

UNIXサーバ用プロセッサ : SPARC64 V

SPARC64 V Processor for UNIX Servers

あらまし

SPARC64 Vは、メインフレームの高性能と高信頼の技術に基づいて、PRIMEPOWERのプロセッサとして開発された。SPARC V9に準拠し、UNIXサーバにおけるデファクトスタンダードOSであるSolaris Operating Environmentが走行する。実環境での性能に着目した設計となっており、高スループットの2階層キャッシュを搭載し、4命令同時発行、6演算同時投入のスーパースカラ方式で、演算実行からメモリリクエストまで拡張されたフルレンジのアウトオブオーダー実行機能を備える。そして、消費電力は小さく、システムに搭載される実環境で最高の性能を発揮する。また、RAMの1ビットエラーの全救済やデータパスの保護などメインフレームに匹敵する最先端のRAS機能を装備し、ミッションクリティカルな用途に適応する。

本稿では、SPARC64 VマイクロアーキテクチャおよびRAS機能について概説する。

Abstract

SPARC64 V is a newly developed processor for Fujitsu's PRIMEPOWER series of UNIX servers. It leverages high-performance and high-reliability technology that has been developed for mainframe systems. It conforms to the SPARC V9 architecture and supports Solaris Operating Environment, the de facto standard UNIX server OS. Tuned for high performance in the real-world enterprise computing environment, SPARC64 V implements a high throughput two-level cache; a 4-issue, 6-dispatch super-scalar pipeline; and an out of order execution of most types of instructions, including load/store requests. SPARC64 V achieves both a low power consumption and a high performance. It also satisfies mission critical application requirements with the latest mainframe level RAS features such as recovery of any 1-bit error in RAM and parity protection of most non-RAM data paths. This paper describes the micro-architecture overview of SPARC64 V.



井上 愛一郎 (いのうえ あいいちろう)
プロセッサ開発統括部第二開発部 所属
現在、SPARC64 VおよびGS21のCPUの
開発に従事。

まえがき

エンタープライズサーバシステムは、社会経済機構の基盤として重責を担っている。この心臓部であるプロセッサは、高性能であるとともに、高信頼性を達成していることが求められる。そこで、SPARC64 Vは、この要求に対応して、これまでの第4世代のSPARCプロセッサ⁽¹⁾の限界を超える性能と信頼性を実現するために、長年にわたるメインフレームの開発で蓄積した技術⁽²⁾を導入して、第5世代のプロセッサとして開発した。

本稿では、高性能を発揮できるSPARC64 Vマイクロアーキテクチャ、および高信頼性を実現したRAS機能について概説する。

開発のねらい

SPARC64 Vは、富士通のUNIXサーバであるPRIMEPOWERに搭載することを前提として開発した。したがって、PRIMEPOWERのエンタープライズサーバシステムとしての魅力を最大にすることが開発のねらいであり、以下の方針で開発した。

(1) SPARC V9完全互換設計

SPARC V9⁽³⁾に準拠し、アプリケーションの互換性を維持する。オペレーティングシステムのレベルに至る更に高度な互換性を追求するため、サン・マイクロシステ

ムズ社と富士通で共同開発された共通仕様JPS (Joint Programming Specification)^{(4),(5)}に適合する。

(2) システム搭載時の性能に着目した高性能設計

消費電力、チップサイズ、システム接続バス方式などのシステム実装要件を満足しつつ、システムのポテンシャルを最大限に引き出す。メモリアクセスの負荷と演算パワーのバランスを取り、強力なキャッシュをチップ上に搭載し、高いマシンサイクルで最高水準の単体性能を実現するとともに、マルチプロセッサのスケラビリティを確保する。

(3) 最先端のRAS機能の実現

ミッションクリティカルなアプリケーションに対応するため、データインテグリティ保証、エラー訂正・回復・記録などあらゆる面でRAS機能を充実し、メインフレームに匹敵するRAS機能を装備する。

