SPARC Enterprise のアーキテクチャー M シリーズ編

メインフレームクラスの信頼性・柔軟性の提供

ホワイトペーパー

マニュアル番号 : C120-E636-03
2011 年 9 月

Copyright © 2007, 2011 富士通株式会社 All rights reserved.

本書には、オラクル社および / またはその関連会社により提供および修正された技術情報が含まれています。

オラクル社および / またはその関連会社、および富士通株式会社は、それぞれ本書に記述されている製品および技術に関する知的所有権を所有または管理しています。これらの製品、技術、および本書は、著作権法、特許権などの知的所有権に関する法律および国際条約により保護されています。

本書およびそれに付属する製品および技術は、その使用、複製、頒布および逆コンパイルを制限するライセンスのもとにおいて頒布されます。オラクル社および / またはその関連会社、および富士通株式会社およびそのライセンサーの書面による事前の許可なく、このような製品または技術および本書のいかなる部分も、いかなる方法によっても複製することが禁じられます。本書の提供は、明示的であるか黙示的であるかを問わず、本製品またはそれに付随する技術に関するいかなる権利またはライセンスを付与するものでもありません。本書は、オラクル社および富士通株式会社の一部、あるいはそのいずれかの関連会社のいかなる種類の義務を含むものでも示すものでもありません。

本書および本書に記述されている製品および技術には、ソフトウェアおよびフォント技術を含む第三者の知的財産が含まれている場合があります。これらの知的財産は、著作権法により保護されているか、または提供者からオラクル社および / またはその関連会社、および富士通株式会社へライセンスが付与されているか、あるいはその両方です。

GPL または LGPL が適用されたソースコードの複製は、GPL または LGPL の規約に従い、該当する場合に、お客様からのお申し込みに応じて入手可能です。オラクル社および / またはその関連会社、および富士通株式会社にお問い合わせください。

この配布には、第三者が開発した構成要素が含まれている可能性があります。

本製品の一部は、カリフォルニア大学からライセンスされている Berkeley BSD システムに由来しています。UNIX は、米国およびその他の国におけるオープン・グループの登録商標です。

Oracle と Java は Oracle Corporation およびその関連企業の登録商標です。

富士通および富士通のロゴマークは、富士通株式会社の登録商標です。

すべての SPARC 商標は、SPARC International, Inc. のライセンスを受けて使用している同社の米国およびその他の国における登録商標です。SPARC 商標が付いた製品は、オラクル社および / またはその関連会社が開発したアーキテクチャーに基づくものです。

SPARC64 は、Fujitsu Microelectronics, Inc. および富士通株式会社が SPARC International, Inc. のライセンスを受けて使用している同社の商標です。

その他の名称は、それぞれの所有者の商標または登録商標です。

United States Government Rights - Commercial use. U.S. Government users are subject to the standard government user license agreements of Oracle and/or its affiliates and Fujitsu Limited and the applicable provisions of the FAR and its supplements.

免責条項 : 本書または本書に記述されている製品や技術に関してオラクル社、富士通株式会社および / またはそのいずれかの関連会社が行う保証は、製品または技術の提供に適用されるライセンス契約で明示的に規定されている保証に限ります。このような契約で明示的に規定された保証を除き、オラクル社、富士通株式会社および / またはそのいずれかの関連会社は、製品、技術、または本書に関して、明示、黙示を問わず、いかなる種類の保証も行いません。これらの製品、技術、または本書は、現状のまま提供され、商品性、特定目的への適合性または第三者の権利の非侵害の黙示の保証を含みそれに限定されない、明示的であるか黙示的であるかを問わない、なんらの保証も、かかる免責が法的に無効とされた場合を除き、行われないものとします。このような契約で明示的に規定されていないかぎり、オラクル社、富士通株式会社および / またはそのいずれかの関連会社は、いかなる法理論のもとの第三者に対しても、その収益の損失、有用性またはデータに関する損失、あるいは業務の中断について、あるいは間接的損害、特別損害、付随的損害、または結果的損害について、そのような損害の可能性が示唆されていた場合であっても、適用される法律が許容する範囲内で、いかなる責任も負いません。

本書は、「現状のまま」提供され、商品性、特定目的への適合性または第三者の権利の非侵害の黙示の保証を含みそれに限定されない、明示的であるか黙示的であるかを問わない、なんらの保証も、かかる免責が法的に無効とされた場合を除き、行われないものとします。

目次

1 序章 1

2 SPARC Enterprise M シリーズ 2

2.1 性能概要 2

2.2 エントリーレベルシステム : SPARC Enterprise M3000 5

2.3 ミッドレンジシステム : SPARC Enterprise M4000/M5000 6

2.3.1 SPARC Enterprise M4000 6

2.3.2 SPARC Enterprise M5000 7

2.4 ハイエンドシステム : SPARC Enterprise M8000/M9000 8

2.4.1 SPARC Enterprise M8000 8

2.4.2 SPARC Enterprise M9000（32 CPU 構成）. 9

2.4.3 SPARC Enterprise M9000（64 CPU 構成）. 9

2.5 商用コンピューティングおよび科学コンピューティングのニーズへの対応 10

3 システムアーキテクチャー 11

3.1 システムコンポーネントの概要 11

3.1.1 システムインターコネクト 11

3.1.2 SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ 12

3.1.3 メモリ 12

3.1.4 メモリミラーリング 13

3.1.5 システムクロック 15

3.1.6 PCI Express および PCI-X テクノロジー 15

3.1.7 サービスプロセッサ - システム監視機構 16

3.1.8 電源と冷却 16

3.1.9 オペレーターパネル 18

4 システムバスのアーキテクチャー - Jupiter インターコネクト 20

4.1 インターコネクトアーキテクチャー 20

4.1.1 SPARC Enterprise M3000 のシステムインターコネクトアーキテクチャー 21

4.1.2 SPARC Enterprise M4000 のシステムインターコネクトアーキテクチャー 21

4.1.3 SPARC Enterprise M5000 のシステムインターコネクトアーキテクチャー 22

4.1.4 SPARC Enterprise M8000/M9000 のインターコネクトアーキテクチャー 23

4.2 システムインターコネクトの信頼性機能 24

4.3 スケーラブルなパフォーマンス 24

5 SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ 26

5.1 次世代のプロセッサテクノロジー 26

5.1.1 チップマルチプロセッシング（CMP）. 26

5.1.2 垂直マルチスレッディング（VMT）. 27

5.1.3 同時マルチスレッディング（SMT）. 27

5.2 SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサのアーキテクチャー 27

5.2.1 キャッシュシステム 28
5.2.2 SPARC64 VI プロセッサのマイクロアーキテクチャーの概要 30
5.2.3 SPARC64 VII/SPARC64 VII+ プロセッサのマイクロアーキテクチャーの概要 31
5.2.4 SPARC64 VII/SPARC64 VII+ プロセッサの命令処理 33
5.2.5 命令フェッチ部 33
5.2.6 命令実行部 34
5.3 RAS 機能（Reliability、Availability、Serviceability）. 35
5.3.1 内蔵 RAM の信頼性および可用性 36
5.3.2 内蔵レジスタと演算器の信頼性機能 37
5.3.3 同期一括更新方式と命令リトライ 37
5.3.4 サービス性の向上. 38

6 I/O サブシステム 39
6.1 I/O サブシステムのアーキテクチャー 39
6.1.1 SPARC Enterprise M3000 の I/O サブシステム. 39
6.1.2 SPARC Enterprise M4000/M5000 の I/O サブシステム. 40
6.1.3 SPARC Enterprise M8000/M9000 の I/O サブシステム. 41
6.2 内蔵 I/O 43
6.3 PCI ボックス 43

7 信頼性、可用性、保守性 45
7.1 冗長ホットスワップコンポーネント 45
7.2 パーティショニング機能 46
7.2.1 拡張システムボード（XSB）. 46
7.2.2 CPU の混在 49
7.3 Dynamic Reconfiguration（DR 機能）. 51
7.4 高度な信頼性機能 51
7.5 エラー検出、診断、復旧 51

8 システム管理 53
8.1 システム監視機構 53
8.1.1 冗長 XSCF 54
8.1.2 DSCP ネットワーク 54
8.1.3 XSCF 制御パッケージ 54
8.1.4 ロールベースのシステム管理. 55
8.1.5 Enhanced Support Facility 55
8.1.6 Systemwalker Centric Manager 55

9 Oracle Solaris 10 57
9.1 性能 57
9.2 可用性 58
9.3 セキュリティ 60
9.4 仮想化とリソース管理. 61

10 まとめ 62

1. 序章

企業におけるテクノロジー依存は、これまでになく大きくなっています。現在、製品設計から受注処理まで、コンピュータシステムはすべての業務に重要な役割を果たしています。多くの場合、ビジネスの成功は、IT サービスの継続的な可用性に依存します。かつてはデータセンターの一部でのみ必要とされていたメインフレームクラスの信頼性と保守性が、現在は企業全体のシステムに不可欠です。さらに、停電中もデータセンターのサーバに電力を供給し、サービスを稼働させ続けることが、重視されています。

最も優先度が高いのは可用性ですが、コストを予算内に抑え、操作の親しみやすさを保つ必要もあります。ネットワーク型サービスをできる限り効率的かつ経済的に提供するため、**企業**は統合と仮想化という戦略によって、すべての IT 資産を最大限に活用しようと試みています。その結果、現代における IT システムの要件は、処理能力だけではありません。仮想化機能が組み込まれ、サーバの利用を最適化する関連ツール、テクノロジー、プロセスを備えた高い柔軟性を持つサーバが必要なのです。さらに、新しいコンピュータインフラストラクチャーは、技術や教育に対する現在の投資の保護にも役立たなければなりません。

富士通の SPARC Enterprise M シリーズは、非常に信頼性が高く、管理しやすい、垂直スケーラブルシステムで、従来のメインフレームのメリットを多く備えていながら、関連するコストや複雑さ、そしてベンダーロックインがありません。実際、これらのサーバは、メインフレームクラスのシステムアーキテクチャーをオープンシステムの価格で提供します。SPARC Enterprise M シリーズは、1 ～ 64 個のプロセッサによるシンメトリックマルチプロセシング（SMP）のスケーラビリティ、4 TB のメモリサブシステム、高スループットの I/O アーキテクチャーを備えており、統合ワークロードが必要とする力仕事を簡単に実行します。さらに、これらのサーバは強力な Oracle Solaris 10 オペレーティングシステム（Oracle Solaris 10 OS）を動作させるとともに、優れた仮想化技術を採用しています。SPARC Enterprise M シリーズは、パーティショニング機能、拡張システムボード（XSB）、Dynamic Reconfiguration（DR 機能）、Oracle Solaris Containers を提供することによって、オープンシステムのコンピュータプラットフォームにメインフレームクラスの高度なリソース制御を行います[1]。

1.SPARC Enterprise M3000 では、パーティショニング機能、拡張システムボード、Dynamic Reconfiguration 機能はサポートしていません。

2. SPARC Enterprise M シリーズ

莫大なコンピュータパワー、回復力に富むシステムアーキテクチャー、柔軟なリソース制御機能、Oracle Solaris 10 の高度な機能が富士通の SPARC Enterprise M シリーズに統合され、クラス最高のエンタープライズプラットフォームとなっています。もう 1 つのメリットとして、SPARC Enterprise M シリーズ（図 2-1 参照）は前世代の富士通サーバを超えるパフォーマンスを提供しながら、ソフトウェアや教育、データセンターの実務における既存の投資を無駄にしない明確なアップグレードパスを持っています。SPARC Enterprise M シリーズを利用することによって、IT 組織はさらに強力なインフラストラクチャーを作成し、ハードウェアの利用を最適化し、アプリケーションの可用性を高めて、結果的に運用コストとリスクを低減できます。

図 2-1. スケーラブルなパワー、信頼性、柔軟性を企業にもたらす SPARC Enterprise M シリーズ

2.1 性能概要

SPARC Enterprise M シリーズは、パワー、信頼性、柔軟性をもたらすような多くの特性を共有しています。SPARC Enterprise M4000/M5000/M8000/M9000 は、バランスのとれた非常にスケーラブルな SMP 設計を採用しています。SMP 設計は、最新世代の SPARC64 プロセッサを利用して、高速で低レイテンシの新しいシステムインターコネクトによってメモリと I/O に接続し、アプリケー

ションに並外れたスループットをもたらします。また、これらのサーバは、計画的なダウンタイムおよび計画外のダウンタイムを短縮するよう設計されており、機能停止を回避して復旧時間を短縮するための優れた信頼性、可用性、保守性を発揮する機能を備えています。高度な CPU 統合とデータパスの完全性、メモリの拡張 ECC とメモリのミラーリング、エンドツーエンドのデータ保護、ホットスワップ可能なコンポーネント、障害耐性のある電源オプション、ハードウェアの冗長性などの設計機能が、このサーバの信頼性を高めています。

SPARC Enterprise M4000/M5000/M8000/M9000 の構成は、高い柔軟性を誇ります。管理者はパーティショニング機能を使用して、物理的には単一の SPARC Enterprise M4000/M5000/M8000/M9000 を電気的に分離された複数のパーティションに分割し、各パーティションで Oracle Solaris の個別インスタンスを動作させることができます。1 つのパーティショニング機能内でのハードウェアまたはソフトウェア障害は、障害が発生したリソースを共有していない限り、他のドメインで動作しているアプリケーションには影響を及ぼしません。

そのため、Dynamic Reconfiguration によって、重要なシステムを中断することなく、パーティショニング機能間でハードウェアリソースを割り当て直すことができます。SPARC Enterprise M4000/M5000/M8000/M9000 は拡張システムボードの技術を使用してリソース制御を 1 歩前進させており、CPU、メモリ、I/O コンポーネントなどサブシステムボードのリソースをパーティショニング機能に割り当てることができます。このきめ細かなリソース制御は拡張システムボード（XSB）技術によって、リソースの利用をさらに最適化します。

さらなる付加価値として、SPARC Enterprise M シリーズが提供する一連のコンピュータパワーは、多くのクラスに求められる垂直方向のスケーラビリティレベルを実現し、企業がタスクに適切なシステムをマッチングさせることができるようにします。ラックマウント式の SPARC Enterprise M3000 は、SPARC Enterprise M シリーズの多くの特性を備えており、シリーズ共通の操作性、管理の容易性のメリットを共有しているエントリーレベルサーバです。エントリーレベルクラス最高の CPU パワー、回復力に富むシステムアーキテクチャー、Oracle Solaris 10 OS の高度な機能が 2 U ラック筐体にコンパクトに統合され、高性能・高品質なエコロジーサーバとなっています。

ラックマウント式の SPARC Enterprise M4000/M5000 は経済的で強力、かつ信頼性のあるサーバで、エントリーレベルおよびミッドレンジのシステム要件（表 2-1 参照）に適しています。SPARC Enterprise M8000/M9000 は、ハイエンドコンピューティングに必要な莫大な処理能力を提供します（表 2-1 参照）。

表 2-1. SPARC Enterprise M3000/M4000/M5000 の特徴

	SPARC Enterprise M3000	SPARC Enterprise M4000	SPARC Enterprise M5000
筐体	• 2 ラックユニット	• 6 ラックユニット	• 10 ラックユニット
SPARC64 VI プロセッサ	—	• 2.15 GHz • 5 MB の L2 キャッシュ • 最大 4 CPU チップ（2 コア）	• 2.15 GHz • 5 MB の L2 キャッシュ • 最大 8 CPU チップ（2 コア）
SPARC64 VII プロセッサ	• 2.52 GHz/2.75 GHz • 5 MB の L2 キャッシュ • 1 CPU チップ（4 コア／2 コア）	• 2.4 GHz/2.53 GHz • 5 MB/5.5 MB の L2 キャッシュ • 最大 4 CPU チップ（4 コア）	• 2.4 GHz/2.53 GHz • 5 MB/5.5 MB の L2 キャッシュ • 最大 8 CPU チップ（4 コア）
SPARC64 VII+ プロセッサ	• 2.86 GHz • 5 .5 MB の L2 キャッシュ • 1 CPU チップ（4 コア／2 コア）	• 2.66 GHz • 11 MB の L2 キャッシュ • 最大 4 CPU チップ（4 コア）	• 2.66 GHz • 11 MB の L2 キャッシュ • 最大 8 CPU チップ（4 コア）
メモリ	• 最大 64 GB • 8 の DIMM スロット	• 最大 256 GB • 32 の DIMM スロット	• 最大 512 GB • 64 の DIMM スロット
内蔵 I/O スロット	• 4 つの PCI Express	• 4 つの PCI Express • 1 つの PCI-X	• 8 つの PCI Express • 2 つの PCI-X
外部 I/O シャーシ	• —	• 最大 2 ユニット	• 最大 4 ユニット
外部オンボードインターフェース	• 4 つのギガビットイーサネットポート • 1 SAS ポート	• 2 つのギガビットイーサネットポート	• 1 IOU につき 2 つのギガビットイーサネットポート
内蔵ストレージ	• SAS • 最大 4 ドライブ	• SAS • 最大 2 ドライブ	• SAS • 最大 4 ドライブ
パーティショニング機能	• 1	• 最大 2	• 最大 4

表 2-2. SPARC Enterprise M8000/M9000（32 CPU 構成）/M9000（64 CPU 構成）の特徴

	SPARC Enterprise M8000	SPARC Enterprise M9000（32CPU 構成）	SPARC Enterprise M9000（64CPU 構成）
筐体	• 1 筐体	• 1 筐体	• 2 筐体
SPARC64 VI プロセッサ	• 2.28 GHz/2.4 GHz • 5/6 MB の L2 キャッシュ • 最大 16 CPU チップ（2 コア）	• 2.28 GHz/2.4 GHz • 5/6 MB の L2 キャッシュ • 最大 32 CPU チップ（2 コア）	• 2.28 GHz/2.4 GHz • 5/6 MB の L2 キャッシュ • 最大 64 CPU チップ（2 コア）
SPARC64 VII プロセッサ	• 2.52 GHz/2.88 GHz • 6 MB の L2 キャッシュ • 最大 16 CPU チップ（4 コア）	• 2.52 GHz/2.88 GHz • 6 MB の L2 キャッシュ • 最大 32 CPU チップ（4 コア）	• 2.52 GHz/2.88 GHz • 6 MB の L2 キャッシュ • 最大 64 CPU チップ（4 コア）
SPARC64 VII+ プロセッサ	• 3.0 GHz • 12 MB の L2 キャッシュ • 最大 16 CPU チップ（4 コア）	• 3.0 GHz • 12 MB の L2 キャッシュ • 最大 32 CPU チップ（4 コア）	• 3.0 GHz • 12 MB の L2 キャッシュ • 最大 64 CPU チップ（4 コア）
メモリ	• 最大 1 TB • 128 の DIMM スロット	• 最大 2 TB • 256 の DIMM スロット	• 最大 4 TB • 512 の DIMM スロット
内蔵 I/O スロット	• 32 の PCI Express	• 64 の PCI Express	• 128 の PCI Express
外部 I/O シャーシ	• 最大 8 ユニット	• 最大 16 ユニット	• 最大 16 ユニット
外部オンボードインターフェース	• 1 IOUA[a] につき、2 つのギガビットイーサネットポート	• 1 IOUA[a] につき、2 つのギガビットイーサネットポート	• 1 IOUA[a] につき、2 つのギガビットイーサネットポート
内蔵ストレージ	• SAS • 最大 16 ドライブ	• SAS • 最大 32 ドライブ	• SAS • 最大 64 ドライブ
パーティショニング機能	• 最大 16	• 最大 24	• 最大 24

a. IOUA はオプションカードです。

2.2 エントリーレベルシステム : SPARC Enterprise M3000

SPARC Enterprise M3000 は、2 ラックユニット（RU）のサイズで 1 つのプロセッサチップと最大 64 GB のメモリをサポートし、マザーボードには SPARC64 VII/SPARC64 VII+（2 コア／4 コア）プロセッサチップを搭載しています。SPARC Enterprise M3000 の I/O サブシステムは、4 つの内蔵 PCI Express ショートスロット、4 台の内蔵磁気ディスクユニット、1 台の DVD ドライブユニットを搭載するほか、外部ストレージまたはテープドライブユニットを取り付けるための外部 SAS ポートを備えています。電力供給は 2 台の電源ユニット、冷却は 2 台のファンユニットで行われます。図 2-2 に、SPARC Enterprise

