Sun Fire™ E25K/E20K システム

製品概要

Sun Microsystems, Inc.
www.sun.com

Part No. 817-6850-12
2006 年 6 月, Revision A

コメントの送付: http://www.sun.com/hwdocs/feedback

原典: *Sun Fire E25K/E20K Systems Overview Manual*
Part No: 817-4136-13 v2
Revision A

目次

図目次

表目次

はじめに

このマニュアルでは、Sun Fire™ E25K/E20K システムの概要を説明し、キャビネット、システム、構成、動的なシステムドメインの構成、システムボードと、信頼性、可用性、および保守性機能について説明します。

マニュアルの構成

このマニュアルは、以下の章で構成されています。

第 1 章では、システムとボード、最大構成、およびインターコネクトアーキテクチャーについて説明します。

第 2 章では、構成例、インタードメインネットワーキング、ドメイン保護、およびドメインの障害分離について説明します。

第 3 章では、システムのエラー保護を定義し、冗長コンポーネントおよびシステム回復について説明します。また、システムコントローラ技術とシステムの並行保守機能についても説明します。

第 4 章では、システムの中心である Sun™ Fireplane interconnect アセンブリについて説明します。

第 5 章では、システム内のコンポーネントについて説明します。

用語集。

関連マニュアル

用途	タイトル
サイト計画	『Sun Fire E25K/E20K システムサイト計画の手引き』
設置	『Sun Fire E25K/E20K Systems Read Me First』(英語版)
設置	『Sun Fire E25K/E20K システム概要』
設置	『Sun Fire E25K/E20K システム開梱の手引き』
設置	『Sun Fire E25K/E20K システムハードウェアの設置と移動の手引き』
保守	『Sun Fire E25K/E20K システムサービスマニュアル』
保守	『Sun Fire E25K/E20K システムサービスリファレンス I 名称一覧』
保守	『Sun Fire E25K/E20K システムサービスリファレンス II コンポーネントの番号』

Sun のオンラインマニュアル

各言語対応版を含む Sun の各種マニュアルは、次の URL から表示、印刷、または購入できます。

`http://www.sun.com/documentation`

Sun 以外の Web サイト

このマニュアルで紹介する Sun 以外の Web サイトが使用可能かどうかについては、Sun は責任を負いません。このようなサイトやリソース上、またはこれらを経由して利用できるコンテンツ、広告、製品、またはその他の資料についても、Sun は保証しておらず、法的責任を負いません。また、このようなサイトやリソース上、またはこれらを経由して利用できるコンテンツ、商品、サービスの使用や、それらへの依存に関連して発生した実際の損害や損失、またはその申し立てについても、Sun は一切の責任を負いません。

Sun の技術サポート

このマニュアルに記載されていない技術的な問い合わせについては、次の URL にアクセスしてください。

```
http://www.sun.com/service/contacting
```

コメントをお寄せください

マニュアルの品質改善のため、お客様からのご意見およびご要望をお待ちしております。コメントは下記よりお送りください。

```
http://www.sun.com/hwdocs/feedback
```

ご意見をお寄せいただく際には、下記のタイトルと Part No. を記載してください。

『Sun Fire E25K/E20K システム製品概要』、Part No. 817-6850-12

米国の輸出規制法について

このサービスマニュアルに記載されている製品および情報は、米国の輸出規制法に従うものであり、その他の国の輸出または輸入に関する法律が適用される場合もあります。核、ミサイル、化学生物兵器、または核の海上での最終使用あるいは最終使用者は、直接的または間接的にかかわらず厳重に禁止されています。米国の通商禁止対象国、または拒否された人物および特別認定国リストにかぎらず、米国の輸出禁止リストに指定されている実体への輸出または再輸出は、厳重に禁止されています。予備の CPU の使用または交換は、米国の輸出法に従って輸出された製品に対する CPU の修理または 1 対 1 の交換に制限されています。米国政府の許可なしに、製品のアップグレードに CPU を使用することは、厳重に禁止されています。

第1章

Sun Fire E25K/E20K システムの概要

この章では、Sun Fire E25K/E20K システムの概要について説明します。

Sun Fire E25K/E20K システムは、バイナリ互換の Solaris™ UNIX® オペレーティングシステムが動作する最新の UltraSPARC® IV Cu CPU および Sun Fireplane interconnect アーキテクチャーを使用します (図 1-1)。また、業界をリードする動的システムドメインと、信頼性、可用性、保守性 (RAS) 機能を採用し、アクティブセンタープレーン技術を使用しています。

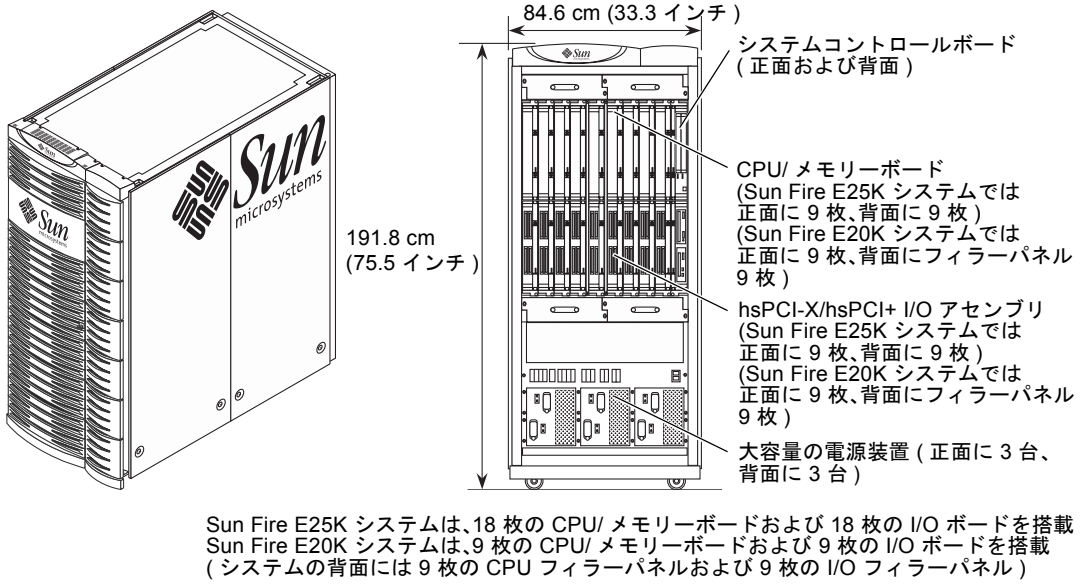


図 1-1 Sun Fire E25K/E20K システム

Sun Fire E25K システムと Sun Fire E20K システムは、基本的には同じシステムです。Sun Fire E25K システムは、CPU/メモリーボードおよび I/O アセンブリをそれぞれ 18 枚まで搭載できます。Sun Fire E20K システムは、CPU/メモリーボードおよび I/O アセンブリをそれぞれ 9 枚まで搭載できます。いずれのシステムも、2 枚のシステムコントロールボード (メイン 1 枚とスペア 1 枚) を搭載しています。

1.1 システムボード

1.1.1 CPU/メモリーボード

各 CPU/メモリーボードは、4 つの CPU を搭載しています。各 CPU は、8 つの DIMM を持つメモリーサブシステムに関連付けられるため、メモリーの帯域幅および容量は、CPU が追加されるごとに増加します。ボードのメモリー容量は、2G バイトの DIMM を使用した場合、64G バイトです。ボード内のメモリーの最大帯域幅は、9.6G バイト/秒です。CPU/メモリーボードは、システムのほかの部分と、4.8G バイト/秒で接続します。

1.1.2 I/O アセンブリ

Sun Fire E25K/E20K システムのホットスワップ PCI アセンブリアーキテクチャー (hsPCI-X または hsPCI+) は、2 つの I/O コントローラを備えています。各コントローラは、各 I/O アセンブリ上に合計 4 つのバス (33 MHz の PCI (Peripheral Component Interconnect) バスが 1 つと、33/66/90 MHz PCI バスが 3 つ) を備えています。そのため、各 I/O アセンブリは、4 つのホットスワップコンポーネント PCI スロットを備えていることになります。Sun Fire I/O アセンブリは、システムのほかの部分と 2.4G バイト/秒で接続します。

1.1.3 システムコントローラ

システムコントローラは、Sun Fire E25K/E20K システムの可用性および保守性技術の中心です。システムコントローラは、システムの構成、起動処理の調整、動的システムドメインの設定、システム環境センサーの監視を行い、エラー検出および診断、回復を処理します。システムには 2 枚のシステムコントロールボードが組み込まれていて、1 枚のボードに障害が発生した場合には、冗長性と自動フェイルオーバーを提供します。

1.1.4 周辺装置

Sun Fire E25K/E20K システムキャビネットには、システムコントローラ周辺装置 (DVD-ROM、DAT ドライブ、およびハードドライブ) を取り付けるスペースはありますが、その他の周辺装置のためのスペースはありません。追加の周辺装置拡張ラックを使用すると、その他の周辺装置を組み込むことができます。

1.2 システム構成

表 1-1 に、Sun Fire E25K/E20K システムの最大構成を示します。

表 1-1 Sun Fire E25K/E20K システムの最大構成

コンポーネント	E25K システムの構成	E20K システムの構成
CPU/メモリーボード	18	9
CPU	72	36
DIMM の数	576	288
メモリー容量 (2G バイトの DIMM を使用)	1152G バイト	576G バイト
Sun Fireplane interconnect	有効	有効
リピータボード	なし	なし
拡張ボード	18	9
ドメイン	18	9
I/O ボード (アセンブリ)	18	9
PCI アセンブリの種類	hsPCI+	hsPCI+
PCI アセンブリの種類	hsPCI-X	hsPCI-X
アセンブリごとの PCI スロット数	4	4
最大の PCI スロット数	72	36
大容量電源装置	6	6
所要電力	24 kW	24 kW
システムコントロールボード	2	2
冗長冷却	あり	あり
冗長 AC 入力	あり	あり
格納装置	Sun Fire E25K/E20K システムキャビネット	Sun Fire E25K/E20K システムキャビネット
格納装置内の周辺装置用スペース	なし	なし

1.3 システムインターコネクト

表 1-2 に、Sun Fire E25K/E20K システムのインターコネクト機能を示します。

表 1-2 Sun Fire E25K/E20K システムのインターコネクト仕様

インターコネクト	仕様
システムクロック	150 MHz
一貫性プロトコル	センタープレーンを介して、各ボードセットをスヌープ
システムアドレスインターコネクト	18 のスヌープバス、 18×18 グローバルアドレスクロスバー、 18×18 グローバル応答クロスバー
CPU/メモリーボードの内部二分帯域幅 (Bisection Bandwidth)	4.8G バイト/秒
CPU/メモリーボードのオフボードデータポート	4.8G バイト/秒
I/O ボードのオフボードデータポート	2.4G バイト/秒
システムデータインターコネクト	18 の 3×3 ボードセットクロスバー、 18×18 グローバルクロスバー
システムの二分帯域幅	43G バイト/秒
ランダムにアクセスした場合を想定した lmbench による (連続ロードでの) 平均応答時間	326 ns

注 – PCI System Architecture, Third Edition (1995、MindShare, Inc.) (ISBN 0-201-40993-3) の付録 A「Glossary」では、スヌープを次のように定義しています。

スヌープ – キャッシュコントローラ以外のエージェントがメモリーにアクセスした場合、キャッシュコントローラはそのトランザクションをスヌープして、現在のマスターがキャッシュ内の情報にアクセスしているかどうかを判断する必要があります。スヌープがヒットした場合、キャッシュコントローラは適切な対応を行なって、キャッシュされた情報の一貫性を確実に保持する必要があります。

1.3.1 Sun Fireplane interconnect アーキテクチャー

Sun Fire E25K/E20K システムは、システムインターコネクトアーキテクチャーとして、UltraSPARC IV Cu CPU 世代で使用される一貫性のある共有メモリープロトコルである Sun Fireplane interconnect を使用します。これは、第 4 世代の共有メモリーインターコネクトです。Sun では、改良を加えたシステムインターコネクトを新世代の CPU それぞれに実装することによって、CPU の性能に応じたシステム性能を引き出します。

Sun Fireplane interconnect アーキテクチャーは、前世代の UPA (Ultra Port Architecture) を発展させたものです。システムクロックレートは 50% 向上し、100 MHz から 150 MHz になりました。クロックごとのスヌープは、0.5 から 1 に倍増しました。両方を合わせると、スヌープの帯域幅は 3 倍の、毎秒 1 億 5 千万アドレスになります。

また、Sun Fireplane interconnect アーキテクチャーでは、ポイントツーポイントのディレクトリ一貫性プロトコルの新しいレイヤーを設けました。このプロトコルは、単一のスヌープバスが提供できる帯域幅以上の能力を必要とするシステムで使用されます。これによって、複数のスヌープバス間の一貫性を保持することができます。

図 1-2 に、Sun Fire E25K システムの Sun Fireplane interconnect アーキテクチャーを示します。ボードの図は、実際のボード上の接続を示していますが、わかりやすくするため、スイッチとコントローラチップは省略しました。

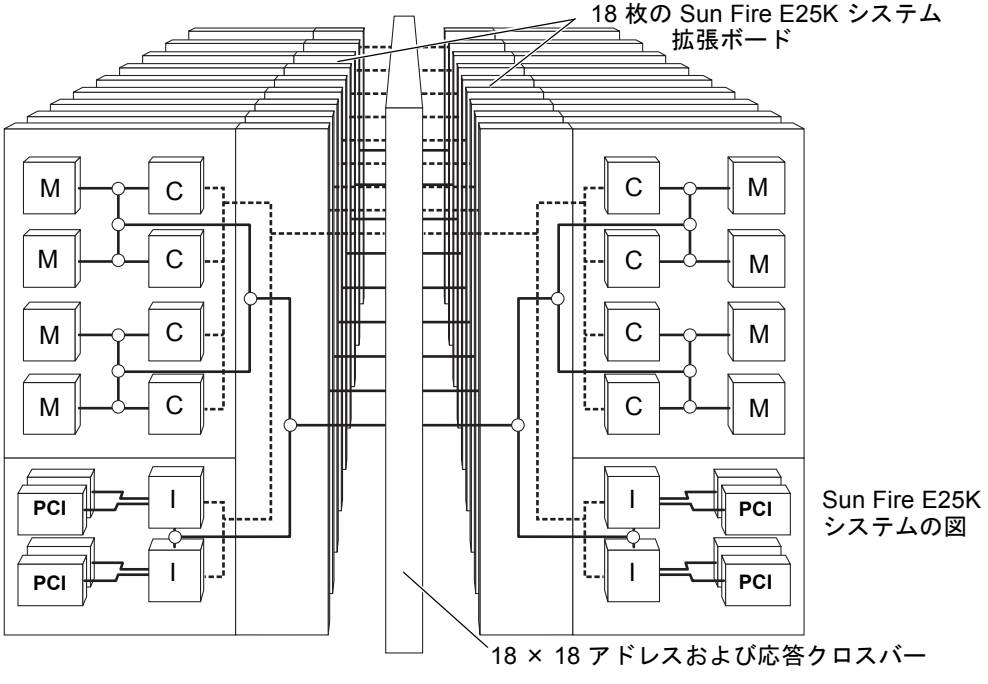


図 1-2 Sun Fireplane interconnect

Sun Fire E25K/E20K システムは、拡張ボードを使用して、CPU/メモリーボード、I/O アセンブリ、および Sun Fireplane interconnect ポート間の 3×3 スイッチを実装します。Sun Fire E25K/E20K システムは、Sun Fireplane interconnect に 3 つの 18×18 クロスバーを備えてアドレス、応答、およびデータに対応しているため、アドレストラフィックがデータトラフィックを妨げることはありません。Sun Fire E25K/E20K システムの Sun Fireplane interconnect の最大帯域幅は 43G バイト/秒です。

1.3.2 アドレスインターコネクト

図 1-2 の点線は、スヌープアドレスバスを示します。スヌープはすべてのシステムクロックごとに発生します。Sun Fire E25K/E20K システムでは、ボードセットごとに異なるスヌープアドレスバスがあります。ボードセットとは、CPU/メモリーボード、I/O アセンブリ、および拡張ボードを組み合わせたものです。ボードセット間の一貫性は、一貫性プロトコルのポイントツーポイント (ディレクトリ) 部分を使用して保持されます。