SPARC64 Vのマイクロアーキテクチャ

SPARC64 Vは、アウトオブオーダー実行機構を備えたスーパースカラプロセッサで、命令フェッチ部、命令実行部、2次キャッシュ (L2\$) 部、システムインタフェース部で構成される。SPARC64 Vの構成を図-1に示す。

以下で、はじめに命令フェッチ部を、それから、命令実行部を説明し、最後にキャッシュをまとめて説明する

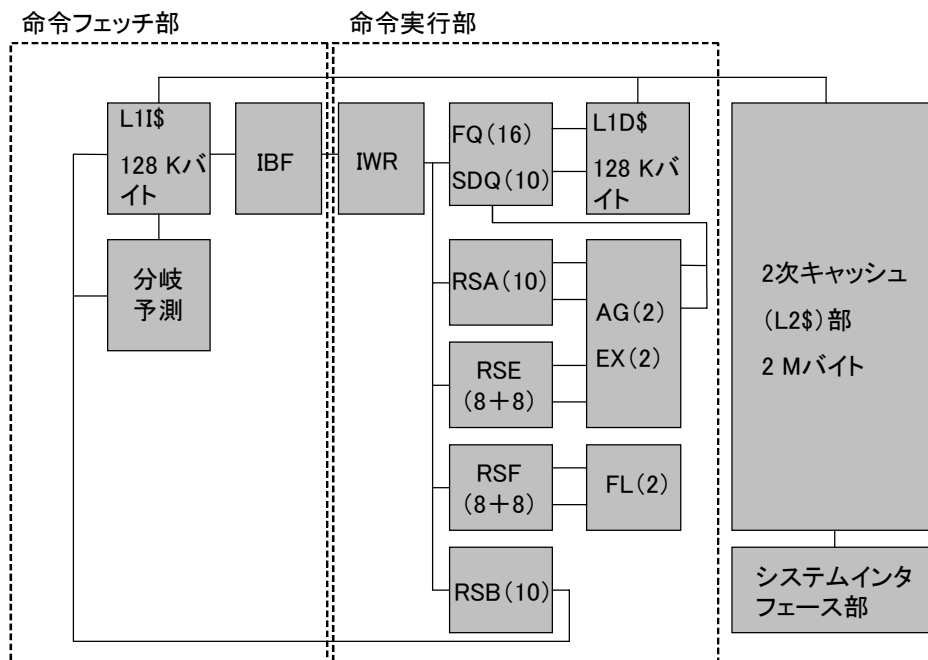


図-1 SPARC64 Vの構成
Fig.1-Configuration of SPARC64 V.

とともに、システムインタフェース、そのほかについて説明する。

命令フェッチ部

命令フェッチ部は、分岐予測に従って将来実行すると予測される命令列をなぞるように命令バッファIBF (Instruction BuFfer) に取り込む。命令フェッチは一度に32バイトすなわち8命令を取り出し、命令実行の2倍のスループットを確保して、IBFのデカップリング効果により、パイプライン損失を抑えて、大容量キャッシュの利益を享受する。

分岐予測のために、大容量 (16 Kエントリ、4ウェイセットアソシアティブ) のブランチヒストリを装備する。さらに一種のグローバル履歴テーブルであるWGHT (Write-cycle-driven Global History Table) とサブルーチンの戻りアドレスを記憶するリターンアドレススタックを装備し、ブランチヒストリの本体と連携して精度の高い予測動作を行う。

命令実行部

命令実行部は、デコード発行・投入・読み出し・演算・更新・コミットの六つのステージで構成されている。SPARC64 Vのパイプラインの概要を図-2に示す。SPARC64 Vは、それぞれ独立して同時に実行ができる六つの演算器を備え、演算のオペランドが用意できたものからプライオリティを取って演算器に投入する、いわゆるデータドリブンのアウトオブオーダー実行制御方式である。

(1) デコード発行ステージ

IBFに蓄えた命令は4命令分をまとめてIWR (Instruction Word Register) にセットする。デコード発行ステージでは、IWR上の命令をデコードして、実

