



Sun Fire™ E6900/E4900 システム 製品概要

Sun Microsystems, Inc.
www.sun.com

Part No. 817-5849-10
2004 年 3 月, Revision A

コメント送付: <http://www.sun.com/hwdocs/feedback>

Copyright 2004 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, CA 95054 U.S.A. All rights reserved.

米国 Sun Microsystems, Inc. (以下、米国 Sun Microsystems 社とします)は、本書に記述されている製品に採用されている技術に関する知的所有権を有しています。これら知的所有権には、<http://www.sun.com/patents>に掲載されているひとつまたは複数の米国特許、および米国ならびにその他の国におけるひとつまたは複数の特許または出願中の特許が含まれています。

本書およびそれに付属する製品は著作権法により保護されており、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。サン・マイクロシステムズ株式会社の書面による事前の許可なく、本製品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。

本製品のフォント技術を含む第三者のソフトウェアは、著作権法により保護されており、提供者からライセンスを受けているものです。

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 システムに基づいていることがあります。UNIX は、X/Open Company Limited が独占的にライセンスしている米国ならびに他の国における登録商標です。

本製品は、株式会社モリサワからライセンス供与されたリュウミン L-KL (Ryumin-Light) および中ゴシック BBB (GothicBBB-Medium) のフォント・データを含んでいます。

本製品に含まれる HG 明朝 L と HG ゴシック B は、株式会社リコーがリョービマジクス株式会社からライセンス供与されたタイプフェースマスタをもとに作成されたものです。平成明朝体 W3 は、株式会社リコーが財団法人日本規格協会 文字フォント開発・普及センターからライセンス供与されたタイプフェースマスタをもとに作成されたものです。また、HG 明朝 L と HG ゴシック B の補助漢字部分は、平成明朝体 W3 の補助漢字を使用しています。なお、フォントとして無断複製することは禁止されています。

Sun、Sun Microsystems、AnswerBook2、docs.sun.com、Sun Fire は、米国およびその他の国における米国 Sun Microsystems 社の商標もしくは登録商標です。サン・ロゴマークおよび Solaris は、米国 Sun Microsystems 社の登録商標です。

すべての SPARC 商標は、米国 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商標または登録商標です。SPARC 商標が付いた製品は、米国 Sun Microsystems 社が開発したアーキテクチャーに基づくものです。

OPENLOOK、OpenBoot、JLE は、サン・マイクロシステムズ株式会社の登録商標です。

ATOK は、株式会社ジャストシステムの登録商標です。ATOK8 は、株式会社ジャストシステムの著作物であり、ATOK8 にかかる著作権その他の権利は、すべて株式会社ジャストシステムに帰属します。ATOK Server/ATOK12 は、株式会社ジャストシステムの著作物であり、ATOK Server/ATOK12 にかかる著作権その他の権利は、株式会社ジャストシステムおよび各権利者に帰属します。

本書で参照されている製品やサービスに関しては、該当する会社または組織に直接お問い合わせください。

OPEN LOOK および Sun Graphical User Interface は、米国 Sun Microsystems 社が自社のユーザーおよびライセンス実施権者向けに開発しました。米国 Sun Microsystems 社は、コンピュータ産業用のビジュアルまたはグラフィカル・ユーザーインタフェースの概念の研究開発における米国 Xerox 社の先駆者としての成果を認めるものです。米国 Sun Microsystems 社は米国 Xerox 社から Xerox Graphical User Interface の非独占的ライセンスを取得しており、このライセンスは米国 Sun Microsystems 社のライセンス実施権者にも適用されます。

U.S. Government Rights-Commercial use. Government users are subject to the Sun Microsystems, Inc. standard license agreement and applicable provisions of the FAR and its supplements.

本書は、「現状のまま」をベースとして提供され、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、なんらの保証も行われぬものとします。

本書には、技術的な誤りまたは誤植の可能性があります。また、本書に記載された情報には、定期的に変更が行われ、かかる変更は本書の最新版に反映されます。さらに、米国サンまたは日本サンは、本書に記載された製品またはプログラムを、予告なく改良または変更することがあります。

本製品が、外国為替および外国貿易管理法 (外為法) に定められる戦略物資等 (貨物または役務) に該当する場合、本製品を輸出または日本国外へ持ち出す際には、サン・マイクロシステムズ株式会社の書面による承諾を得ることのほか、外為法および関連法規に基づく輸出手続き、また場合によっては、米国商務省または米国所轄官庁の許可を得ることが必要です。

原典:	Sun Fire E6900/E4900 Systems Overview Part No: 817-4119-10 Revision A
-----	---



Adobe PostScript

目次

はじめに vii

1. Sun Fire 製品概要 1-1
 - 1.1 標準機能 1-1
 - 1.2 マシン構成 1-4
 - 1.2.1 Sun Fire E6900 システム 1-4
 - 1.2.2 Sun Fire E4900 システム 1-6
2. システムの機能および特徴 2-1
 - 2.1 パーティションおよびドメイン 2-2
 - 2.1.1 パーティション 2-3
 - 2.1.2 ドメイン 2-3
 - 2.2 信頼性 2-4
 - 2.2.1 エラー確率の低減 2-4
 - 2.2.2 誤り訂正符号を使用したエラーの修正 2-4
 - 2.2.2.1 データインターコネクットの誤り訂正符号による保護 2-5
 - 2.2.3 修正不可能なエラーの検出 2-5
 - 2.2.3.1 複数ビットデータエラー 2-5
 - 2.2.3.2 アドレスエラー 2-5
 - 2.2.3.3 システムタイムアウトエラー 2-6

2.2.3.4	電源修正障害	2-6
2.2.4	環境検知	2-6
2.2.4.1	温度	2-6
2.2.4.2	電源サブシステム	2-6
2.3	可用性	2-7
2.3.1	Sun Fire システムの可用性の目標	2-7
2.3.2	Sun Fire システムの高可用性	2-7
2.3.2.1	冷却装置	2-8
2.3.2.2	AC 電源スイッチ	2-8
2.3.2.3	ECC	2-8
2.3.3	回復機能	2-8
2.3.3.1	DC 電源	2-9
2.3.3.2	論理ボード	2-9
2.3.3.3	プロセッサ	2-9
2.3.3.4	メモリー	2-9
2.3.4	冗長コンポーネント	2-10
2.4	保守性	2-10
2.4.1	マシンの保守性	2-11
2.4.1.1	DC 電源装置の取り外しおよび取り付け	2-11
2.4.1.2	ファントレーの取り外しおよび取り付け	2-11
2.4.1.3	ドメインの分離	2-11
2.4.2	非並行保守	2-11
2.4.3	遠隔保守	2-12
3.	ハードウェアの概要	3-1
3.1	標準操作	3-1
3.2	データインターコネクト	3-3
3.3	コンソールバスインターコネクト	3-4