1.3.3 データインターコネクト

図 1-2 の実線は、データパスを示します。この実線の交差部分にある小さな円は、3 ポートスイッチを示します。CPU/メモリーボードには、CPU またはメモリーユニットとオフボードポート間の 3 レベルの 3×3 スイッチがあります。CPU/メモリーボードのオフボードの帯域幅は 4.8G バイト/秒です。I/O アセンブリの帯域幅は 2.4G バイト/秒です。

1.4 動的システムドメイン

Sun Fire E25K/E20K システムの各ドメインには、1 枚以上の CPU/メモリーボードおよび 1 枚以上の I/O アセンブリが含まれます。各ドメインは、Solaris オペレーティングシステムインスタンスを実行し、個々に周辺装置およびネットワーク接続を備えています。ドメインの再設定は、ほかのドメインの操作を中断せずに行うことができます。ドメインは、次の目的に使用できます。

- 新規アプリケーションの評価
- オペレーティングシステムの更新
- さまざまな部門のサポート
- 修復またはアップグレードのためのボードの取り外しと再取り付け

次に、フル構成された Sun Fire E25K システムを、3 つのドメインに分割して、3 種類の機能を処理する例を示します。

- ドメイン 1 は、OLTP (Online Transaction Processing) を実行するように設定します。これは、それぞれ 4 つの CPU を搭載した 8 枚のボードを持つ 32 CPU のドメインです。
- ドメイン 2 は、DSS (Decision Support Software) を実行するように設定します。これも、それぞれ 4 つの CPU を搭載した 8 枚のボードを持つ 32 CPU のドメインです。
- ドメイン 3 は、開発者用のドメインとして設定します。これは、それぞれ 4 つの CPU を搭載した 2 枚のボードを持つドメインです。

負荷の変更が要求されると、ボードは自動的にドメイン間で移行されます。

Sun Fire E25K システムでは、最大 18 のドメインを持つことができます。Sun Fire E20K システムでは、最大 9 のドメインを持つことができます。ドメインは、インターコネクト ASIC (Application-Specific Integrated Circuit) によって、それぞれが分離されています。

1.5 信頼性、可用性、および保守性

信頼性、可用性、および保守性 (RAS) は、ビジネスに不可欠なアプリケーションを展開しているユーザーにとって重要な要件です。Sun Fire E25K/E20K システムは、業界をリードする RAS 機能に基づいて構築されています。この節では、RAS を向上させる主な機能のいくつかについて説明します。

1.5.1 集積回路の信頼性

- **起動時の診断。**主要な Sun Fire E25K/E20K システムの ASIC は、すべて電源投入時に組み込み自己診断 (BIST) を行います。システムクロックレートでランダムなパターンを適用し、組み合わせ論理によって高い確率で障害を検出します。電源投入時自己診断 (POST) は、システムコントローラから制御され、まず分離された論理ブロックをテストします。その後、POST はシステムをさらに使用してテストを続けます。障害が検出されたコンポーネントは、電気的に Sun Fireplane interconnect から分離されます。その結果、この自己診断に合格し、エラーのない状態で操作できる論理ブロックだけを使用して、システムが起動されます
- **UltraSPARC IV Cu CPU 内の内部 SRAM 保護。**CPU がより高密度になり、コア電圧がより低くなると、SRAM のセルは宇宙線によってビット反転を起こしやすくなります。大部分の内部 SRAM に対するシングルビットエラーは、検出および回復が可能です。
- **外部 SRAM 保護**。すべての外部 SRAM は、ECC (Error-Correcting Code) によって保護されます。これには、CPU の外部キャッシュデータおよび Sun Fire E25K/E20K システムの一貫性ディレクトリキャッシュが含まれます。

1.5.2 インターコネクトの信頼性

- **アドレスインターコネクト保護。**Sun Fire E25K/E20K システムのアドレスバスおよび制御信号は、シングルビットエラーを検出するためにパリティー保護されています。また、Sun Fireplane interconnect のアドレスクロスバーおよび応答クロスバーは、シングルビットエラーを訂正し、ダブルビットエラーを検出するために、ECC によって保護されています。
- **データインターコネクト保護。**すべてのシステムデータパスは ECC によって保護され、データが破壊される前にシングルビットエラーを訂正し、ダブルビットエラーを検出します。ECC は、CPU または I/O コントローラが書き込みコマンドを実行するときに生成されます。追加されたビットは、インターコネクトを介して転送先まで送信されます。メモリーサブシステムはエラーの検査または修正は行わず、追加の記憶ビットを提供するだけです。データはメモリーから読み出されるときに検査され、必要に応じて受信側の CPU または I/O コントローラによって訂正されます。障害を分離するため、データがチップからチップへ渡され

るときにパリティーも検査されます。データスイッチ ASIC も ECC を検査します。ECC パターンには検出の終了した DRAM チップ障害を使用しますが、その訂正はできません。

1.5.3 耐障害の冗長性

これらのサブシステムでの障害は、可用性を損ないません。

- **N+1 の冗長性。**AC 電源入力および大容量電源装置、冷却ファンはすべて、N+1 冗長性を使用した耐障害性を備えています。これらサブユニットの 1 つに障害が発生すると、システムを停止することなく、その他のコンポーネントがシステムの操作を継続します。
- **実行中のフェイルオーバー。**システムコントロールボードは、2 枚一組で構成されます。1 枚が動作し、もう 1 枚がホットスペア用です。システムコントローラの CPU またはクロック生成ロジックに障害が発生した場合、システムを停止することなく、障害の発生したボードからもう 1 枚のボードに制御が切り替わります。

1.5.4 障害発生後の再構成

- **自動システム回復。**適切に構成されたシステムは、障害発生後常に再起動します。システムコントローラが障害を特定し、障害の発生した CPU、メモリー、またはインターコネクトコンポーネントを除いてシステムを再構成し、オペレーティングシステムを再起動します。
- **障害発生後のインターコネクトの再構成。**システムインターコネクトに障害が発生したあと、システムは障害の発生したインターコネクトコンポーネントを分離し、システムの帯域幅の半分が使用可能な状態で再起動します。3 つのクロスバーはドメインごとに、フルモードから縮退モードの間で別々に再構成できます。

1.5.5 保守性

- **システムコントローラ。**システムコントロールボードは、RAS 技術の中心です。SC CPU ボードは、UltraSPARC-IIi システムを組み込んだ、既成の SPARCengine® Netra 2140 6U cPCI ボードです。このボードは、Solaris ソフトウェアおよびシステム管理ソフトウェアを実行します。システムコントローラは、JTAG (Joint Test Action Group) によってマシン内の主なチップのレジスタにアクセスし、マシンの状態を継続的に監視します。問題が検出されると、システムコントローラはどのハードウェアに障害が発生したかを判断し、そのハードウェアが交換されるまで、アクセスされないようにします。

- **コンソールバス。**コンソールバスは、システムコントローラがシステムのアドレスバスおよびデータバスの完全性に依存することなく、マシンの内部動作にアクセスできるようにするためのセカンダリバスです。これによって、システムコントローラは、システムの動作の継続を妨げる障害が発生しても動作できます。コンソールバスは、パリティー保護されています。
- **環境監視。**システムコントローラは、温度、ファンの動作、電源装置の性能など、システムの安定性に関する主な測定値に基づいて、キャビネットの環境を監視します。
- **並行保守性。**ファン、大容量電源装置、およびシステムボードは、すべてホットスワップ対応のコンポーネントです。実行中のシステムからの取り外しおよび交換ができます。
- **動的システムドメイン。**動的システムドメインによって、動作中のドメインに対する修復されたボードやアップグレードされたボードの追加、取り外しができます。

第2章

動的システムドメイン

Sun Fire E25K/E20K システムでは、動的ドメインを構成できます。この章では、動的ドメインについて説明します。

Sun Fire E25K システムは、最大 18 の動的システムドメインに動的に分割できます。Sun Fire E20K システムは、最大 9 の動的システムドメインに動的に分割できます。各ドメインは、それぞれ、Solaris OS の特定のインスタンスを実行するときに使用する起動ディスク、ディスク記憶装置、ネットワークインタフェース、および I/O インタフェースを備えています。CPU ボードおよび I/O アセンブリは、動作中のドメインに対して個別に追加、取り外しができます。

ドメインは、アプリケーションサーバー、Web サーバー、データベースサーバーなどのソリューションの個々のパートを実行して、サーバー統合を実現するために使用されます。ドメインは、ほかのドメインのハードウェア障害またはソフトウェア障害から、ハードウェア保護されています。

2.1 ドメインの構成

各システムボード (スロット 0 およびスロット 1 のボード) は、個別に動作中のドメインに対する追加および取り外しができます。このため、CPU およびメモリー資源は、ディスク記憶装置およびネットワーク接続を妨げることなく、ドメインからドメインへと移動できます。Sun Fire E25K システムでは、各ドメインは 1 枚の I/O アセンブリを必要とするため、ドメイン数は最大で 18 となります。Sun Fire E20K システムでは、各ドメインは 1 枚の I/O アセンブリを必要とするため、ドメイン数は最大で 9 となります。

1 つのボードセットの中の 2 枚のシステムボードが別々のドメインにある場合、このボードセットは「分割拡張ボードセット」と呼ばれます。この拡張ボードは、各システムボードに対してトランザクションを分けます。図 2-1 に、2 つのドメインに分割されたボードセットの構成例を示します。ドメイン内のボードを物理的に近接させる必要はありません。

分割拡張ボードセットハードウェアは 2 つのドメイン間で共有されるため、このボードセットに障害が発生すると、両方のドメインが停止します。たとえば、フル構成されたシステムを 9 つのボードセットを持つ 2 つのドメインに分割すると、分割しない場合に比べて、すべて分割された拡張ボードセットでは、平均故障間隔 (MTBF : Mean Time Between Failure) が約 5% 長くなります。また、分割拡張ボードセットを介したメモリーアクセスは、2 システムクロック (13 ns) 遅くなります。すべての拡張ボードセットが分割されている場合、ほかのボードセットにアクセスして読み込みを行なったときの応答時間は、約 6% 増加します。

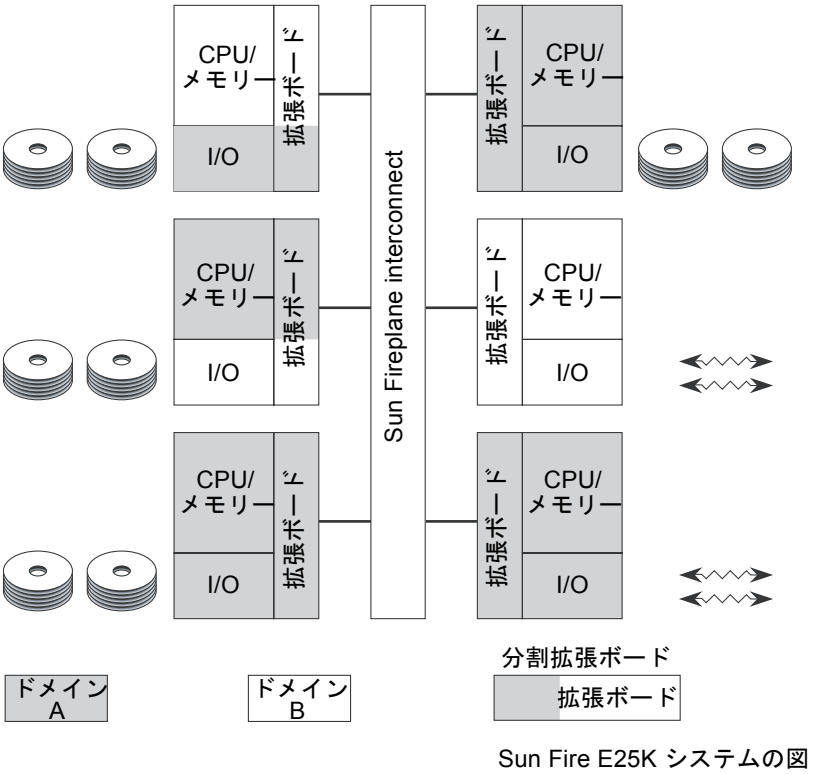


図 2-1　　分割ボードセットを含むドメイン構成の例

2.2 ドメイン保護

一次ドメイン保護は、トランザクションが最初に検出されたとき、AxQ (Address eXtender Queue) ASCI でドメインの妥当性について各トランザクションを検査することで実現されます。Sun Fire E25K システムでは、SDI (System Data Interface) チップも、有効な宛先に対する最大 36 のシステムボードへのデータ転送要求を選別できます。また、各 Sun Fireplane interconnect アービタ (データ、アドレス、応答) は、最大 18 の拡張ボードへの要求を選別します。Sun Fire E20K システムでは、SDI チップは、有効な宛先に対する最大 18 のシステムボードへのデータ転送要求を選別できます。各 Sun Fireplane interconnect アービタ (データ、アドレス、応答) は、最大 9 の拡張ボードへの要求を選別します。これは、AXQ および SDI チップに含まれるほかのドメイン保護機構で二重に検査されます。

AXQ で違反エラーが検出されると、AXQ は、そのエラー操作を、存在しないメモリーへの要求と同様に処理します。これによって、マップされた一貫性プロトコル信号をアサートせずに要求が再発行され、Solaris ソフトウェアが 1 つのプロセスから別のプロセスへ実行を切り替えます。Sun Fireplane interconnect の違反エラーによって、違反しているドメインのドメインストップが発生します。このエラーは、一次保護機構の障害を示しているからです。

2.3 ドメインの障害分離

ドメインは、ほかのドメインのソフトウェア障害またはハードウェア障害から保護されています。特定のドメインに割り当てられているプロセッサまたはメモリーのハードウェアに障害が発生した場合は、そのドメインだけが影響を受けます。複数のドメインが共有するハードウェアに障害が発生した場合は、そのハードウェアを共有するドメインだけが影響を受けます。

2 つのドメインで共有されるハードウェアの例として、一方のドメインに CPU/メモリーボードが構成され、もう一方のドメインに関連する I/O アセンブリが構成されている場合を考えてみます。分割拡張ボード上のロジックは、この 2 つのドメイン間で共有されます。分割拡張ボードまたは Sun Fireplane interconnect への制御配線上の障害は、関連する 2 つのドメインだけの障害の原因になります。システムクロックジェネレータ、Sun Fireplane interconnect チップなど、広域的に共有されるハードウェアの障害は、すべてのドメインの障害の原因になります。

制御配線上のパリティーエラーや ASIC 障害などの致命的なエラーでは、ドメインストップが発生します。拡張ボードから Sun Fireplane interconnect のアービタチップへのステアリング信号は、パリティー保護されています。パリティーエラーが発生す

ると、Sun Fireplane interconnect アービタの複数のコピーが同期を取れなくなります。そのため、このようなパリティーエラーでは、ただちにドメインのドメインストップが発生します。

Sun Fireplane interconnect を介して送信されるパケットの修正可能なシングルビットエラーなど、重大ではないエラーの場合は、レコードストップが発生します。レコードストップによって ASIC 内の履歴バッファが凍結されるので、ドメインが動作している状態で、JTAG を使用して障害に関する情報を走査できます。

分割拡張トランザクション (ボード 0 とボード 1 が異なるドメインにある) では、アービタの同期を保持して、エラーが複数のドメインに影響しないように対処する必要があります。このようなトランザクションでは、応答時間が 2 サイクル余分にあるので、アービタの 1 つがそれ自身の正しいバージョンのステアリングを処理する前に、すべてのアービタがステアリングパリティーエラーを検出できます。分割拡張ボードセットをできるだけ少なくしてシステムを構成すると、性能が向上します。

データアービタ ASIC からデータ MUX ASIC への Sun Fireplane interconnect のステアリング信号はパリティー保護されています。データマルチプレクサ (MUX) チップでは、ステアリングに従って処理する前にエラーを照合することはできません。そのため、こうした局所的な配線上のパリティーエラーによって、複数のドメインまたはすべてのドメインでドメインストップが発生する可能性があります。