行に必要な資源を決定する。そして、空きがあれば、その資源とともに、命令識別子 (IID: Instruction IDentification) を割り当てて命令を発行する。同時に発行できる命令には資源の割り当てや命令種類の組合せの制限がなく、また4命令分の空きがない場合にも可能な分だけを発行できる。

(2) 投入ステージ

SPARC64 Vは、整数演算用リザベーションステーションRSE (Reservation Station for Execution) と浮動小数点演算用リザベーションステーションRSF (Reservation Station for Floating point) を装備する。RSEとRSFは、いずれも二つのキューに分かれており、演算のオペランドが用意できたものの中で1番古い (Oldest Ready) ものを、それぞれ選択して演算器に投入する。さらに、アドレス生成用リザベーションステーションRSA (Reservation Station for Address generation) を装備し、こちらは単一のキューで、投入準備の完了したリクエストの中から二つを選択してアドレス生成演算器AG (Address Generation Unit) に投入する。これらを合わせて、四つの演算と二つのアドレス生成を同時に投入できる。整数系・浮動小数点系の演算を最大16個ずつ、また、アドレス生成は10個をキューイングできる。

(3) 読み出しステージ

すべての演算器が並列動作できるように、整数レジスタ (GPR) は8リードポート、浮動小数点レジスタ (FPR) は6リードポートを備える。GPRは8ウィンドウレジスタを実装し、GPRの一部のコピーをJWR (Joint Work Register) に用意して、ハードウェアがバックグラウンドで切り替える。

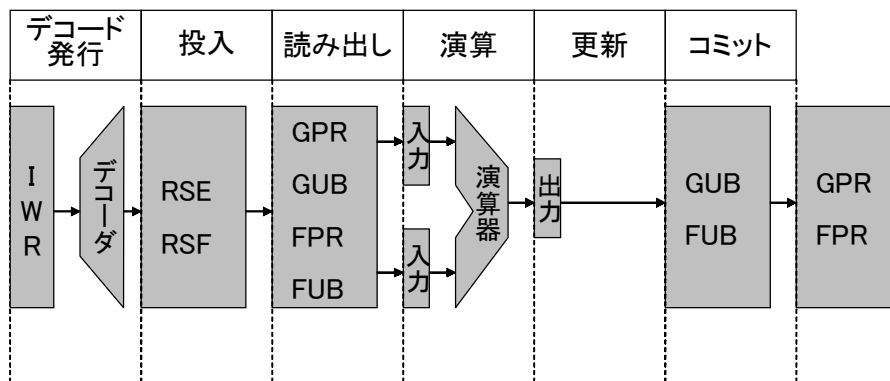


図-2 SPARC64 Vのパイプライン
Fig.2-Pipeline structure of SPARC64 V.

(4) 演算ステージ

SPARC64 Vは、SPARC V9命令に加え、VIS 2.0 (Visual Instruction Set Version 2.0)、FMADD (Multiply & Add) をサポートする。整数演算器 (EX) 浮動小数点演算器 (FL) は、それぞれ2組装備し、頻度の高い演算はどちらでも実行できる。FMADDもその例外ではない。アドレス生成演算器AG (Address Generation Unit) も2組装備する。

依存関係のある命令間ではデータフォワーディングを行い、連続して実行する。

(5) 更新ステージ

演算結果はレジスタ更新バッファ、GUB (GPR Update Buffer) またはFUB (FPR Update Buffer) に格納する。

(6) コミットステージ

IIDに対応して64個のCSE (Commit Stack Entry) がある。実行結果をプログラマブルな資源に反映する制御はCSEが行う。アウトオブオーダーで実行して、GUB、FUBに格納した結果は、コミットステージで命令実行順序に従って、GPRやFPRやメモリに格納する。PC (Program Counter) をはじめとする制御レジスタも、コミットステージで一括して更新する。これを同期一括更新方式と言い、正確な割り込みを保証し、いつでも実行中の処理をキャンセルできる。同時にコミットする命令数は最大4である。