- 4. Sun Fire システムのコンポーネント 4-1
 - 4.1 CPU/メモリーボード 4-1
 - 4.2 I/O アセンブリ 4-1
 - 4.2.1 PCI I/O 4-1
 - 4.3 リピータボード 4-2
 - 4.4 システムコントローラボード (バージョン 2、拡張メモリー付き) 4-2
 - 4.4.1 冗長システムコントローラ 4-4
 - 4.4.2 ドメインの仮想キースイッチ 4-4
 - 4.4.3 Solaris コンソール 4-4
 - 4.4.4 仮想時刻 4-4
 - 4.4.5 環境の監視 4-4

はじめに

このマニュアルでは、Sun Fire™ E6900/E4900 システムについて説明します。このマニュアルの内容は次のとおりです。

- Sun Fire E6900/E4900 システムのマシン構成
- ハードウェアの概要
- システムのコンポーネント
- 信頼性、可用性、および保守性の特徴

書体と記号について

書体または記号	意味	例
AaBbCc123	コマンド名、ファイル名、ディレクトリ名、画面上のコンピュータ出力、コード例。	.login ファイルを編集します。 ls -a を実行します。 % You have mail.
AaBbCc123	ユーザーが入力する文字を、画面上のコンピュータ出力と区別して表します。	マシン名 % su Password:
<i>AaBbCc123</i> またはゴシック	コマンド行の可変部分。実際の名前や値と置き換えてください。	rm <i>filename</i> と入力します。 rm ファイル名 と入力します。
『 』	参照する書名を示します。	『Solaris ユーザーマニュアル』
「 」	参照する章、節、または、強調する語を示します。	第 6 章「データの管理」を参照。 この操作ができるのは「スーパーユーザー」だけです。
\	枠で囲まれたコード例で、テキストがページ幅をこえる場合に、継続を示します。	% grep `^#define \ XV_VERSION_STRING`

関連マニュアル

用途	タイトル
設置	『Sun Fire E6900/E4900 システム設置マニュアル』
操作	『Sun Fire キャビネット設置およびリファレンスマニュアル』 『Sun Fire E6900/E4900 システムの概要』 『Sun Fire E6900/E4900 システムサービスマニュアル』

Sun のオンラインマニュアル

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

<http://www.sun.com/documentation>

Sun 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

弊社では、マニュアルの改善に努力しており、お客様からのコメントおよびご忠告をお受けしております。コメントは下記よりお送りください。

<http://www.sun.com/hwdocs/feedback>

コメントにはマニュアルの Part No. (817-5849-10) とタイトルを記載してください。

第1章

Sun Fire 製品概要

この章では、Sun Fire E6900 および Sun Fire E4900 システムの機能と収容できるコンポーネント数について説明します。

このサーバーファミリでは、初歩的なレベルからハンドエンドレベルまでのサーバー機能を提供します。Sun Fire E6900 システムの 19 インチ型キャビネットには、内蔵周辺装置を取り付けるための空間があります。これ以外のシステムは、業界標準の 19 インチ型キャビネットに取り付けるか、あるいは Sun Fire システムキャビネットに搭載しておくかを自由に選択できます。Sun Fire システムキャビネットには、1 つの Sun Fire E4900 システムを搭載できます。

1.1 標準機能

これらのシステムの標準機能は、次のとおりです。

- 業界標準の 19 インチ型ラックに搭載可能 (Sun Fire E4900)
- 最大 24 の CPU のサポート
- 最大 32 の PCI I/O スロットのサポート
- 広範囲の冗長性
- システムコントローラ
- 複数のドメインのサポート
- ハードウェアの並行保守
- 共通コンポーネント
- 冗長電源・冷却装置
- 9.6G バイトのバス帯域幅

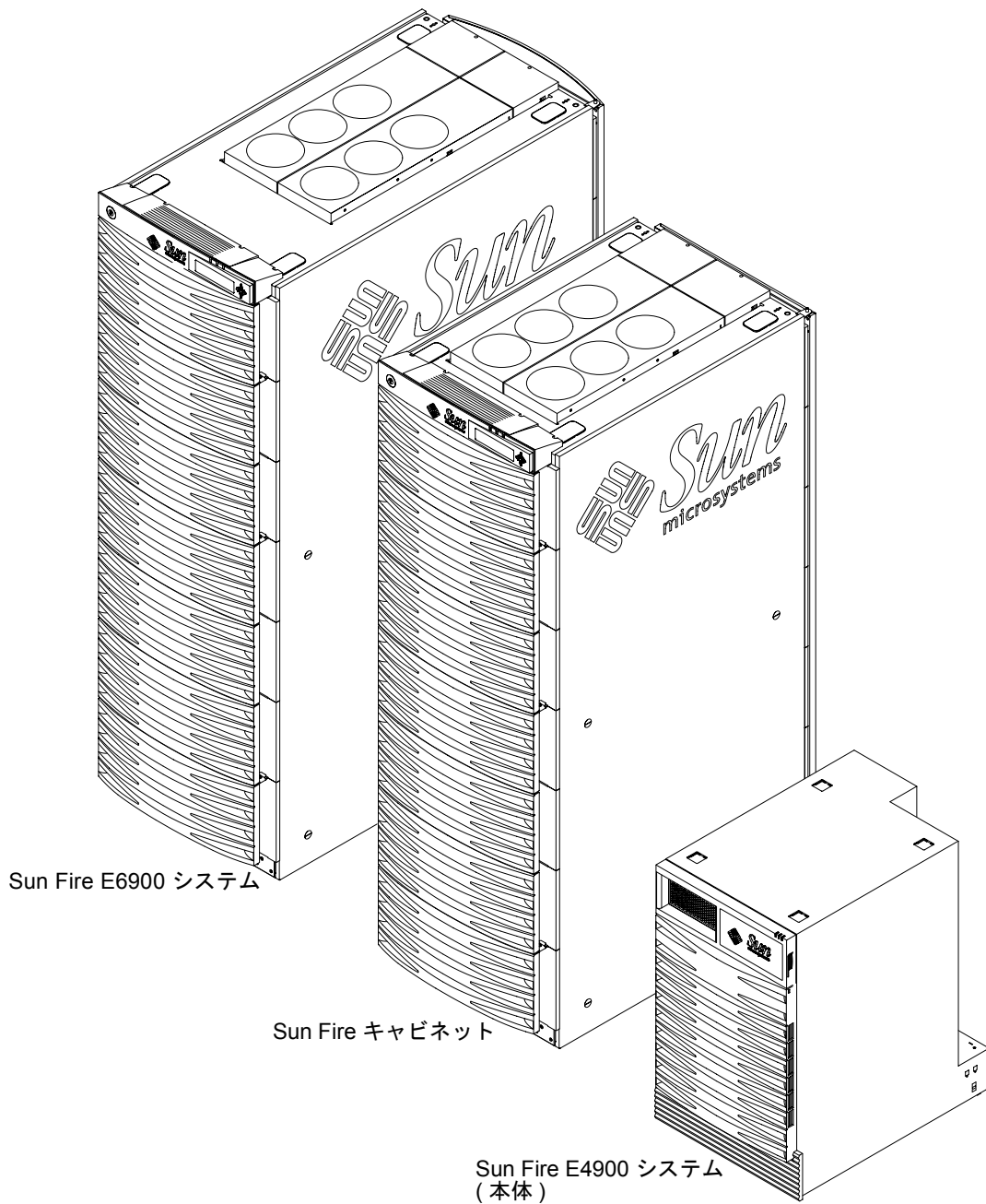


図 1-1 Sun Fire システムおよび Sun Fire キャビネット

Sun Fire E6900 および E4900 システムは、次の共通のコンポーネントを使用します。

- CPU/メモリーボード
- CPU プロセッサ
- メモリー DIMM
- PCI I/O アセンブリ
- PCI I/O カード
- システムコントローラボード (バージョン 2、拡張メモリー付き)
- リピータボード