第3章

信頼性、可用性、および保守性

信頼性、可用性、および保守性 (RAS) は、システムの継続稼働と保守時間の短縮を行う能力を評価および測定する基準です。システムの信頼性は、障害を低減し、データの完全性を保証します。保守性は、コンポーネントのアップグレードが必要な場合や障害が発生した場合に、保守のために電源を切断している時間を短くします。障害を回避する高い信頼性と、障害からすばやく回復できる迅速な保守性を組み合わせることによって、高い可用性を実現できます。システムの可用性とは、システムがサポートする機能およびアプリケーションへのアクセス可能性を持続することです。この章では、サポートされる機能およびアプリケーションについて説明します。

3.1 SPARC CPU のエラー保護

この CPU は、図 3-1 に示すように、外部キャッシュ SRAM を ECC (Error Correction Code) 保護し、主な内部 SRAM 構造をパリティー保護しています。ブロック図の P と E の文字はそれぞれ、パリティーの生成および検査と、受け取ったユニットによる ECC の生成、検査、および訂正を意味します。内部キャッシュ構造のパリティーエラーはソフトウェアによって訂正されるので、障害発生後の正しい操作が保証されます。

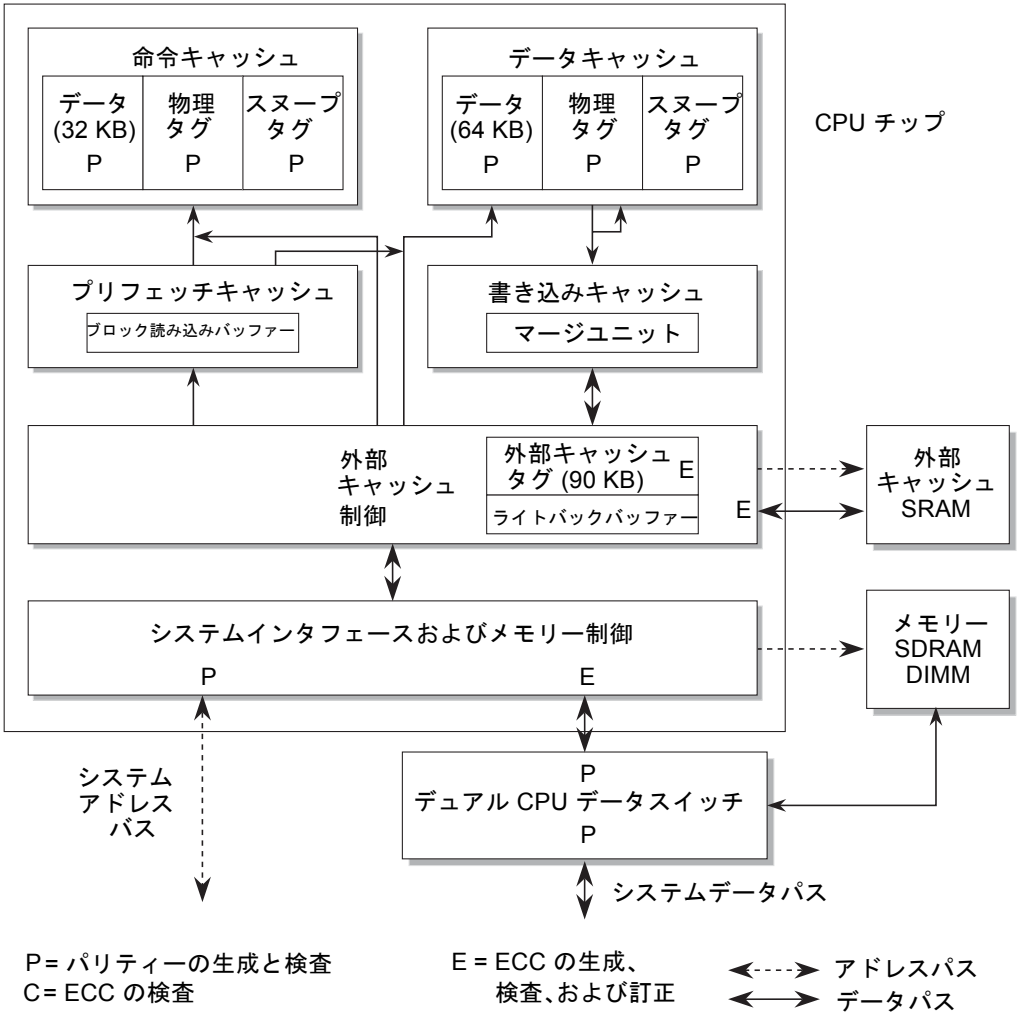


図 3-1　　CPU のエラー検出および訂正

外部キャッシュデータは、8 つの高速 (4 ns) SRAM にあります。シングルビットエラー訂正とダブルビットエラー検出コードが、64 バイト幅のキャッシュラインを保護します。データキャッシュまたは命令キャッシュ中に発生したエラーは、ソフトウェアのフラッシュおよび無効化によって回復されます。システムデータトランザクションの間に発生したエラーは、ハードウェアによって訂正されます。

Sun Fire E25K/E20K システムでは、CPU とアドレスリピータ間のアドレスバス接続はパリティー保護されます。

CPU は、すべての送信データブロックに対してパリティーおよび ECC の両方を生成します。パリティーは、受け取るデュアル CPU データスイッチが検査します。ECC は、転送パスのすべてのデータスイッチユニットが検査します。ECC は、データブロックを受け取った CPU によって検査および訂正されます。

3.2 システムインターコネクトのエラー保護

図 3-2 に、アドレスおよびデータインターコネクトのさまざまなポイントでの保護方法を示します。ブロック図の P、E、C の文字はそれぞれ、パリティーの生成および検査、ECC の検査、受け取ったユニットによる ECC の生成および検査、訂正を意味します。点線はアドレスインターコネクトを示し、実線はデータインターコネクトを示します。

3.2.1 アドレスインターコネクトのエラー保護

Sun Fireplane interconnect アドレスバスには、3 つのパリティーエラービットがあります。バスレベルの保護に加えて、Sun Fire E25K/E20K システムの Sun Fireplane interconnect のアドレスおよび応答クロスバーは、Sun Fireplane interconnect 全体のアドレストランザクションを ECC 保護します。ECC は、シングルビットアドレスエラーを訂正し、ダブルビットエラーを検出します。アドレスパリティーまたは訂正不可能な ECC エラーは、影響を受ける動的システムドメイン内の実行を停止します。

3.2.2 データインターコネクトのエラー保護

すべてのデータインターコネクトトランザクションは、64 バイト幅のデータブロックを移動します。システム装置は、データを発信するときに、装置からの書き込み、または装置の読み取りに対する応答で、ECC を生成します。データを受信したときは、ECC を検査して、シングルビットエラーを訂正します。そのため、すべてのデータはメモリーおよびデータパスエラーから保護されます。

3.2.3 データインターコネクトのエラー分離

データを受信したときに、システム装置が ECC だけを検査するのでは、エラーの原因を診断することは困難です。装置がメモリーへの書き込みに対して問題のある ECC を生成すると、エラーはほかの装置によって検出されますが、エラーの原因を分離することは困難です。エラーの原因を分離するには、次の 2 つの追加の検査を行います。

- 個々の二地点間データ通信は、パリティーによって保護されます。これは、図 3-2 の P で示された箇所です。
- ECC は、各システム装置に送受信されるとき、レベル 1 のデータスイッチによって検査されます。これは、図 3-2 の E で示された箇所です。

データスイッチが実行する ECC 検査は、ほとんどの場合、ECC エラーの原因を特定できます。ECC エラーの訂正が困難になるのは、装置が不正な ECC をメモリーに書き込む場合です。これらのエラーは、かなりあとになってから、ほかの装置がこの場所を読み取ったときに検出されます。問題のある装置の書き込み側が、多くの場所に不正な ECC を書き込み、それらが多くの装置によって読み取られるため、本当のエラーは 1 つの問題のある装置の書き込み側で発生しているのに、エラーが多くの場所にあるように見えます。

データスイッチ ASIC は、各デバイスからほかのデバイスへのすべてのデータの送受信で ECC を検査するため、エラーの発生元を分離できます。たとえば、問題のある装置の書き込み側が別のボード上のメモリーに不正な ECC を書き込むと、2 つのデータスイッチで ECC エラーが検出されます。方向およびトランザクションタグ情報は、どの CPU の組がエラーの発生元で、どの装置が不正な ECC 装置書き込みの対象であるかを特定できます。

問題のある装置の書き込み側が、そのローカルメモリーに不正な ECC を書き込んだ場合、データはデータスイッチを通りません。そのため、同じ CPU またはもう一方の装置が、不正な ECC を持つデータを読み取るまで、問題のある装置の書き込み側は検出されません。どちらの場合でも、ECC エラーの原因は、DCDS (Dual CPU Data Switch) を共有する CPU の組に分離されます。同じ CPU がデータを読み取った場合は、そのボード上のデータスイッチはエラーを検出しません。これは、データがローカル CPU または DCDS によって破壊されたことを示します。別の CPU の組がデータを読み取った場合は、データはデータスイッチを通り、特定の DCDS または関連する CPU から発生したエラーと同様に、ECC エラーが検出されます。

3.2.4 コンソールバスのエラー保護

コンソールバスは、システムコントローラがプライマリデータバスおよびプライマリアドレスバスの健全性に依存することなく、マシンの内部動作にアクセスできるようにするためのセカンダリバスです。これによって、システムコントローラは、システムの主動作の継続を妨げる障害が発生しても動作できます。コンソールバスの動作は、すべてのドメインに共通で、パリティー保護されています。

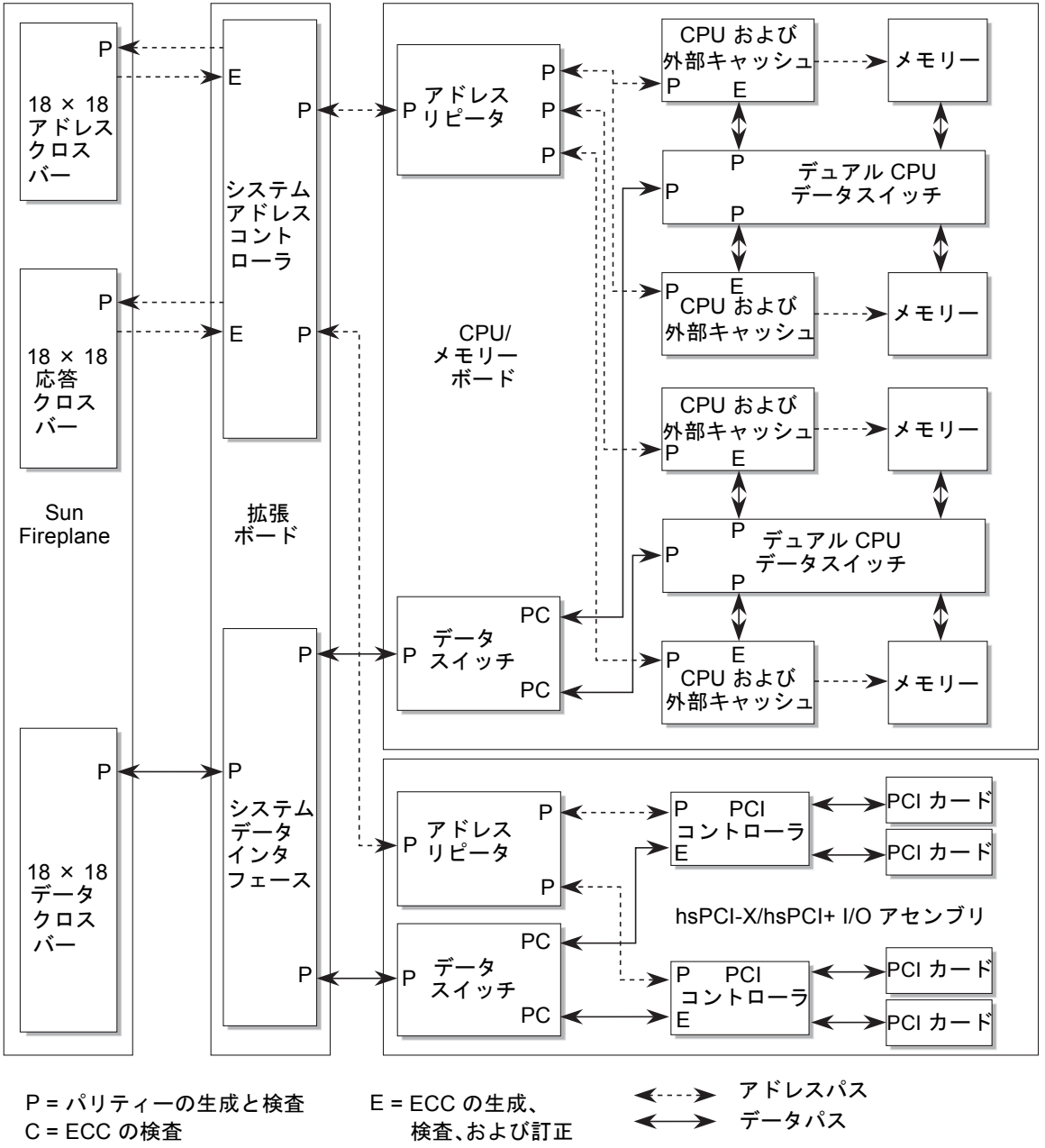


図 3-2　　インターコネクト ECC およびパリティーチェック

3.3 冗長コンポーネント

システムの可用性は、冗長コンポーネントを構成できることで大きく向上しました。必要に応じて、システム内のすべてのホットスワップコンポーネントを冗長構成にすることができます。各システムボードは、独立した運用が可能です。Sun Fire E25K/E20K システムは、複数のシステムボードで構築されるため、構成されたボードのサブセットで運用できます。

冗長システムコンポーネントは次のとおりです。

- CPU/メモリーボード
- I/O アセンブリ
- PCI カード
- システムコントロールボード
- システムクロックソース
- 大容量電源装置
- ファントレー

3.3.1 冗長 CPU/メモリーボード

Sun Fire E25K システムは、最大 18 枚の CPU/メモリーボードで構成できます。Sun Fire E20K システムは、最大 9 枚の CPU/メモリーボードで構成できます。各ボードは、最大 4 つの CPU と、それに関連するメモリーバンクを搭載しています。各 CPU/メモリーボードは独立した運用が可能で、動作中のシステムからホットスワップで取り外して、システムドメイン間で移動できます。システムは、構成されたボードのサブセットで運用できます。

3.3.2 冗長 I/O アセンブリ

Sun Fire E25K システムは、最大 18 枚の I/O アセンブリ (hsPCI-X または hsPCI+) で構成できます。Sun Fire E20K システムは、最大 9 枚の I/O アセンブリで構成できます。各アセンブリは、最大 4 枚の PCI カードをサポートします。I/O アセンブリは、動作中のシステムからホットスワップで取り外して、システムドメイン間で移動できます。

3.3.3 冗長 PCI カード

ホットスワップ交換手順に従って、カードの交換を可能にする特別なカセットを使用すると、Sun Fire E25K/E20K システムの PCI I/O アセンブリに標準の PCI カードを搭載できます。また、システムに複数の周辺装置を接続して、冗長コントローラおよびチャネルを使用可能にすることもできます。ソフトウェアは、複数のパスを保持し、プライマリパスに障害が発生した場合に、代替パスに切り換えることができます。

3.3.4 冗長システムコントロールボード

Sun Fire E25K/E20K システムは、2 枚のシステムコントロールボードを備えています。組み込まれている CPU で実行されるシステムコントローラソフトウェアは、もう一方のシステムコントローラを検査し、状態情報をコピーして、動作中のシステムコントローラに障害が発生した場合に、ほかのシステムコントローラに自動的にフェイルオーバーすることを可能にします。

また、システムは、メインシステムコントロールボードと、ホットスワップによる交換が可能な代替システムコントロールボードを搭載しています。メインシステムコントロールボードは、システムのすべてのシステムコントローラ資源を提供します。メインシステムコントロールボードでハードウェアまたはソフトウェアの障害が発生した場合、またはメインシステムコントロールボードからほかの装置へのハードウェアコントロールパス (コンソールバスインタフェース、Ethernet インタフェース) で障害が発生した場合には、システムコントローラのフェイルオーバーソフトウェアが自動的に予備のシステムコントロールボードへのフェイルオーバーをトリガーします。予備のシステムコントロールボードは、メインシステムコントロールボードの役割を引き継ぎ、メインシステムコントローラのすべての作業を引き受けます。システムコントローラのデータファイル、構成ファイル、ログファイルは、両方のシステムコントロールボードに複製されます。