M3000 の正面図と背面図を示します。

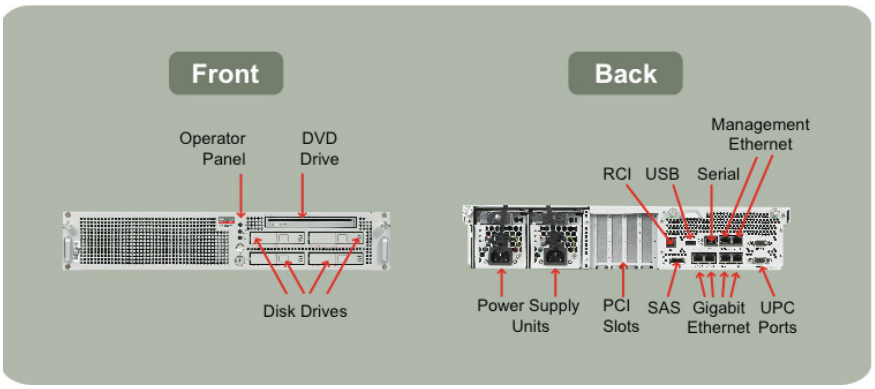

図 2-2. SPARC Enterprise M3000 の筐体の図

2.3 ミッドレンジシステム : SPARC Enterprise M4000/M5000

SPARC Enterprise M4000/M5000 は、エンタープライズクラスの機能を備え、費用効果の優れた高性能コンピューティングプラットフォームです。これらのミッドレンジサーバは、データセンターで発生する負荷に確実に対応し、基幹業務を支援するよう設計されています。

2.3.1 SPARC Enterprise M4000

SPARC Enterprise M4000 は、6 ラックユニット（RU）のサイズで最大 4 つのプロセッサチップ、256 GB のメモリ、最大 2 つのパーティションをサポートし、4 つの内蔵 PCI Express ショートスロット、1 つの内蔵 PCI-X ショートスロット、2 台の内蔵磁気ディスクユニット、1 台の DVD ドライブユニット、テープドライブユニット（オプション）も備えています。電力供給は 2 台の電源ユニット、冷却は 4 台のファンユニットで行われます。図 2-3 に、SPARC Enterprise M4000 の正面図と背面図を示します。

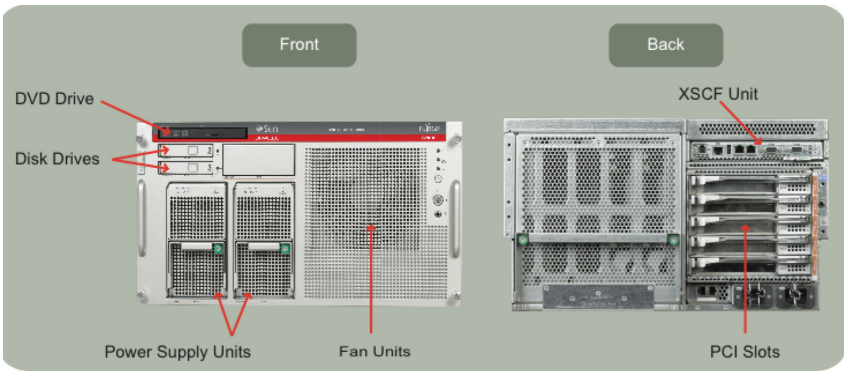

図 2-3. SPARC Enterprise M4000 の筐体の図

2.3.2 SPARC Enterprise M5000

SPARC Enterprise M5000 は、10 ラックユニット（RU）のサイズで最大 8 つのプロセッサチップ、512 GB のメモリ、最大 4 つのパーティションをサポートし、8 つの内蔵 PCI Express ショートスロット、2 つの内蔵 PCI-X ショートスロットのほか、4 台の内蔵磁気ディスクユニット、1 台の DVD ドライブユニット、テープドライブユニット（オプション）も備えています。電力供給は 4 台の電源ユニット、冷却は 4 台のファンユニットで行われます。図 2-4 に、SPARC Enterprise M5000 の正面図と背面図を示します。.

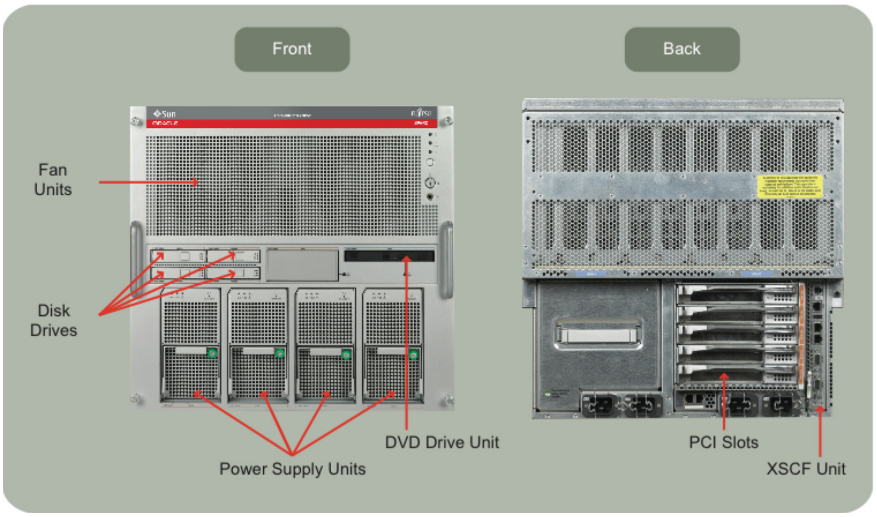

図 2-4. SPARC Enterprise M5000 の筐体の図

2.4 ハイエンドシステム : SPARC Enterprise M8000/M9000

SPARC Enterprise M8000/M9000 は、優れたパフォーマンスを通じて最大級の負荷にも対応するよう設計されています。これらのサーバでは、メインフレームの信頼性、スーパーコンピュータに採用されることの多い高度なパフォーマンステクノロジー、オープンシステム環境が融合されており、信頼性とスループットの高い柔軟なソリューションの構築を支援します。

2.4.1 SPARC Enterprise M8000

SPARC Enterprise M8000 は、エンタープライズシステムの筐体に搭載します。最大で 4 つの CPU メモリユニット（CMU）と 4 つの I/O ユニット（IOU）をサポートし、完全構成時には、16 つのプロセッサチップ、1 TB のメモリ、32 つの内蔵 PCI Express ショートスロットを備え、16 のパーティションに分割することができます。また、最大で 16 台の内蔵磁気ディスクユニット、1 台の DVD ドライブユニット、およびテープドライブユニット（オプション）をサポートします。電力供給は 9 台の電源ユニット、冷却は 12 台のファンユニットで行われます。図 2-5 に、SPARC Enterprise M8000 の正面図と背面図を示します。

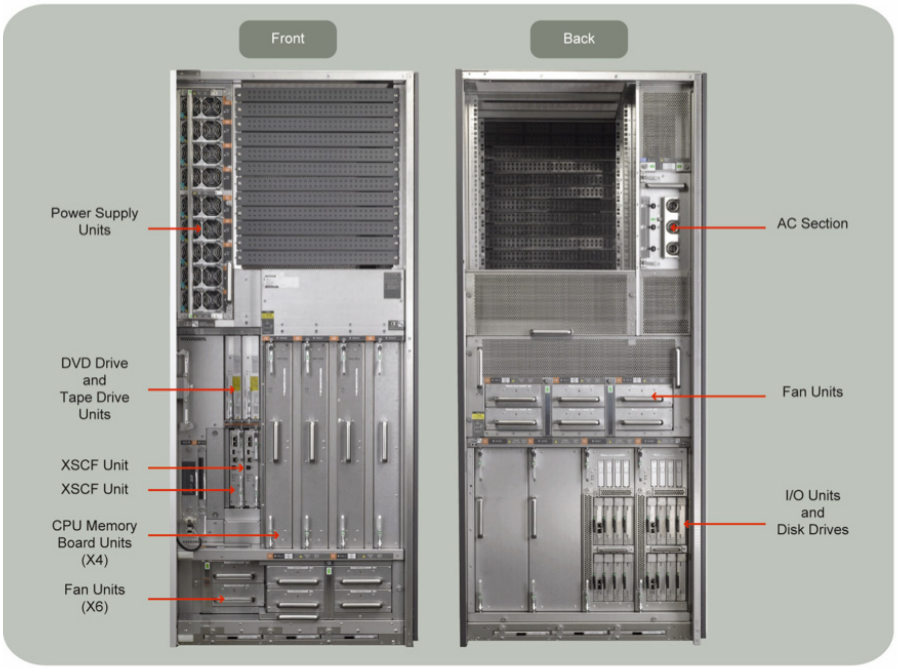

図 2-5. SPARC Enterprise M8000 の筐体の図

2.4.2 SPARC Enterprise M9000（32 CPU 構成）

SPARC Enterprise M9000（32 CPU 構成）は、エンタープライズシステムの筐体に搭載します。最大で 8 つの CMU と 8 つの IOU をサポートし、完全構成時には、32 つのプロセッサチップ、2 TB のメモリ、64 つの内蔵 PCI Express ショートスロットを備え、24 のパーティションに分割することができます。また、最大で 32 台の内蔵磁気ディスクユニット、1 台の DVD ドライブユニット、およびテープドライブユニット（オプション）をサポートします。SPARC Enterprise M9000（32 CPU 構成）の電力供給は 15 台の電源ユニット、冷却は 16 台のファンユニットで行われます。図 2-6 に、M9000（32 CPU 構成）の正面図と背面図を示します。

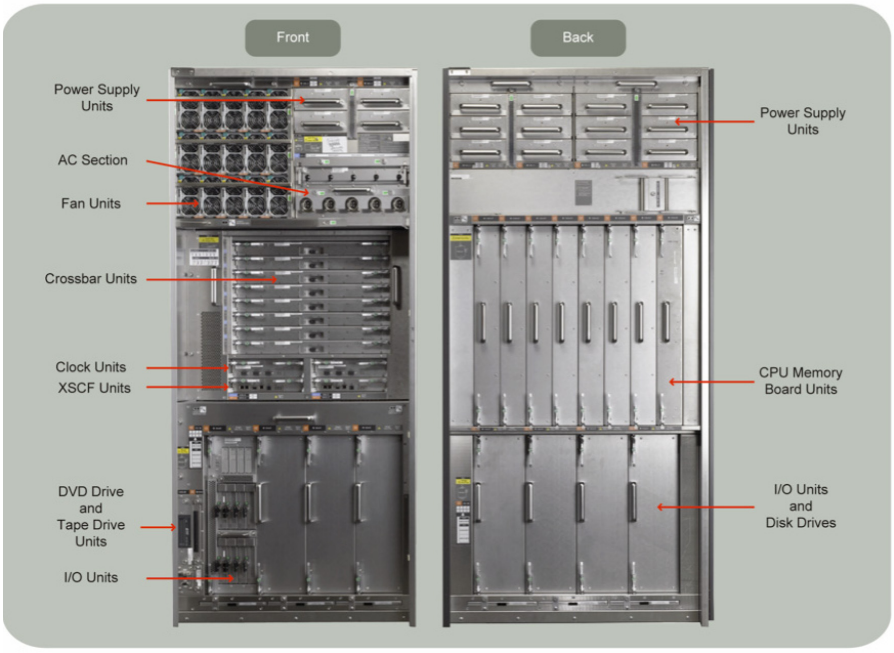

図 2-6. SPARC Enterprise M9000（32 CPU 構成）の筐体の図

2.4.3 SPARC Enterprise M9000（64 CPU 構成）

既存の基本筐体に拡張筐体を追加して SPARC Enterprise M9000（64 CPU 構成）を構築すると、最大で 16 つの CMU と 16 つの IOU をサポートできます。完全構成時には、64 つのプロセッサチップ、4 TB のメモリ、128 つの内蔵 PCI Express ショートスロットを備え、24 のパーティションに分割することができます。また、最大で 64 台の内蔵磁気ディスクユニット、2 台の DVD ドライブユニット、2 台のテープドライブ（オプション）をサポートします。電力供給は 30 台の電源ユニット、冷却は 32 台のファンユニットで行われます。図 2-7 に、SPARC Enterprise M9000（64 CPU 構成）の正面図を示します。

図 2-7. SPARC Enterprise M9000（64 CPU 構成）の筐体の図

2.5 商用コンピューティングおよび科学コンピューティングのニーズへの対応

広範なコンピューティング環境に適応する SPARC Enterprise M シリーズは、商用コンピューティングの負荷に対応する可用性を実現するための機能に加え、ハイパフォーマンスコンピューティング（HPC）で要求される高い処理能力を備えています（表 2-3 参照）。

表 2-3. 広範なエンタープライズアプリケーションに利点をもたらすSPARC Enterprise M シリーズの能力と柔軟性

SPARC Enterprise M3000/M4000/M5000	SPARC Enterprise M8000/M9000
• サーバ統合 • 業務処理（ERP、CRM、OLTP、バッチ） • データベース • 意志決定支援 • データマート • ウェブサービス • システムとネットワークの管理 • アプリケーション開発 • 科学エンジニアリング	• サーバ統合 • 業務処理（ERP、CRM、OLTP、バッチ） • データベース • 意志決定支援 • データウェアハウス • IT インフラストラクチャー • アプリケーションサービス • 演算能力集約型科学エンジニアリング

3. システムアーキテクチャー

ワークロードの増大と、より少ないコストでより大きな効果を得るという要望の増大に絶えず直面しているため、IT 組織は、少数かつ強力なシステムで処理要件を満たすことによって経済的メリットを提供できます。SPARC Enterprise M シリーズのシステムインターコネクト、プロセッサ、メモリサブシステム、I/O サブシステムが連携して、サーバ統合ニーズにいつでも取り組める、スケーラブルなハイパフォーマンスプラットフォームを作成します。これらのサーバを利用すると、複数のプロジェクトを単一のプラットフォームに集約し、より安価にアプリケーションの実行を促進できます。

3.1 システムコンポーネントの概要

SPARC Enterprise M シリーズの設計は、高信頼性、卓越したパフォーマンス、真の SMP のスケーラビリティを実現することに特に重点を置いています。SPARC Enterprise M シリーズ内におけるすべてのサブシステムの特性と機能は、この目標を目指しているのです。SPARC Enterprise M シリーズの高帯域幅のシステムバス、強力な SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ、高密度メモリオプション、高速 PCI Express （PCIe）および PCI eXtended（PCI-X）拡張スロットは、エンタープライズアプリケーション用の信頼できるスケーリングだけでなく、高レベルの稼働時間とスループットをもたらします。

3.1.1 システムインターコネクト

メインフレーム技術を基盤とする Jupiter システムインターコネクトは、SPARC Enterprise M シリーズのハイレベルのパフォーマンス、スケーラビリティと信頼性を促進します。SPARC Enterprise M3000 内の単一のシステムコントローラー、および、SPARC Enterprise M4000/M5000/M8000/M9000 内の複数のシステムコントローラーとクロスバーユニットが、CPU、メモリ、I/O サブシステム間をエンドツーエンドで接続します。コンポーネント間に複数のバスルートを持つため、パフォーマンスが向上し、障害による切り替え発生時にもシステムオペレーションを継続できます。実際、SPARC Enterprise M シリーズで使用されているシステムインターコネクトは、737 GB/ 秒のピーク時帯域幅を実現し、富士通の前世代ハイエンドサーバの 5.5 倍以上のシステムスループットを供給します。各 SPARC Enterprise M シリーズのシステムインターコネ

クトに関する他の詳細な技術情報は、*“ システムバスのアーキテクチャー - Jupiter インターコネクト（20 ページ）”* に記載されています。

3.1.2 SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ

SPARC Enterprise M3000 は、富士通が開発した SPARC64 VII および SPARC64 VII+ プロセッサをサポートし、一方、SPARC Enterprise M4000/M5000/M8000/M9000 は、SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサを利用することができます。マルチコア、2 スレッドの SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサの設計は、高レベルの信頼性と性能を実現するメインフレームシステム開発の数十年の経験に基づいています。2 コア、2 スレッドの SPARC64 VI プロセッサは 90 nm テクノロジーを利用しており、一方、SPARC64 VII/SPARC64 VII+ プロセッサは、65 nm テクノロジー採用により、チップに 4 コア、2 スレッド、最大 12MB の L2 キャッシュメモリを備え、クロック速度も向上しています。いずれのプロセッサも、消費電力は 160 W 以下です。さらに、SPARC Enterprise M4000/M5000/M8000/M9000 では、同一パーティション内、または同一のシステムボードにおいて SPARC64 VI、SPARC64 VII、SPARC64 VII+ プロセッサの混在する構成をサポートすることにより、柔軟性を増大させ、投資保護を維持します。SPARC64 プロセッサに関するその他の詳細技術情報は、*“SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ（26 ページ）”* に記載されています。

3.1.3 メモリ

SPARC Enterprise M シリーズのメモリサブシステムは、システムのスケーラビリティとスループットを高めます。実際、SPARC Enterprise M9000 は、最大 4 TB のメモリを搭載できます。SPARC Enterprise M3000 は、2 ウェイメモリインタリーブを備えた DDR-II DIMM をサポートします。SPARC Enterprise M4000/M5000/M8000/M9000 は 8 ウェイメモリインタリーブを備えた DDR-II DIMM を使用してシステムパフォーマンスを強化しています。単一バンク内で複数の DIMM サイズはサポートされていませんが、システムボード内で異なるメモリサイズを搭載することができます。使用可能な DIMM サイズは、1 GB、2 GB、4 GB、8 GB です。メモリサブシステムの詳細を表 3-1 に説明します。

表 3-1. SPARC Enterprise M シリーズのメモリサブシステム仕様

	SPARC Enterprise M3000	SPARC Enterprise M4000	SPARC Enterprise M5000	SPARC Enterprise M8000	SPARC Enterprise M9000（32CPU 構成）	SPARC Enterprise M9000（64CPU 構成）
最大メモリ容量	• 64 GB	• 256 GB	• 512 GB	• 1 TB	• 2 TB	• 4 TB
DIMM スロット	• 8	• 32	• 64	• 最大 128	• 最大 256	• 最大 512
バンクサイズ	• 4 DIMM	• 4 DIMM	• 4 DIMM	• 8 DIMM	• 8 DIMM	• 8 DIMM
バンクの数	• 2	• 8	• 16	• 最大 16	• 最大 32	• 最大 64

SPARC Enterprise M シリーズのメモリサブシステムは、パフォーマンスに加えて、信頼性を念頭に置いて作成されています。メインメモリに保存されているすべてのデータに対して、ECC 保護が実装されています。また、次の拡張機能によって早期診断と障害分離を強化して、システムの完全性を維持し、アプリケーションの可用性を向上させます。

- メモリパトロール - メモリパトロールは、メモリにエラーがないか定期的にスキャンする機能です。この予防機能によって、システムまたはアプリケーションのエラーが発生する前にメモリの障害領域を使用しないようにできるため、システムの信頼性が向上します。
- 拡張 ECC - SPARC Enterprise M シリーズのメモリ拡張 ECC 機能は、1 ビットエラー訂正を行ったり、メモリデバイスの障害によって生じる場合があるバースト読み込みエラーなどのイベントにもかかわらず処理を継続したりできます。この機能は、IBM のチップキル技術に似ています。