1.2 マシン構成

次の 2 つのマシン構成があります。

- Sun Fire E6900 システム
- Sun Fire E4900 ラック搭載型システム

1.2.1 Sun Fire E6900 システム

Sun Fire E6900 システムは、6 つの CPU/メモリーボード、4 つの I/O アセンブリ、4 つのリピータボード、および 2 つのシステムコントローラボードをサポートしています。リピータボードは 4 つありますが、論理的には 2 つの冗長リピータになります (2 つのボードで 1 つの論理リピータを構成)。図 1-2 に、Sun Fire E6900 システムキャビネットの正面図および背面図を示します。表 1-1 に、Sun Fire E6900 システムの機能を示します。

表 1-1 Sun Fire E6900 システムの機能

機能	数量または説明
CPU/メモリーボード	6
CPU	24
最大メモリー	192 の DIMM ソケット
I/O アセンブリ	4 (PCI)
システムコントローラボード (バージョン 2)	2
リピータボード	4
ドメイン	4 (最大)
電源装置	6
所要電力	200 ~ 240 VAC
冗長冷却	あり
冗長 AC 入力	あり
内蔵周辺装置	なし (ただし、キャビネット内に周辺装置 オプション用の空間あり)
パッケージ	Sun Fire E6900 キャビネット

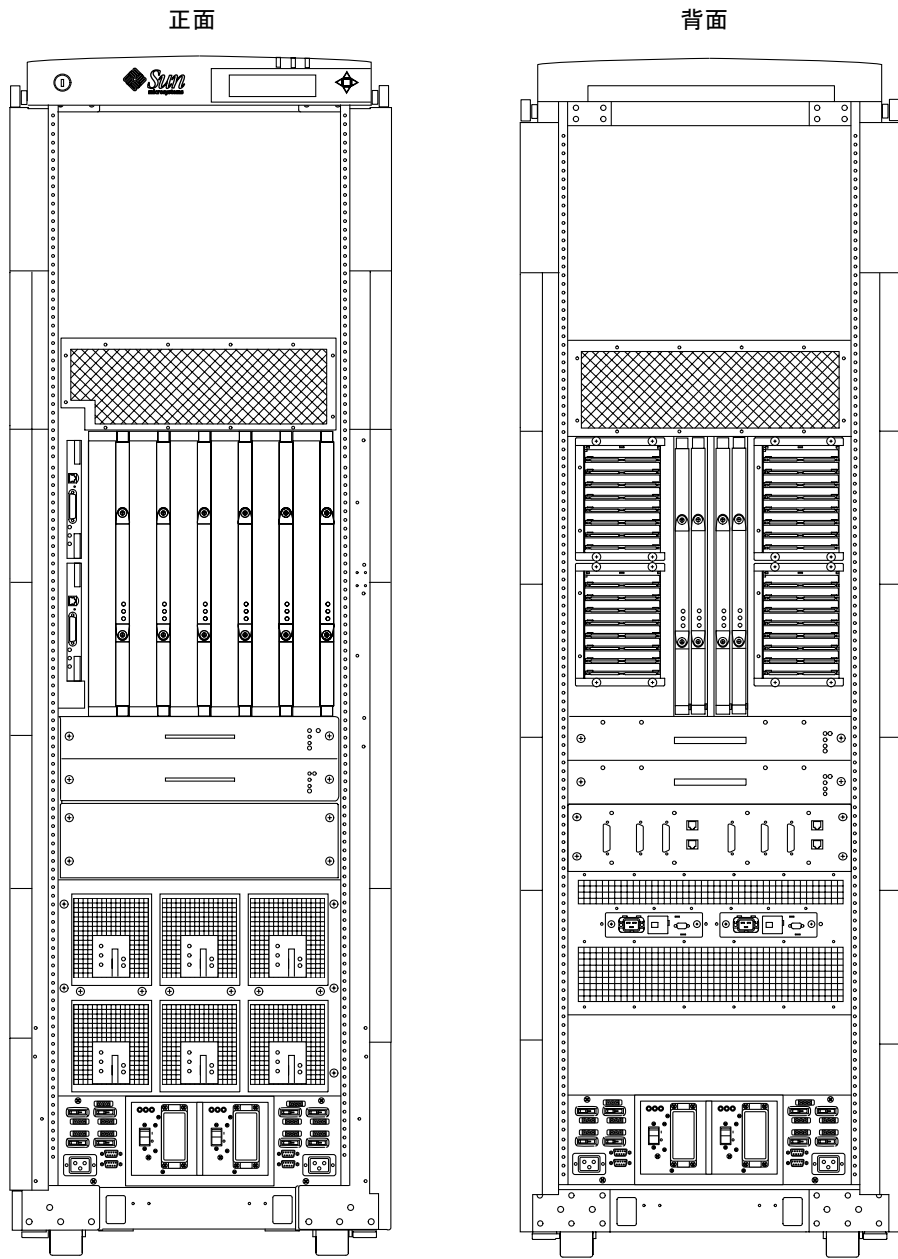


図 1-2 Sun Fire E6900 システムキャビネットの正面図および背面図

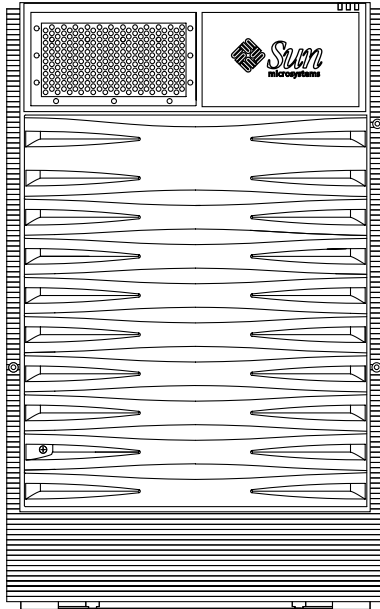
1.2.2 Sun Fire E4900 システム

Sun Fire E4900 システムは、3 つの CPU/メモリーボード、2 つの I/O アセンブリ、2 つのリピータボード、および 2 つのシステムコントローラボードをサポートしています。図 1-3 に、Sun Fire E4900 システムの正面図および背面図を示します。表 1-2 に、Sun Fire E4900 システムの機能を示します。

表 1-2 Sun Fire E4900 システムの機能

機能	数量または説明
CPU/メモリーボード	3
CPU	12
最大メモリー	96 の DIMM ソケット
I/O アセンブリ	2 (PCI)
システムコントローラボード (バージョン 2)	2
リピータボード	2
ドメイン	2 (最大)
電源装置	3
所要電力	200 ~ 240 VAC
冗長冷却	あり
冗長 AC 入力	なし
内蔵周辺装置	なし
パッケージ	ラック搭載型、または Sun Fire キャビネットに搭載済み