3.3.5 冗長システムクロック

Sun Fire E25K/E20K システムには、冗長システムクロックがあります。1 つのシステムコントロールボードのシステムクロックに障害が発生すると、障害の発生したシステムコントロールボードを交換するために停止するまで、そのクロックラインのコンシューマはもう 1 枚のシステムコントロールボードからクロック資源を取得し続けます。

3.3.6 冗長電源

Sun Fire E25K/E20K システムのキャビネットは、6 つの 4-kW デュアル AC-DC 電源装置を使用します。各 AC 電源装置には 2 本の電源ケーブルがつながっているので、それぞれ別の電源に接続できます。これらの装置は、入力電源を N+1 冗長の 48 VDC に変換します。そのため、必要に応じて、システムは障害の発生した電源装置があっても動作を続行できます。

システムの動作中に電源装置を交換できます。電力は、個別の DC 回路遮断器を介して個々のシステムボードに供給されます。各ボードセットは独自のオンボード電圧コンバータを持ち、48 VDC からオンボードの論理コンポーネントが必要とするレベルに変圧します。DC/DC コンバータに障害が発生した場合は、その特定のシステムボードだけに影響します。

3.3.7 冗長ファン

システムボードの上部に 4 つ、下部に 4 つのファントレーがあります。各ファントレーには 2 層の 6 インチファンが取り付けられています。ファンには、標準と高速の 2 種類の速度があります。過熱したコンポーネントが感知された場合は、すべてのファンが高速に切り替わります。1 つのファンに障害が発生した場合は、トレーの対応する層の冗長ファンが高速に切り替わります。ファンは N+1 の冗長性を持つので、システムは障害の発生したファンがあっても動作できます。ファントレーは、システムの動作中にホットスワップできます。

3.4 再構成可能な Sun Fireplane interconnect

Sun Fire E25K/E20K システムは、Sun Fireplane interconnect に、アドレス、応答、およびデータ用の 3 つの独立クロスバーを実装しています。Sun Fireplane interconnect には、20 の ASIC があります。これは、システム内で唯一のホットスワップできない論理コンポーネントです。障害の発生した Sun Fireplane interconnect ASIC は動作中のシステムから取り外せないため、3 つの Sun Fireplane interconnect クロスバーをそれぞれ個別に縮退モードで構成できます。縮退モードは、各システムドメインで個々に構成できます。

3.5 自動システム回復

適切に構成されたシステムは、障害発生後常に再起動します。システムコントローラが障害を特定し、障害の発生した CPU、メモリー、I/O、またはインターコネクトコンポーネントを除いてシステムを再構成し、オペレーティングシステムを再起動します。

システムコントローラは、明らかな重大エラービットを持つ部品だけを構成します。使用しているマシンまたは別のマシンによってすでに障害を検出されている FRU (現場交換可能ユニット) は、使用しないでください。

3.5.1 組み込み自己診断

ASIC の組み込み自己診断 (BIST) ロジックは、同じシステムクロックレートで擬似ランダムパターンを適用し、組み合わせ論理によって高い確率で障害を検出します。ローカルの BIST は各 ASIC 内で動作し、ASIC が正しく動作していることを確認します。インターコネクト組み込み自己診断 (IBIST) はインターコネクトテストを実行して、ASIC がインターコネクトを介して通信できることを確認します。ローカルの BIST は、既知のテストデータを相互に送信する各 ASIC のインタフェースに依存します。

3.5.2 電源投入時自己診断

電源投入時自己診断 (POST) は、最初に各論理ブロックを個別に評価し、徐々に範囲を広げてシステムを評価します。障害が発生したコンポーネントは、Sun Fireplane interconnect から切り離されます。その結果、この自己診断に合格し、エラーのない状態で操作できる論理ブロックだけを使用して、システムが起動されます。

ローカルの POST は各 CPU で実行され、システム POST はシステムコントローラで実行されます。

3.6 システムコントローラ

システムコントローラは、Sun の可用性技術の中心です。このコントローラには、UltraSPARC-IIi システムを組み込んだ、既成の SPARCengine Netra 2140 6U cPCI ボードが搭載されています。このボードは、Solaris ソフトウェアおよびシステム管理ソフトウェアを実行します。

システムコントローラは、JTAG (Joint Test Action Group) によってマシン内の主なチップのレジスタにアクセスし、マシンの状態を継続的に監視します。問題が検出されると、システムコントローラはどのハードウェアに障害が発生したかを判断し、そのハードウェアが交換されるまで、使用されないようにします。

システムコントローラの主な機能は次のとおりです。

- システムの設定および起動処理の調整によるシステムの構成
- システムパーティションおよびドメインの設定
- システムクロックの生成
- システム全体の環境センサーの監視
- エラーの検出と診断および回復
- プラットフォームコンソール機能およびドメインコンソールの提供
- システムログを介して syslog ホストに至るメッセージのルーティングを提供

3.6.1 コンソールバス

コンソールバスは、システムコントローラがシステムアドレスバスおよびデータバスの健全性に依存することなく、システムの内部動作にアクセスできるようにするためのセカンダリバスです。これによって、システムコントローラは、システムの動作の継続を妨げる障害が発生しても動作できます。システムコントローラは、パリティー保護されています。

3.6.2 環境監視

システムコントローラは定期的にシステムの環境センサーを監視し、起こりうる状況を事前に警告します。これによって、マシンは正常に停止されて、システムへの物理的損傷およびデータ破壊を防ぎます。

監視される環境項目は次のとおりです。

- 電源の状態
- 電圧
- ファンの速度
- 温度
- 装置の障害
- 装置の有無

3.7 並行保守性

Sun Fire E25K/E20K システムのもっとも重要な保守性機能は、「並行保守」としてシステムボードをオンラインのまま交換することです。これは、動作中のシステムを停止せずに、マシンのさまざまな部品を保守できる機能です。障害の発生しているコンポーネントは、FRU を明示する障害ログによって特定されます。Sun Fireplane interconnect、電源センタープレーン、ファンバックプレーン、電源モジュールを除き、システムのすべてのボードおよび電源装置は、システムの動作中にホットスワップ交換手順を使用して取り外しおよび取り付けができます。システムを停止する必要はありません。また、メインシステムの動作を妨げることなく、現在動作中のシステムコントロールボードを交換したり、冗長システムコントロールボードに制御を切り替えることができます。

停止時間なしでこれらを修復する機能は、より高い可用性の実現に大きく貢献します。また、オンラインでのシステムの修復が可能なので、設置されたハードウェアをアップグレードできる利点もあります。ユーザーが、メモリーや予備の I/O コントローラの追加を希望することがあります。これらの操作をオンラインの状態で実行できるため、影響を受けるシステムボードを一時的にサービスから除外することによるほんの短時間 (そしてわずかな) 性能の低下だけで済みます。

並行保守は、次のハードウェア設備の機能です。

- すべての Sun Fireplane interconnect 接続はポイントツーポイントのため、システムを動的に再構成することによって、システムボードを論理的に分離できます。
- Sun Fire E25K/E20K システムは、分散型 DC 電力システムを使用しています。各システムボードは独自の電源装置を備え、各システムボードへの電源投入または切断を個々に行えるようになっています。
- オフボードの Sun Fireplane interconnect に接続する ASIC にはすべて、ループバックモードがあり、システムに動的に再構成される前に、システムボードを検査できます。

3.7.1 システムボードの動的再構成

システムボードを動作中のシステムから取り外して交換することを「動的再構成」と言います。たとえば、CPU の 1 つに障害の発生したボードでも、システムに構成できます。システムを停止させずにモジュールを交換するため、動的再構成はボードをシステムから分離し、ホットスワップ手順を使用してボードを交換できるようにします。この動的再構成の操作には、3 つの手順があります。

- 動的切り離し
- ホットスワップ
- 動的接続

動的再構成によって、現在システムが使用していないボードが、システムに資源を供給することが可能になります。この機能をホットスワップ交換とともに使用すると、システムを停止しないでアップグレードしたり、1 つのドメインから別のドメインへ資源を移動したりできます。また、システムによって構成解除され、その後ホットスワップおよび修復または交換された障害のあるモジュールを交換する場合にも使用できます。

動的構成解除および再構成は、システムコントローラを使用して作業できるシステム管理者 (または保守プロバイダ) が行います。次に、構成変更およびホットスワップ交換の手順を示します。

1.Solaris オペレーティング環境のスケジューラに、対象のボードで新しい処理を起動しないように通知します。一方、実行中の処理および入出力操作は終了し、メモリーの内容をほかのメモリーバンクにふたたび書き込みます。

2.代替 I/O パスへのスイッチオーバーが行われるので、I/O アセンブリが取り外されても、システムはデータへのアクセスを続行します。

3.システム管理者は、構成解除されたシステムボードを手動でシステムから取り外して、ホットスワップ操作を実行します。取り外し処理は、システムコントローラによって制御されるので、システム管理者はソフトウェアの指示に従います。

4.取り外したシステムボードを、修復、交換、またはアップグレードします。

5.新しいボードをシステムに挿入します。

6.スワップされたシステムボードは、挿入時に、オペレーティングシステムによって動的に再構成されます。入出力を元の状態に切り替えると、スケジューラは新しい処理を割り当て、メモリーへの書き込みが始まります。

動的再構成をホットスワップ交換とともに使用して、Sun Fire E25K/E20K システムの修復およびアップグレードを行うと、ユーザーの不利益を最小限に抑えることができます。オンサイトでシステムボードを交換することによって、ハードウェアのホットスワップ交換に要する時間を分単位にまで低減できます。

ハードウェアの動的再構成およびホットスワップ交換のもう 1 つの利点は、オンラインの状態でシステムのアップグレードを実行できることです。たとえば、ユーザーが追加のシステムボードを購入した場合、動作を中断せずに、そのボードもシステムに追加できます。

3.7.2 システムコントロールボードセットの取り外しおよび取り付け

システムクロックを供給していないホットスペアのシステムコントロールボードセットは、動作中のシステムから取り外すことができます。

3.7.3 大容量電源装置の取り外しおよび取り付け

大容量の 4-kW デュアル AC/DC 電源装置は、システムを停止せずにホットスワップできます。交換中は、残りの電源装置がシステムに電力を供給します。

3.7.4 ファントレーの取り外しおよび取り付け

ファンに障害が発生した場合、システムコントロールは別の層にある対応するファンを高速動作に切り替えて、減少した通気を補います。このような状況下でも、障害が発生したファンアセンブリの保守が終わるまで、システムは正常に動作するように設計されています。ファントレーは、システムを中断せずにホットスワップできます。

3.7.5 遠隔保守

予定にない再起動やエラーログ情報を、自動的にユーザーの保守管理部門に電子メールで報告する機能をオプションで使用できます。すべてのシステムコントローラには遠隔アクセス機能があり、システムコントローラに遠隔ログインできます。この遠隔接続によって、すべてのシステムコントローラの診断を実行できます。Solaris ソフトウェアがほかのシステムボードで動作している間、構成解除されたシステムボード上で遠隔またはローカルの診断を実行できます。

第4章

システムインターコネクト

この章では、Sun Fireplane interconnect について説明します。

図 4-1 に、Sun Fire E25K/E20K システムのインターコネクトの全体図を示します。図中の小さな数値は、インターコネクトの各レベルの最大のデータ帯域幅を示します。

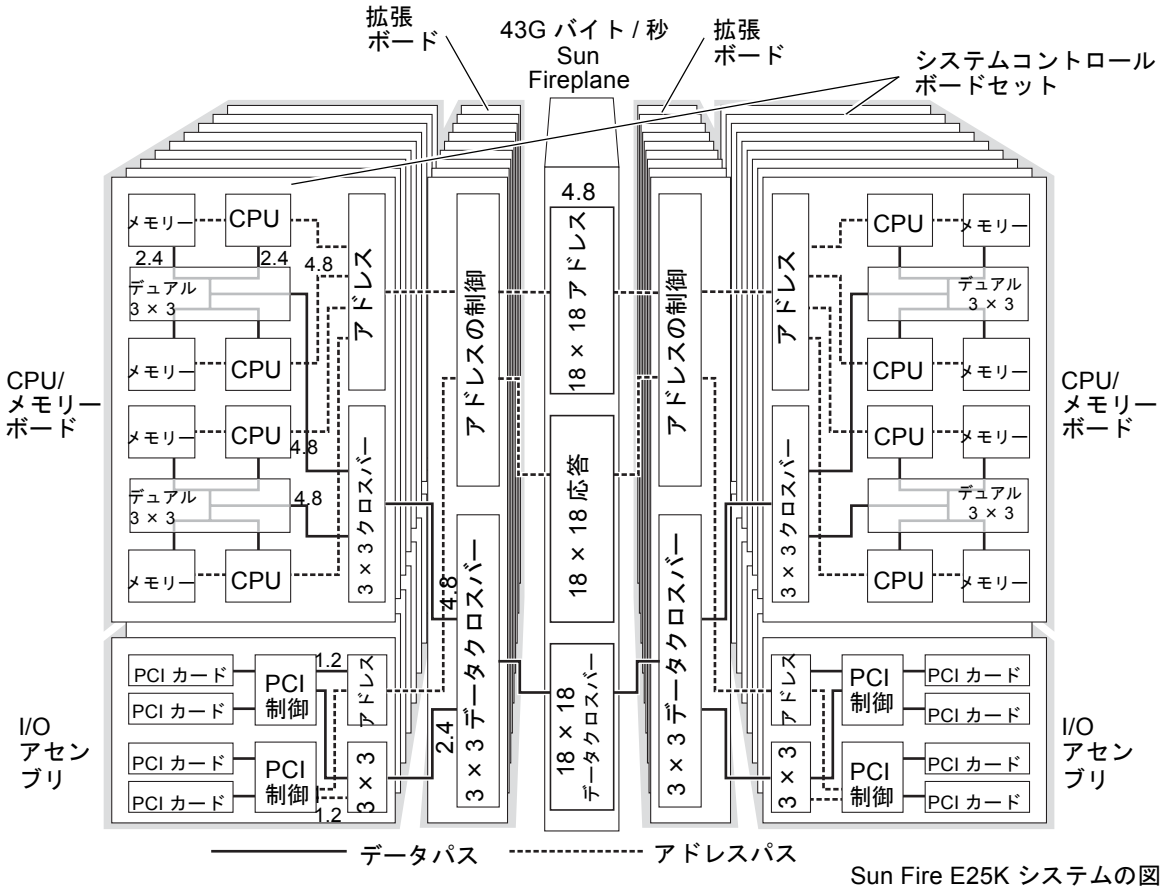


図 4-1　Sun Fire E25K/E20K システムのインターコネクト

4.1 データ転送インターコネクトのレベル

Sun Fire E25K/E20K システムのインターコネクトは、いくつかの物理層に実装されます (図 4-2)。大規模なサーバーの機能ユニット (CPU/メモリーユニット、I/O コントローラ) のすべてを直接接続することは、物理的な装置の配置を考えると現実的ではありません。サーバーのシステムインターコネクトは、チップがボードに接続し、ボードが Sun Fireplane interconnect に接続するレベルの階層として実装されます。同じボード上のコンポーネント間では、別のボードのコンポーネントとの間に比べて接続が多いため、応答時間は短く、帯域幅は高くなります。