3.1.4 メモリミラーリング

SPARC Enterprise M4000/M5000/M8000/M9000 は、メモリミラーリング機能をサポートしています。メモリミラーリングは、高可用性機能で、可用性要件が厳しいアプリケーションの実行に適しています。SPARC Enterprise M4000/M5000/M8000/M9000 上でメモリミラーリングモードが有効になっている場合、メモリサブシステムは書き込み時にデータを二重化し、読み出し時にデータをメモリミラーの各サイドと比較します。バスレベルまたは DIMM レベルでエラーが発生した場合、もう一方のメモリバスと代替 DIMM セットを使用して、通常のデータ処理を継続します。SPARC Enterprise M4000/M5000 では、共通メモリアクセスコントローラー（MAC）の Application Specific Integrated Circuit（ASIC）を使用して、同じメモリモジュール内でメモリをミラーリングしています（図 3-1 および図 3-2 参照）。

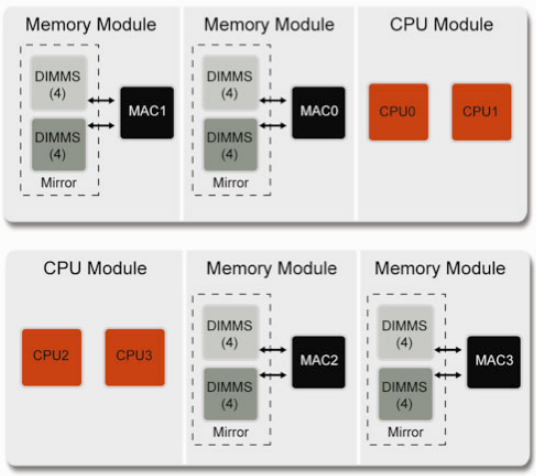

図 3-1. SPARC Enterprise M4000 のメモリミラーリングアーキテクチャー

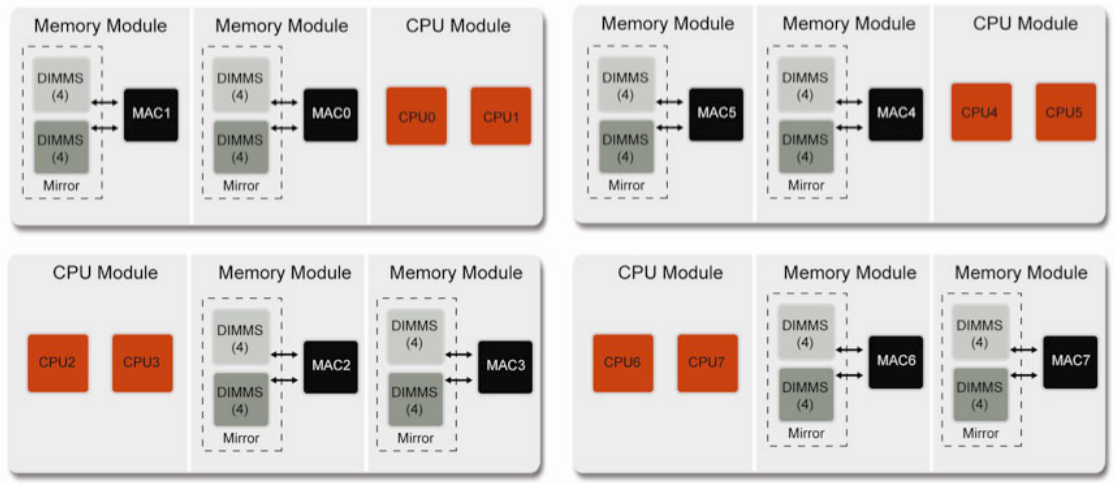

図 3-2. SPARC Enterprise M5000 のメモリミラーリングアーキテクチャー

SPARC Enterprise M8000/M9000 では、隣接した複数の MAC ASIC にわたってメモリをミラーリングし、信頼性を高めています（図 3-3 参照）。しかし、Quad-XSB 構成では、対になった DIMM が、SPARC Enterprise M8000/M9000 の異なる Quad-XSB に分割されます。そのため、Quad-XSB 構成では、メモリミラーはサポートされません。

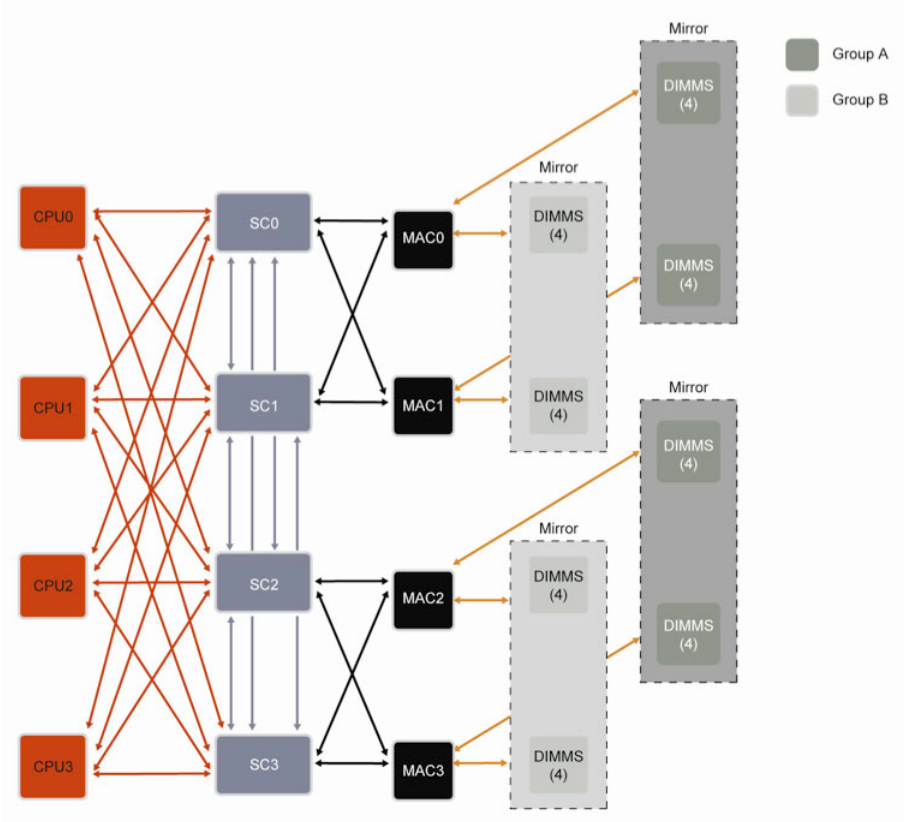

図 3-3. SPARC Enterprise M8000/M9000 のメモリミラーリングアーキテクチャー

3.1.5 システムクロック

システムクロックの実装は、SPARC Enterprise M シリーズのモデルによって異なりますが、各サーバは信頼性を念頭に置いて設計されています。特に、SPARC Enterprise M8000 のクロックチップは、内部が冗長構成となっています。さらに可用性を強化し、保守を容易にするために、SPARC Enterprise M9000 には、クロック信号源も 2 つ実装されており、クロックチップとシステムボード間は冗長化されています。そのため、片方のルートで障害が発生しても、もう片方のルートでシステムを再起動させることができます。

3.1.6 PCI Express および PCI-X テクノロジー

SPARC Enterprise M シリーズは、PCI バスを使用して、I/O サブシステム内の高速データ転送を実現します。SPARC Enterprise M シリーズは、PCI Express 拡張カードをサポートするため、PCI Express 物理レイヤー（PCIe PHY）ASIC を使用して、PCI Express プロトコルを実装します。PCI Express テクノロジーは、もとの PCI テクノロジーのピーク時データ転送速度を倍増させ、スループットは 2.5 GB/ 秒に達します。事実、PCI Express は、Fibre Channel、Infiniband、

Gigabit Ethernet などの高速インターコネクトを組み込むために開発されました。SPARC Enterprise M4000/M5000/M8000/M9000 では、外部デバイスへの高速なアクセスを実現するために PCI-X 拡張カードもサポートしています。PCI-X は、既存の PCI カードとの下位互換性がありますが、帯域幅を増やして、64 ビットデバイスに対して最大 1 GB/ 秒のデータ転送を可能にしています。SPARC Enterprise M シリーズの I/O サブシステムに関する他の詳細な技術情報は、*"I/O サブシステム（39 ページ）"* を参照してください。

3.1.7 サービスプロセッサ - システム監視機構

コンピュータシステムの管理を単純化することによって、ホスティングされているアプリケーションの可用性がより高まります。このことを念頭に置いて、SPARC Enterprise M シリーズは、システム監視機構（XSCF）を備えています。XSCF は、サーバから独立した専用プロセッサで構成され、XSCF 制御パッケージ（XCP）を実行して、リモート監視機能と管理機能を提供します。このサービスプロセッサは、環境センサーを定期的に監視し、潜在的なエラー状態を事前に警告し、必要に応じて予防的システム保守手順を実行します。実際、サーバに電力が供給されている間は、システムがアクティブでない場合でも、XSCF は常にプラットフォームを監視します。

XCP は、パーティショニング機能構成、監査管理、ハードウェア制御機能、ハードウェアステータスの監視、レポート、処理、自動診断とドメイン復旧、キャパシティオンデマンド操作、XSCF フェイルオーバーサービスを可能にします。XSCF と XCP に関する他の詳細な技術情報は、*" システム管理（53 ページ）"* を参照してください。

3.1.8 電源と冷却

SPARC Enterprise M シリーズは、電源と冷却に別々のモジュールを使用します。システム全域に配置されたセンサーが、プロセッサと主要 ASIC の温度や、いくつかの場所の気温を測定します。電源および冷却サブシステムにおけるハードウェアの冗長性は、環境モニタリングと一体となって、電源またはファンが故障している状態でも、サーバを動作させ続けます。

3.1.8.1 ファンユニット

SPARC Enterprise M シリーズは、完全に冗長なホットスワップファン機能を一次冷却システムとして使用します（表 3-2 および表 3-3 参照）。1 つのファンが故障した場合、XSCF が障害を検出し、残りのファンを高速動作に切り替えて、エアフローの減少を補います。SPARC Enterprise M シリーズは、このような状況下で正常に動作し、故障したファンユニットを修理、交換するのに十分な時間を与えます。ファンユニットは、アプリケーションの処理を中断することなく交換できます。

3.1.8.2 電源

冗長な電源および電源ケーブルを使用することによって、SPARC Enterprise M シリーズの障害耐性は増します（表 3-2 および表 3-3 参照）。冗長なホットスワップ電源によって電力を供給し、電源障害時もサーバの動作を継続できます。電源ユニットはホットスワップ可能なので、システムの動作中に減設および交換できます。

SPARC Enterprise M8000/M9000 は、三相電源ユニットを用意しています。三相電源を持つモデルは、中性線と各相を接続するスター接続と、各相を接続するデルタ接続の 2 つの構成が可能です。

表 3-2. SPARC Enterprise M3000/M4000/M5000 の電源および冷却仕様

	SPARC Enterprise M3000	SPARC Enterprise M4000	SPARC Enterprise M5000
ファンユニット	• 2 つのファンユニット • 2 つの 80 mm ファン • 1+1 冗長性	• 4 つのファンユニット • 2 つの 172 mm ファン • 2 つの 60 mm ファン • いずれかのタイプは冗長	• 4 つのファンユニット／2 つのファングループ • 4 つの 172 mm ファン • 1 つのファングループに 2 つのファンユニット • 1 つのファングループごとに 1 つの冗長ファン
電源	• AC: 505 W、DC: 510 W（2.86 GHz）（最大） • 2 つのユニット • 1+1 冗長性 • 交流単相・直流電源のどちらかを選択	• 1692 W（2.66 GHz）（最大） • 2 つのユニット • 1+1 冗長性 • 単相	• 3270 W（2.66 GHz）（最大） • 4 つのユニット • 2+2 冗長性 • 単相
電源ケーブル	• 2 本の電源ケーブル • 1+1 冗長性の電源ケーブル	• 2 本の電源ケーブル • 1+1 冗長性の電源ケーブル	• 4 本の電源ケーブル • 2+2 冗長性の電源ケーブル

表 3-3. SPARC Enterprise M8000/M9000 の電源および冷却仕様

	SPARC Enterprise M8000	SPARC Enterprise M9000（32CPU 構成）	SPARC Enterprise M9000（64CPU 構成）
ファンユニット	• 12 のファンユニット • 4 つの 172 mm ファン • 8 つの 60 mm ファン • N+1 冗長性	• 16 のファンユニット • 16 の 172 mm ファン • N+1 冗長性	• 32 のファンユニット • 32 の 172 mm ファン • N+1 冗長性

	SPARC Enterprise M8000	SPARC Enterprise M9000（32CPU 構成）	SPARC Enterprise M9000（64CPU 構成）
電源	• 10,500 W（3.0 GHz）（最大） • 9 ユニット • N+1 冗長性	• 20,220 W（3.0 GHz）（最大） • 15 ユニット • N+1 冗長性	• 40,440 W（3.0 GHz）（最大） • 30 ユニット • N+1 冗長性
オプション	• 単相 • 三相 • デュアルグリッド	• 単相 • 三相 • デュアルグリッド	• 単相 • 三相 • デュアルグリッド
電源ケーブル	• 3 本の電源ケーブル（1 系統受電） • 6 本の電源ケーブル（2 系統受電） • 2 本の電源ケーブル（三相）	• 5 本の電源ケーブル（1 系統受電） • 10 本の電源ケーブル（2 系統受電） • 2 本の電源ケーブル（三相）	• 10 本の電源ケーブル（1 系統受電） • 20 本の電源ケーブル（2 系統受電） • 4 本の電源ケーブル（三相）

3.1.8.3 2 系統受電

企業はデータセンター内のほとんどの要素を制御できますが、停電の発生を予見することは困難です。停電は、IT オペレーションにとって壊滅的な結果をもたらすことがあります。そのようなインシデントの影響を減らすことができるよう、SPARC Enterprise M シリーズは、2 系統受電が可能です。これらのサーバの AC 電源サブシステムは完全に二重化されていて、2 つの独立した外部 AC 電源から選択的に受電できます。2 系統受電と冗長電源の使用によって、システムの可用性が高まり、単一電力網の障害後でも、サーバの動作は影響を受けないままでいられます。

3.1.9 オペレーターパネル

SPARC Enterprise M シリーズは、サーバの状態を表示し、サーバ識別情報とユーザー設定情報を保存し、運用モードと保守モードを切り替え、ドメインの電源を投入するため、オペレーターパネルを採用しています（図 3-4 参照）。サーバの起動中、フロントパネルの LED ステータスインジケーターで、XSCF とサーバの動作が確認できます。

図 3-4. SPARC Enterprise M シリーズのオペレーターパネル

4. システムバスのアーキテクチャー - Jupiter インターコネクト

多数の CPU を持つハイエンドシステムがスケーラビリティを実現できるのは、すべてのプロセッサがアプリケーションのパフォーマンスに実際に貢献している場合のみです。幅広いアプリケーション群に対して、ほぼ直線的なスケーラビリティと高速で予測可能なパフォーマンスを提供できるかどうかは、主としてシステムバスの能力に依存します。SPARC Enterprise M シリーズは、コンポーネント間に莫大な帯域幅と一貫した低レイテンシをもたらすよう設計されたシステムインターコネクトを利用しています。Jupiter システムバスは、バランスがとれ、予測可能なパフォーマンスをアプリケーションのワークロードに提供することによって、IT オペレーションにメリットをもたらします。

4.1 インターコネクトアーキテクチャー

Jupiter インターコネクトの設計は、SPARC Enterprise M シリーズの全体的なパフォーマンスを最大化します。このシステムバスは、パケット交換技術を利用したポイントツーポイント接続として実装されていて、複数のデータストリームを転送することにより、応答時間を高速化します。パケット交換を利用すると、インターコネクトがバス上の「デッド」サイクルを排除して、システム全域にわたり、はるかに高スループットで機能できます。すべての経路は一方向の非競合パスで、各方向にアドレス、データ、コントロール、ECC が多重化されています。

SPARC Enterprise M4000/M5000/M8000/M9000 上のインターコネクトアーキテクチャーにあるシステムコントローラーでは、ローカル CPU、メモリ、I/O サブシステム、インターコネクトパスの間のトラフィックを方向付けます。SPARC Enterprise M8000/M9000 では、システムバスはシステムボード間のクロスバースイッチとして実装されていて、すべてのコンポーネント間でレイテンシ時間が一定の高スループットのデータ転送をサポートします。パフォーマンスを向上させるため、SPARC Enterprise M4000/M5000 のマザーボード上にあるメモリの物理アドレスや、SPARC Enterprise M8000/M9000 の CMU は、その同じボード上にあるすべてのシステムコントローラーにわたって均等に分散されています。

4.1.1 SPARC Enterprise M3000 のインターコネクトアーキテクチャー

SPARC Enterprise M3000 のシステムは、単一のマザーボードの中に含まれています。このアーキテクチャーにおいては、Jupiter システムコントローラー (JSC) という 1 つの ASIC が、メモリアドレスコントローラーおよびシステムコントローラーの両方の機能を実行します。JSC は、CPU、メモリ DIMM、I/O コントローラー (PCI Express ブリッジ) に接続しています。SPARC Enterprise M3000 のアーキテクチャー図を、図 4-1 に示します。

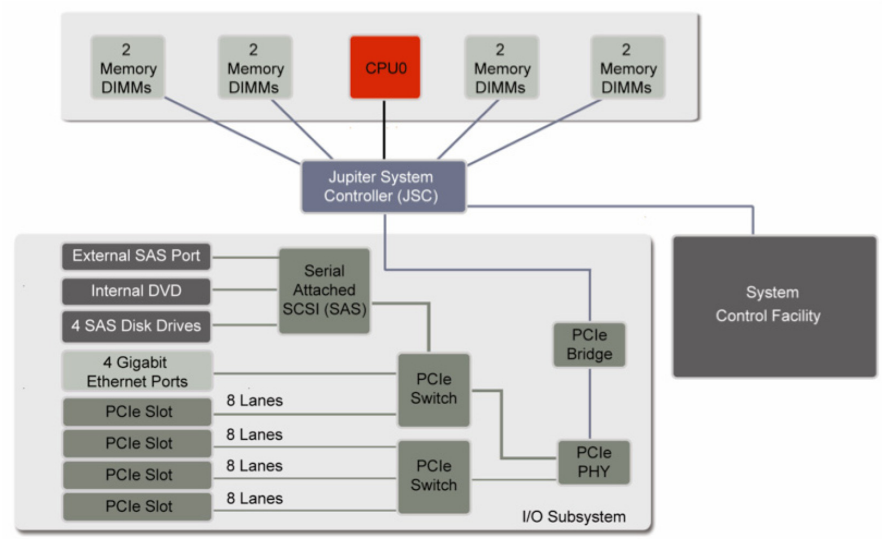

図 4-1. SPARC Enterprise M3000 のインターコネクト図

4.1.2 SPARC Enterprise M4000 のインターコネクトアーキテクチャー

SPARC Enterprise M4000 のシステムは、2 つのシステムコントローラーを備えた単一のマザーボード内に実装されています。2 つのシステムコントローラーは、CPU モジュール、メモリアクセスコントローラー、IOU とだけでなく、システムコントローラーどうしも接続しています（図 4-2 参照）。