正面



背面

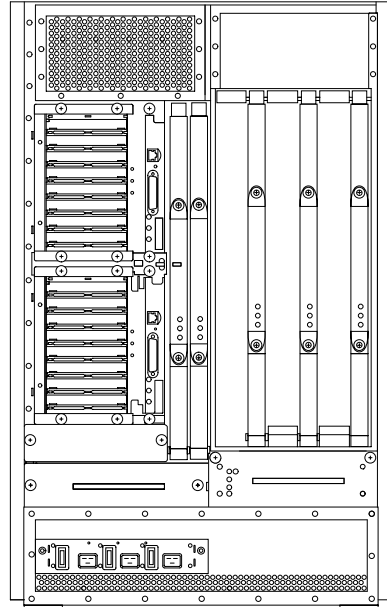


図 1-3 Sun Fire E4900 システムの正面図および背面図

第2章

システムの機能および特徴

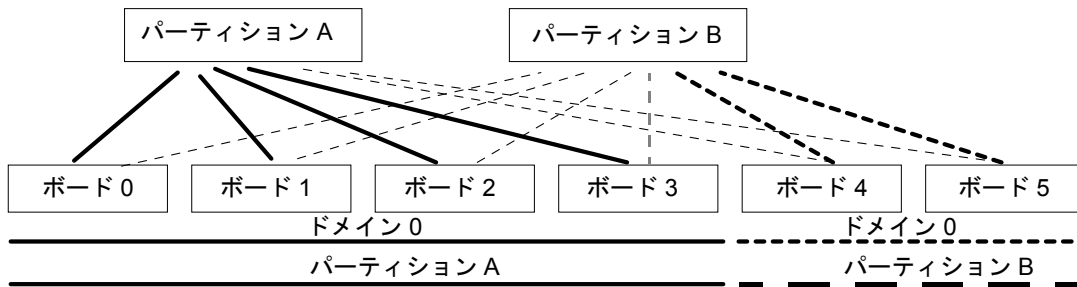
Sun Fire E6900/E4900 システムの主な機能には、システムのパーティションへの分割、ドメインの作成などがあります。これらの機能によって、より高い信頼性、可用性、および保守性が提供され、稼動時間の向上が実現されます。このシステムの機能および特徴は、次のとおりです。

- パーティション — 1つのシステムが論理的に2つの異なるシステムとして動作できるようにする機能
- ドメイン — 1つのパーティション内に論理的に独立した複数のセクションを作成し、各ドメインが個々のオペレーティングシステムを実行できるようにする機能
- 信頼性 — ハードウェアおよびソフトウェアの設計段階で組み込まれた保護機能、選択したコンポーネントの品質、および製造プロセスの品質 (ESD による保護、クリーンルームなど)
- 可用性 — 顧客のシステムが本番作業を実行できる時間の割合
- 保守性 — システムの修復時間 (停止時間) の最小化

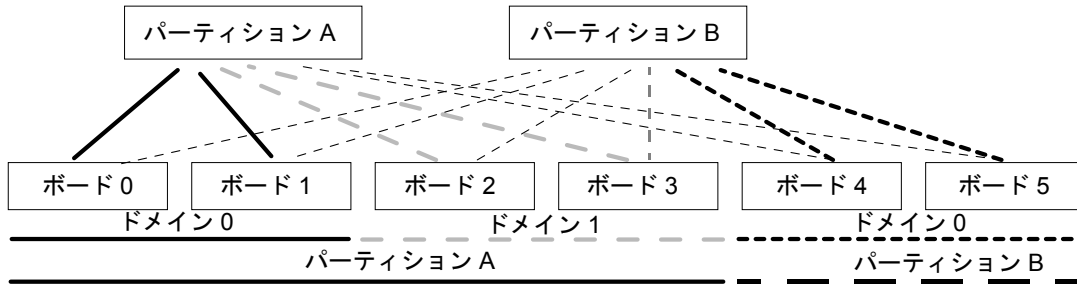
2.1 パーティションおよびドメイン

Sun Fire システムでは、システムをパーティションおよびドメインに分割できます。パーティションおよびドメインを使用することによって、物理的に 1 つのシステムに複数の独立した論理システムを作成し、各システムでオペレーティングシステムを実行することができます。パーティションとドメインは、柔軟性と独立性の点だけが異なります。

パーティション A および B には、個別のリピータボードがある



1 つのパーティションは、最大 2 つのドメインを持つことができる



- ⋯⋯ アクティブドメイン接続
- アクティブドメイン接続
- — アクティブドメイン接続
- - - アクティブでない論理接続

図 2-1 Sun Fire E6900 システムのパーティションおよびドメイン

2.1.1 パーティション

1 つの物理的な Sun Fire E6900 システムは、2 つのパーティションに分割できます。1 つのパーティションのボードともう 1 つのパーティションのボードとの間の接続は、すべて使用不可になります。システムは、論理的に 2 つの異なるシステムとして動作します。

パーティションが Sun Fire E6900 システムの物理的な半分に割り当てられた場合、各パーティションに関連付けられた電源プレーンも分離されます。Sun Fire E6900 システムは、各パーティション用の 1 セットのリピータボードを論理的に分離することによって、2 つのパーティションに分割できます。Sun Fire E4900 システムも 2 つのパーティションをサポートしています。

Sun Fire E6900 システムでは、各パーティションに最大 2 つのドメイン (全体で最大 4 つのドメイン) を持つことができます。Sun Fire E4900 システムでは、単一パーティションの場合は 2 つのドメインをサポートし、パーティションが 2 つである場合は各パーティションに 1 つのドメインのみをサポートします。

2.1.2 ドメイン

Sun Fire システムは、論理的に複数のドメインに分割できます。各ドメインは 1 つ以上のシステムボードで構成されるため、1 つのドメインには 1 ~ 24 のプロセッサを含めることができます。各ドメインは、独自のオペレーティングシステムインスタンスを実行し、独自の周辺装置およびネットワーク接続を持ちます。ドメインの設定は、同じシステムにある、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
- 複数のドメインの構成による、異なる部門のサポート

本番作業が残りの (通常は大きい方の) ドメイン上で継続されている間、ドメイン間での悪影響を及ぼす相互作用は起こりません。そのため、本番作業を中断することなくアプリケーションの正確性を評価できます。評価作業が完了すると、再起動せずにシステムを論理的に結合し直すことができます (ドメインを使用する場合、物理的な変更はありません)。このため、問題が発生した場合でも残りのシステムは影響を受けません。

Sun Fire E6900 システムでは、最大 4 つのドメインを持つことができます。Sun Fire E4900 システムでは、最大 2 つのドメインを持つことができます。Solaris™ オペレーティングシステムの各インスタンスは、個々のドメインで実行されます。ドメインが、相互に依存したり対話したりすることはありません。