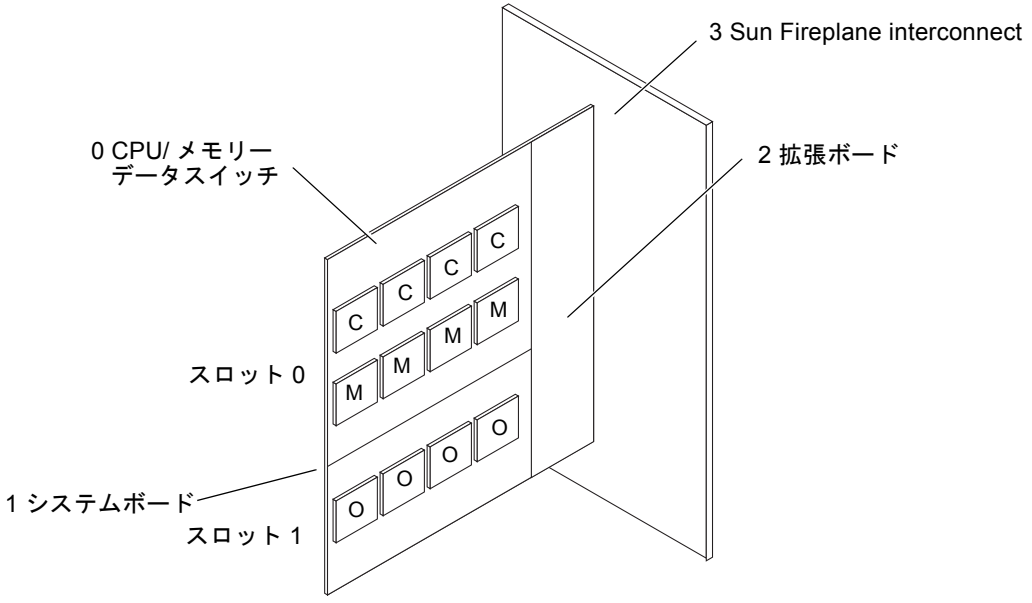


図 4-2　Sun Fire E25K/E20K システムのデータ転送インターコネクトのレベル

システムは、2 つの個別のインターコネクトを備えています。1 つはアドレスインターコネクト、もう 1 つはデータ転送インターコネクトです (表 4-1)。

- アドレスインターコネクトは、次の 3 レベルの階層になっています。

  A 各 CPU/メモリーボードまたは I/O アセンブリのアドレスリピータは、そのボード上のデバイスからのアドレス要求を収集し、それらを拡張ボードのシステムアドレスコントローラに転送します。

  B 各ボードセットの拡張ボードは、一貫性のある毎秒 1 億 5 千万スヌープの帯域幅を持つスヌープアドレスバスを備えています。

  C 18×18 の Sun Fireplane interconnect アドレスおよび応答クロスバーには、最大で毎秒 13 億の要求および 13 億の応答を処理できる帯域幅があります。

- データ転送インターコネクトには、図 4-2 に示すとおり、4 レベルのクロスバーの階層があります。

  0 2 つの CPU/メモリーの組は、3 つの 3×3 スイッチによってボードレベルのクロスバーに接続されます。

  1 各 CPU/メモリーボードは、システムポートと 2 組の CPU の間に 3×3 クロスバーを持っています。各 I/O ボードは、システムポートと 2 つの PCI バスコントローラの間に 3×3 クロスバーを持っています。

  2 各拡張ボードは、Sun Fireplane interconnect ポートと 2 つのシステムボードの間に 3×3 クロスバーを持っています。

  3 18×18 の Sun Fireplane interconnect データクロスバーは、18 ボードセットそれぞれに対して 4.8G バイト/秒のポートを持ち、合計で 43G バイト/秒の帯域幅を持ちます。

Sun Fire E25K/E20K システムは、2 枚のボードを Sun Fireplane interconnect ポートに接続する、追加のインターコネクトレベルを持っています。このインターコネクトは拡張ボードです。

表 4-1 インターコネクトレベル

インターコネクト	レベル	説明
アドレス インターコネクト	A ボードセット :	スヌープバスセグメント
	B 拡張ボード :	スヌープバスセグメント
	C Sun Fireplane interconnect :	ポイントツーポイントトランザクションのための 2 つの 18 ポートスイッチ
データ転送 インターコネクト	0 CPU/ メモリー :	2 つの 3 ポートスイッチ
	1 ボードセット :	3 ポートスイッチ
	2 拡張ボード :	3 ポートスイッチ
	3 Sun Fireplane interconnect :	18 ポートスイッチ

Sun Fire E25K/E20K システムでは、またぐ必要のあるロジックのレベルが少ないため、同じボード上のメモリーへの応答時間がもっとも短くなります。

4.2 アドレスインターコネクト

Sun Fire E25K/E20K システムのアドレスインターコネクトには、次の 3 レベルのチップがあります (図 4-3)。

- **ボードセットレベル。**アドレスリピータは、オンボードの CPU および I/O コントローラからのアドレストランザクションを収集し、それらに対してブロードキャストします。
- **拡張ボードレベル。**システムアドレスコントローラのレベル B のアドレスリピータは、2 枚のボードからアドレス要求を収集し、それらに対してブロードキャストします。Sun Fireplane interconnect のアドレスおよび応答クロスバーを介して、ほかの拡張ボードへグローバルアドレストランザクションを送信します。
- **Sun Fireplane interconnect レベル。**18×18 Sun Fireplane interconnect アドレスおよび応答クロスバーは、18 のシステムアドレスコントローラを接続します。

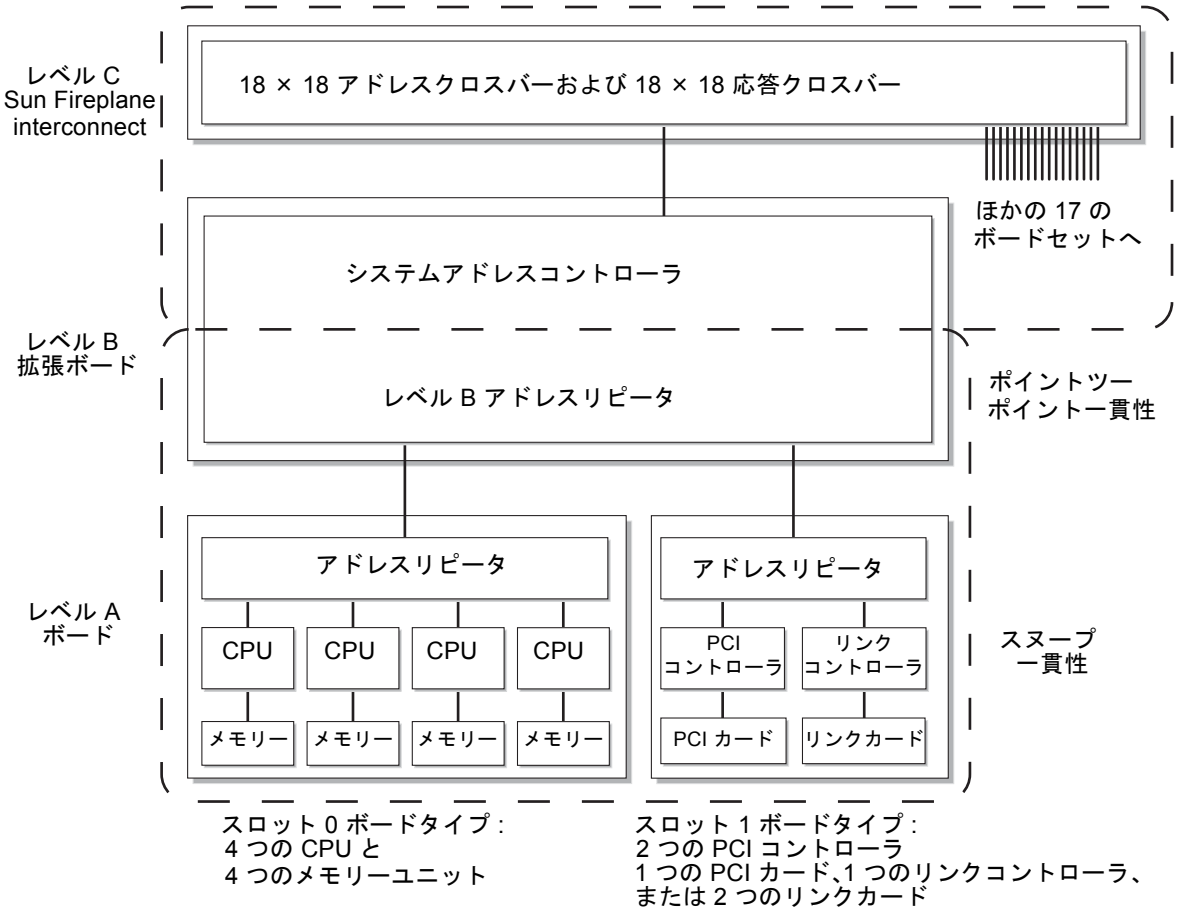


図 4-3　　アドレスインターコネクトレベル

アドレスは、CPU から別のボードのメモリーコントローラに到着するまでに、5 つのチップを通過します。同じボード上のメモリーに送信されるアドレスは、Sun Fireplane interconnect のアドレス帯域幅を消費しません。

4.3 データインターコネクト

Sun Fire E25K/E20K システムのデータインターコネクトには、次の 4 レベルのチップがあります (図 4-4 を参照)。

レベル 0 − CPU/メモリーレベル。5 ポートのデュアル CPU データスイッチは、2 組の CPU/メモリーをボードデータスイッチに接続します。CPU およびメモリーユニットは、それぞれ 2.4G バイト/秒で接続され、もう 1 つの CPU およびメモリーユニットと 4.8G バイト/秒のボードデータスイッチへの接続を共有します。

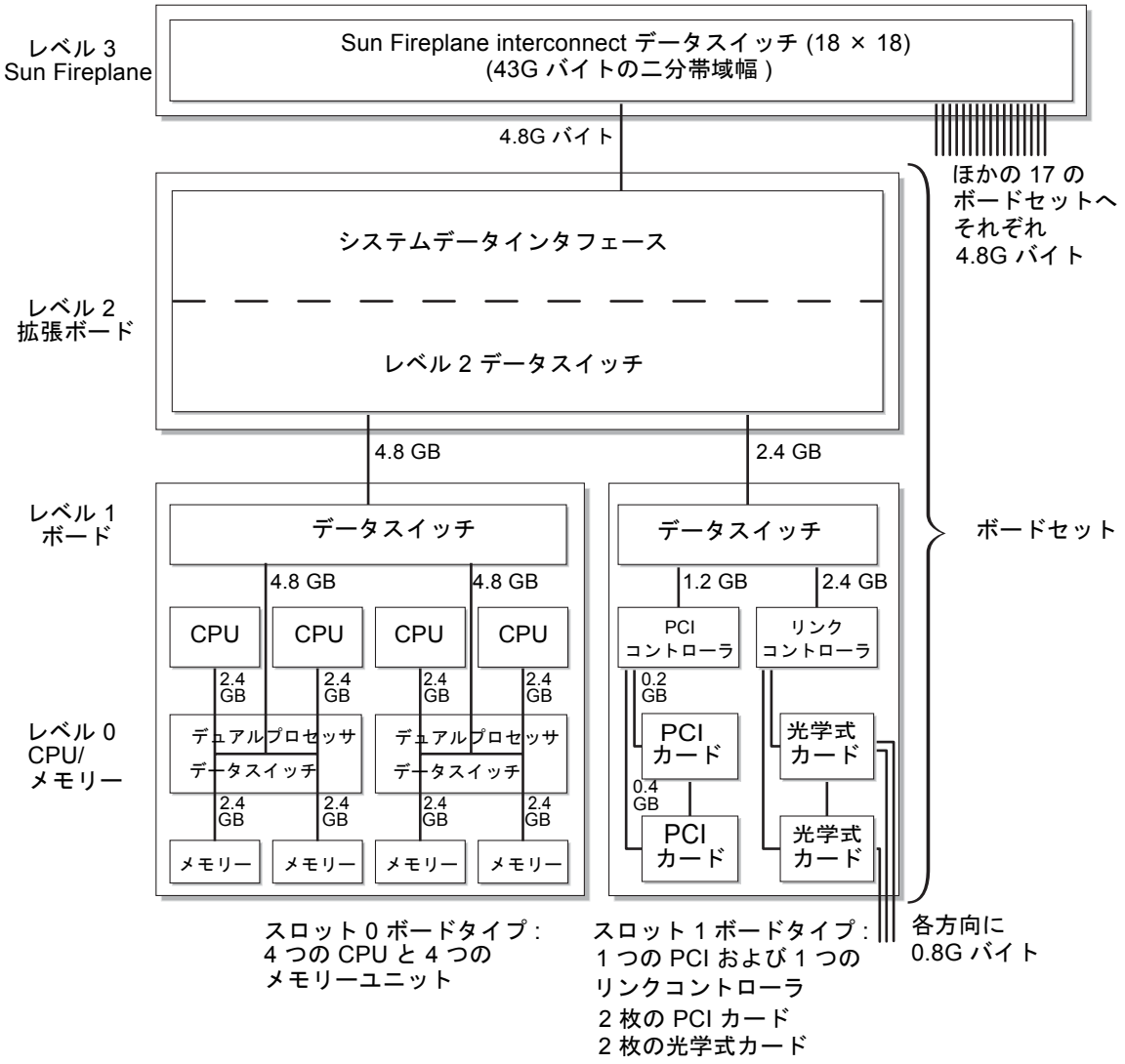
レベル 1 − ボードレベル。3 ポートのボードデータスイッチは、オンボードの CPU または I/O インタフェースを、拡張ボードデータスイッチに接続します。スロット 0 ボードには 4.8G バイト/秒のスイッチがあり、スロット 1 ボードには 1.2G バイト/秒および 2.4G バイト/秒のスイッチがあります。

レベル 2 − 拡張ボードレベル。3 ポートのシステムデータインタフェースは、2 枚のボードをシステムデータクロスバーに接続します。スロット 0 ボード (4 つの CPU およびメモリー) は 4.8G バイト/秒で接続し、スロット 1 ボード (hsPCI-X または hsPCI+) は 2.4G バイト/秒で接続します。

レベル 3 − Sun Fireplane interconnect レベル。18×18 の Sun Fireplane interconnect クロスバーは 32 バイト幅で、43G バイト/秒のシステム二分帯域幅があります。

データは、1 枚のボードのメモリーから別のボードの CPU に到着するまでに、7 つのチップを通過します。同じボードセットのメモリーへのアクセスは、Sun Fireplane interconnect のデータ帯域幅を消費しません。

図 4-4 の中の数字は、各レベルの最大帯域幅を示します。すべてのデータパスは双方向になっています。各パスの帯域幅は、機能ユニットへのトラフィックと機能ユニットからのトラフィックの間で共有されます。



G バイトの数字は、インターコネクトの各部分の最大帯域幅です。

図 4-4　データインターコネクトレベル

4.4 インターコネクトの帯域幅

この節では、Sun Fire E25K/E20K システムのインターコネクトの応答時間および帯域幅を数字で示します。帯域幅とは、データのストリームが送信されるレートです。表 4-2 に、インターコネクトの実装によって制限される最大メモリー帯域幅を示します。メモリーは、1 枚のボード上に 4 つあるメモリーユニットで 16 ウェイインタリーブされることを想定しています。

表 4-2 インターコネクトの最大帯域幅

メモリーアクセス	Sun Fire E25K システムのメモリー帯域幅	Sun Fire E20K システムのメモリー帯域幅
リクエスタと同じ CPU	9.6G バイト/秒×ボードセット数 18 ボードセットで最大 172.8G バイト/秒	9.6G バイト/秒×ボードセット数 9 ボードセットで最大 86.4G バイト/秒
リクエスタと同じボード	6.7G バイト/秒×ボードセット数 18 ボードセットで最大 120.6G バイト/秒	6.7G バイト/秒×ボードセット数 9 ボードセットで最大 60.3G バイト/秒
リクエスタと異なるボード	2.4G バイト/秒×ボードセット数 18 ボードセットで最大 43.2G バイト/秒	2.4G バイト/秒×ボードセット数 9 ボードセットで最大 21.6G バイト/秒
ランダムなデータ位置	47.0G バイト/秒	23.5G バイト/秒

同じボードの最大帯域幅：すべてのメモリーアクセスが、リクエスタと同じボード上のメモリーに行われる場合に発生します。

同じボードの最大帯域幅は、ボード 1 枚につき 9.6G バイト/秒です。次のいずれかの場合に、最大帯域幅となります。

- すべての CPU が、自身のローカルメモリーにアクセス
- すべての CPU が、対になっている CPU のメモリーにアクセス
- 2 つの CPU が、ローカルメモリーと、ボードの残り半分にある 2 つのアクセスメモリーにアクセス