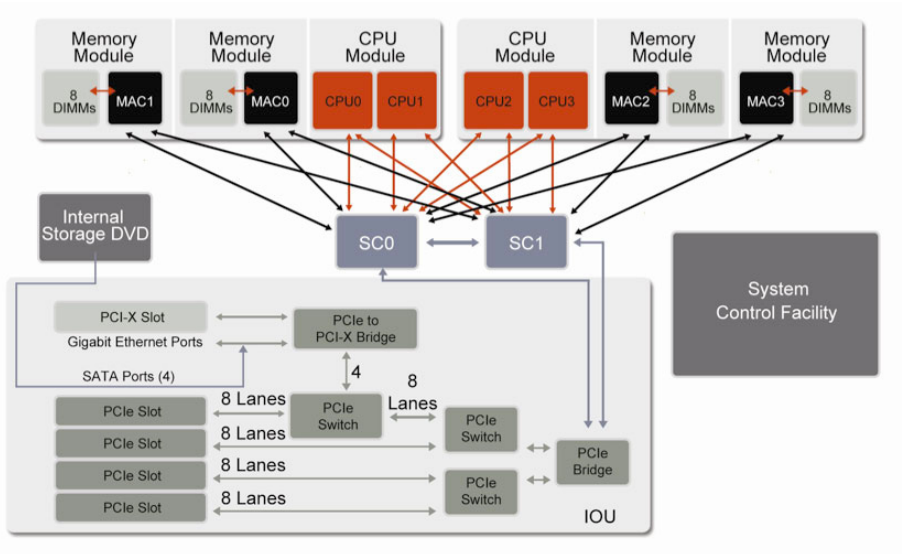

図 4-2. SPARC Enterprise M4000 のインターコネクト図

4.1.3 SPARC Enterprise M5000 のインターコネクトアーキテクチャー

SPARC Enterprise M5000 のシステムは、単一のマザーボード内に実装されていますが、2 つの論理システムボードを採用しています。SPARC Enterprise M4000 の設計と同様に、論理システムボードにはそれぞれ 2 つのシステムコントローラーがあり、CPU モジュール、メモリアクセスコントローラー、IOU とだけでなく、システムコントローラーどうしも接続しています。さらに、システムコントローラーはそれぞれ、もう一方の論理システムボード上の対応するシステムコントローラーに接続しています（図 4-3 参照）。

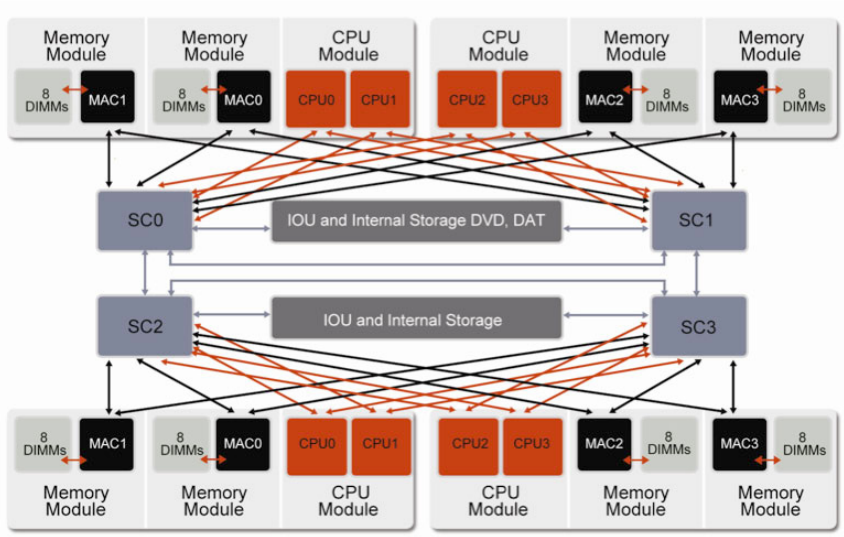

図 4-3. SPARC Enterprise M5000 のインターコネクト図

4.1.4 SPARC Enterprise M8000/M9000 のインターコネクトアーキテクチャー

SPARC Enterprise M8000/M9000 は、共通のクロスバーに接続する複数のシステムボードを採用しています。システムボードにはそれぞれ 4 つのシステムコントローラーがあります。各システムコントローラーは、すべての CPU モジュールに接続します。帯域幅を増やすため、各メモリアクセスコントローラーは 2 つのシステムコントローラーに接続し、各システムコントローラーはシステムボード内にある他のシステムコントローラーすべてに接続します。システムコントローラーは各クロスバーユニットにも接続し、他のシステムボードへのデータ転送を可能にしています（図 4-4 参照）。

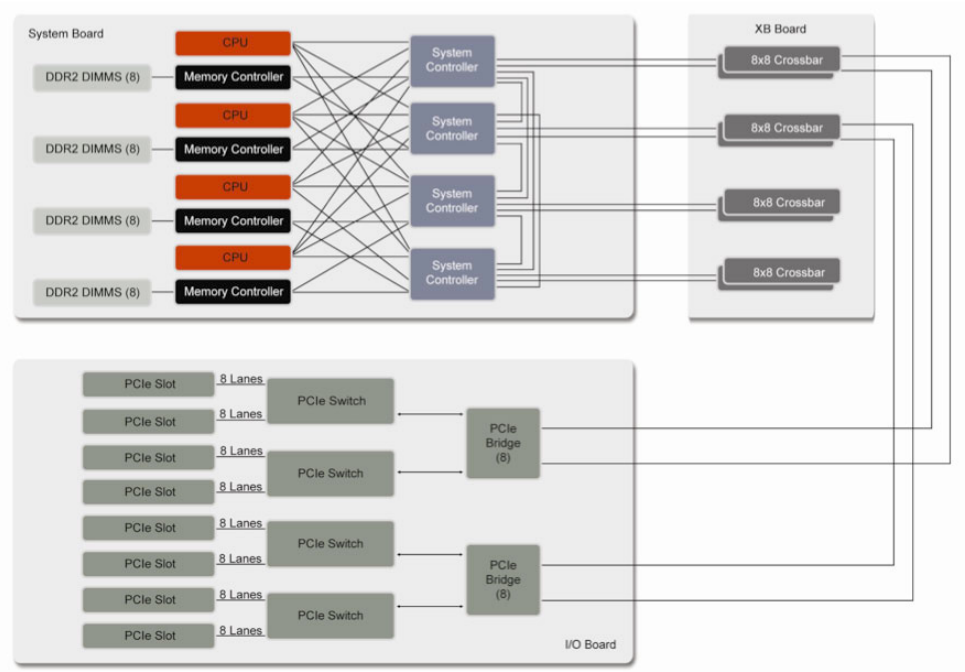

図 4-4. SPARC Enterprise M8000/M9000 のインターコネクト図

4.2 システムインターコネクトの信頼性機能

SPARC Enterprise M シリーズのシステムインターコネクトに組み込まれている冗長性機能と信頼性機能が、サーバの安定性を高めます。Jupiter インターコネクトは、アドレス・データとともに ECC で保護されているので、すべてのシステムバス上およびメモリ内の、データの損失または破損を防ぎます。CPU、メモリアクセスコントローラー、または I/O コントローラー内で 1 ビットデータエラーが検出されると、ハードウェアがデータを訂正し、転送を実行します。SPARC Enterprise M8000/M9000 では、特定のバスを縮退するか、もしくはクロスバースイッチを縮退するか、自動的に最適な選択を行うことができます。インターコネクト内のハードウェア障害はめったに起きませんが、もし障害が発生した場合には、システム再起動時に生き残っているバスを使用し、障害のあるクロスバーを切り離して、オペレーションを再開します。

4.3 スケーラブルなパフォーマンス

Jupiter システムインターコネクトの帯域幅の広さと設計全体が、SPARC Enterprise M シリーズのスケーラブルなパフォーマンスに貢献します。理論上のピーク時システムスループット、スヌープ帯域幅、I/O 帯域幅の数値およびストリームベンチマーク結果を表 4-1 に示します。

SPARC Enterprise M4000/M5000/M8000/M9000 では、CPU、メモリアクセスコントローラー、I/O コントローラーは、データ転送用の高速ブロードバンドス

イッチであるシステムコントローラーと直接接続しています。結果として、個々のコンポーネント間で比較的均等なレイテンシを維持できます。コンポーネントを追加しても、処理能力とレイテンシは悪化しません。実際に、SPARC Enterprise M8000/M9000 の場合、実装されているクロスバーインターコネクトによって、サーバにシステムボードを追加するたびに、インターコネクトの帯域幅は増加します。

表 4-1. SPARC Enterprise M シリーズの理論上のピーク時システム帯域幅および I/O 帯域幅、スヌープ帯域幅、ストリームベンチマーク結果

	理論上のピーク時システム帯域幅[a] (GB/ 秒)	スヌープ帯域幅 (GB/ 秒)	ストリームベンチマーク トライアド結果 (GB/ 秒)	ストリームベンチマーク コピー結果 (GB/ 秒)	理論上のピーク時 I/O 帯域幅[b] (GB/ 秒)
SPARC Enterprise M3000	20	N/A	5.1	6.4	4
SPARC Enterprise M4000	32	129	12.7	12.5	8
SPARC Enterprise M5000	64	129	25.2	24.8	16
SPARC Enterprise M8000	184	245	69.6	60.3	61
SPARC Enterprise M9000 (32 CPU 構成)	368	245	134.4	114.9	122
SPARC Enterprise M9000 (64 CPU 構成)	737	245	227.1	224.4	244

a. 理論上のピーク時システム帯域幅は、バスの幅と、システムコントローラーとメモリアクセスコントローラー間のバスの周波数を掛けて算出します。
b. 論理上のピーク時 I/O 帯域幅は、バスの幅と、システムコントローラーと PCI ブリッジ間のバスの周波数を掛けて算出します。

5. SPARC64 VI および SPARC64 VII/ SPARC64 VII+ プロセッサ

SPARC Enterprise M3000 は、富士通製の 4 コアまたは 2 コアの SPARC64 VII/ SPARC64 VII+ プロセッサをサポートしています。SPARC Enterprise M4000/ M5000/M8000/M9000 では、富士通製の 2 コアの SPARC64 VI プロセッサと 4 コアの SPARC64 VII/SPARC64 VII+ プロセッサをサポートしています。これらのプロセッサは革新的なマルチコアテクノロジーとマルチスレッドテクノロジーを採用しており、信頼性を高めるさまざまな機能が提供されます。また、どちらのプロセッサも SPARC V9 レベル 2 に準拠しているため、現存する数千種類ものソフトウェアアプリケーションに対するサポートの提供も容易です。SPARC Enterprise M シリーズで SPARC64 VI プロセッサおよび SPARC64 VII/ SPARC64 VII+ プロセッサを使用することにより、極めて高い信頼性、アプリケーション選択の幅、優れた処理能力が実現します。

5.1 次世代のプロセッサテクノロジー

この 10 年間に、プロセッサのアーキテクチャーには大きな変化がもたらされました。CPU の動作クロックが上昇するにつれて、パフォーマンスに対する効果は小さくなり、電力消費と発熱の問題も生じることをシステム設計の技術者が発見したためです。演算能力の向上計画では、チップマルチプロセッシング（CMP）、垂直マルチスレッディング（VMT）、同時マルチスレッディング（SMT）テクノロジーなどの革新的な手法を取り入れることが現在の主流となっています。

5.1.1 チップマルチプロセッシング（CMP）

チップマルチプロセッシングテクノロジー（CMT）は、複数の物理コアが単一のプロセッサモジュールに統合されたアーキテクチャーです。各物理コアは、どの時点でも、マルチスレッドアプリケーションの 1 つの実行スレッドをそれぞれ他のコアから独立して実行しています。このテクノロジーを利用すると、多くの場合、2 コアプロセッサのパフォーマンスは 1 コアモジュールに対して最大 2 倍になります。各クロックサイクルで複数の命令を処理できることがパフォーマンス上の大きな利点ですが、従来の CPU 間通信と比較してチップ間の距離が短く、バスが高速であることも性能向上の要因となっています。

5.1.2 垂直マルチスレッディング（VMT）

垂直マルチスレッディング（VMT）テクノロジーでは、単一の物理コアで複数のスレッドを処理することが可能になります。各コアは、オペレーティングシステムでは仮想 CPU として表示されます。同一コア上では、複数のスレッドが時分割方式で実行されます。いずれかの特定の時点に注目すると、常に 1 つのみが実行されています。スレッドは、アイドル状態の場合、またはキャッシュミスが発生してメインメモリのデータを待機している場合は実行されません。スレッドの切り替えは、L2 キャッシュのミス、ハードウェアタイマー、割り込み、特定のスレッド制御命令などのイベントでも発生します。VMT では、こうした方法でプロセッサの使用率を最大限まで高め、キャッシュミスによる影響を効率的に緩和することでシステムのパフォーマンスを改善します。システムのスレッド数がコアの数を超えると、VMT が自動的に有効になり、パフォーマンスが向上します。

5.1.3 同時マルチスレッディング（SMT）

同時マルチスレッディング（SMT）テクノロジーは、マルチスレッド対応のコアにおいて、複数のスレッドの同時実行をサポートするテクノロジーです。ソフトウェアの観点から見ると、各スレッドは独立しています。この方式によるマルチスレッド化は、コンピューティングリソースを二重化することで実現しています。命令バッファ、リザーベーションステーション、キャッシュなど、いくつかのリソースはスレッド間で共有されたままになります。SMT では、コンテキストスイッチングの時間が不要になるほか、単一コア内のスレッド間で命令パイプラインが円滑に共有されます。どちらのスレッドも実行待ち状態となっている場合は、スレッドがスーパースケーラー命令の発行サイクルを交互に実行し、必要に応じて機能ユニットを共有します。

5.2 SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサのアーキテクチャー

SPARC64 シリーズプロセッサが現在広く受け入れられている主な要因は、設計の全面刷新ではなく着実な機能強化と変更を継続したことにあります。その目標を達成するため、SPARC64 VI および SPARC64 VII/SPARC64 VII+ の CPU モジュールでは、パイプラインに大規模な変更を加えることなく、定評があり信頼性の高い SPARC64 V コアを再利用しています。SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサでは、スループットの向上を目指して、マルチコアのマルチスレッド対応アーキテクチャーを採用しています。SPARC64 VI プロセッサでは VMT テクノロジーを採用し、2 つのコアを内蔵してコアごとに 2 つのスレッドを提供します。SPARC 64 VII/SPARC64 VII+ プロセッサでは、コアごとに 2 つのスレッドを処理可能な 4 コアアーキテクチャーを採用し、SMT テクノロジーの利用によってパフォーマンスがさらに向上し

ます。表 5-1 に、SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサの仕様の詳細を示します。

表 5-1. SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサの仕様

	SPARC VI プロセッサ	SPARC VII プロセッサ	SPARC VII+ プロセッサ
速度	• 2.15 GHz • 2.28 GHz • 2.4 GHz	• 2.4 GHz • 2.52 GHz • 2.53 GHz • 2.75 GHz • 2.77 GHz • 2.88 GHz	• 2.66 GHz • 2.86 GHz • 3.0 GHz
アーキテクチャー	• 2 コア • SPARC V9 • sun4u • 90 nm プロセステクノロジー	• 4 コア／2 コア [a] • SPARC V9 • sun4u • 65 nm プロセステクノロジー	• 4 コア／2 コア [a] • SPARC V9 • sun4u • 65 nm プロセステクノロジー
L1 キャッシュ	• コアごとに 128 KB の L1 I キャッシュ • コアごとに 128 KB の L1 D キャッシュ	• コアごとに 64 KB の L1 I キャッシュ • コアごとに 64 KB の L1 D キャッシュ	• コアごとに 64 KB の L1 I キャッシュ • コアごとに 64 KB の L1 D キャッシュ
L2 キャッシュ	• 5 MB （2.15 GHz、2.28 GHz） • 6 MB（2.4 GHz） • 10 ウェイ連想型（2.15、2.28 GHz） • 12 ウェイ連想型（2.4 GHz） • 256 バイトラインサイズ • ECC タグおよびデータ保護	• 5 MB （2.4 GHz、2.52 GHz[a]、2.75 GHz、2.77 GHz） • 5.5 MB（2.53 GHz） • 6 MB（2.52 GHz[b] 、2.88 GHz） • 12 ウェイ連想型 • 256 バイトラインサイズ • ECC タグおよびデータ保護	• 5.5 MB （2.86 GHz） • 11 M（2.66 GHz） • 12 MB（3.0 GHz） • 12 ウェイ連想型 • 256 バイトラインサイズ • ECC タグおよびデータ保護
POWER	• 120 W（通常） • 150 W（最大）	• 135 W（通常） • 150 W（最大）	• 135 W（通常） • 160 W（最大）

a.SPARC Enterprise M3000
b.SPARC Enterprise M8000/M9000

5.2.1 キャッシュシステム

図 5-1 に示すように、SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサは 2 層構造のキャッシュメモリを実装しています。いずれのプロセッサも、中容量のプライマリキャッシュ（L1 キャッシュ）と大容量のセカンダリキャッシュ（L2 キャッシュ）を備え、L1 キャッシュは命令専用のキャッシュ（L1 I キャッシュ）とオペランド専用のキャッシュ（L1 D キャッシュ）で構成されます。L1 D キャッシュは 4 バイトアドレス境界によって 8

個のバンクに分割され、2 つのオペランドアクセスを同時に実行できます。この L1 キャッシュでは、キャッシュインデックスに仮想アドレスを利用し、キャッシュタグに物理アドレスを利用します。この方式は VIPT（virtually indexed physically tagged）と呼ばれるものです。VIPT のシノニム問題（同一の物理アドレスに 2 つの仮想アドレスのインデックスが付加される）に対処するため、SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサの L2 キャッシュでは、L1 キャッシュのシノニムエントリーの作成が抑止されます。

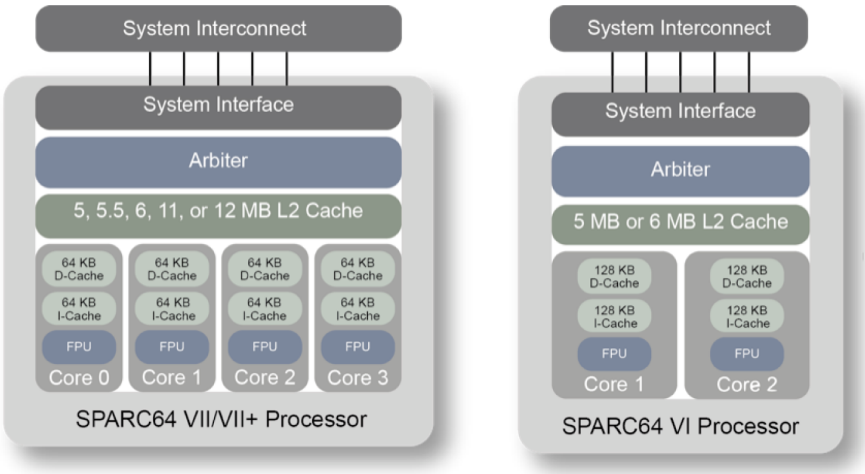

図 5-1. SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサに組み込まれたキャッシュ構造

SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサでは、L2 キャッシュがすべてのコアにわたって共有されます。L2 キャッシュから読み出したデータを L1 キャッシュに送信するバスでは、2 コアごとに 32 バイト幅が提供され、L1 D キャッシュから L2 キャッシュにデータを送信するバスではコアごとに 16 バイト幅が提供されます。

L1 キャッシュと L2 キャッシュのキャッシュ更新ポリシーは、いずれもライトバックです。キャッシュされたデータに対する更新内容は、システムメモリではなくキャッシュのみに書き込まれます。結果として、1 つの階層のキャッシュを更新するだけでストア操作を完了できます。キャッシュの位置にあるデータは、キャッシュラインが別のメモリ位置に再割り当てされた時点でシステムメモリに書き戻されます（ライトバックされます）。ストア操作の実行頻度が高い場合は、ライトバック方式を採用することでキャッシュ間のトラフィックとメモリアクセストラフィックが削減され、パフォーマンス上の利点を得られます。

ライトバック方式では、最新のデータが常にキャッシュに保持されます。したがって、1 つのキャッシュで障害が発生してもシステムの動作全体に影響するため、特別な対策が必要です。SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサは、プロセッサで障害が発生したおそれのある場合、エラーの状態を追跡して、キャッシュからシステムメモリへのライトバックを強制的に実行する高度な機能を備えています。これらの機能の詳細については、この章の *"RAS 機能（Reliability、Availability、Serviceability）"* を参照してください。