Sun Fire E6900 システム上の単一のパーティションは、2 つのドメインに分割できます。パーティションとは異なり、ドメインはリピータボードを共有します。各ドメインは、全システムバスのアドレス帯域幅の半分を使用します。

2.2 信頼性

Sun Fire システムの信頼性の特性は、次の 4 つのカテゴリに分類されます。

- エラー確率の低減
- 誤り訂正符号 (ECC) を使用したエラーの修正
- 修正できないエラーの検出
- 環境的な要因の検知

2.2.1 エラー確率の低減

すべての ASIC は、最悪の場合の温度、電圧、周波数、および通気の組み合わせを想定して設計されています。ASIC は高いレベルで論理統合されているので、コンポーネントおよびインターコネクトの数が減少しています。

分散型電源システムによって、電源装置の性能および信頼性が向上します。

ハードウェアに障害が発生したあとの電源投入による再起動時には、次の広範囲な自己診断が行われ、Sun Fire システムの主な論理ブロックがすべて検査されます。

- すべての ASIC 内に組み込まれた自己診断ロジック
- システムコントローラボードから制御される電源投入時自己診断 (POST)。最初に各論理ブロックが個別に評価され、徐々に範囲を広げてシステムが評価されます。障害が発生したコンポーネントは、電氣的にセンタープレーンから分離されます。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。

すべての I/O ケーブルにはポジティブロック機構および留め具サポートがあり、偶発的な切り離しを防止しています。

2.2.2 誤り訂正符号を使用したエラーの修正

Sun Fire システムには、確実にエラーから回復できる多くのサブシステムが含まれています。サブシステムに多くの接続がある場合、障害発生の可能性は高くなります。エラー発生の確率をもっとも高いサブシステムは、誤り訂正符号を使用したシングルビットエラー訂正を使用することによって、一時的なエラーから保護されています。

2.2.2.1 データインターコネクトの誤り訂正符号による保護

局所データクロスバーおよびメモリーサブシステムからのすべてのデータパスは、誤り訂正符号で保護されています。これらのサブシステムで検出されたシングルビットデータエラーは、受信する UltraSPARC™ s400 モジュールによって修正され、エラー発生ログを記録するためにシステムに通知されます。

メモリーサブシステムではエラーの検査または修正は行われませんが、追加の記憶ビットが提供されます。Sun Fire のデータバッファチップは、誤り訂正符号を使用して障害の分離を支援します。

修正可能なエラーがインターコネクトによって検出された場合、システムコントローラにエラーが通知され、障害をインターコネクトシステム内のシングルネットに分離できるだけの情報が保存されます。エラーを含むデータはインターコネクトを介してそのまま送信され、エラーが報告されます。

メモリーエラーはソフトウェアによって記録されるため、欠陥のある DIMM は定期保守時に特定および交換できます。

2.2.3 修正不可能なエラーの検出

ほとんどの内部システムバスは、いくつかの形式の冗長検査機構によって保護されています。これによって、不良データの転送が検出され、不良データの伝播は通知なしで防止されます。すべての修正不可能なエラーは、エラー状態になります。回復するには、オペレーティングシステムを自動的に再起動させる必要があります。

2.2.3.1 複数ビットデータエラー

複数ビットの ECC エラーは、受信ポートによって検出されオペレーティングシステムに通知されるため、影響を受けるプロセスによっては、障害がシステム全体に及ぶことを回避できます。

インターコネクトへの外部キャッシュの読み取りで発生したパリティエラーは、複数ビットの ECC データエラーとなり、他の複数ビットエラーとして処理されます。

2.2.3.2 アドレスエラー

アドレスのインターコネクトで検出されたシングルビットまたは複数ビットのエラーは回復不可能であるため、重大なオペレーティングシステムエラーとなります。

2.2.3.3 システムタイムアウトエラー

ポートコントローラまたはメモリーコントローラによって検出されたタイムアウトエラーは、トランザクションが消失したことを示します。そのため、タイムアウトは常に回復不可能です。

2.2.3.4 電源修正障害

Sun Fire システムは、信頼性の高い分散型電源システムを使用しています。システム内の各入出力サブシステム、CPU/メモリーボード、システムコントローラボード、またはリピータボードには、そのボード専用の DC/DC コンバータがあり、各電圧用に複数のコンバータがあります。DC/DC コンバータに障害が発生すると、システムコントローラに通知されます。障害を報告したシステムボードは、システムから構成解除されます。障害が発生した際に行われていたシステム動作の継続は保証されません。

2.2.4 環境検知

システムシャーシの環境では、温度、通気、電源装置の性能など、システムの安定性に対する主な測定値が監視されます。システムコントローラは常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、システムへの物理的損傷およびデータ破壊が防止されます。

2.2.4.1 温度

システムの内部温度は、復旧メカニズムによって主要な位置で監視されます。温度の読み取り値に基づいて、システムは発生する可能性がある問題を管理者に通知するか、あるいは正常な停止を開始することができます。また、即時にシステムの電源を切断することもできます。

2.2.4.2 電源サブシステム

定期的な健全性検査を使用可能にすると、Sun Fire システムはさらに検知を行い、システムの信頼性を向上させます。DC 電圧は、システム内の主要な位置で監視されます。各電源装置からの DC 電流が監視され、システムコントローラに報告されます。CPU 電源制御は、システムを停止せずに過熱しているすべての CPU を停止します。

2.3 可用性

企業全体でユーザーがただちに情報を入手できることを目標とする組織にとって、高可用性は必須です。これは、特に Sun Fire システムなどの大規模な共用資源システムでは重要です。

2.3.1 Sun Fire システムの可用性の目標

Sun Fire システムの信頼性、可用性、および保守性 (RAS) の目標は、顧客データの完全性を保護し、可用性を最大にすることです。次の 3 つの領域に焦点を絞ります。

- 問題の検出および分離 — 問題を特定し、その問題が伝播されないようにする
- 許容性および回復 — 異常なシステムの動作を吸収および修正するか、あるいは動的に回避する
- 冗長性 — 重大なコンポーネントの二重化

ハードウェアレベルでのデータの完全性を保証するため、データはすべて誤り訂正符号 (ECC) で保護され、制御バスはディスク上のデータに対するパリティチェックによって保護されます。これらの検査によって、エラーが含まれていることが確認されます。

エラーに対する許容性のために、Sun Fire システムには回復機能が設計され、縮退モードでもシステム動作の継続が保証されます。対象多重処理システムであるため、1 つ以上のプロセッサが使用不可の場合でも Sun Fire システムは機能できます。問題からの回復時、障害を特定するためにシステムがすばやく検査され、停止時間を最小限に抑えることができます。システムを冗長ハードウェアで構成すると、停止時間を短縮できます。

2.3.2 Sun Fire システムの高可用性

Sun Fire システムの特徴によって、システムの可用性が通常の市販レベルから高可用性レベルに引き上げられます。これらの特徴は、次のようにグループ化されます。

- 耐障害機能 — すべての障害はそれぞれ、すべてユーザーに対して完全に透過的です。耐障害性のあるシステムの領域では、ユーザーは性能や機能を低下させずに作業できます。
- 回復機能 — この機能によって、障害が発生しリソースが減少していても、処理およびデータへのアクセスを継続できます。通常、この機能では、システムの再起動が必要です。
- 保守性の機能 — この機能によって、障害発生時の修復時間が短縮またはなくなります。