同じボードの最小帯域幅は、ボード 1 枚につき 4.8G バイト/秒です。これは、4 つすべての CPU がボードの残り半分にあるメモリーにアクセスする場合に発生します。メモリーが 16 ウェイインタリーブされた場合 (通常のケース)、同じボードの最大帯域幅は、ボード 1 枚につき 6.7G バイト/秒です。

オフボードの帯域幅：オフボードのデータパスは、32 バイト幅×150 MHz で、4.8G バイト/秒になります。この帯域幅は、ボードの CPU から送信された要求と、ほかの CPU から受信するメモリーの要求の両方に対応するため、ボードあたりの二分帯域幅は 2.4G バイト/秒に 2 分割されます。

4.5 インターコネクトの応答時間

応答時間とは、単一のデータ項目がメモリーから CPU に送信される時間のことです。応答時間には何種類かの算出、測定方法があります。次に、2 つの応答時間について説明します。

- ピン間の応答時間 : インターコネクトの論理サイクルから算出します。CPU のデータ処理には依存しません。
- 連続ロードの応答時間 : lmbench ベンチマークのカーネルによって測定します。

これらの応答時間は、1 つの CPU がメモリーにアクセスする場合の最良の例を表します。

ピン間の応答時間は、CPU がアドレスを要求してから、CPU へのデータ転送が完了するまでに要する、インターコネクト論理設計上のクロックを計算することによって算出します (表 4-3 および表 4-4 を参照)。

表 4-3 メモリー内のデータのピン間の応答時間

メモリーの位置	クロック数	CDC[1] ヒット	応答時間が増加する条件[2]
同じボード (リクエスタのローカルメモリー)	180 ns、27 クロック	—	
同じボード (同じデュアル CPU データスイッチ上の別の CPU)	193 ns、29 クロック	—	
同じボード (データスイッチのもう一方の側)	207 ns、31 クロック	—	
その他のボード	333 ns、50 クロック	あり	2、3
	440 ns、66 クロック	なし	3

1 一貫性ディレクトリキャッシュ (Coherency Directory Cache)

2 条件 1 データがスロット 1 (I/O またはデュアル CPU ボード) から着信 1 サイクル 7 ns
条件 2 データをスロット 1 (I/O またはデュアル CPU ボード) へ送信 2 サイクル 13 ns
条件 3 アドレスが共有ボードから着信または共有ボードセットへ送信 2 サイクル 13 ns
条件 4 スレーブアドレスが共有ボードセットから着信または共有ボードへ送信 2 サイクル 13 ns
条件 5 CDC のミスでホーム応答が共有ボードセットから着信または共有ボードセットへ送信 2 サイクル 13 ns
条件 6 CDC のミスでスレーブ応答が共有ボードセットから着信または共有ボードセットへ送信 2 サイクル 13 ns

表 4-4　キャッシュ内のデータのピン間の応答時間

キャッシュの位置	クロック数	CDC[1] ヒット	応答時間が増加する条件[2]
リクエスタボード (Sun Fire E25K/E20K システム : ホームボードセットのリクエスタ)	280 ns、42 クロック	—	
ホームボード	407 ns、61 クロック	あり	1、2、3
	440 ns、66 クロック	なし	3、5
その他のボード	473 ns、71 クロック	あり	1、2、3、4
	553 ns、83 クロック	なし	3、4、6

1 一貫性ディレクトリキャッシュ (Coherency Directory Cache)

2 条件 1 データがスロット 1 (I/O またはデュアル CPU ボード) から着信　1 サイクル　7 ns
条件 2 データをスロット 1 (I/O またはデュアル CPU ボード) へ送信　2 サイクル　13 ns
条件 3 アドレスが共有ボードから着信または共有ボードセットへ送信　2 サイクル　13 ns
条件 4 スレーブアドレスが共有ボードセットから着信または共有ボードへ送信　2 サイクル　13 ns
条件 5 CDC のミスでホーム応答が共有ボードセットから着信または共有ボードセットへ送信　2 サイクル　13 ns
条件 6 CDC のミスでスレーブ応答が共有ボードセットから着信または共有ボードセットへ送信 2 サイクル　13 ns

第5章

システムコンポーネント

この章では、Sun Fire E25K/E20K システム内の主なコンポーネントについて説明します (図 5-1)。

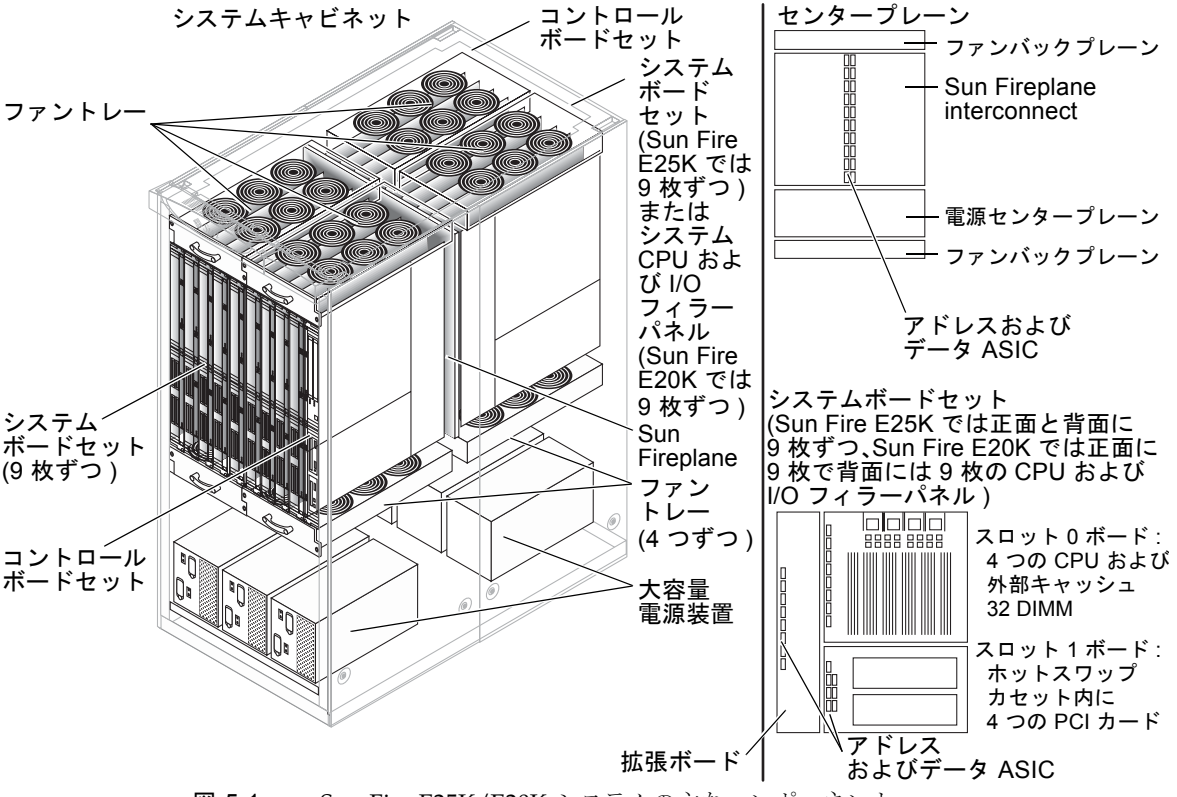


図 5-1　Sun Fire E25K/E20K システムの主なコンポーネント

5.1 キャビネット

Sun Fire E25K/E20K システムは、1 つのシステムキャビネットと、ユーザーの選択による 1 つ以上の I/O 拡張ラックの、合計 2 つ以上の空冷式キャビネットで構成されます (図 5-2)。システムキャビネットには、CPU/メモリーとシステム制御用周辺装置が含まれます。

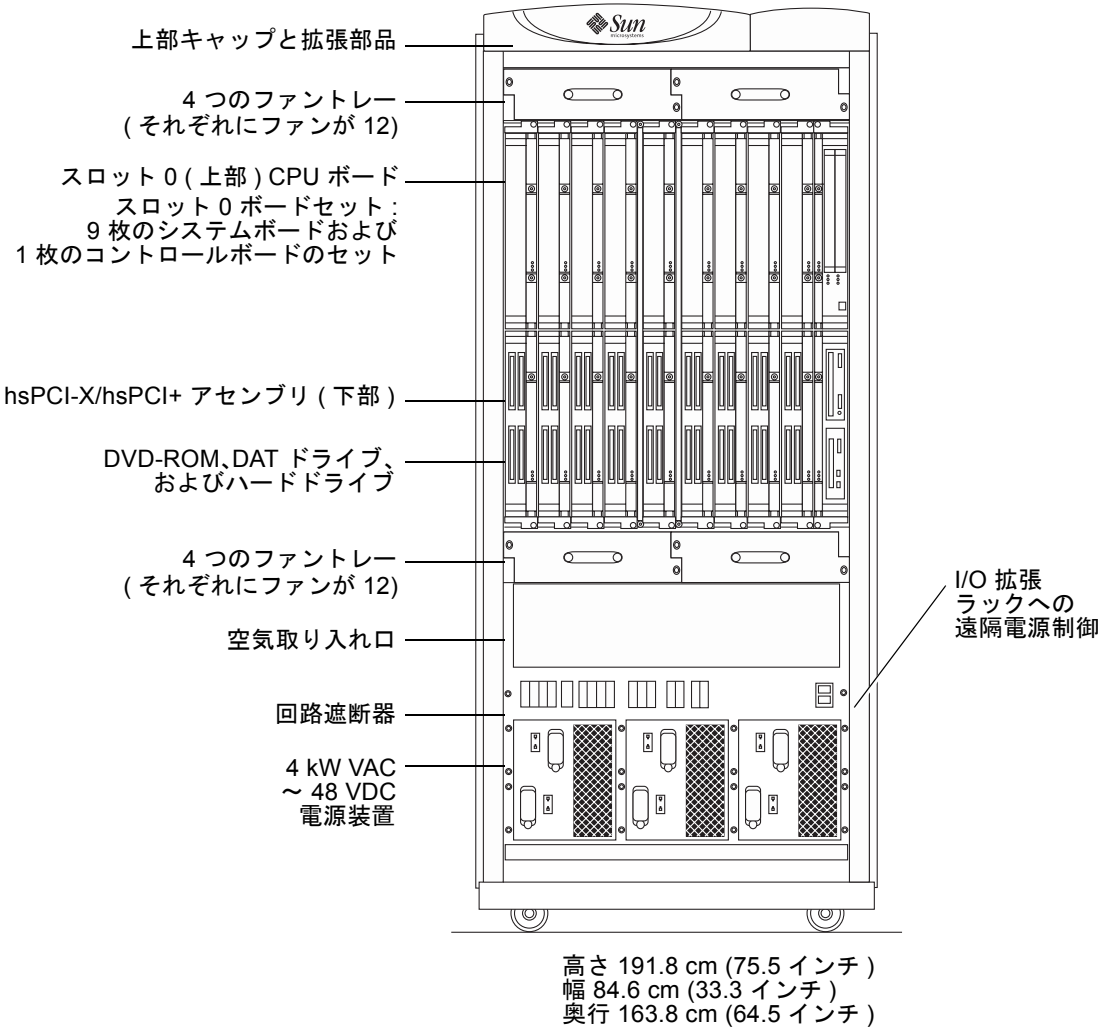


図 5-2　Sun Fire E25K/E20K システムのキャビネット ― 正面図

システムキャビネットは、8 つのファントレー、6 つの大容量電源装置、2 つのシステムコントロールボードセットで構成されます。コントロールボードセットは、RAS 機能を実行します (5-11 ページの 5.3.2 節「コントローラボードセット」を参照)。

Sun Fire E25K システムでは、システムごとに最大 18 のシステムボードセットを構成でき、システムボードセットの数によって CPU の数と記憶容量が決まります。Sun Fire E20K システムでは、システムごとに最大 9 のシステムボードセットを構成でき、システムボードセットの数によって CPU の数と記憶容量が決まります (5-7 ページの 5.3.1 節「システムボードセット」を参照)。

Sun Fire E25K システムのフル構成されたキャビネットの重量は、1,121.7 kg (2,467.8 ポンド) です。Sun Fire E20K システムのフル構成されたキャビネットの重量は、987.0 kg (2,141.0 ポンド) です。

5.1.1 システムの電源

Sun Fire E25K システムは、200 ～ 240 VAC、周波数 47 ～ 63 Hz の単相電力で動作します。システムキャビネットには、12 の 30 A 回路が必要で、それらは通常 2 つの別々の電源に接続します。設置場所の電源ソケットは、北米および日本では NEMA L6-30P で、その他の国々では IEC 309 です。システムと建物の電源ソケットをつなぐ電源ケーブルは、システムに付属しています。

システムキャビネットは、6 つのデュアル入力 4-kW デュアル AC/DC 大容量電源装置を使用します。各装置には 2 本の電源ケーブルが接続します。これらの装置は、入力電源を 48 VDC に変換します。システムは、大容量電源装置に障害が発生しても動作可能で、その大容量電源装置はシステムの動作中に交換できます。

電力は、別々の DC 回路遮断器を使用して個々のボードに供給されます。各ボードは独自のオンボード電圧コンバータを持ち、48 VDC からオンボードの論理コンポーネントが必要とするレベルへの変圧を行います。DC/DC コンバータに障害が発生した場合は、そのシステムボードだけに影響があります。

5.1.2 システムの冷却

Sun Fire E25K/E20K システムの動作環境に対する制限は、次の項目だけです。

- 温度 : 10 ～ 35°C (50 ～ 90°F)
- 湿度 : 20 ～ 80%
- 高度 : 3,048 m (10,000 フィート) 以下

フル構成のシステムの電力消費量は 24 kW で、空調の負荷は Sun Fire E25K システムでは約 81,352 BTU/時、Sun Fire E20K システムでは約 44,081 BTU/時です。構成が小さい規模になると、消費電力は少なくなります。

1 台の Sun Fire E25K システムまたは 1 台の Sun Fire E20K システムの放熱に対応するには、システム装置の下に有孔タイルを置く必要があります。各タイルは、毎分 17.0 立方メートル (600 立方フィート) の冷却用空気を通気する必要があります。フル構成のシステムキャビネットは、並べて設置できます。詳細は、『Sun Fire E25K/E20K システムサイト計画の手引き』を参照してください。

空気は、システムキャビネットの下部、正面、および背面にある空気取り入れ口から入り、上部に抜けます。システムボードの上に 4 つ、下に 4 つのファントレーがあります。ファンには 2 種類の速度があり、通常は高速で動作します。コンポーネントが過熱すると、ファンは高速に切り替わります。ファンに障害が発生してもシステムは動作可能で、そのファントレーは、システムの動作中にホットスワップできます。

5.2 センタープレーン

図 5-3 に、Sun Fire E25K/E20K システムの片側のボードおよびファントレーが、どのようにファンバックプレーン、電源センタープレーン、および Sun Fireplane interconnect に接続しているかを示します。

スロット 0 ボードおよびスロット 1 ボードは、拡張ボードの付いたシステムキャリアプレートに接続し、システムキャリアプレートは Sun Fireplane interconnect に接続します。このユニットをボードセットと呼びます (5-6 ページの 5.3 節「システムボード」を参照)。

Sun Fire E25K システムでは、9 枚のシステムボードセットは、システムキャリアプレートおよび拡張ボードとともに Sun Fireplane interconnect の両側のスロット 0 ～ 8 (正面) およびスロット 9 ～ 17 (背面) に接続します。Sun Fire E20K システムでは、9 枚のシステムボードセットは、システムキャリアプレートおよび拡張ボードとともに Sun Fireplane interconnect の正面のスロット 0 ～ 8 に接続し、スロット 9 ～ 17 (背面) には 9 枚の CPU および I/O フィラーパネルが取り付けられます。2 つのシステムコントローラボードセット (システムコントロールボードおよびシステムコントロール周辺装置ボード) は、どちらのシステムでも、システムコントロールキャリアプレートおよびセンタープレーンサポートボードとともに Sun Fireplane interconnect の両側の、スロット SC0 (正面) およびスロット SC1 (背面) に接続します。すべてのボードセットへの電源は、Sun Fireplane interconnect の下部にある電源センタープレーンから供給されます。