5.2.2 SPARC64 VI プロセッサのマイクロアーキテクチャーの概要

SPARC64 VI プロセッサの設計では、SPARC64 V プロセッサの高いパフォーマンスと信頼性を維持しながら、スループットの大幅な向上を実現しています。このパフォーマンス向上を達成することを目指して、SPARC64 VI プロセッサでは CMP テクノロジーと VMT テクノロジーを組み合わせて実装しています。このプロセッサは 2 つの物理コアで構成され、コアごとに 2 つの VMT スレッドをサポートしているため、4 つのスレッドを並列実行できます。オペレーティングシステムでは、各スレッドは仮想プロセッサとして表示されます。例えば、Oracle Solaris で psradm コマンドを使用すると、必要に応じてそれぞれの仮想 CPU をスペア、オンライン、またはオフラインに設定できます。

同一の物理コアに属する 2 つのスレッドは、ALU や命令パイプラインなど、コアのほとんどのリソースを共有します。一方、2 つの物理コアで共有されるのは L2 キャッシュとシステムインターフェースのみです。粗粒度マルチスレッディング技術が採用され、遅延の長いイベントが発生するまでは 1 つのスレッドがコアの全リソースを占有します。スレッドの切り替えが発生するのは、具体的には L2 キャッシュミスが発生した場合、または所定の時間が経過した場合です。この手法では、ブロックされていないスレッドをスケジューリングすることでキャッシュミスによる影響を緩和しながら、公平性も維持されるため、すべてのスレッドが円滑に処理を実行できます。

複数のスレッドをこの手法でサポートするには、汎用レジスタ（GPR）、浮動小数点レジスタ（FPR）、プログラムカウンター（PC）および制御（状態）レジスタを二重化する必要があります。カレントウィンドウレジスタ（CWR）と呼ばれる GPR のコピーによって、1 サイクルというレジスタファイルアクセス時間が実現しています。また、GPR から CWR までを高速かつ高帯域幅のパスで接続することにより、レジスタウィンドウの移動またはスレッドの切り替えが必要となった場合の処理が高速化されています。

CMP と VMT という革新的なテクノロジーによって、特にマルチスレッド処理のパフォーマンスを強化する一方、シングルスレッド処理のスループットも損

なわれていません。SPARC64 VI プロセッサによるシングルスレッド処理のパフォーマンスは、1.35 GHz で動作する SPARC64 V プロセッサの約 2 倍となっています。SPARC64 VI プロセッサでは、浮動小数点積和演算の速度が向上した改良型のコア、ミスの発生率を低減するために 2 倍の容量となったトランスレーションルックアサイドバッファ（TLB）など、その他の面も強化され、整数演算アプリケーションと浮動小数点演算アプリケーションのパフォーマンスがいずれも向上します。

5.2.3 SPARC64 VII/SPARC64 VII+ プロセッサのマイクロアーキテクチャーの概要

SPARC64 VII/SPARC64 VII+ プロセッサのコアパイプラインの基本構造は、SPARC64 VI プロセッサと同一です。ただし、SPARC64 VII/SPARC64 VII+ プロセッサでは、VMT テクノロジーではなく SMT テクノロジーを利用してマルチスレッディングを実装しています。図 5-2 に示すように、SPARC64 VI プロセッサでは VMT テクノロジーを利用して 2 つのスレッドを並列実行します。アクティブになるのは常に 1 つのスレッドのみです。

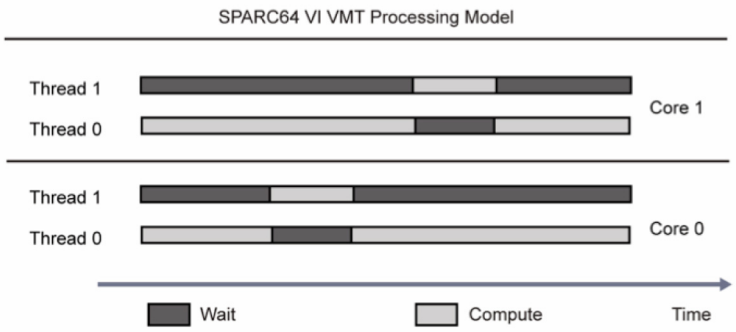

図 5-2. SPARC64 VI プロセッサで利用されている VMT 処理モデルでは、アクティブになるのはコアごとに常に 1 つのスレッドのみ

VMT モデルの場合、処理を他方のスレッドに切り替えるには、特定のトリガーが発生する必要があります。SPARC VII/SPARC64 VII+ プロセッサでは、SMT テクノロジーを実装したことで、各コアの両方のスレッドを同時に実行できるようになります（図 5-3）。このため、SPARC VII/SPARC64 VII+ プロセッサはスループットとパフォーマンスの大幅な向上を達成しうる潜在能力を秘めています。

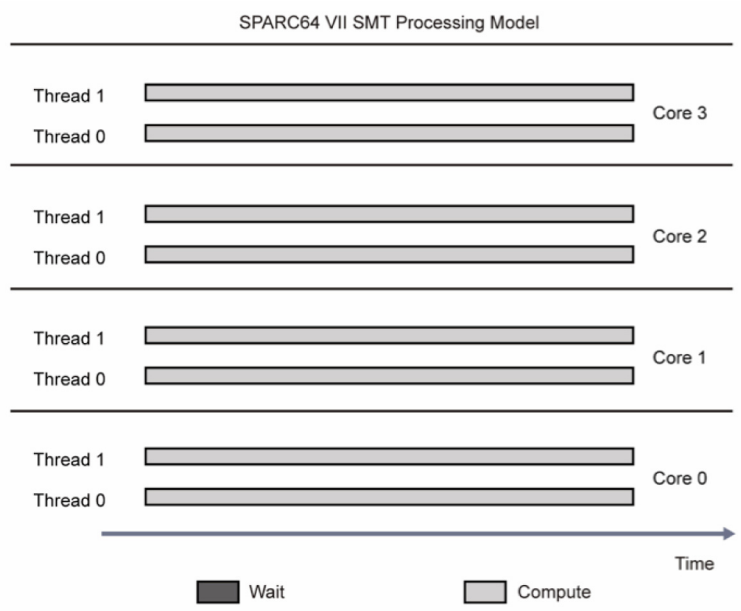

図 5-3. SPARC64 VII/SPARC64 VII+ プロセッサに実装された SMT テクノロジーではコア内の全スレッド同時実行が可能

SMT では 2 つのスレッドを同時に実行できますが、これらのスレッドは依然として同一のパイプラインコアを共有しています。SPARC64 VII/SPARC64 VII+ プロセッサは、スレッド間の干渉を最小限に抑えてパフォーマンスを最適化するように動作します。例えば、命令のフェッチステージ、デコードステージ、およびコミットステージで、いずれかのスレッドを各サイクル内で選択します。この手法は、実行中のスレッドの処理が停止スレッドによって阻害されることを避けるのに役立ちます。また、2 つのスレッドを単一のコア内で実行する一方で、各スレッドに個別に割り当てられるハードウェアリソースは分離された状態のままになります。さらに、1 つのスレッドがアイドル状態になったときは、実行中のスレッドが共有リソース全体を利用できます。これは、シングルスレッド処理のパフォーマンスを向上するための機能です。

SPARC64 VII/SPARC64 VII+ プロセッサは、次の高度な機能も提供します。[1] ：

- 整数積和命令
  整数積和命令は、符号なし 8 バイト整数値を保持する倍精度浮動小数点レジスタ内のデータに対して、乗算と加算の複合命令を実行するものです。積和演算の効率が上昇することで、一部アプリケーションでパフォーマンスが向上します。

1.SPARC64 VII/SPARC64 VII+ プロセッサの高度な機能を利用するには、特定のコンパイラーバージョン、オペレーティングシステムレベル、システム設定値が必要となる場合もあります。詳細については、システムのドキュメント（プロダクトノートおよびコンパイラーの関連ドキュメント）を参照してください。

- Shared Context

  Shared Context は、命令またはデータを格納するための仮想アドレス空間を、２つ以上のプロセスで共有することを可能にする機能です。この機能を利用すると、共有の TLB エントリーを複数のプロセスで使用して、一連の実行可能ファイル、親和型共有メモリ（ISM）セグメント、または動的親和型共有メモリ（DISM）セグメントのメモリマッピングにアクセスできるため、オーバーヘッドと TLB ミスが低減します。共有コンテキストを実装すると、仮想マシン（VM）、データベース、およびその他の並列処理アプリケーションのパフォーマンス向上に役立ちます。

5.2.4 SPARC64 VII/SPARC64 VII+ プロセッサの命令処理

以下にマイクロアーキテクチャーの詳細について述べます。

SPARC64 VII/SPARC64 VII+ プロセッサのコアは、図 5-4 に示すように、命令フェッチ部と命令実行部を含んでいます。命令フェッチ部には命令用の 1 次キャッシュ（L1 I キャッシュ）を含み、実行部にはオペランド用の 1 次キャッシュ（L1 D キャッシュ）を含んでいます。

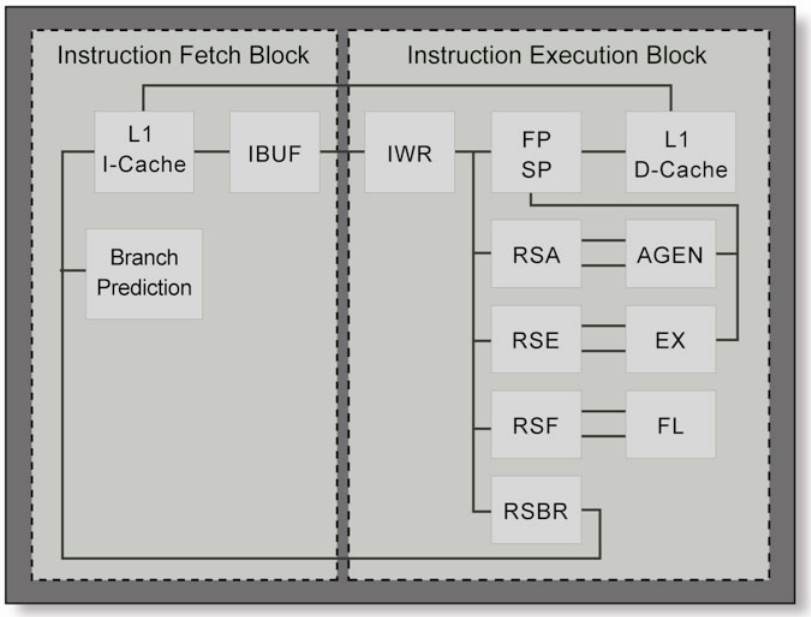

図 5-4. SPARC64 VII/SPARC64 VII+ プロセッサコアの機能図

5.2.5 命令フェッチ部

命令フェッチ部は、命令実行部とは独立して動作し、命令列を命令バッファ IBUF （Instruction BUFfer）に取り込みます。IBUF の容量は 256 バイトであり、最大 64 命令の保持が可能です。両スレッドが動作しているときは、命令バッファをスレッドごとに均等に分割して使用します。

IBUF 内に保持されている命令は、分岐予測に従って実行するための準備が整っています。命令実行が滞った場合にも、IBUF がいっぱいになるまで命令フェッチは停止しません。逆に、キャッシュ・ミス等で命令フェッチ操作が停止した場合にも、命令は IBUF からの取り出しを継続し、ソフトウェアの実行は、IBUF に命令がある限り、進められます。

命令フェッチは毎サイクル起動でき、一度に 8 命令を取り出します。重要なことは、命令実行のスループットはサイクル当たり最大 4 命令であるということです。2 倍のスループットで、命令フェッチと IBUF は、L1 I- キャッシュによって課せられるレイテンシを隠蔽し、システムパフォーマンスの向上に寄与することができます。

5.2.6 命令実行部

命令フェッチ部で IBUF に格納された命令は、Instruction Word Register (IWR) によって、サイクル当たり 4 命令ずつ検索されます。命令実行部内で、これらの命令は、デコード、発行、実行、コミットされます。

5.2.6.1 命令デコードおよび発行

命令デコードおよび発行ステージでは、IWR 上の 4 つの命令を同時にデコードして、各種リザベーション・ステーション、フェッチポート・ストアポート、レジスタ更新バッファなどの実行に必要な資源を決定します。必要なリソースが割り当て可能であれば、0 ～ 63 の範囲の命令識別子 (IID、Instruction Identification) が割り当てられ、命令が発行されます。この識別スキームにより、インフライト命令の最大数は 64 命令であり、また、両スレッド動作時は各スレッドの最大命令数は 32 命令です。同じサイクルにどちらかのスレッドからの命令デコードを行い、交互にスレッドを切り替えます。

命令を発行するとその現在の IWR を開放します。IWR の、いずれかのスロットのリザベーション・ステーションのような資源の割り当て制限はなく、また、命令種類の組み合わせの制限もありません。このため、資源の空きさえあれば命令を発行可能です。4 命令分の空きがない場合にも、プログラム順序の命令から可能な分だけを発行します。このように命令発行の停滞を極力排除した方式により、どのようなバイナリコードであっても常に高い並列度を確保しています。

5.2.6.2 命令実行

SPARC64 VII/SPARC64 VII+ プロセッサはデコードされた命令をリザベーションステーションに登録します。整数演算用リザベーション・ステーション (RSE) と浮動小数点演算用リザベーション・ステーション (RSF)、また、分岐命令用リザベーション・ステーション (RSBR) とロードストア命令のアドレス計算用リザベーション・ステーション (RSA) が、用意されています。RSE と

RSF は、演算器に対応して、それぞれ RSEA、RSEB、RSFA、RSFB の、2 つのキューに分かれています。リザベーション・ステーションに格納された各々の命令は、ソースオペランドが用意できたものから、個々のリザベーション・ステーションごとに順次対応する演算器に投入します。したがって 4 つの演算を同時に投入可能です。

一般に、リザベーション・ステーションにある命令のうち、投入準備が整っていて最も古い命令が選択されます。しかし、ロード命令がレジスタを更新するように設定され、レジスタがソースオペランドとして演算に使用される場合には、命令は、ロード命令の結果が得られる以前に投機的に投入されます。投機的投入の成功は、演算ステージ（EX）で判断されます。投機的投入より、キャッシュアクセスのパイプラインのレイテンシを隠蔽して演算器の使用効率を高めています。

5.2.6.3 命令コミット

同時にコミットできる命令数は最大 4 です。命令コミットステージは 2 つのスレッドで共用されており、サイクルごとにどちらか一方のスレッドを選択しコミット処理を行います。アウト・オブ・オーダーで実行した結果は、当初は GPR Update Buffer (GUB) および FPR Update Buffer (FUB) 作業レジスタに格納され、ソフトウェアからアクセスできない状態になっています。

プログラムの順序性を保証するために、GPR や FPR などのレジスタやメモリは、コミットステージでプログラム命令順序に従って（イン・オーダーで）更新します。さらに、PC 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します。これにより、正確な割り込みを保証し、いつでも実行中の処理のキャンセルが可能となっています。この同期一括更新方式の活用は、分岐予測ミスに関わるやり直しを簡単にし、信頼性の向上に寄与します。

5.3 RAS 機能（Reliability、Availability、Serviceability）

SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサモジュールの設計は、改良された障害回避およびエラー修正機能を実現することにより、システムの信頼性を強化させています。実際に、SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサ上の領域のかなりの部分を、CPU 内のエラー検出およびデータ修正のために使用しています。RAM ユニットは、ECC 保護されているか、または二重化されており、ほとんどのラッチおよび演算器はパリティ保護されています。SPARC64 プロセッサの強力な RAS 機能により、1 ビットエラーを確実に検出し、訂正または再実行します。エラーによっては、キャッシュをウェイ単位で縮退させたり、コア単位で縮退させたりします。
SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサのその他の信頼性機能には、エラーのマーキング、命令リトライ、予防保守のサポートが含まれます。具体的には、メモリの読み出しデータがマルチビットエラーを有する場

合、エラーの発生箇所を識別するための特別のマークがデータに書き込まれ、ECC シンドロームは、障害の正確なソースを識別するための貴重な情報を提供する、特別な値となります。また、ハードウェアエラーが検出されると、現在実行中の命令はキャンセルされ、一時的なエラーから回復するために自動的に再実行されます。SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサによって生成されたエラーデータは、サービスプロセッサにも送付され、予防保守をサポートします。

以上のように、SPARC64 VII/SPARC64 VII+ プロセッサは、メインフレームに匹敵する RAS を実現しています。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 の基本的な事項を徹底しています。これにより SPARC64 VII/SPARC64 VII+ プロセッサは、ミッションクリティカルな UNIX サーバのプロセッサとして、高度な信頼性、可用性、サービス性、データ保全性を提供しています。

5.3.1 内蔵 RAM の信頼性および可用性

SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは、高度なデータ完全性をサポートする信頼性および可用性の機能を提供します。表 5-2 は、SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサのエラー検出およびエラー訂正方法を明らかにしています。

表 5-2. 高いレベルのデータ完全性をサポートするエラー検出と訂正方法

種別		エラー検出方法 保護方法	エラー訂正方法
L1 命令キャッシュ	データ	パリティ	無効化および再読み込み
	タグ	パリティ＋二重化	二重化データの再書き込み
L1 データキャッシュ	データ	SECDED ECC	ECC による 1 ビットエラー訂正
	タグ	パリティ＋二重化	二重化データの再書き込み
L2 キャッシュ	データ	SECDED ECC	ECC による 1 ビットエラー訂正
	タグ	SECDED ECC	ECC による 1 ビットエラー訂正
命令 TLB		パリティ	無効化
データ TLB		パリティ	無効化
ブランチヒストリ		パリティ	分岐予測失敗からの回復

SECDED：Single Error Correction Double Error Detection

- L1 キャッシュ、L2 キャッシュ、TLB はウェイ単位での縮退が可能です。SPARC64 VII/SPARC64 VII+ プロセッサは、L1 キャッシュ、L2 キャッシュ、および TLB をウェイ単位に分割する、セットアソシアティブ方式を実装し

ているため、粒度の細かい縮退が可能です。エラーの発生回数を機能単位ごとにカウントし、単位時間当たりのエラー回数が上限値を越えると縮退して、それ以降はそのウェイを使用しないようにします。縮退はハードウェアで自動的に行われ、キャッシュに保存されたデータのコヒーレンシを継続する以下の処理が実行されます。L2 キャッシュに書き戻す動作。L1D キャッシュの縮退されるウェイのダーティラインを書き戻します。

- メモリに書き戻す動作。L2 キャッシュの縮退されるウェイのダーティラインを書き戻します。ウェイ縮退は、ソフトウェアにダメージを与えることなく実行され、ソフトウェアは処理速度の低下を除いて動作に影響を受けません。

5.3.2 内蔵レジスタと演算器の信頼性機能

信頼性をさらに高めるため、SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは、レジスタや演算器もエラー保護しています。これらの機能の要約を、表 5-3 に示します。

SPARC64 VII/SPARC64 VII+ プロセッサは、整数アーキテクチャーレジスタの ECC 保護を採り入れています。エラー発生時には ECC 回路でエラーを訂正します。浮動小数点アーキテクチャーレジスタや他のレジスタは、パリティビットによって保護されています。また、演算器の中にパリティ予測回路、剰余チェック回路などが実装され、出力結果へとパリティを伝播させます。万一パリティエラーを検出した場合は、次に述べるとおりハードウェアが自動的に命令を再実行し回復を試みます。

表 5-3. SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサにおける、内蔵レジスタと演算器のエラー検出およびデータ保護

種別		エラー検出方法 保護方法
レジスタ	整数レジスタ	パリティないし SECDED ECC
	浮動小数点レジスタ	パリティ
	PC、PSTATE	パリティ
	演算入出力レジスタ	パリティ
演算器	加減算、除算、シフト、グラフィックス演算	パリティ予測
	乗算	パリティ予測 + 剰余チェック

5.3.3 同期一括更新方式と命令リトライ

SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは同期一括更新方式を採用しています。エラーが検出されると、その時点で実行途中の処理をすべてキャンセルします。コミットまでの間の途中結果は廃棄可能であり、エラーに遭遇することなく完了した命令が更新した結果だけがプログラマブルな資源に残ります。したがって、エラーによるプログラマブルな資源の破壊を防止す