2.3.2.1 冷却装置

Sun Fire システムには、冗長冷却装置があります。1つのファンに障害が発生した場合、残りのファンが自動的に加速されるため、周囲温度が仕様の最高値になってもシステムの操作を継続することができます。したがって、1つのファンに障害が発生した場合でも操作を中断する必要はありません。ファンは、システムの動作中でも交換できます。この場合も、可用性メトリックに悪影響はありません。もちろん、Sun Fire システムには、包括的で失敗のない温度監視機能があり、冷却装置に障害が発生した場合にコンポーネントの温度が超過しないようにできます。

2.3.2.2 AC 電源スイッチ

Sun Fire システムには、最大 4 つの独立した 30A 単相冗長転送スイッチ (RTS) を使用して、AC 電力が供給されます。各 RTS モジュールは、2 つまたは 3 つの 2,200 W の大容量の DC 電源装置に電力を供給します。

AC 接続は、顧客が準備した回路遮断器で制御する必要があり、高いレベルの可用性が必要な場合は、分離された電源グリッド上に確立できます。任意でサン以外のバッテリーバックアップ電源を使用して、ユーティリティーに障害が発生した場合に AC 電力を供給することができます。

2.3.2.3 ECC

Sun Fire システムでは、関連付けられたプロセッサの代わりにデータバッファによって、データエラーの検出、修正、報告のいずれか、またはそのすべてが実行されます。また、インターコネクトで発生するデータエラーも検出されます。このエラーが発生すると、ASIC は記録停止状態になります。この ASIC は、記録停止状態を検出および開始します。その後これらの履歴バッファおよび記録停止状態ビットが読み込まれ、オフライン診断で使用されます。

2.3.3 回復機能

回復機能によって、障害が発生しリソースが減少していても処理およびデータへのアクセスを継続できます。通常、これらの機能には、システムの再起動が必要になります。可用性の式では、この再起動の時間は修復時間として計算されます。

2.3.3.1 DC 電源

Sun Fire 論理 DC 電源システムは、システムボードレベルのモジュールです。各システムボードには、回線保護回路を使用して 56 V の大容量 DC が供給されます。56 V の電圧は、複数の小規模な DC/DC コンバータを使用して、ボードに必要な一定の低電圧に変換されます。DC/DC コンバータに障害が発生した場合、特定のシステムボードだけに影響があります。大容量の DC 電源装置は、特定のシステム構成に必要な数だけ構成する必要があります。標準の冗長構成では、最大 3 つのシステムボード用に 3 つの DC 電源装置があり、Sun Fire E6900 システムでは、最大 6 つのシステムボード用に 6 つの DC 電源装置があります。

2.3.3.2 論理ボード

システムコントローラボードには、クロックソースおよび緊急停止ロジックに加えて、システムコントローラインタフェースがあります。冗長性の目的で、システムに 2 つのシステムコントローラボードを任意に構成できます。

リピータ、CPU/メモリーボード、および入出力サブシステムには、アドレスリピータ、システムデータコントローラ、システムデータクロスバー、その他すべての ASIC に電力を供給する DC/DC コンバータが搭載されています。1 つのリピータボードに障害が発生した場合、システムは縮退モードで操作を継続します。このモードでは、4 つのアドレスバスおよびデータバスのうち 2 つが使用されます。

2.3.3.3 プロセッサ

UltraSPARC s400 プロセッサ、デュアルデータスイッチ、外部キャッシュ SRAM、または関連付けられたサポート ASIC に障害が発生した場合、障害が発生したプロセッサは、POST の設定手順によってシステムのほかの部分から分離されます。構成内で少なくとも 1 つのプロセッサが機能しているかぎり、システムは動作します。

2.3.3.4 メモリー

POST でのメモリーサブシステムの評価が完了すると、障害が発生したメモリーバンクが特定されます。その後、POST は、メモリーコントローラのアドレス照合論理の高い構成可能特性を利用して、信頼性のあるメモリーバンクだけを使用しメモリー構成を再構成します。

2.3.4 冗長コンポーネント

Sun Fire システムの機能によって冗長コンポーネントを構成できるため、顧客先で発生する障害の平均時間、および顧客に対するシステムの可用性がどちらも向上します。必要に応じて、システム内のすべてのコンポーネントを冗長に構成できます。各システムボードは、個別に操作できます。Sun Fire システムは、複数のシステムボードで構築され、本質的に、構成されたボードのサブセットが使用できる場合だけ操作できます。

基本システムボードに加えて、次の冗長構成可能なコンポーネントがあります。

- システムコントローラボード (バージョン 2)
- リピータボード
- AC 電源スイッチ
- DC 電源装置
- 周辺装置コントローラおよびチャネル

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。ソフトウェアでは、複数のパスが保持され、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。

システムコントローラは、コンソールでのインタフェースを持つワークステーションを使用して制御されます。必要に応じて、冗長システムコントローラおよびインタフェースを構成できます。

2.4 保守性

修復時間を短縮するために、Sun Fire システムには多くの保守機能および補助機能が設計されています。これらは、Sun Fire システム管理者および保守プロバイダによって使用されます。

いくつかの機能によって、強制的な停止時間を予定せずに保守を行うことができます。障害の発生したコンポーネントは、現場交換可能ユニット (FRU) を明確に特定できる形式で障害ログに示されます。適切に構成されたシステム内のすべてのボードおよび電源装置は、停止時間を予定せずに、システムの操作中に取り外したり取り付けたりすることができます。

2.4.1 マシンの保守性

コネクタは、ボードを逆に取り付けることを防ぐような形になっています。システムの内部を取り扱うために、特別な工具は必要ありません。これは、適切な安全機関で定義されているように、キャビネット内のすべての電圧が超低電圧 (ELV) であると認識されているためです。

Sun Fire システムの構成には、ジャンパは必要ありません。これによって、新規かアップグレード、またはその両方のシステムコンポーネントの取り付けがたいへん容易になります。システムコントローラおよびリピータボードに必要な特別なスロット以外、スロットに依存するものではありません。

Sun Fire システムの冷却システム設計には、RAS を向上させる機能が含まれています。あらゆる場所に、標準準拠の部品およびコンポーネントを使用しています。FRU および構成部品は、できるだけ工具を使用せずに、すばやく簡単に交換できるように設計されています。

2.4.1.1 DC 電源装置の取り外しおよび取り付け

56 VDC の電源装置は、システムを中断せずにホットスワップできます。この場合、システムが出荷時から電源装置の冗長用に構成されていることが前提になります。

2.4.1.2 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、残りの動作中のファンはシステムコントローラによって高速操作に設定され、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。

2.4.1.3 ドメインの分離

Sun Fire システムには、インターコネクトドメイン機能があり、システムボードを別々のドメインに割り当てることができます。たとえば、あるドメインが本番稼働している間、別のドメインでは実験的に次のバージョンのオペレーティングシステムを実行するか、あるいは障害が発生している可能性がある不良ボードの動作テストを本番と同様の作業で行うことができます。

