドウェアバリア機構を有する。これについては後述する。

メモリアクセスの低レイテンシ化、高スループット化のためメモリコントローラを内蔵した。メモリ帯域は理論ピーク値64 Gバイト/秒である。

また「京」専用のインターコネクトチップと高速シリアルI/Oで結合し、チップ間通信のスループットを確保している。

命令拡張 HPC-ACE

HPC-ACEはSPARC-V9アーキテクチャに対する科学技術計算向けの拡張命令セットである。この開発に当たっては富士通ソフトウェア開発部門とともに多数のHPCアプリケーションの解析に基づいて行った。

(1) レジスタ数の拡張

SPARC-V9における浮動小数点レジスタの数は32本であり、HPCアプリケーションには十分な本数ではない。しかしレジスタ数を増やすにも32ビット長のSPARCアーキテクチャでは命令長が不足し不可能であった。この課題解決のためHPC-ACEではSXAR (Set eXtended Arithmetic Register) と

いう前置命令を新設した。SXAR命令は直後の最大2命令に対して、レジスタのアドレッシングの拡張などを行う。レジスタアドレスを3ビット拡張して浮動小数点レジスタ本数をSPARC-V9の8倍、256本まで指定可能にした（図-3）。

コンパイラはこの大容量レジスタを用いてソフトウェアパイプラインなどの最適化を行い、アプリケーションが持つ命令レベルの並列性を最大限に引き出す。HPCの代表的なベンチマークの一つである姫野ベンチマークでは1.65倍の性能向上となっている。

(2) SIMD演算，ロード，ストア

SIMDは一つの命令で複数のデータ処理を並列実行する技術である。HPC-ACEはSIMD技術を採用し、一つの命令で二つのFMA演算を実行する。さらに複素数の乗算を高速化するためのSIMD演算もサポートしている。またロード命令とストア命令もSIMD実行が可能となっている。ロード命令は倍精度のとき8バイトアライン、単精度のとき4バイトアラインでペナルティなくSIMD処理を行う。

(3) セクタキャッシュ機構

HPC-ACEではソフトウェア制御可能なキャッシュ機構（セクタキャッシュ）を開発した。従来のキャッシュはソフトウェアから制御を行うことができない。たとえユーザには再利用頻度の高いデータと分かっている、ハードウェアがほかのデータをキャッシュに登録する際にキャッシュから掃き出してしまう性能向上の妨げになる場合

があった。これに対しセクタキャッシュ機構は、キャッシュを二つの領域（セクタ）に分け、再利用頻度の高いデータをほかとは別のセクタに登録することがソフトウェアから可能になる。ユーザが再利用頻度の高いデータをキャッシュに保持し続けるよう制御することで性能向上に寄与する。

(4) 三角関数sin，cosの高速化命令

三角関数のsin，cosの高速化命令を追加した。従来は多数の命令を組み合わせて処理を行っているが、専用命令化により命令数を削減したことで5倍以上高速化する。

(5) 条件付き実行

if文を含むループを効率良く処理するためには条件分岐命令をなくすことが必要である。そのためHPC-ACEでは条件付き実行命令を追加した。具体的には新規の比較命令で比較結果を浮動小数点レジスタに書き込み、その比較結果に基づいて条件付き実行命令を処理する。条件付き実行命令には、浮動小数点レジスタ間のデータ転送と浮動小数点レジスタからメモリへのストアを用意した。これらの命令を組み合わせて条件分岐命令を取り除くことで、コンパイラはif文を含むループに対してソフトウェアパイプラインなどによる最適化が可能になる。

(6) 除算，平方根近似

逆数近似値を求める命令を追加した。これにより除算，平方根のパイプライン処理を可能にしており、レジスタ数の拡張と合わせた効果でおおよそ4倍のスループット向上となっている。

これら (1) から (6) の機能はいずれも周波数を上げることなく性能向上を可能とし、SPARC64 VIIIfxの電力あたりの性能の向上に大きく寄与している。

VISIMPACT

VISIMPACTのためのハードウェア機構について説明する。

(1) 共有2次キャッシュ

SPARC64 VIIIfxは6 Mバイトの2次キャッシュを備え、全8コアで共有する。コア間でのデータ共有を容易にすることで一つのプロセスを複数のコアで効率良く並列処理することが可能となる。