るばかりではなく、エラー検出後にハードウェアで命令リトライを行うことが可能となります。滞っている処理をいったん捨てて、はじめからやり直すことができるので、一時的なハングアップの場合においても、回復の可能性があります。

図 4-5 に示すように、命令リトライはエラーをトリガーとして、自動的に起動します。リトライ中は、1 命令ずつ実行することで正常実行の可能性を上げ、正常に実行完了すると、通常の実行状態に自動的に復帰します。この間ソフトウェアに介入は不要であり、命令リトライが成功すれば、エラーのソフトウェアへの影響はありません。命令リトライは、閾値に達するまで繰り返されます。閾値を超えた場合には、プロセッサはエラーの内容を記録し、異常が発生したことをオペレーティングシステムに通知して、以降の処理をオペレーティングシステムに依頼します。

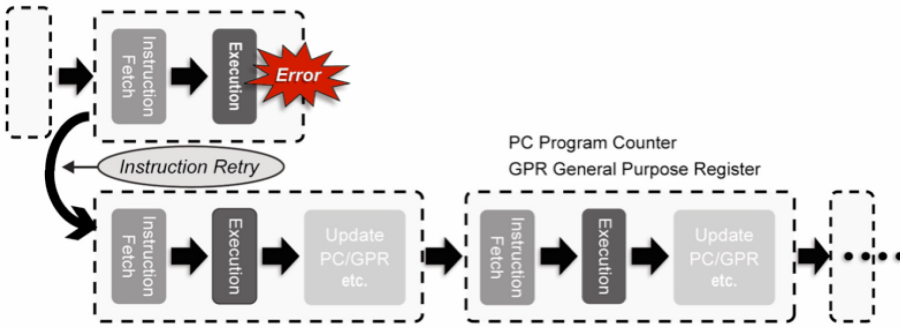

図 5-5. SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは、可用性向上のため自動命令リトライ処理を実装する

5.3.4 サービス性の向上

SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは、さまざまなエラーのチェッカを装備しています。これらのプロセッサはエラーをモニターし、エラー発生時にはその情報を XSCF (eXtended System Control Facility) に送信します。この通知により、XSCF は、エラーログを収集し、解析します。SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサのエラー通知機能を活用することにより、システムは、運用を継続しながら故障箇所と故障種別を迅速かつ正確に特定することができます。結果として、システムは、予防保守に有用な情報を提供し、サービス性を向上させます。

6. I/O サブシステム

ビジネスオペレーションのすべての面で増大するコンピュータシステムへの依存は、かつてないほど大量の情報を保存し、処理する必要性を伴っています。このような大量のデータセットを効果的に移動し、操作するには、強力な I/O サブシステムが不可欠です。SPARC Enterprise M シリーズは非常に優れた I/O の拡張とパフォーマンスによって、システムを拡張し、新たなデータ保存ニーズにすばやく対応できます。

6.1 I/O サブシステムのアーキテクチャー

SPARC Enterprise M シリーズの I/O サブシステムのパフォーマンスの鍵となるのは、PCI 技術の使用です。PCI Express ブリッジが、メインシステムと、PCI-X スロット、PCI Express スロット、内蔵ドライブなど I/O ユニットのコンポーネントを接続します。PCI Express バスも、内蔵 PCI スロットを使用したり PCI ボックスを接続したりすることによって、外部 I/O デバイスをサポートします。

PCI Express および PCI-X アダプターカードのホットプラグを容易にするため、SPARC Enterprise M4000/M5000/M8000/M9000 は、PCI カセットを利用しています。PCI カードは PCI Hot Plug をサポートしています。管理者は、PCI カードを PCI カセットに搭載し、稼働中のサーバの内蔵 PCI スロットまたは PCI ボックスに挿入できます。

6.1.1 SPARC Enterprise M3000 の I/O サブシステム

SPARC Enterprise M3000 の I/O サブシステムを、図 6-1 に示します。SPARC Enterprise M3000 のマザーボード上に実装される 1 つの PCI Express ブリッジが、すべての I/O コンポーネントをシステムコントローラーに接続します。I/O サブシステムは、4 つの PCI Express スロットおよび 1 つの外部 SAS ポートを提供することにより、外部 I/O デバイスの追加をサポートします。外部 SAS ポートは、SAS テープデバイスまたはストレージデバイスを接続するために使用することができます。

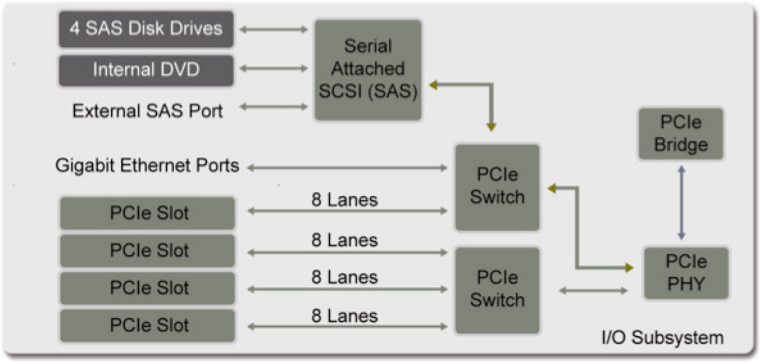

図 6-1. SPARC Enterprise M3000 の I/O サブシステムのアーキテクチャー

SPARC Enterprise M3000 のオンボード SAS コントローラーは Oracle Solaris OS の *raidctl* ユーティリティを使用した RAID 1（ミラーリング）をサポートします。

6.1.2 SPARC Enterprise M4000/M5000 の I/O サブシステム

SPARC Enterprise M4000 は 1 つの IOU をサポートしていますが、SPARC Enterprise M5000 は 2 つの IOU をサポートしています。1 つの PCI Express ブリッジが各 IOU をシステムコントローラーに接続し、PCI Express から PCI-X へのブリッジは SPARC Enterprise M4000/M5000 上のオンボード PCI-X スロットの収容をサポートします。SPARC Enterprise M4000 の 1 つの IOU には、4 つの PCI Express スロットと 1 つの PCI-X スロットがあります（図 6-2 参照）。SPARC Enterprise M5000 の 2 つの IOU には、全部で 8 つの PCI Express スロットと 2 つの PCI-X スロットがあります（図 6-3 参照）。さらに、PCI ボックスを使用すると、SPARC Enterprise M4000/M5000 上で使用できる PCI スロットの数が増加します。

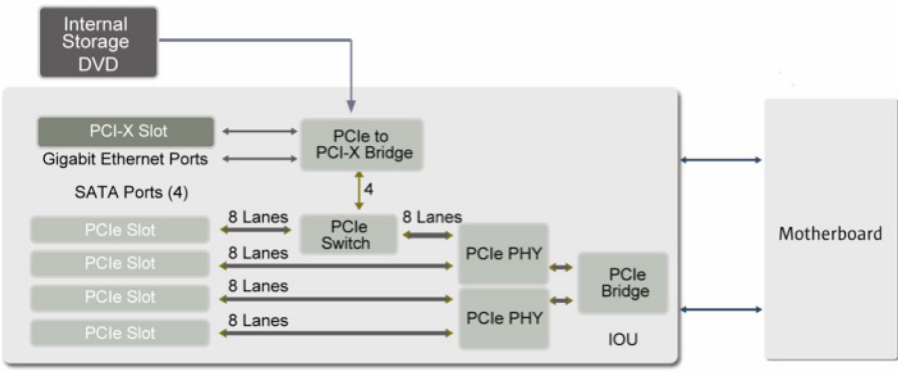

図 6-2. SPARC Enterprise M4000 の I/O サブシステムのアーキテクチャー

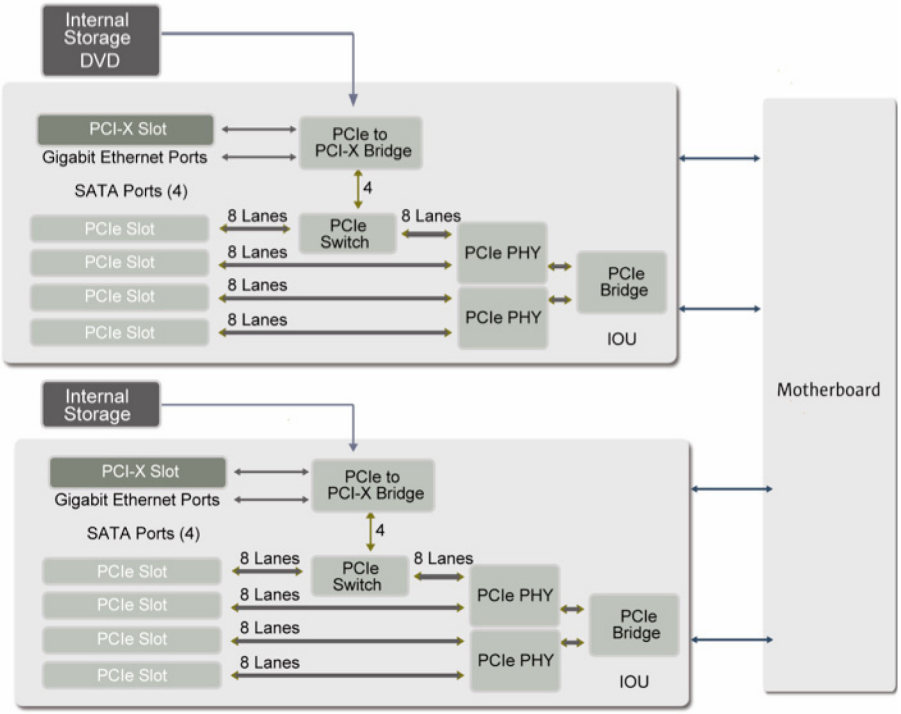

図 6-3. SPARC Enterprise M5000 の I/O サブシステムのアーキテクチャー

6.1.3 SPARC Enterprise M8000/M9000 の I/O サブシステム

SPARC Enterprise M8000/M9000 では 1 枚のシステムボードと 1 台の IOU を組み合わせて使用することができます。2 つの PCI Express ブリッジが、各システムボード上の IOU とクロスバースイッチを接続します。各 PCI Express ブリッジは、システムボード上の 4 つの PCI Express スロットとの通信も制御します（図 6-4 参照）。

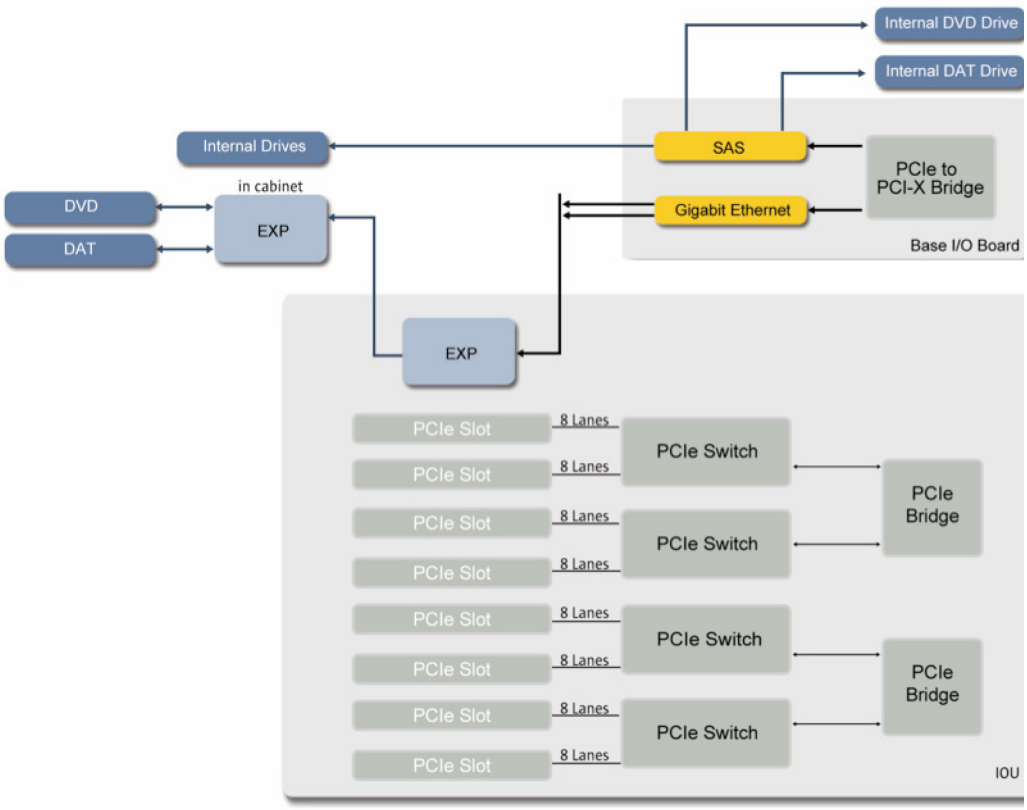

図 6-4. SPARC Enterprise M8000/M9000 の I/O サブシステムのアーキテクチャー

SPARC Enterprise M8000/M9000 の IOU には、8 つの PCI Express スロットがあります。これらのサーバの PCI スロットの総数は、実装しているシステムボードの数に依存します。SPARC Enterprise M8000/M9000 の内蔵 PCI Express スロットの最大数を表 6-1 に一覧表示します。さらに、PCI ボックスを SPARC Enterprise M8000/M9000 に追加して、使用できる PCI スロットの総数を増やすことができます。

表 6-1. SPARC Enterprise M8000/M9000 の内蔵 PCI スロット数

サーバ	内蔵 PCI Express スロットの最大数
SPARC Enterprise M8000	32
SPARC Enterprise M9000（32 CPU 構成）	64
SPARC Enterprise M9000（64 CPU 構成）	128

6.2 内蔵 I/O

SPARC Enterprise M3000 は、1 つの内蔵 DVD ドライブと 4 つの内蔵 Serial Attached SCSI（SAS）2.5 インチハードディスクドライブ、および 2.5 インチ solid state drive（SSD）をサポートしています。SPARC Enterprise M3000 はまた、1 つの外部 SAS ポートをサポートしており、SAS テープデバイスまたはストレージデバイスを接続できます。SAS ポートは 2 レーン構成で、合計最大 6 Gb/ 秒の帯域幅をサポートしています。

ディスクデバイスとテープデバイスは SPARC Enterprise M4000/M5000 に直接組み込まれています。SPARC Enterprise M8000/M9000 では、アドオンのベース I/O カードを使用することで内蔵デバイスにアクセスできます。SPARC Enterprise M4000/M5000/M8000/M9000（基本筐体）は、1 つの内蔵 DVD ドライブと 1 つのオプション内蔵 DAT テープドライブをサポートしています。拡張筐体を持つ SPARC Enterprise M9000 は、2 つの内蔵 DVD ドライブと、2 つのオプション内蔵 DAT テープドライブをサポートしています。SPARC Enterprise M4000/M5000 は、複数の内蔵 SAS 2.5 インチハードディスクドライブをサポートしています。SPARC Enterprise M8000/M9000 はそれぞれ、複数の内蔵 SAS 2.5 インチハードディスクドライブ（または 2.5 インチ SSD）をサポートしています。

6.3 PCI ボックス

SPARC Enterprise M4000/M5000/M8000/M9000 は、I/O 接続を追加するため、オプションの PCI ボックスの接続をサポートしています。PCI ボックスは 4 RU ラックにマウント可能なデバイスで、最大 2 つの IOU と 6 つの PCI Express スロットまたは PCI-X スロットを収容します。PCI カセットを使用することにより、外部 I/O シャーシは、ホットプラグ PCI カードの活性交換をサポートします。

ホストに実装された I/O リンクカードは、SPARC Enterprise M4000/M5000/M8000/M9000 の PCI ボックスへの接続を提供し、ホストから管理、制御をサポートします。I/O リンクカードは、ローハイトのカッパーカードまたはフルハイトのファイバーカードとして使用でき、4 GB/ 秒の帯域幅の 8 レーンの PCI Express バスが 1 つあります。SPARC Enterprise M4000/M5000/M8000/M9000 の増設ユニットのアーキテクチャーは、高スループットの I/O パフォーマンスを実現し、多くの種類の PCI Express カードの最大データ速度と、追加 PCI Express カードからのバーストトラフィックをサポートします（図 6-5 参照）。

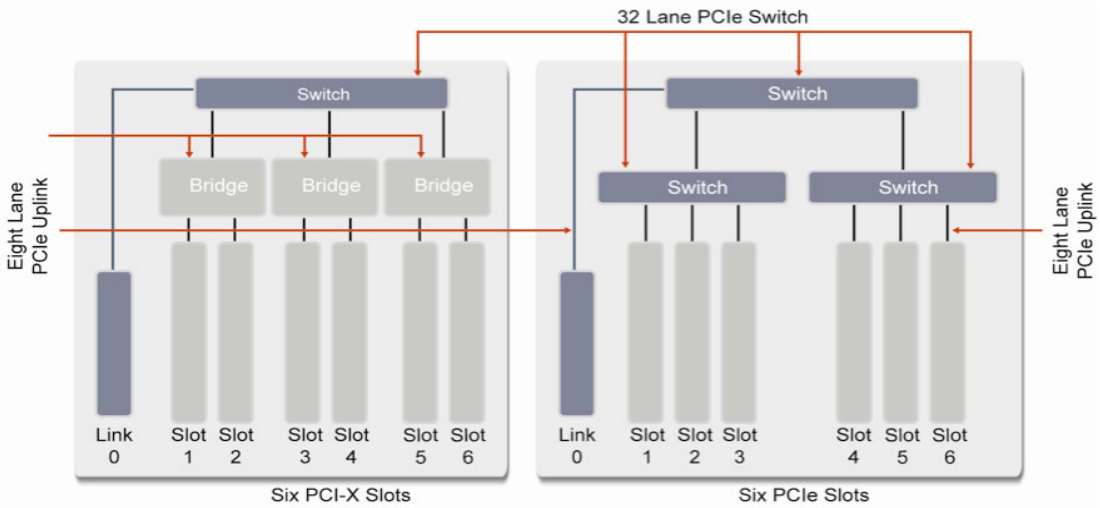

図 6-5. PCI ボックスのアーキテクチャー図

PCI ボックスは、I/O リンクカードを内蔵 PCI Express スロットに挿入し、ケーブルを使用して I/O リンクカードを接続することによって、SPARC Enterprise M4000/M5000/M8000/M9000 に追加します。I/O リンクカードは、ローハイトのカッパーリンクカードキットか、フルハイトのファイバーリンクカードキットを選択します。SPARC Enterprise M4000/M5000/M8000/M9000 は、表 6-2 に示すように、複数の PCI ボックスの接続をサポートします。

表 6-2. PCI ボックスを使用して広範な拡張をサポートする SPARC Enterprise M4000/M5000/M8000/M9000

サーバ	PCI ボックスの最大数	PCI スロットの最大数
SPARC Enterprise M4000	2	25
SPARC Enterprise M5000	4	50
SPARC Enterprise M8000	8	112
SPARC Enterprise M9000（32 CPU 構成）	16	224
SPARC Enterprise M9000（64 CPU 構成）	16	288

PCI ボックスの管理を容易にするため、SPARC Enterprise M4000/M5000/M8000/M9000 の XSCF ファームウェアに以下の管理機能が含まれています。XSCF のコマンドにより、これらの管理情報にアクセスすることができます。

- PCI Express スロットの電源投入時に PCI ボックスと FRU を検出
- 環境情報、電圧情報、ステータス情報を収集
- PCI ボックスのエラーデータを記録