2.4.2 非並行保守

非並行保守を行うには、システム全体の電源を切る必要があります。

2.4.3 遠隔保守

すべてのシステムコントローラボード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。オペレーティングシステム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での診断を実行できます。

第3章

ハードウェアの概要

Sun Fire システムは、メモリー共有型対称マルチプロセッサ (SMP) で構成される製品ファミリです。

この章では、Sun Fire システムの次の機能について説明します。

- 標準操作 (単純 SMP および OS 機能)
- インターコネクト (OS 起動および RAS 機能でのインターコネクトの詳細)
- コンソールバスインターコネクト (システムコントローラボードによるシステム制御方法の詳細)

3.1 標準操作

標準操作とは、単純に、オペレーティングシステムを標準機能で実行する SMP の標準操作を示します。標準操作は、インターコネクトバスを介して接続された CPU/メモリー装置および入出力装置で構成されます。データインターコネクトは、実際にはクロスバススイッチですが、論理的にはバスです。図 3-1 にこれを示します。

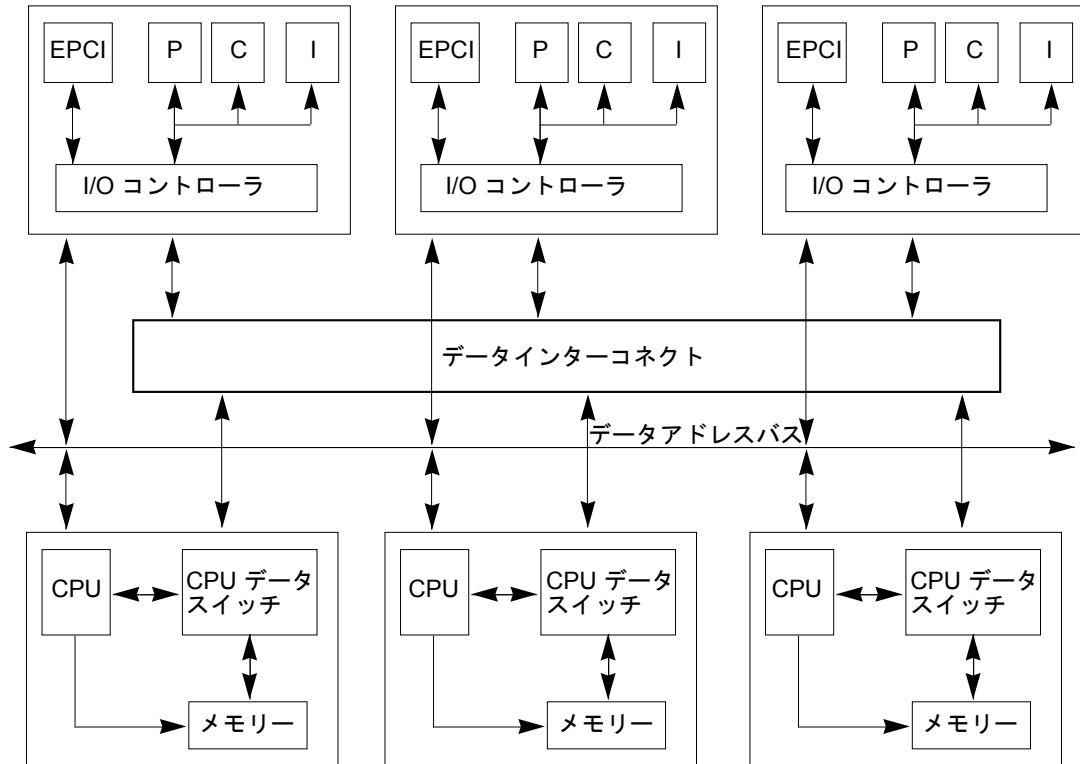


図 3-1 Sun Fire E6900/E4900 システムの標準操作

3.2 データインターコネクト

Sun Fire システムの標準操作は単純なバスのインターコネクトでの操作と似ていますが、実際は 2 つのレベルのリピータまたはスイッチを持つポイントツーポイント交換型インターコネクトです。

このスイッチでは、次のような複雑な機能を実行できます。

- 完全に分離されたパーティションへのシステムの分割
- 論理的に分離されたドメインへのパーティションの分割

オペレーティングシステムを起動し前述の機能を実行するには、システムコントローラがスイッチインターコネクトの論理構造を認識する必要があります。

Sun Fire E6900 システムには、CPU/メモリーボード用の 6 つのスロットがあります。Sun Fire E4900 システムには、CPU/メモリーボード用の 3 つのスロットがあります。各 CPU/メモリーボードには、最大 4 つの UltraSPARC s400 CPU を搭載できます。CPU にはメモリーコントローラも含まれており、各 CPU は 1 つのメモリーバンクで最大 8 つの DIMM をサポートできます。

Sun Fire E6900 システムには、I/O アセンブリ用の 4 つのベイがあります。Sun Fire E4900 システムには、I/O アセンブリ用の 2 つのベイがあります。Sun Fire E6900/E4900 システムは、PCI をサポートします。各 PCI I/O アセンブリには 2 つの I/O コントローラがあり、各 I/O コントローラには 66 MHz の PCI バスと 33 MHz の PCI バスが 1 つずつあります。

Sun Fire E6900 システムは、前世代のシステムの信頼性、保守性、および可用性 (RAS) を大幅に改善するように設計されています。Sun Fire システムは、どのようなハードウェア障害からも回復するように設計されています。システムが冗長電源装置用に構成されている場合は、電源装置障害などが発生してもユーザーに影響を及ぼすことなく障害から回復できます。CPU 障害などでは、障害から回復するには再起動が必要となり、ユーザーに影響を及ぼしますが、適切に構成されたシステムではどのようなハードウェア障害からも必ず回復できます。

アドレスバスとデータバスの扱いはわずかに異なります。アドレスバスには、完全に冗長な 2 つのリピータがあります。アドレスリピータ (AR) 機能は 2 つの AR 間でビットスライスされているため、絶対アドレスリピータバスには 2 つのリピータボードが必要です。Sun Fire E6900 システムでは、標準操作用にデータバスが 4 つのすべてのリピータボード間でビットスライスされています。データ帯域幅が縮小してもすべての機能が保持されるように、任意で 1 組のリピータボードを「ダブルポンプ」モードで使用できます。

リピータボードにはアクティブな装置があります。センタープレーンは比較的保守が困難であるため、Sun Fire E6900/E4900 システムは、アクティブな装置がセンタープレーン上に存在しないように設計されています。

3.3 コンソールバスインターコネクト

システムコントローラ (SC) は、コンソールバスを使用して、システムのほかの部分にあるレジスタの読み取りおよび書き込みを行います。一度にコンソールバスのマスターになることができるのは 2 つのシステムコントローラのうちの 1 つだけです。各システムコントローラはコンソールバスハブ (CBH) に接続され、2 つの CBH が使用するコンソールバスに対してアービトレーションを行います。