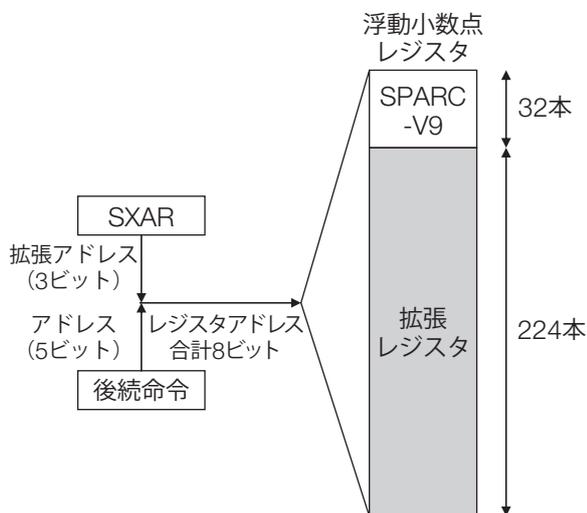


図-3 SXAR命令によるレジスタアドレス拡張

(2) ハードウェアバリア機構

SPARC64 VIIIfxはコア間の高速同期処理のためハードウェアバリア機構を備える。一つのプロセッサを複数のコアで並列実行させたとき、コア間で待合せ（同期処理）を行う場合がある。通常のプロセッサでは同期処理をソフトウェアで行うが、SPARC64 VIIIfxは専用のハードウェアにより10倍以上高速化している。同期処理のオーバヘッドが大きく削減されたことで、小さいループも複数コアによる並列処理での高速化が可能となる。

低消費電力

SPARC64 VIIIfxはゲート長の長いトランジスタを使用し、また冷却方式に水冷を採用してジャンクション温度を30℃まで下げることでリーク電力をチップ全体の電力の10%まで抑えている。

また各ラッチに徹底したクロックゲーティングを行い、電力削減に効果的な回路の組み方や制御方式に変更して動作時に消費するダイナミック電力を削減した。

その結果、128 GFLOPSという高性能ながら、チップばらつきの平均で58 Wという低消費電力を実現した。これは電力あたりの性能で当社前機種種のSPARCプロセッサに対し6倍以上にまで達する。

高信頼性

SPARC64 VIIIfxは、メインフレームやUNIXサーバの開発で培った高信頼性技術を備える。

プロセッサは非常に微細なトランジスタで構成されており、宇宙線の衝突などで信号が変化する可能性がある。このような間欠エラーの場合も誤動作することなく処理を続けるために、エラーが発生した命令をハードウェアで自動的に再実行する命令リトライ機構を備えている。またプロセッサ内の全てのRAMおよび固定小数点、浮動小数点レジスタの1ビットエラーはハードウェアで訂正処理を行う。プログラム実行に関連する部分についてはエラー検出コードで保護し、データ保全性を

確保している。

これらの技術によりプロセッサを8万個以上接続したシステムの安定稼働を実現する。

む す び

SPARC64 VIIIfxの開発は、著者らにとってまさに挑戦のプロジェクトであった。この開発に当たってはプロセッサ開発部門に加えて、ソフトウェア開発部門、研究所などのメンバが結集し、富士通の総力を挙げての開発であった。長年にわたって培ってきたプロセッサ技術を継承しつつ、新たな技術開発を行うことでスーパーコンピュータとして発展させることができたと考えている。

このプロセッサを搭載したスーパーコンピュータ「京」が、今後様々な分野の課題解決に貢献することを期待する。

参考文献

- (1) T. Maruyama et al. : SPARC64 VIIIfx: A New-Generation Octocore Processor For PETASCALE Computing. *IEEE Micro*, Vol.30 Issue2, p.30-40 (2010).
- (2) SPARC International : The SPARC Architecture Manual (Version 9).
<http://www.sparc.org/standards/SPARCV9.pdf>
- (3) 富士通 : SPARC Joint Programming Specification (JPS1) Commonality.
<http://jp.fujitsu.com/solutions/hpc/brochures/>
- (4) 富士通 : SPARC JPS1 Implementation Specification SPARC64 V.
<http://jp.fujitsu.com/solutions/hpc/brochures/>
- (5) 富士通 : SPARC64 VIIIfx Extensions.
<http://jp.fujitsu.com/solutions/hpc/brochures/>
- (6) 井上愛一郎 : UNIXサーバ用プロセッサ : SPARC64 V. *FUJITSU*, Vol.53, No.6, p.450-455 (2002).
- (7) 丸山拓巳ほか : SPARC64プロセッサの過去・現在・未来. *FUJITSU*, Vol.61, No.6, p.526-530 (2010).

著者紹介



吉田利雄 (よしだ としお)
次世代テクニカルコンピューティング
開発本部LSI開発統括部 所属
現在、プロセッサコアの開発に従事。



菅 竜二 (かん りゅうじ)
エンタプライズサーバ事業本部プロ
セッサ開発統括部 所属
現在、プロセッサコアの開発に従事。



本藤幹雄 (ほんどう みきお)
次世代テクニカルコンピューティング
開発本部PAプロジェクト 所属
現在、プロセッサおよびシステムの性
能評価に従事。



杉崎 剛 (すぎざき ごう)
エンタプライズサーバ事業本部プロ
セッサ開発統括部 所属
現在、プロセッサの開発に従事。