7. 信頼性、可用性、保守性

計画的および計画外のダウンタイムの短縮は、IT サービスにとって非常に重要です。システム設計においては、主要サービスの可用性に影響を及ぼすことなく、障害耐性、迅速な修理、さらに短時間で拡張できるような仕組みが盛り込まれている必要があります。特に、複雑なネットワークコンピューティングソリューションと厳しい高可用性要件をサポートするよう設計された SPARC Enterprise M シリーズには、冗長ホットスワップシステムコンポーネント、設計全体にわたる診断機能およびエラー復旧機能、組み込みリモート管理機能があります。このような信頼性あるサーバの高度なアーキテクチャーを使用すると、アプリケーションの可用性、多くのタイプのハードウェア障害からの迅速な復旧、システムオペレーションの単純化、企業のコスト低減を高いレベルで実現できます。

7.1 冗長ホットスワップコンポーネント

今日の IT 組織は、ノンストップで動き続けるビジネスオペレーションへの対応に苦慮しています。ネットワーク化されたグローバル経済では、昼夜を問わず収益機会を得ることができるため、計画的なダウンタイムを短縮し、場合によってはまったくなくさざるを得ません。このような要望に応えるため、SPARC Enterprise M シリーズは組み込み型の冗長ホットスワップハードウェアを採用して、個々のコンポーネントの障害またはシステム構成の変更による中断を緩和できるようにします。実際に、これらのシステムは、ほとんどの場合はユーザーやシステム機能に影響を及ぼすことなく、ハードウェア障害から復旧できます。

SPARC Enterprise M シリーズは、電源ユニットとファンユニットの冗長構成と活性交換を採用しています。また、SPARC Enterprise M4000/M5000/M8000/M9000 では PCI ホットプラグによる I/O カードの活性交換が可能です。さらに、SPARC Enterprise M8000/M9000 では、複数の CPU やメモリを構成するためのオプションとしても利用できる、CMU と IOU の活性交換が可能です。管理者は、SPARC Enterprise M シリーズのホットスワップディスクドライブと、ディスクミラーリングソフトウェアを組み合わせて、冗長な内蔵ストレージを作成できます。ハイエンドサーバには冗長ホットスワップサービスプロセッサもサポートし、SPARC Enterprise M9000 には、縮退可能なクロスバーユニットと冗長なクロック制御ユニットがあります。障害が発生した場合、二重化されているこれらのコンポーネントによって、オペレーションの継続をサポートします。コンポーネントとエラーのタイプによって、システムは縮退モードで動

作を継続することもあれば、障害を自動的に診断し、関連コンポーネントをシステム構成から自動的に切り離して、リブートすることもあります。さらに、SPARC Enterprise M シリーズ内のホットスワップハードウェアは、サービスを高速化し、システムを停止する必要なく、コンポーネントの交換または増設を単純化できます。

7.2 パーティショニング機能

コストと管理上の負荷を軽減するため、多くの企業はサーバの統合への関心を高めています。しかし、複数のアプリケーションを単一サーバ上にホスティングする安全性と有効性を高めるツールが必要です。SPARC Enterprise M4000/M5000/M8000/M9000 でパーティショニング機能（Dynamic Domains）を利用すると、IT 企業は単一の大規模システムを複数の障害分離型サーバに分割しでき、それぞれが Oracle Solaris オペレーティングシステムの個別インスタンスを動作させます。適切に構成すれば、1 つのドメイン（パーティション）内のハードウェア障害またはソフトウェア障害を分離し、他のドメインの動作に影響を及ぼさないようにできます。単一サーバプラットフォーム内の各ドメインは、異なるバージョンの Oracle Solaris OS でも動作させることができるため、この技術は新規アプリケーションまたは変更したアプリケーションの実稼働前テストに非常に有効です。サーバごとのパーティショニング機能の最大数を表 7-1 に示します。

表 7-1. SPARC Enterprise M4000/M5000/M8000/M9000 のパーティショニング機能の制限

サーバ	ドメインの最大数
SPARC Enterprise M4000	2
SPARC Enterprise M5000	4
SPARC Enterprise M8000	16
SPARC Enterprise M9000	24

7.2.1 拡張システムボード（XSB）

パーティショニング機能は、非常に効果的な統合ツールで、リソースを理想的に分離できます。このハイレベルな分離を実現するため、富士通が開発した前世代のサーバは、システムボード全体を、1 つのドメインに割り当て可能な最小単位と指定していました。このようなハイレベルな分離に恩恵を受ける企業がある一方、現在のワークロードにより正確に一致するコンピューティングパワーを備えた、より数多くのドメインを作成する能力から恩恵を受けることができる企業もあります。これらのニーズの満たすため、SPARC Enterprise M4000/M5000/M8000/M9000 は、拡張システムボード（XSB）を提供しています。

物理的なシステムボードを使用するには、ボード上のハードウェアリソースを分割し、拡張システムボードとして再構成して、パーティショニング機能に割り当てます。拡張システムボードには、2 つの種類があります。システムボード全体で構成される XSB は、Uni-XSB と呼ばれます。もう 1 つは、論理的に 4 つに分割されているシステムボードまたはマザーボードで、Quad-XSB と呼ばれます。次の図は、SPARC Enterprise M4000/M5000/M8000/M9000 の論理的な分割線を表したものです（図 7-1、図 7-2、図 7-3、図 7-4、図 7-5 参照）。

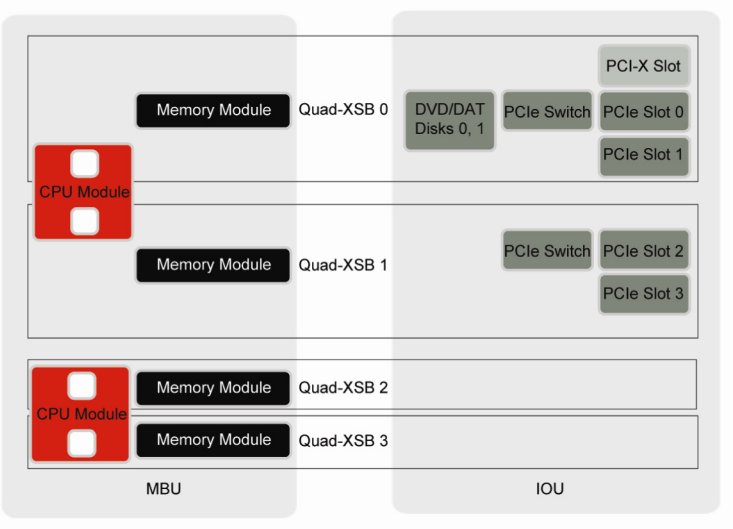

図 7-1. SPARC Enterprise M4000 の Quad-XSB 構成

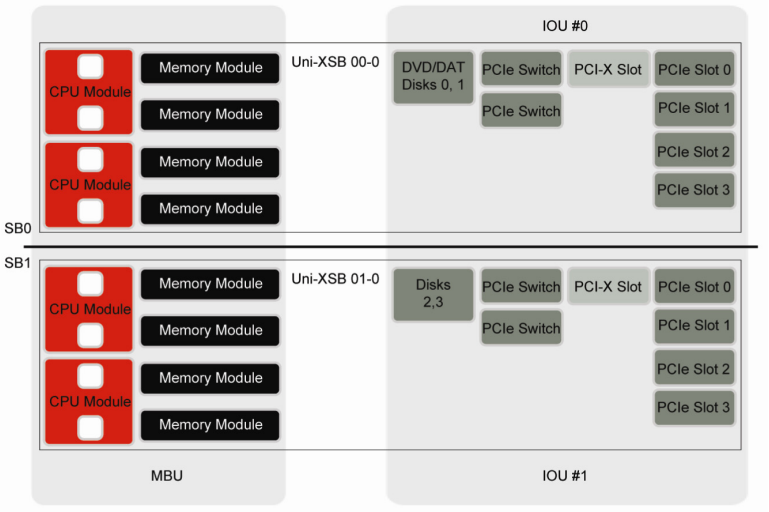

図 7-2. SPARC Enterprise M5000 の Uni-XSB 構成

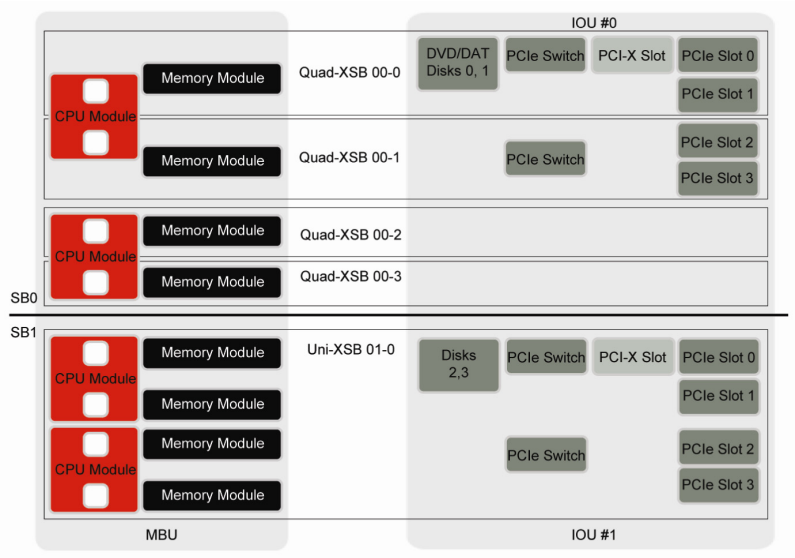

図 7-3. SPARC Enterprise M5000 の Uni-XSB および Quad-XSB 構成

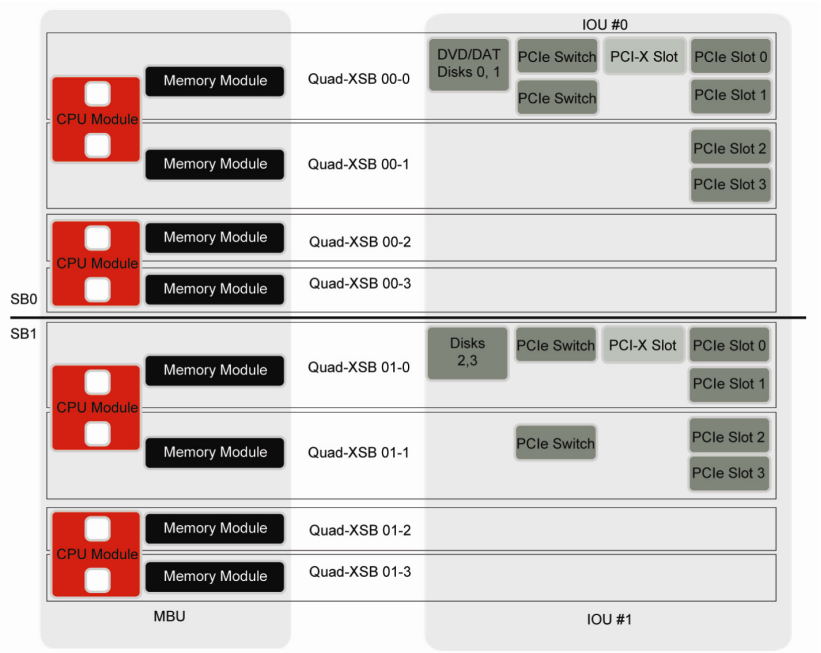

図 7-4. SPARC Enterprise M5000 の Quad-XSB 構成

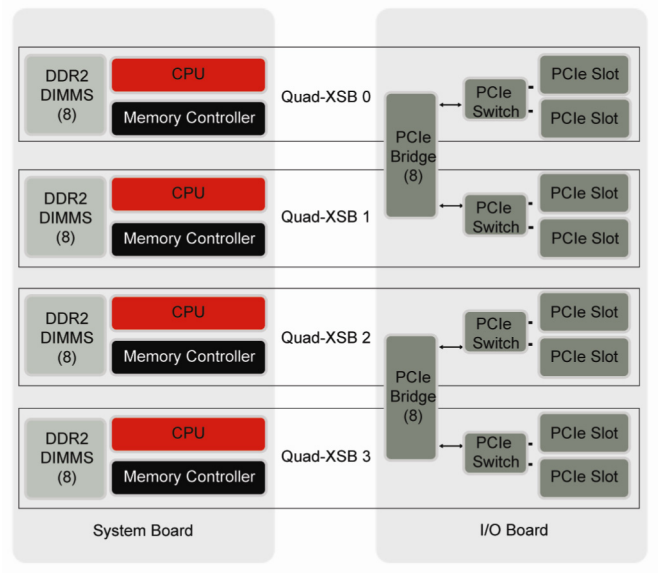

図 7-5. SPARC Enterprise M8000 および SPARC Enterprise M9000 のシステムボードの Quad-XSB 構成

拡張システムボードを使用すると、リソースを粒度の細かいサブシステムボードとしてパーティショニング機能に割り当てることができます。パーティショニング機能は、Uni-XSB と Quad-XSB のあらゆる組み合わせで構成でき、企業が高度な資産割り当てを実行できるようにします。1 つのドメインに含める XSB の正確な数とタイプを決定するには、障害分離の必要性とリソース使用率を最大化するという要望のバランスをとる必要があります。XSB に加え、I/O ユニットに接続されている DVD デバイスと DAT デバイスも、パーティショニング機能に割り当てることができます。パーティショニング機能と XSB を使用すると、分離されたセキュアなデータとプログラムを顧客に提供し続けながら、ハードウェアリソースの使用をさらに最適化できます。

7.2.2 CPU の混在

SPARC Enterprise M4000/M5000/M8000/M9000 構成において SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサの混在をサポートすることは、一層の投資保護を提供するとともに、ソリューションの柔軟性をさらに高めます。図 7-6 に示すように、SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサは、物理システムボード、個々のパーティションに、混在させて搭載することができます。

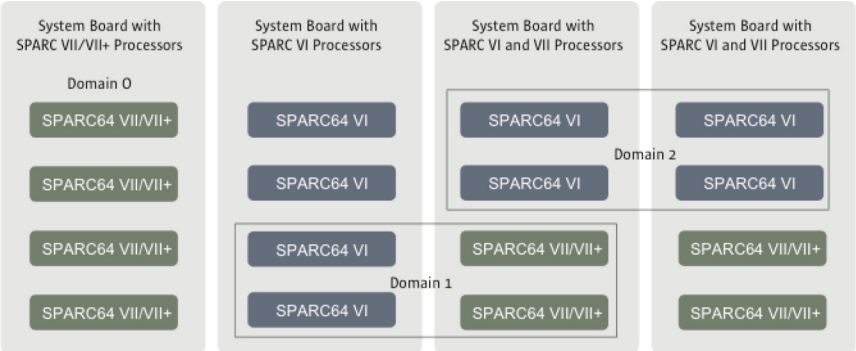

図 7-6. システム構成内の SPARC64 VI および SPARC64 VII/SPARC64 VII+ プロセッサの混在をサポートする SPARC Enterprise M4000/M5000/M8000/M9000

SPARC Enterprise M4000/M5000 用の SPARC64 VII+ プロセッサには 11 MB の L2 キャッシュ、SPARC Enterprise M8000/M9000 用の SPARC64 VII+ プロセッサには 12 MB の L2 キャッシュが搭載されています。この L2 キャッシュ容量を実現するには、以下 2 つの条件が満たされなければなりません。

1. システムボード上の 4 つのプロセッサはすべて SPARC64 VII+ プロセッサであること。4 つのうち 1 つでも SPARC64 VI または SPARC64 VII プロセッサであってはなりません。
2. SPARC Enterprise M4000/M5000 の場合マザーボードは、MBU_B 以降、SPARC Enterprise M8000/M9000 の場合 CMU が CMU_C 以降であること。

新しい MBU または CMU を必要とする理由は、旧版上の SC ASIC では、最大 6 MB の L2 キャッシュにしかアクセスできないためです。新版のボードにはアップデートされた SC ASIC が搭載されており、最大 12 MB の L2 キャッシュにアクセスすることができます。SPARC64 VII+ プロセッサは、古い版数のボードに搭載可能ですが、その L2 キャッシュは半減します (SPARC Enterprise M4000/M5000 上の L2 キャッシュの 5.5 MB、SPARC Enterprise M8000/M9000 上の L2 キャッシュの 6 MB)。もし、SPARC64 VI または VII プロセッサが、2 つの SPARC64 VII+ プロセッサを搭載済みの新しい MBU_B または CMU_C に搭載されると、SPARC64 VII+ プロセッサの L2 キャッシュは半減します (SPARC Enterprise M4000/M5000 上の L2 キャッシュの 5.5 MB、SPARC Enterprise M8000/M9000 上の L2 キャッシュの 6 MB)。これらのルールは、プロセッサの物理的な搭載にのみ適用され、ドメインに対する XSB の論理的な割り当てには適用されません。

7.3 Dynamic Reconfiguration（DR 機能）

Dynamic Reconfiguration（DR 機能）は、管理者が SPARC Enterprise M4000/M5000/M8000/M9000 をオフラインにすることなくリソースをシフトできるようにして、パーティショニング機能に付加価値をもたらします。この技術を使用すると、サーバがアプリケーションの実行を継続しつつ、管理者による保守、ライブアップグレード、システムハードウェアリソースへの物理的変更に役立ちます。Dynamic Reconfiguration によって、重要なシステムを中断することなく、ハードウェア構成を複数同時に変更することもできます。

動作しているシステムから CPU、メモリ、I/O サブシステムなどのコンポーネントを減設または増設する機能によって、システムのダウンタイムを短縮できます。Dynamic Reconfiguration を使用すると、ハードウェア構成の変更後にシステムをリブートする必要がなくなるため、保守とアップグレードが簡単になります。

7.4 高度な信頼性機能

SPARC Enterprise M4000/M5000/M8000/M9000 のコンポーネントにある高度な信頼性機能が、このプラットフォーム全体の安定性を高めます。例えば、SPARC Enterprise M4000/M5000/M8000/M9000 には複数のシステムコントローラーがあり、SPARC Enterprise M8000/M9000 には縮退可能なクロスバースイッチがあって、システムバス内に冗長性をもたらします。サーバアーキテクチャー内のコンポーネントの数と複雑さを減少させることは、信頼性に貢献します。さらに、高度な CPU 統合と、保証されているデータパスの完全性が、SPARC64 VI プロセッサおよび SPARC64 VII/SPARC64 VII+ プロセッサによる自律的なエラー復旧を実現し、是正措置を開始する時間を短縮することによって、稼働時間を増やします。

Oracle Solaris による予測的セルフヒーリングソフトウェアは、SPARC Enterprise M シリーズの信頼性をさらに高めます。Oracle Solaris による予測的セルフヒーリングの実装により、CPU およびメモリを定期的に監視できます。エラーの特性に応じて、スレッド、コア、CPU 全体のいずれかを自動的にオフラインにして、永続的な CPU ソフトエラーを解決できます。さらに、Memory Page Retirement 機能は、特定メモリ DIMM へのデータアクセスの複数回の訂正に合わせて、メモリページをオフラインで事前に取得する能力をサポートします。
.