第4章

Sun Fire システムのコンポーネント

Sun Fire E6900/E4900 システムは、多くの同じコンポーネントを共有します。これらのコンポーネントには、CPU/メモリーボード、I/O アセンブリ、リピータボード、およびシステムコントローラボードがあります。

4.1 CPU/メモリーボード

Sun Fire E6900/E4900 システムは、同じ CPU/メモリーボードを使用します。このボードは、最大で 4 つの UltraSPARC s400 CPU モジュール、8 つのメモリーバンク (各 CPU に 2 つずつ)、および合計 32 の DIMM (各バンクに 4 つずつの DIMM ソケット) をサポートします。

バンクにある DIMM は、すべて容量やサイズが同じである必要があります。同一ボード上で DIMM のサイズが異なるようにしてください。

4.2 I/O アセンブリ

Sun Fire E6900/E4900 システムは、PCI I/O アセンブリをサポートします。

4.2.1 PCI I/O

I/O アセンブリは論理的にも物理的にも、Sun Fire E6900/E4900 システムで同じです。基本的な PCI I/O アセンブリには、標準 PCI (33 MHz) デバイスボード用の 6 つのスロットに加え、PCI-66 (66 MHz) デバイスボード用の 2 つのスロットがあります。

4.3 リピータボード

Sun Fire E6900/E4900 システムは、回復やアップグレードが以前のシステムより容易かつ迅速に行えるように設計されています。これは、リピータボード上に有効な ASIC を搭載したためです。2 つのリピータボードをシステムに取り付けると、1 つのボードに障害が発生しても 2 つ目のボードから代替バスが使用できます。

リピータボードには、信頼性のための冗長性と高い帯域幅の 2 つの機能があります。システムは、1 つのリピータボードだけで動作できます。リピータボードはスイッチとして機能し、複数の CPU/メモリーボードと I/O ボードに接続します。リピータボードには、アドレスリピータ (AR)、Sun Fire システムデータコントローラ (SDC)、およびデータクロスバー (DX) の 3 つのコンポーネントがあります。

通常の操作では、Sun Fire E6900 システムに、10 個のバス (6 つの CPU および 4 つの入出力) への配信に使用される 4 つのリピータボードがあります。1 つのリピータボードに障害が発生した場合、システムは、隣接する 1 組のリピータボードを使用して縮退モードで動作を続けることができます。その場合は、データ幅が半分にカットされ、2 つのリピータボードが 10 個のバスの経路を指定します。

Sun Fire E4900 システムでは、2 つのリピータボードしかサポートしていないため、2 つのリピータボードが一緒に動作して、5 つのバス (3 つの CPU バスおよび 2 つの入出力バス) の経路を指定します。1 つのリピータボードに障害が発生した場合、データ幅が半分にカットされ、1 つのリピータボードが 5 つのバスの経路を指定できません。

4.4 システムコントローラボード (バージョン 2、拡張メモリー付き)

システムコントローラボードには、システムクロックおよびサービスプロセッサが含まれています。ボード上のプロセッサは microSPARC IIep で、独自の POST/OBP フラッシュ PROM およびメモリーを搭載しています。また、プロセッサは 33 MHz の PCI バスも備えており、そのバス上には 2 つの装置があります。また、システムコントローラボードには、10/100BASE-T Ethernet 接続およびさまざまな装置用の Ebus インタフェースがあります。たとえば、TOD/NVRAM 装置、追加の NVRAM 領域用のフラッシュ PROM、OS コードを保持する大規模フラッシュ PROM、1 つの 16552 デュアルシリアルポート装置などがあります。

各システムに1つのシステムコントローラボードが必要です。Sun Fire E6900 には、冗長用の2つのシステムコントローラが取り付けられています。Sun Fire E4900 システムには、必要に応じて、2つ目のシステムコントローラボードを取り付けることができます。2つのシステムコントローラボードがあれば、一方のシステムコントローラボードに障害が発生した場合でも、もう一方のシステムコントローラボードによって、メインシステムの動作を妨害することなくシステムを制御できます。

システムコントローラボードの主な機能は、次のとおりです。

- システムの設定および起動処理の統合
- 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
- エラーメッセージの分析と記録および修正処理
- システムパーティションおよびドメインの設定
- システムコンソール機能の提供
- ドメインの仮想キースイッチの提供
- 各ドメインのネットワーク Solaris ソフトウェアコンソールの提供
- 各ドメインの仮想時刻の提供

システムコントローラには、ドメイン A コンソール、ドメイン B コンソール、ドメイン C コンソール、ドメイン D コンソール、およびシステムコントローラシェルの5つのポートがあります。システムコントローラシェルでは、次の機能が提供されます。

- 設定の制御
- 環境状態
- ドメインの再構成
- 電源グリッドの電源の投入および切断
- システムコントローラのパスワードの変更
- その他の汎用システムコントローラ機能

システムコントローラソフトウェアは、次の操作によって、システムの起動処理を行います。

- ハードウェアの構成
- ドメインの設定
- コンポーネント (システムボード、電源装置、ファンなど) の電源の投入および切断
- コンポーネントの評価
- ドメインの構築

システムコントローラソフトウェアでは、システムの構成を変更するためのツールが提供されています。また、このツールによってエラーを記録することもできます。

4.4.1 冗長システムコントローラ

Sun Fire E6900/E4900 システムに 2 つのシステムコントローラボードを取り付ける場合、2 つ目のボードは冗長システムコントローラボードになります。各システムコントローラボードは、他方のシステムコントローラボードの健全性および状態を検査できます。メインシステムコントローラボード (SC0) に障害が発生した場合、冗長システムコントローラボード (SC1) がシステムの動作を妨害せずに操作を引き継ぎます。

4.4.2 ドメインの仮想キースイッチ

システムコントローラは、各ドメインに仮想キースイッチを提供します。keyswitch コマンドを使用して、各ドメインの仮想キースイッチの位置を制御します。

4.4.3 Solaris コンソール

システムコントローラは、各ドメインに Solaris ソフトウェアコンソールを提供します。システムコントローラボードの Ethernet ポートまたはシリアルポートは、Solaris ソフトウェアコンソール用の物理接続です。シリアルポートは、一度に 1 つのコンソールしかサポートできません。一方、Ethernet ポートは、同時に多数のコンソールをサポートできます。システムコントローラはこれらの物理接続を多重化して、各ドメインおよびシステムコントローラにコンソールサービスを提供します。

4.4.4 仮想時刻

システムコントローラボード上には、Sun Fire システムの TOD/NVRAM チップがあります。システムコントローラは、この物理的な TOD チップを多重化して、各ドメインおよびシステムコントローラボードに TOD サービスを提供します。また、システムコントローラは、メインシステムコントローラボードと冗長システムコントローラボード間で TOD を同期化します。

4.4.5 環境の監視

Sun Fire システムには、温度、電圧、および電流を監視する多くのセンサーがあります。システムコントローラは、これらの装置を定期的にポーリングします。しきい値を超えると、システムコントローラはさまざまなコンポーネントを停止して損傷を防ぎます。