分岐命令、ロード命令、ストア命令

(1) 分岐命令実行

分岐命令は、10エントリの分岐用リザーベーションステーションRSB (Reservation Station for Branch) で分岐確定、分岐予測情報の更新、分岐予測失敗時の再命令フェッチなど一連の処理を行う。分岐命令はほかの命令と並列に実行される。

(2) ロード命令・ストア命令実行

ロード命令やストア命令は、16エントリのフェッチキューFQ (Fetch Queue) でキャッシュを制御して実行する。ストア命令の場合には、10エントリのストアデータキューSDQ (Store Data Queue) で、ストアデータの書き込みを制御する。ロード命令、ストア命令とも、同時に二つのアクセス要求を受け付け、キャッシュミスなどで滞った場合は追い越し処理が可能である。

キャッシュ

SPARC64 Vのキャッシュは、中容量の1次キャッシュ (L1\$) と大容量の2次キャッシュ (L2\$) の2階層

で構成され、いずれも、ストアインもしくはスワップと呼ばれるストア制御方式である。SPARC64 Vのキャッシュを図-3に示す。L1\$は、命令専用のL1ISとオペランド (データ) 専用のL1DSで構成され、いずれも容量128 Kバイト、2ウェイセットアソシアティブのキャッシュ構成を持ち、ブロックサイズは64バイトである。L1DSは、4バイトアドレス境界によって、八つのバンクに分割されており、バンクごとに別のアドレスでアクセスできる。L2\$は、容量2 Mバイト、2または4ウェイセットアソシアティブのキャッシュ構成を持ち、ブロックサイズは64バイトである。L2\$のデータ部は省電力化している。L2\$とL1\$の間は32ないしは16バイト幅のバスでつながれ、プロセッサコアの命令実行の最大スループットに匹敵するスループットを確保している。

システムインタフェース

SPARC64 Vは、インタコネクタアーキテクチャUPA (Unified Port Architecture) を物理的にも制御プロトコルの面でも拡張したシステムインタフェースを備える。

バスクロックは周波数250 MHzを超え、DDR (Double Data Rate) 方式により、ピークで8 Gバイト/秒以上の最大データ転送速度を達成する。また、発行した要求が完了する前に次の要求を発行できる数、すなわちアウトスタンディング数を大幅に強化し、識別子によって順序制約を緩和して、システム動作の並列度を向上させる。

プリフェッチ機構

(1) 連鎖プリフェッチ (Chained pre-fetch) 機構

連鎖プリフェッチ機構は、連続したアドレスに対する

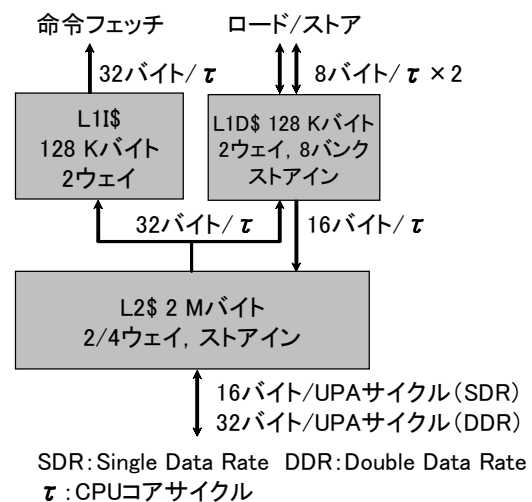


図-3 SPARC64 Vのキャッシュ
Fig.3-Cache structure of SPARC64 V.