7.5 エラー検出、診断、復旧

SPARC Enterprise M シリーズは、障害を早期に訂正し、周辺コンポーネントが繰り返しダウンしないようにする、重要な技術を採用しています。サーバのハードウェアサブシステム内にあるエラー検出機能および復旧機能が、本質的

に信頼性を向上させるアーキテクチャーの進歩を増幅します。最終的に、次の機能が連携して、アプリケーションの可用性を高めます。

- エンドツーエンドのデータ保護機能が、システム全体でエラーを検出し、訂正して、データの完全性を保証します。
- これらのサーバは、最先端の障害分離機能を使用してコンポーネント境界内のエラーを切り離し、コンポーネント全体ではなく関連チップのみをオフラインにできます。チップ単位でエラーを切り離して安定性を向上させ、最大コンピュータパワーの可用性を維持します。この機能は、CPU、メモリ、アクセスコントローラー、クロスバー ASIC、システムコントローラー、I/O ASIC に適用されます。
- 常に環境をモニタリングし、すべての適切な環境状態とエラー状態の履歴ログを提供します。
- ホストウォッチドッグ機能は、ドメインのオペレーティングシステムなどの、ソフトウェアの動作を定期的に確認します。この機能は、XSCF ファームウェアも使用して、エラー通知機能と復旧機能を起動します。
- CPU リソースの動的縮退機能として、プロセッサ障害の検出、分離、復旧の各段階の処理を提供します。この機能は、Dynamic Reconfiguration を使用して、動作中のアプリケーションを中断させることなく、CPU リソースを運用システムに動的に再構成することを可能とします。
  CPU の動的縮退はすべてのサーバで可能ですが、動的な交換は SPARC Enterprise M8000/M9000 でサポートされます。また、ドメインに対する DR 処理は SPARC Enterprise M4000/M5000/M8000/M9000 でサポートされます。
- 定期的なコンポーネントステータスチェックを実行して、多くのシステムデバイスの状態を判断し、障害の兆候を検出します。復旧メカニズムを起動して、システムおよびアプリケーションの障害を防ぎます。
- エラーロギング、多段階アラート、電子的な FRU 識別情報、システム障害の LED インジケーターは、迅速な問題解決に貢献します。

8. システム管理

ほとんどの組織にとって、サーバシステムを直接操作してローカルなシステム管理を行うことは、もはや現実的ではありません。24 時間のシステムオペレーション、災害復旧ホットサイト、地理的に分散した組織が、システムのリモート管理の必要性を生み出しています。富士通サーバの、多くの利点の 1 つである、完全自動データセンターのサポートは、サポートスタッフがネットワークにアクセスしてあらゆる場所で作業できるようになります（有償）。SPARC Enterprise M シリーズの設計は、強力なシステム監視機構（eXtended System Control Facility : XSCF）、XSCF 制御パッケージ、富士通のシステム管理ソフトウェアと合わせて、ハードウェアへの物理アクセスを伴わない、ほぼすべてのタスクを管理者がリモートで実行し、制御できるようにします。これらの管理ツールとリモート機能によって管理上の負荷を軽減し、時間を節約し、運営費を削減します。

8.1 システム監視機構

システム監視機構は、SPARC Enterprise M シリーズのリモート監視機能および管理機能の中心です。XSCF は、サーバシステムから独立した専用プロセッサで構成されており、XSCF 制御パッケージを実行します。XSCF とサーバ間の通信には、Domain to Service Processor Communication Protocol （DSCP）を使用します。DSCP プロトコルは、サービスプロセッサと各ドメイン間の TCP/IP ベースまたは PPP ベースの専用通信リンク上で動作します。サーバに入力電力が供給されている間は、すべてのドメインが非アクティブな場合でも、XSCF は常にシステムを監視します。

XSCF は環境センサーを定期的に監視し、潜在的なエラー状態を事前に警告し、必要に応じて予防的システム保守手順を実行します。例えば、システムを物理的に損傷させる可能性がある温度状態に対応して、XSCF はサーバのシャットダウンを開始できます。サービスプロセッサ上で動作している XSCF 制御パッケージを使用すると、管理者はプラットフォーム自体と同様に、ドメインをリモートで制御し、監視できます。

管理者は、XSCF へのネットワーク接続またはシリアル接続を使用して、ネットワーク上のどこからでもサーバを効果的に管理できます。サービスプロセッサへのリモート接続は、オペレーティングシステムから切り離されて動作し、システムコンソールを完全に制御し、権限を付与します。

8.1.1 冗長 XSCF

SPARC Enterprise M8000/M9000 上では、1 つの XSCF をアクティブとして、もう 1 つの XSCF をスタンバイとして構成します。2 つのサービスプロセッサ間の XSCF ネットワークは、システム管理情報を交換できます。フェイルオーバー発生時、サービスプロセッサは既に同期していて、すぐに役割を変更できます。

8.1.2 DSCP ネットワーク

Domain to Service Processor Communication Protocol サービスは、サービスプロセッサと各ドメイン間にセキュアな TCP/IP ベースおよび PPP ベースの通信リンクを提供します。このリンクがないと、XSCF はドメインと通信できません。サービスプロセッサは、リンクの XSCF 側に DSCP サービス専用の IP アドレスを 1 つ、リンクの各ドメイン側に IP アドレスを 1 つ必要とします。複数の XSCF を持つシステムでは、スタンバイ XSCF はドメインと通信しません。XSCF のフェイルオーバー発生時、新たにアクティブになった XSCF は、フェイルオーバーしたサービスプロセッサの IP アドレスを引き継ぎます。

8.1.3 XSCF 制御パッケージ

XSCF 制御パッケージを使用すると、SPARC Enterprise M シリーズと個々のパーティショニング機能を迅速かつ効果的に制御し、監視できます。XSCF 制御パッケージは、コマンドラインインターフェース（CLI）とウェブブラウザユーザーインターフェースを提供し、管理者とオペレーターがシステムコントローラー機能にアクセスできるようにします。固有の管理機能を持ち、パスワード保護されたアカウントも、ドメインコンソールにシステムセキュリティをもたらします。XSCF と個々のドメイン間の通信は、Secure Shell （SSH）と Secure Socket Layer （SSL）を基盤とする、暗号化された接続をサポートしており、XSCF 制御パッケージを使用して入力されたコマンドをセキュアにリモート実行できます。

XSCF 制御パッケージは、次の主要サーバ機能に対するインターフェースとなります。

- Dynamic Reconfiguration タスクを実行して、ドメインが中断なくアプリケーションを実行し続けている間に、実装されているシステムボードとオペレーティングシステムを論理的に接続または切断する。
- ドメインを管理し、Uni-XSB ユニットと Quad-XSB ユニットからなる論理システムボードを作成する。
- XSCF とドメイン間における、インタラクションの記録などの監査を管理する。

- SPARC Enterprise M4000/M5000/M8000/M9000 内のコンポーネントへの電力を監視および制御する。
- 提示されたハードウェア情報の解釈、温度上昇または電源の問題など差し迫った問題の通知、システム管理インターフェースへのアクセス。
- Oracle Solaris の障害管理アーキテクチャーと統合し、正確な障害診断と予測的障害分析によって可用性を向上させる。
- OpenBoot PROM （OBP）や Power-On Self-Test （POST）などの診断プログラムを実行し、監視する。
- Fujitsu Capacity on Demand 操作を実行し、ステージング機能を提供して、追加された処理リソースを後でアクティブにする。
- SPARC Enterprise M8000/M9000 上の冗長 XSCF 構成に障害がないか監視し、必要に応じて自動フェイルオーバーを実行する。

8.1.4 ロールベースのシステム管理

XSCF 制御パッケージを使用すると、複数のシステム管理者が、複数の自律ドメインを個別に管理し、単一の SPARC Enterprise M シリーズプラットフォーム内で連携させることができます。この管理ソフトウェアは、グループに編成された複数のユーザーアカウントをサポートします。各グループに異なる権限を割り当てます。権限を使用して、物理コンポーネント、ドメイン、またはドメイン内の物理コンポーネントなど、特定のハードウェアセットに関する特定の行動セットを実行できます。さらに、多数のドメイン上で、複数の異なる権限を保持できます。

8.1.5 Enhanced Support Facility

Enhanced Support Facility は、XSCF と連携して SPARC Enterprise M シリーズの運用管理性・保守性を向上させるソフトウェアです。本体装置の構成、状態の表示や故障情報の表示が行えます。また、OS 上からディスクや電源、PCI カード等の状態を監視し、トラブルが発生すると検知した情報を XSCF に通知します。他にも、システム情報を一括採取や、/etc/systems ファイルの設定チェック、サーバの電源投入・切断のスケジュール運用、ディスクの活性保守手順の表示などが可能です。

8.1.6 Systemwalker Centric Manager

Systemwalker Centric Manager は、システム運用のライフサイクル（導入 / 設定～監視～復旧～評価）に従い、ソフトウェア資源の配付、システムやネットワークの集中監視、リモートからのトラブル復旧などの優れた機能で運用管理作業を軽減します。また、このライフサイクル管理によりマルチプラットフォーム環境やインターネット環境など、最新のビジネス環境におけるシステ

ムの統合管理、運用プロセスの標準化（ITIL）、運用セキュリティの統制を支援します。

9. Oracle Solaris 10

ミッションクリティカルなビジネス目標を懸けて、企業は、パフォーマンス、可用性、安全性、ハードウェア資産の使用率を最適化できる、堅牢な運用環境を必要とします。他に類を見ない Oracle Solaris 10 は、多くの革新的技術を提供し、IT 組織がオペレーションを改善し、SPARC Enterprise M シリーズの潜在能力を完全に実現できるようにします。

9.1 性能

組織は、ハードウェアプラットフォームのパワーを有効に利用する必要があります。Oracle Solaris はプロセッサ数の増加に応じたほぼ直線的なスケーラビリティを容易にします。また、富士通における最大のサーバの、物理的なメモリ制限をもはるかに超えるメモリアドレス指定能力をサポートします。Oracle Solaris の高度な機能によって、IT 組織は潜在的なソフトウェアチューニング機会を識別し、raw システムスループットを最大化できます。

- Solaris Dynamic Tracing フレームワーク（DTrace）は強力なツールで、Java virtual machine（JVM）で動作しているものも含め、アプリケーションとカーネルのアクティビティに対して、真のシステムレベルのビューを提供します。DTrace ソフトウェアは、動作しているオペレーティングシステムのカーネルや、アクティブなアプリケーションを安全に計測します。カーネルをリブートしたり、ソフトウェアを再コンパイルまたは再起動したりすることはありません。この機能を使用すると、管理者は正確で簡潔な情報をリアルタイムで見て、アプリケーション実行におけるパターンとトレンドを強調できます。DTrace が提供する動的計測によって、問題を診断する時間が数日および数週間から数分および数時間に短縮され、データ駆動型の修正をより高速に行えます。
- Oracle Solaris の TCP/IP スタックは、非常にスケーラブルで、最適化されているため、パケットの処理に必要な命令数を削減して、オーバーヘッドを低減します。この技術は多数の接続もサポートし、サーバネットワークのスループットを CPU およびネットワークインターフェースカード（NIC）の数に比例して増加させることができます。Oracle Solaris 10 のネットワークスタックを利用して、アプリケーションの効率性とパフォーマンスを大幅に向上させることができます。
- Oracle Solaris 10 のメモリ処理システムは、アプリケーションが仮想メモリにより効率的にアクセスできるよう、複数のページサイズをサポートしている

ため、大量のメモリを集中的に使用するアプリケーションのパフォーマンスが向上します。さらに、Oracle Solaris 10 の Memory Placement Optimization（MPO）は、システム内のバランスを引き続き十分に維持しつつ、データがメモリ内でできる限り近くに保存されることを保証します。MPO は、ビジネスワークロードにおけるパフォーマンスを 20 %、一部の高パフォーマンスコンピューティングのワークロードにおいては 50 % 向上させることができます。

- Oracle Solaris のマルチスレッド実行モデルは、Oracle Solaris サーバがスケーラブルなパフォーマンスを実現できるようにする上で、重要な役割を果たします。Oracle Solaris のスレッディング機能はリリースごとに改善され、その結果、再コンパイルすることなく、既存アプリケーションのパフォーマンスと安定性が向上しています。

9.2 可用性

ハードウェア障害およびアプリケーション障害をすばやく診断し、切り離し、復旧する能力は、ビジネスオペレーションを停止させないというニーズを満たすために最も重要です。Oracle Solaris に昔からある機能が、システムの自己回復を実現します。例えば、カーネルメモリスクラバーは、物理メモリを常にスキャンし、1 ビットエラーを訂正して、この問題が訂正不可能な 2 ビットエラーになる可能性を減少させます。Oracle Solaris 10 は、Oracle Solaris Fault Manager および Oracle Solaris Service Manager 技術を導入して、自己回復の面で大きく飛躍しています。このソフトウェアを使用すると、ビジネスクリティカルなアプリケーションと必須のシステムサービスは、ソフトウェア障害、主要ハードウェアコンポーネントの故障、およびソフトウェアの設定ミスの問題が発生した場合に、中断することなく継続できます。

- Oracle Solaris Fault Manager ソフトウェアは、システムの障害を自動診断し、サービスの中断を回避できるよう自己回復行動を開始することによって、複雑さを減少させます。Oracle Solaris Fault Manager 診断エンジンは、受信したエラーストリームから認識できるパターンを見つけたら、障害診断を行います。エラーの識別に続いて、Oracle Solaris Fault Manager は、特定の障害への対処方法を知っているエージェントに情報を提供します。障害が発生する前に、問題のコンポーネントをシステム構成から除外できます。そして障害発生時には、この機能が自動復旧とアプリケーションの再起動を開始します。例えば、メモリエラーに対処するよう設計されているエージェントは、特定のチップ障害によって影響を受けるメモリアドレスを判断し、使用可能なメモリプールから影響を受ける位置を削除できます。
- Oracle Solaris Service Manager ソフトウェアは、オペレーティングシステムと一緒にパッケージ化されたサービスのコアセットを、管理者が一貫性ある管理コマンド群で操作できるファーストクラスオブジェクトに変換します。管

理者は、Oracle Solaris Service Manager を使用して、起動、停止、再起動、有効化、無効化、ステータス表示、スナップショットなど、サービスに関する操作を行えます。サービスのスナップショットは、サービスの構成全体を保存し、誤った変更を管理者がすべてロールバックできるようにします。スナップショットは、サービスが開始すると常に自動的に作成され、誤ったエラーを防ぐことによって、リスクを低減できるようにします。Oracle Solaris Oracle Service Manager は、Oracle Solaris Fault Manager に統合されています。低レベルの障害が、動作中のサービスのより高いレベルのコンポーネントに影響を及ぼすことが判明した場合、Oracle Solaris Fault Manager は、適切な行動をとるよう、Oracle Solaris Service Manager に指示できます。

エラー状態の対処に加え、計画的なダウンタイムを効率的に管理することによって、可用性レベルが大幅に向上します。Oracle Solaris Flash および Oracle Solaris Live Upgrade ソフトウェアなど、Oracle Solaris に同梱されているツールを使用すると、より迅速で一貫性あるソフトウェアのインストール、アップグレード、パッチを実行でき、結果的に稼働時間が増加します。

- IT 組織は、Oracle Solaris Flash 機能を使用して、企業ニーズに合わせたオペレーティングシステム構成で、システムをすばやくインストールおよびアップグレードできます。この技術は、アプリケーション、パッチ、パラメーターなど、カスタマイズされた高速インストールイメージをシステム管理者が作成するためのツールを提供します。このインストールイメージは、ハードウェアの最大速度に近いデータ速度でインストールできます。
- Oracle Solaris Live Upgrade 機能は、ディスク上にある複数の Oracle Solaris のインスタンスをアップグレードし、管理する仕組みを提供します。システム管理者は、この技術を使用して、動作している実稼働システムをオフラインにすることなく、新しいオペレーティングシステムをインストールできます。アプリケーションのダウンタイムは、新しい構成のリブートに必要な時間のみです。

9.3 セキュリティ

今日、システムはますます接続されるようになり、利点と課題を生み出しています。グローバルネットワークはより大きな収益機会をもたらしますが、企業はセキュリティに関する懸念事項に細心の注意を払う必要があります。地球上で最もセキュアなオペレーティングシステムである Oracle Solaris 10 は、以前は軍用レベルの Trusted Solaris にのみあった機能を提供します。これらの機能は、政府や金融機関にとって必須の強力な制御を可能にし、セキュリティに関する懸念事項と監査機能の要件に重点を置いているすべての企業にもメリットをもたらします。

- User Rights Management と Process Rights Management は、Oracle Solaris Container 仮想化技術と連動して、複数のアプリケーションが同じドメインをセキュアに共有できるようにします。ユーザーとアプリケーションに、割り当てられた業務を実行するのに必要な最小限の能力のみを付与することによって、セキュリティリスクは減少します。市場の他のソリューションと異なり、これらのセキュリティ強化機能を利用するのに、アプリケーションを変更する必要はありません。
- Oracle Solaris Trusted Extension は、既存の Oracle Solaris 10 のセキュリティポリシーをラベリング機能で拡張したもので、以前は極めて特化したオペレーティングシステムまたはアプライアンスでのみ使用可能でした。この拡張は、商用オペレーティングシステムに真のマルチレベルのセキュリティをもたらすため、固有の規制または情報保護要件を持つ民間団体にとって有益です。
- Oracle Solaris 10 OS の中核を成すのは、プラットフォームを脅威に対して強化する機能です。Oracle Solaris 10 ディストリビューションに含まれるファイアーウォール保護技術は、個々のシステムを攻撃から保護します。さらに、Oracle Solaris 10 のファイルの完全性チェックとデジタル署名バイナリを使用して、管理者はプラットフォームにハッカーの手がまだ触れていないことを確認します。セキュアなリモートアクセス機能も、複数オペレーティングシステムにわたるシステムアクセスの管理を集中させることによって、安全性を高めます。

9.4 仮想化とリソース管理

各 IT 資産の利用を最大化するという経済的なニーズによって、複数のアプリケーションを単一のサーバプラットフォームに統合したいとういう要望があります。仮想化手法は、組織が 1 台のサーバ上の各ドメイン内にあるアプリケーション間に管理およびリソース境界を作成できるようにして、統合戦略をさらに一歩強化します。Oracle Solaris Containers は、仮想化とソフトウェアパーティショニングに対するブレイクスルーアプローチで、Oracle Solaris OS の単一インスタンス内に多数の専用実行環境を作成できるようにするものです。この技術を使用すると、IT 組織はアイドル状態のコンピュータパワーをセキュアで分離されたランタイム環境にすばやく利用し、プロビジョニングして、管理するオペレーティングシステムのインスタンス数を増やすことなく、新たに配備できます。さらに、アプリケーションを個々の Oracle Solaris Container 内にホスティングすると、管理者は統合サーバ内の権限とリソースをきめ細かく制御できます。

さらに、Oracle Solaris Resource Manager は、構造化されたポリシー駆動方式で、コンピューティングリソースを Oracle Solaris Container 内および個々のタスクとユーザーに割り当てることができます。Oracle Solaris のリソース管理機能を使用して、CPU 時間、プロセス、仮想メモリ、接続時間、ログインなどのシステムリソースを、前もってきめ細かく割り当て、制御し、監視すると、サービスレベルはさらに予測可能になります。Oracle Solaris Resource Manager は、ビジネスニーズの変化に伴って、計算リソースを使用するための新しいプロパティを企業が定期的に設定できるようにします。Oracle Solaris Container および Solaris Resource Manager を利用することによって、リソースの使用状況を改善し、ダウンタイムを短縮し、ソリューションコストを低減できます。

10. まとめ

データセンターにおけるスケーラビリティ、信頼性、管理容易性に対する高度な要望をサポートするため、インフラストラクチャーはより簡単に配備、調整、管理できるようになる一方で、パフォーマンスとキャパシティを増加させ続ける必要があります。SPARC64 VI プロセッサもしくは SPARC64 VII/ SPARC64 VII+ プロセッサ、大容量メモリ、信頼できるアーキテクチャー、システム監視機構を備えた SPARC Enterprise M シリーズは、新たなレベルのパワー、可用性、使いやすさを企業にもたらします。パーティショニング機能、拡張システムボード、Dynamic Reconfiguration で提供される高度なリソース管理によって、ハードウェア資産の利用を最適化できるため、SPARC Enterprise M シリーズの価値はさらに高まります。高速でスケーラブルな SPARC Enterprise M シリーズを利用すると、最高のパワーと柔軟性を獲得できます。これは、競争に勝ち、勝ち続けるための戦略的資産です。

参考情報

富士通の SPARC Enterprise M シリーズ、および関連するソフトウェアやサービスの詳細については、当社の営業担当にご連絡いただくか、表 10-1 に記載されているウェブサイトを参照してください。

表 10-1. 関連ウェブサイト

ウェブサイトの URL	説明
http://primeserver.fujitsu.com/sparcenterprise/	SPARC Enterprise
http://jp.fujitsu.com/platform/	富士通コンピュータプラットフォーム一覧
http://jp.fujitsu.com/	富士通

（参考資料）

UNIX サーバ用プロセッサ「SPARC64 V」、Aug 2004, 富士通株式会社