Sun Fireplane interconnect には、システムコントローラボードセット専用のスロットが 2 つ (右側の正面と背面に) あります。システムコントローラボードセットは、電源およびクロック、Sun Fireplane interconnect ASIC 用の JTAG サポートを備え、システムコントロールボードとその関連周辺装置 (DVD-ROM、DAT ドライブ、およびハードドライブ) を搭載します。

ファンバックプレーンの 2 つは Sun Fireplane interconnect の上部に、2 つは電源センタープレーンの下部に搭載されて、8 つのファントレーに電源を供給します。

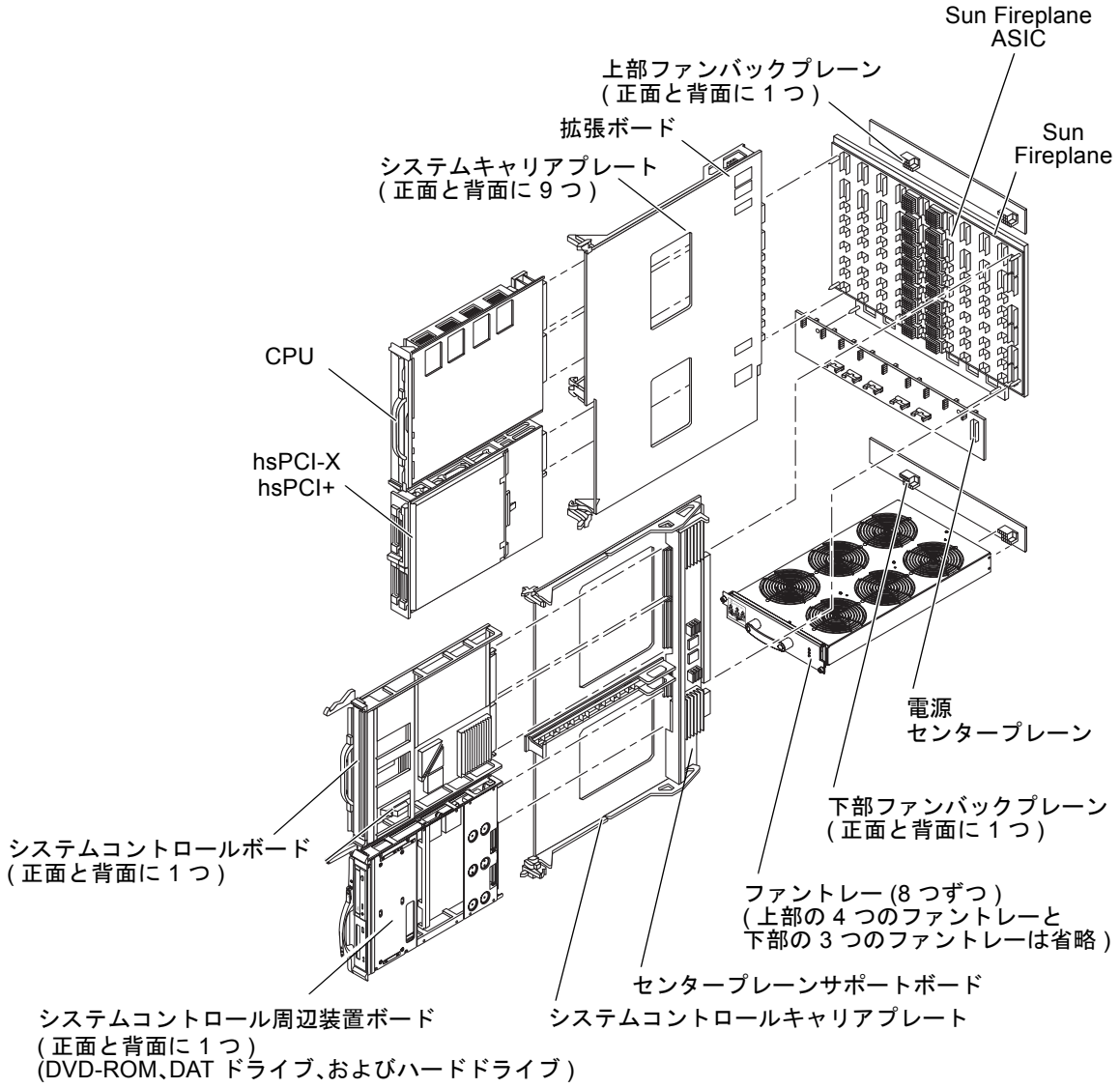


図 5-3　Sun Fireplane interconnect とその他のコンポーネント

5.2.1 Sun Fireplane interconnect

Sun Fireplane interconnect は Sun Fire E25K/E20K システムの中心で、18 ボードセットで 43G バイト/秒の最大帯域幅を提供します。また、Sun Fireplane interconnect は、各ボードセットにコンソールバスおよび Ethernet 接続を提供します。

Sun Fireplane interconnect には、3 つの 18×18 クロスバーがあります。18×18 アドレスクロスバーは、各拡張ボードの AXQ ASIC の間のアドレストランザクション用のパスを提供します。単方向のパスの組が各拡張ボードにつながり、1 つは送信、1 つは受信に使用されます。各アドレストランザクションがアドレスクロスバーを通過するには、2 システムインターコネクトサイクル (13.3 ns) が必要です。

18×18 応答クロスバーは、各拡張ボードの AXQ ASIC 間の返信パスを提供します。各応答メッセージには、種類によって、1 または 2 システムインターコネクトサイクル (6.7 ns または 13.3 ns) が必要です。応答パスは、アドレスパスの半分の幅です。単方向のパスの組が各拡張ボードにつながり、1 つは送信、1 つは受信に使用されます。

18×18 データクロスバーは、各拡張ボードの SDI (System Data Interface) ASIC の間で、キャッシュライン (72 バイト幅) パケットを移動します。各接続は、双方向の 36 バイト幅のパスです。帯域幅は、18 スロット×32 バイトパス×150 MHz を 2 で割って (双方向パスのため)、43.2G バイト/秒になります。これら双方向パスを最大限に利用するため、DMX (Data Multiplexer) ASIC は、受信したデータを待ち行列に入れます。

5.3 システムボード

ボードセットとは、Sun Fireplane interconnect に接続される 3 枚のシステムボードを組み合わせたものです。拡張ボードセットとも呼ばれます。ボードセットには、次の 2 種類があります。

- **システムボードセット**：CPU/メモリー、PCI バスコントローラを備えたボードです (5-7 ページの 5.3.1 節「システムボードセット」を参照)。
- **コントローラボードセット**：電源、クロック、Sun Fireplane interconnect 用の JTAG サポート、システムコントローラボードと関連する周辺装置を備えたボードです (5-11 ページの 5.3.2 節「コントローラボードセット」を参照)。

5.3.1 システムボードセット

システムボードセットとは、拡張ボード、スロット 0 ボード、およびスロット 1 ボードの 3 枚のボードを組み合わせたものです。ボードセット全体を、Sun Fireplane interconnect からホットスワップすることはできません。各コンポーネントは重いので、まずスロット 0 およびスロット 1 ボードを個別に取り外し、そのあと拡張ボードとそのキャリアプレートをホットスワップします。個々のスロット 0 およびスロット 1 ボードは、拡張ボードからホットスワップできます。

スロット 0 ボードは、4.8G バイト/秒のオフボードのデータポートを備えています。Sun Fire E25K/E20K システムでは、CPU は基本的にスロット 0 ボードに搭載されます。また、メモリーは、スロット 0 ボードだけに搭載されます。Sun Fire E25K/E20K システムでは、1 種類のスロット 0 ボードだけが使用されます。

スロット 1 ボードは、2.4G バイト/秒のオフボードのデータポートを備えています。hsPCI-X および hsPCI+ はスロット 1 ボードで、Sun Fire E25K/E20K システムおよび Sun Fire 15K/12K システムに固有のボードです。

5.3.1.1 拡張ボード

拡張ボードは、Sun Fireplane interconnect スロットを拡張する 2:1 MUX として動作し、スロット 0 およびスロット 1 タイプのボードを搭載できます。拡張ボードは、毎秒 1 億 5 千万スヌープを行うレベル 2 のアドレスバスを提供します。拡張ボードの AXQ は、ほかのボードセットへのアドレスを認識し、Sun Fireplane interconnect を介してそれらを送信します。

拡張ボードは、3 ポートのデータスイッチを提供し、スロット 0 ボード、スロット 1 ボードと、Sun Fireplane interconnect 間のデータを配信します。この 3 ポートデータスイッチは、Sun Fireplane interconnect およびスロット 0 ボードに対しては 36 バイト幅で、スロット 1 ボードに対しては 18 バイト幅です。ボードセットは、ほかのボードセットに対して、最大 4.8G バイト/秒で転送できます。

1 枚のシステムボード (スロット 0 またはスロット 1) だけで拡張ボードを使用することもできます。システムボードは、ほかのボードを妨げずに、拡張ボードにホットスワップしてテストし、動作中のシステムに構成できます。拡張ボードは、2 枚のシステムボードを取り外したあとに、ホットスワップおよび取り付けできます。

5.3.1.2 CPU/メモリーボード

CPU/メモリーボードは、スロット 0 ボードです。これは、最大 4 つの CPU と、8 つの外部キャッシュ DIMM を搭載しています。各 CPU は、0、4、または 8 枚の DIMM を制御します。DIMM の最大サイズは 2G バイトで、これを利用するとボード 1 枚につき最大 64G バイトを記憶できます。DIMM は同じサイズにして、同一ボード上で DIMM のサイズが異ならないようにする必要があります。同一ボード上の CPU は、すべて同じ速度にする必要があります。

2 つの CPU/メモリーの組は、レベル 0 のデュアル CPU データスイッチによって、システムのその他の部分に接続されます。各 CPU/メモリーは、最大 2.4G バイト/秒でデータを転送できます。CPU/メモリーユニットの組は、4.8G バイト/秒のポートをレベル 0 のデータスイッチに割り当てます。レベル 1 のデータスイッチは、2 組の CPU を、拡張ボードにつながっているオフボードのデータポートに接続します (図 5-4 を参照)。

5.3.1.3 システムボードセットの例

図 5-4 および図 5-5 に、ボードセット例と、拡張ボードおよび CPU/メモリーボード、PCI ボードで構成されたボードセットの配置を示します。

5.3.1.4 PCI アセンブリ (hsPCI-X または hsPCI+)

I/O アセンブリは、スロット 1 のオプションボードです。各 hsPCI-X アセンブリは、2 つの PCI コントローラを備えており、4 つの PCI スロットを提供します。スロットのうちの 1 つは 33 MHz で、3 つは 33/66/90 MHz です。また、hsPCI+ アセンブリも、2 つの PCI コントローラを備えており、4 つの PCI スロットを提供します。スロットのうちの 1 つは 33 MHz で、3 つは 33/66 MHz です。

カセットを使用すると、業界標準の PCI アセンブリをホットスワップできます。カセットは、カードを取り付けるためのキャリアで、標準の PCI ピンをコネクタに合わせて変換します。

PCI カードを PCI ホットスワップカセットに入れて、そのカセットを PCI アセンブリにホットスワップします。ソフトウェアはこのアセンブリを標準の PCI アセンブリと認識し、システムコントローラが各 PCI スロットの電源投入および切断を行います (図 5-4 を参照)。