リード要求があった場合には、さらに先の連続アドレスに対するアクセスが将来行われる可能性が高いという予測に基づいて、ハードウェアが投機的にリード要求を発行するものである。命令フェッチ用とオペランドアクセス用の連鎖プリフェッチ機構は、それぞれ16系列のストリームに対応可能である。

(2) ソフトウェアプリフェッチ機構

コンパイラによる高速化手法の一つとして、プリフェッチが一般的に行われる。SPARC64 Vでは、キャッシュのどの階層にロードするかなど、きめ細かい制御オプションをサポートすることで、コンパイラによる効果的なスケジューリングを支援する。

SPARC64 VのRAS機能

SPARC64 Vは、確実にエラーを検出すること、エラーの影響範囲を限定すること、回復処理を試みること、エラーの記録を残すこと、ソフトウェアに通知することなどRAS機能の基本的な事項を徹底した。これによりSPARC64 Vは、高度な信頼性、可用性、サービス性、データ安全性を提供する。SPARC64 Vのエラー保護の概要を表-1に示す。

内蔵RAMのRAS機能

SPARC64 Vは、RAMの1ビットエラーは、ソフトウェアの介入なしに、すべてハードウェアで自動的に訂正し、ソフトウェアへの影響は皆無である。

(1) ECCによるエラーの検出と訂正動作

キャッシュのデータ部などは、ECC (Error Checking and Correction) を備えている。エラーが発生した場合は、ECC回路で訂正したデータを以降の処

理に使用し、RAMの固定故障であっても救済する。また、RAMを上書きして間欠エラーを消去する。2次キャッシュL2Sのタグ部は、内容を全ビット反転して再書き込みするリバースライト機能によって固定故障も救済する。

(2) パリティによるエラーの検出と訂正動作

パリティによるエラー保護の場合にも、RAMの1ビットエラーはすべて救済する。例えば、1次キャッシュL1Sのタグ部は二重化されており、いずれかのエラーが発生した場合には、互いに他方に通知して、エラーしていない方のタグ情報に基づいて回復する。また、ブランチヒストリは、RSBで分岐命令の処理を行う過程で予測内容を検査する。エラーがあれば必然的に予測失敗となり、予測失敗の回復動作の過程で、命令実行に対するエラーの影響は取り除かれる。したがって、エラーのビット数にも、また間欠エラーか固定故障のエラー発生状況にもよらず救済する。

(3) 縮退

キャッシュはウェイ単位での縮退を行う。エラーの発生回数をカウントし、単位時間あたりのエラー回数が上限値を超えると縮退する。縮退に際して、L1Sの縮退されるウェイのダーティライン (更新されたライン) をすべてL2Sに書き戻し、L2Sの縮退されるダーティラインをメモリに書き戻すことによってハードウェアでつじつまを合わせるために、ソフトウェアに影響を与えない。

メモリのRAS機能

PRIMEPOWERのメモリは、ECCで保護されている。SPARC64 Vがエラーを検出するとデータを訂正して使用し、修正不能な多ビットエラーは、ソフトウェアに割込みで通知する。

読み込んだメモリに多ビットエラーがある場合には、そのデータに自己識別子をデータに埋め込むエラーマーキングを行う。これにより、エラーの発生元を特定するために有用な情報を提供する。

内蔵レジスタと演算器のRAS機能

(1) 内蔵レジスタのエラー保護

GPRやFPRにパリティを備える。JWRにもパリティを備え、ウィンドウレジスタ本体から再ロードして回復する。また、PCやPSTATE (Processor STATE) レジスタ、割込み表示レジスタなどの制御レジスタにもパリティを備え被害拡大を防止する。

(2) 演算器のエラー保護

演算器は結果のパリティ予測と、乗除算での剰余

表-1 SPARC64 Vのエラー保護

演算器種別 レジスタ種別	エラー保護機構
L2Sデータ	ECC
L2Sタグ	ECC
L1ISデータ	パリティ保護、再書き込み
L1ISタグ	パリティ保護&二重化
L1DSデータ	ECC
L1DSタグ	パリティ保護&二重化
TLB	パリティ保護
ブランチヒストリ	分岐予測内容を検査して回復
レジスタ	パリティ保護
演算器	パリティもしくは剰余チェック

: Translation Lookaside Buffer

チェックの組合せで保護し、GPR、FPRに書き込むまでの経路も含めてパリティ伝播させてエラーを監視する。このほかにも、レジスタアドレス、メモリアドレス、キャッシュインデクス、命令コード、システムインタフェースのアドレスパスやデータパスなども監視の対象としている。ハードウェアのハングアップも、専用ハードウェアで検出する。

同期一括更新方式と命令リトライ

SPARC64 Vは同期一括更新方式を採用しているため、エラーが検出されると、コミットまでの間の途中結果は廃棄して処理をキャンセルし、正常に完了した命令の結果だけがプログラマブルな資源に残る。

このことは同時に命令リトライを行う機会を与える。命令リトライはエラーを契機に起動して1命令ずつ実行する。正常に1命令を完了すると通常の状態に復帰する。ソフトウェアの介入は不要であり、命令リトライが成功すれば、エラーによる影響は皆無である。命令リトライを繰り返しても成功しない場合には、割り込みによってソフトウェアにエラーを通知する。

そのほかのRAS機能

ソフトウェアに割り込む際には、エラーの波及範囲を限定し、これに応じた処理ができる情報を提供する。ダンプ処理を支援するモードも装備する。

また、エラーがどこで見つかったかを表示するエラーリジョンは679ビット、またエラーに関する補助情報は821ビットを備え、さらに、エラー発生時点の時系列の記録を80ビット幅×1,024インデックスのRAM 5個から成るイベント履歴に保存する。これらの情報は、専用のインタフェースでソフトウェアの介入なしに収集され、故障か所と故障種別を特定するための、予防保守に有用な情報を提供する。

SPARC64 Vのシリコンテクノロジー

SPARC64 Vは、130 nmのCMOSプロセスで作られている。トランジスタ数19,100万、信号269ピン、チップサイズ17.8 mm×15.7 mmである。動作周波数は1.3 GHz以上を達成し、消費電力は約50 Wである。

SPARC64 Vチップ写真を図-4に示す。

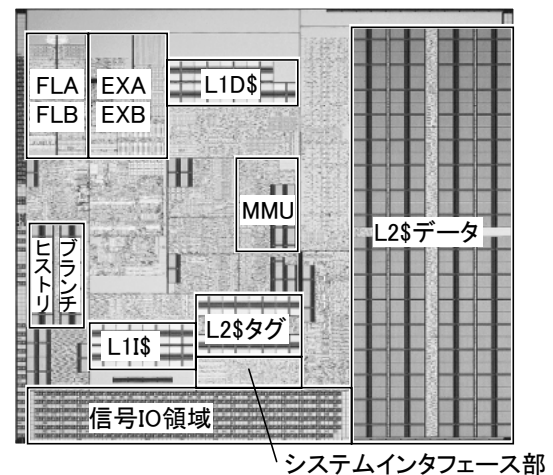


図-4 SPARC64 Vチップ
Fig.4-SPARC64 V chip.

む す び

SPARC64 Vは、高性能技術をすべて投入し、また、これまでのUNIXサーバでは達成し得なかった高度なRAS機能を実現した。消費電力も既存のSPARCプロセッサに比べて小さく、高性能でコンパクトなシステム構築を可能にする。そして、UNIXサーバにおけるデファクトスタンダードOSであるSolaris Operating Environmentが走行し、豊富な品揃えを誇るSolaris Operating Environment上のサーバ用ミドルウェアやアプリケーション群が使用できる。このプロセッサは、ユーザの期待に最も良く応えるものであると確信している。

参考文献

- (1) 引地徹ほか：64ビットRISCプロセッサ：SPARC64 GP . *FUJITSU* , Vol.51, No.4, p.226-231 (2000) .
- (2) 利根廣貞ほか：GS8600のハードウェア . *FUJITSU* , Vol.47, No.2, p.96-108 (1996) .
- (3) SPARC International Inc . : The SPARC Architecture Manual-Version 9 , 1994 .
- (4) Sun Microsystems. Inc. and Fujitsu Limited : SPARC Joint Programming Specification (JPS1) : Commonality .
- (5) Fujitsu Limited : SPARC JPS1: Fujitsu SPARC64 V Implementation Supplement .