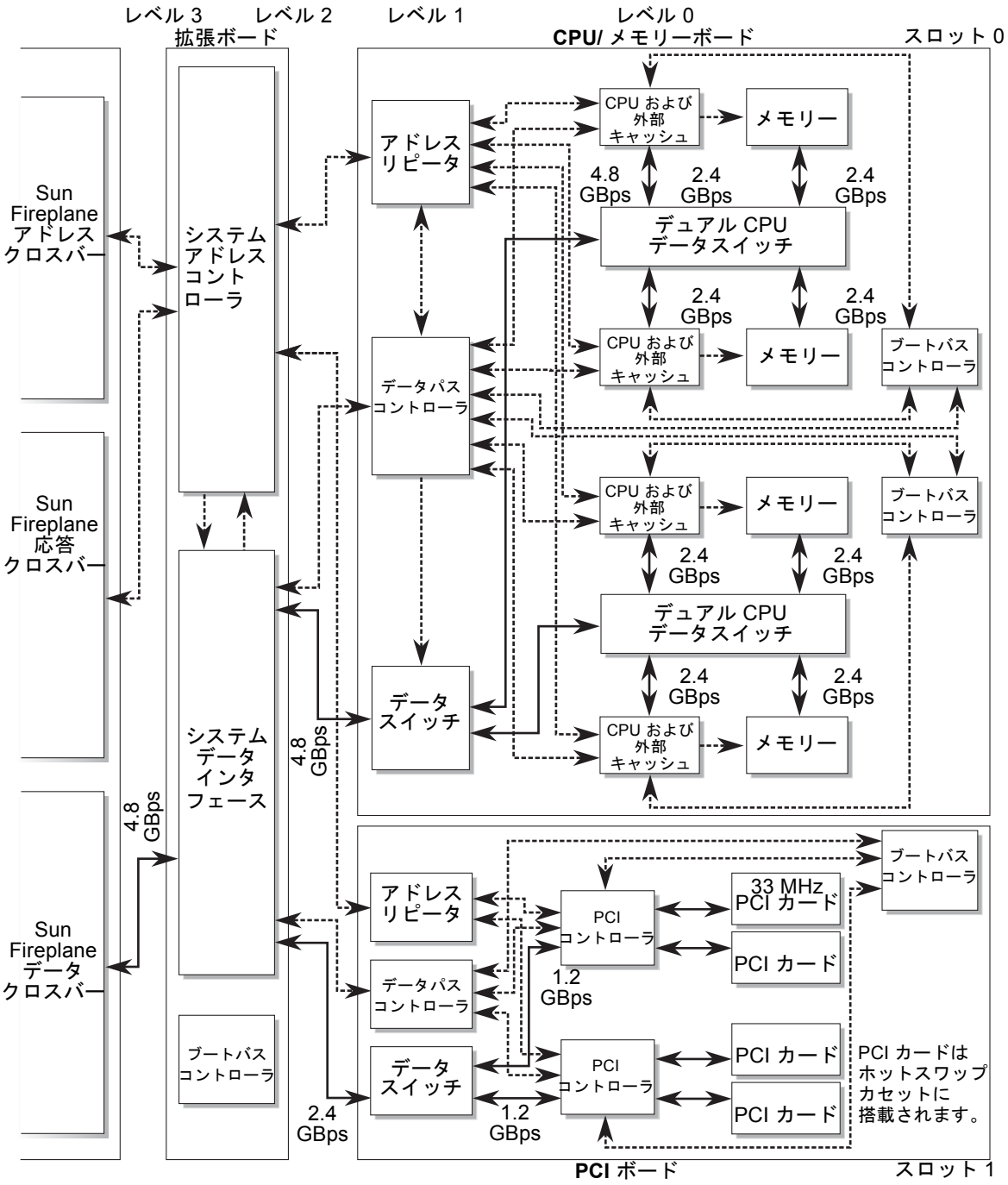


図 5-4　　ボードセットのブロックダイアグラム

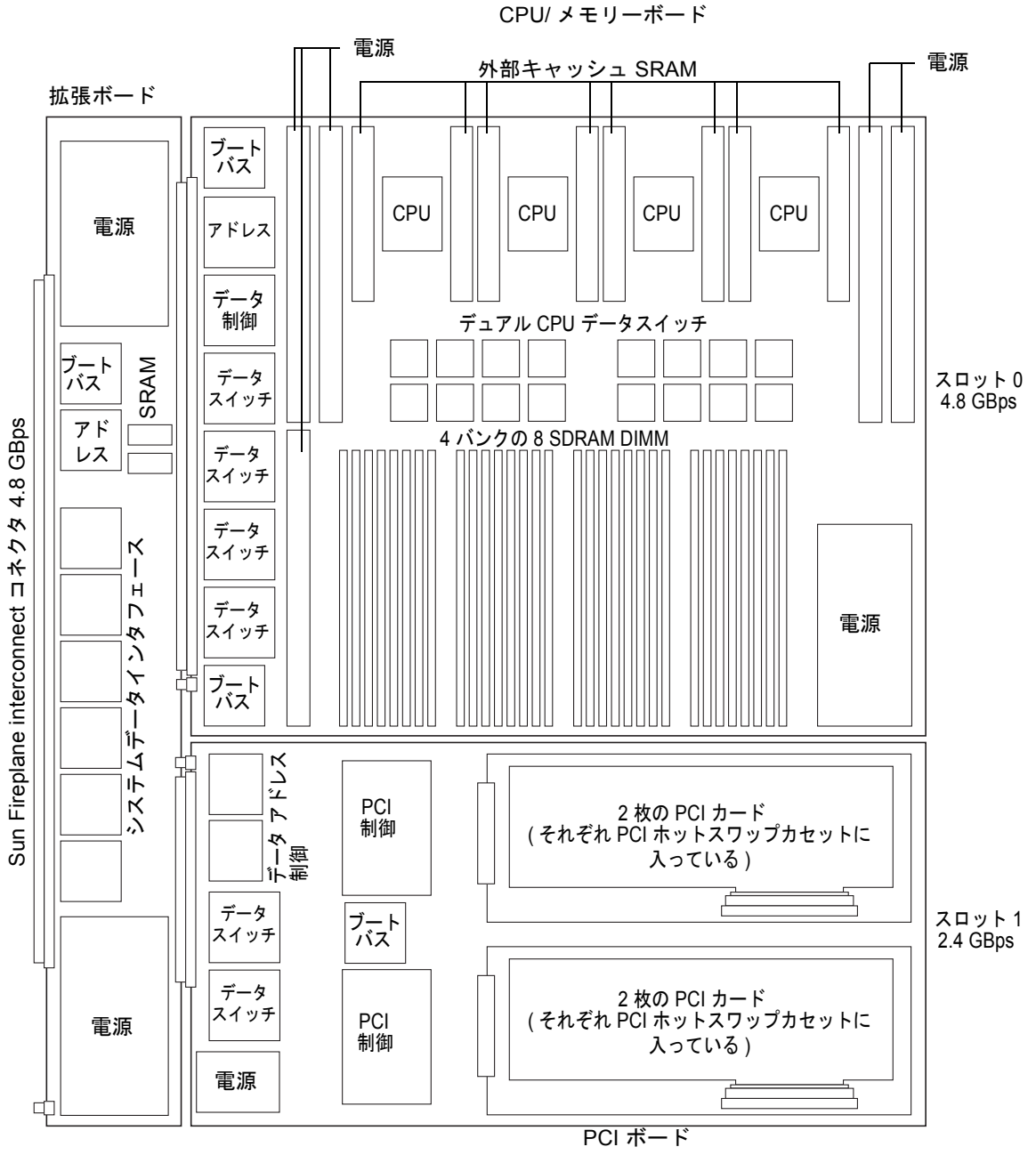


図 5-5　　システムボードセットの配置

5.3.2 コントローラボードセット

コントローラボードセットは、Sun Fire E25K/E20K システムの操作および制御に必要な、重要なサービスと資源を提供します (図 5-6)。

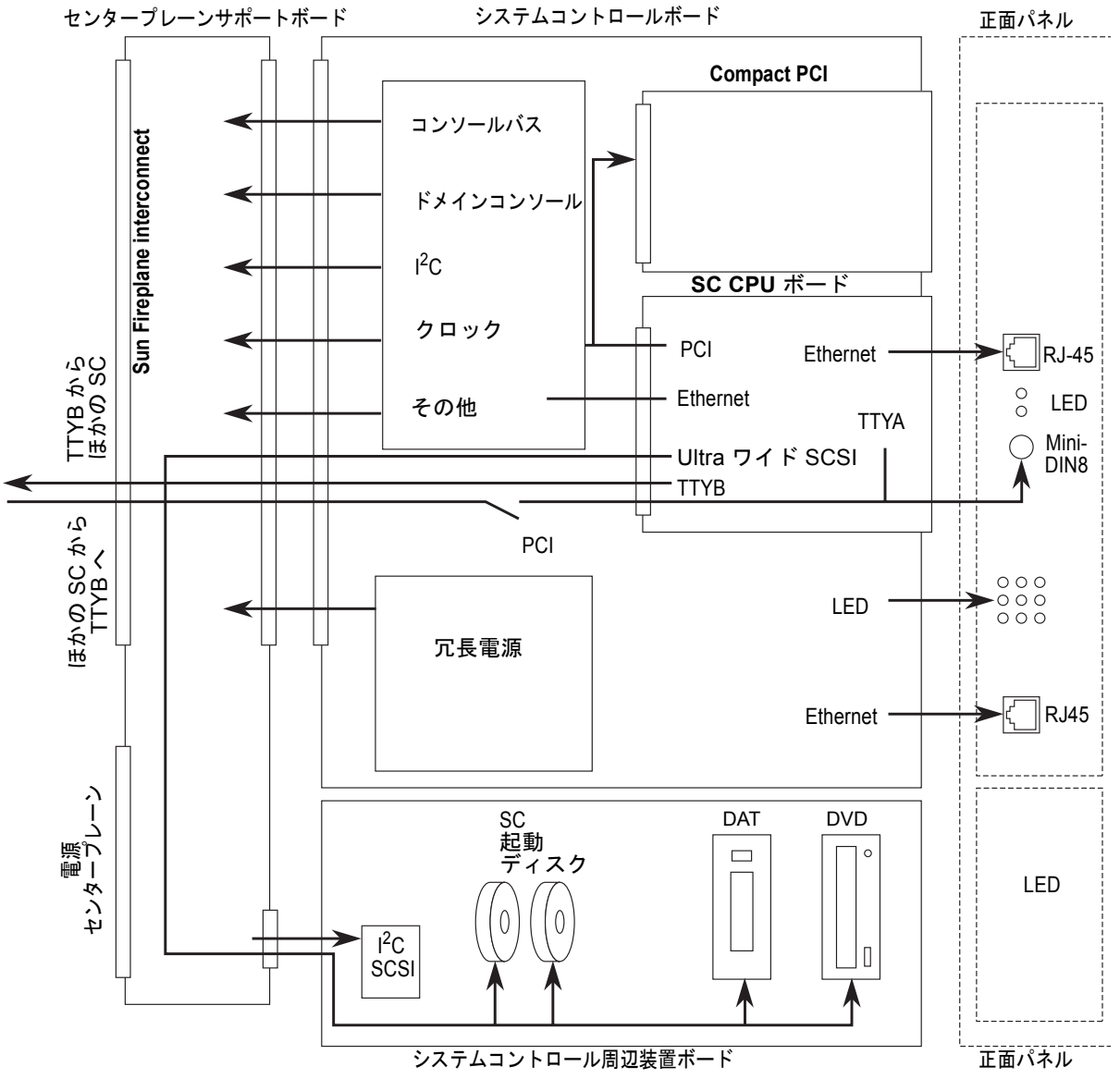


図 5-6　　システムコントローラボードの配置

このボードセットは、3 枚のボードで構成されます。

- **センタープレーンサポートボード**：専用の Sun Fireplane interconnect スロットに接続し、電源、クロック、および Sun Fireplane interconnect の JTAG サポートを行います。拡張ボードと同じサイズです。
- **システムコントロールボード**：センタープレーンサポートボードに接続します。スロット 0 システムボードと同じサイズです。
- **システムコントロール周辺装置ボード**：センタープレーンサポートボードに接続します。スロット 1 システムボードと同じサイズです。この周辺装置ボードは、DVD-ROM、DAT ドライブ、およびハードドライブを備えています。

システムコントロールボードは、2 枚のボードを組み合わせたものです。

- **SC CPU ボード。**SC CPU ボードは、UltraSPARC-IIi システムを組み込んだ、既成の SPARCengine Netra 2140 6U cPCI ボードです。このボードは、Solaris ソフトウェア、システム管理ソフトウェアと、システムの起動、保守、問い合わせに必要なすべての関連アプリケーションを実行します。
- **システムコントロールボード。**このコントロールボードは、Sun Fire E25K/E20K システムに、固有のロジックおよびセンタープレーンサポートボードへの接続を提供します。

システムコントローラボードセットは、Sun Fire E25K/E20K システムの操作および制御に必要な、次の重要なサービスと資源を提供します。

- システムクロック
- システム全体への I^2C バス
- システム全体へのコンソールバス
- SC CPU ボードを経由するシリアル (TTY) ポート
- 2 つのシステムコントローラの間のシリアル (TTY) ポート
- Solaris ソフトウェア、システム管理ソフトウェアと、起動、保守、およびシステムの問い合わせに必要なすべてのアプリケーションを実行するための Netra 2140 Compact PCI
- すべての動的システムドメインコンソールへの排他的アクセス
- DVD-ROM、DAT ドライブ、およびハードドライブをサポートする SCSI
- SC 操作を冗長 SC にフェイルオーバーするための高可用性機能のサポート
- B1 レベルまでセキュリティー保護された管理環境を提供するセキュリティー機能のサポート
- 各拡張 MAN (Management Area Network) のすべての I/O ボードへのプライベート Ethernet 回線の保護

SPARCengine cPCI+ カードは、PCI カードを I/O アセンブリに取り付けるときと同じ方法で、SC の上部に水平に取り付けます。

用語集

C

CDC システムアドレスコントローラ (AXQ) ASIC 内部の一貫性ディレクトリキャッシュ (Coherency Directory Cache)。メモリーの ECC ビットに保存されている最新のメモリータグの状態をキャッシュして、ほかのボードセットのキャッシュラインへのアクセスを高速化する。

CPU/メモリーボード それぞれ 8 枚の DIMM を制御する、4 つの CPU を備えたスロット 0 ボード。CPU/メモリーボードは、4.8G バイト/秒のオフボード帯域幅を持つ。

D

DCDS 2 つの CPU と 2 つのメモリーユニットをデータスイッチ ASIC に接続するデュアル CPU データスイッチ ASIC。

G

G バイト/秒 (GBps) 1 秒ごとに 1G バイトの容量 = 2^{30} = 1,073,741,824 バイト。

H

hsPCI+ アセンブリ 1 つの 33 MHz 標準 PCI カードと、3 つの 33/66 MHz 標準 PCI カードを備えるアセンブリ。PCI カードは、システムの動作中に I/O スロットからホットスワップして、動的再構成ができる。

hsPCI-X アセンブリ 1 つの 33 MHz 標準 PCI カードと、3 つの 33/66/90 MHz 標準 PCI カードを備えるアセンブリ。PCI カードは、システムの動作中に I/O スロットからホットスワップして、動的再構成ができる。

J

JTAG ジョイントテストアクショングループ (Joint Test Action Group)。チップの内部レジスタの連続走査に関する IEEE 標準 (1149.1)。

M

M バイト 1M バイトの容量 = 2^{20} = 1,048,576 バイト。

P

PCI コントローラ ASIC hsPCI-X ボード、hsPCI+ ボード、およびリンクボードで使用され、システムインターコネクトを PCI バスに接続する。

PCI ホットスワップカセット パッシブなホットスワップキャリアで、標準の PCI ピンをコネクタに合わせる。

S

Sun Fire アドレスバス 最大スヌープレートが毎秒 1 億 5 千万スヌープ、またはデータレートが 9.6G バイト/秒のアドレスバス。

Sun Fireplane interconnect
CPU の UltraSPARC IV Cu 世代が使用するインターコネクトアーキテクチャー。システムアドレスクロスバーおよびデータクロスバーを実装する、物理的なアクティブ論理のセンタープレーン。

Sun Fireplane interconnect アーキテクチャー
すべての UltraSPARC IV Cu CPU ベースのシステムで使用される、キャッシュ一貫性プロトコルおよびアドレストランザクションのセット。

Sun Fireplane interconnect データパス
DCDS と DX ASIC の間で使用されるポイントツーポイントデータプロトコル。

U

UltraSPARC CPU
UltraSPARC IV Cu CPU は、CPU/メモリーボードで使用される。

あ

アドレスリピータ (AR) ASIC
スロット 0 およびスロット 1 ボードで使用され、オンボードのシステムアドレスバスを実装する。4 つの CPU (または 2 つの入出力コントローラ) を、拡張ボードのアドレスコントローラに接続する。

応答時間
単一のデータ項目がメモリーから CPU に送信される時間。

応答マルチプレクサ (RMX) ASIC
トランザクションの応答を送信し、各拡張ボードのアドレスコントローラを接続する 18×18 クロスバー。

か

拡張可能な共有メモリー (SSM)
複数のスヌープ一貫性ドメインを接続できるようにするシステムインターコネクトのモード。

拡張ボード　スロット 0 およびスロット 1 ソケットで Sun Fireplane interconnect に接続するボード。

コントロールボード　Sun Fireplane interconnect の 2 つの制御スロットの 1 つに接続する。センタープレーンサポートボード、システムコントロールボード、および周辺装置ボードから構成される。

さ

システムアドレスコントローラ (AXQ) ASIC　スロット 0 およびスロット 1 ボードのアドレスリピータを、Sun Fireplane interconnect のアドレスクロスバーおよび応答クロスバーに接続する。拡張ボードで使用される。

システムコントロールボードセット　センタープレーンサポートボードによって Sun Fireplane interconnect の 2 つのシステムコントロールスロットの 1 つに接続する。このボードセットには、システムコントロールボードおよびシステムコントロール周辺装置ボード (DVD-ROM、DAT ドライブ、ハードドライブ) が含まれる。

システムデータインタフェース (SDI) ASIC　拡張ボードで使用され、スロット 0 およびスロット 1 ボードのデータスイッチを、Sun Fireplane interconnect のデータクロスバーに接続する。

システムボードセット　拡張ボードによって Sun Fireplane interconnect の 18 のシステムスロットの 1 つに接続する。スロット 0 ボードとスロット 1 ボードが含まれる。

自動システム回復 (ASR)　ハードウェア障害が発生した場合に、システムの動作を回復する機能。障害の発生しているハードウェアコンポーネントを特定および分離して、障害の発生したハードウェアコンポーネントを除いた起動可能なシステム構成を構築する。

た

データアービタ (DARB) ASIC　Sun Fireplane interconnect で、18×18 データクロスバーを制御するために使用される。

データスイッチ (DX) ASIC　スロット 0 およびスロット 1 ボードで使用され、オンボードのシステムデータパスをオフボードのシステムデータパスに接続する。

データパス コントローラ (SDC) ASIC スロット 0 およびスロット 1 ボードで使用され、オンボードのシステムデータパスを制御する。コンソールバスを、2 つのオンボードブートバスコントローラにリピートする。

データマルチプレクサ (DMX) ASIC 18×18 データクロスバーで、各拡張ボードのシステムデータインタフェースを Sun Fireplane interconnect に接続する。

電源 電源装置のグループによって電力を供給されるハードウェアコンポーネント。

動的再構成 ユーザーのアプリケーションの続行中に、Solaris オペレーティングシステムを実行しながらボード、電源装置などの装置を起動または停止する処理。

ドメインストップ クライアントドメイン間でエラーを分離するための機能。

ドメインセット SRD とそのクライアントドメインの組み合わせ。

ドメインのリンク インタードメインネットワークから除外されていたドメインをリンクすること。

ドメインのリンク解除 インタードメインネットワークからドメインを除外すること。

は

ブートバスコントローラ (SBBC) ASIC スロット 0 およびスロット 1 ボードで使用され、ボードを初期化するための、PROM バス、JTAG、I^2C 装置へのコンソールバス－スレーブインタフェースを提供する。CPU とともに使用すると、POST コードへの起動バスパスを提供する。

分割拡張ボード ボードセットの 2 つのシステムボードが、異なるドメインにあるもの。

並行保守 動作中のシステムを妨げずに、マシンのさまざまな部品を保守する機能。

ボードセット (拡張) 拡張ボード、スロット 0 ボード、およびスロット 1 ボードの組み合わせ。

ホットスワップ 動作中のシステムへの取り付けおよび取り外しを行なって、動的再構成ができる動作中の装置。

ら

レコードストップ　データパスの修正可能なシングルビットエラーなど、重大ではないエラーの場合にレコードストップが発生する。