

SPARC® Enterprise

T5440 サーバ・アーキテクチャ

マルチスレッドテクノロジーを搭載した
UltraSPARC T2 Plus プロセッサの威力

White Paper

2008 年 10 月

目次

全体概要	3
第1章 マルチスレッド・テクノロジーの進化	4
Web2.0 のビジネスの課題	4
マルチスレッド・テクノロジーの革新	5
SPARC Enterprise T5440 サーバ.....	9
第2章 マルチスレッド・テクノロジーを搭載する UltraSPARC T2 plus プロセッサ	15
UltraSPARC T2 Plus: 世界初の大量スレッド処理 SoC	15
次世代へ進化したマルチスレッド設計	17
第3章 サーバ・アーキテクチャ	25
システムレベルのアーキテクチャ	25
SPARC Enterprise T5440 サーバの概要	31
第4章 エンタープライズ・クラスの実装とソフトウェア	36
システム管理テクノロジー.....	36
サーバ管理ソフトウェア	37
マルチスレッド・テクノロジーのスケラビリティとサポート	38
エンド・ツー・エンドの仮想化テクノロジー	40
障害管理と予測的セルフヒーリング	44
第5章 まとめ	46
詳細情報	47

全体概要

ダイナミックに加速する Web2.0 サービスの開発とアプリケーション配備を提供することは、データセンターの運用においてますます重要な課題になっています。サービスは、小規模から開始でき、急速に拡大できる必要があります。それは、高可用性を維持したまま、3 か月ごとに 2 倍以上のキャパシティに増加することさえあります。IT インフラストラクチャのサポートでは、管理負担を増加せずに、このような膨大なスケーラビリティの要求に応えなければなりません。しかし、データセンターのほとんどは、すでに設置面積と電力（エネルギー費の上昇）の両方の面で非常に厳しい状況におかれています。また、データセンターがエネルギー消費や汚染を減らす役割への新たな評価もあります。このような情報は、ユーザや企業の経営者のどちらにとっても可視性が高まっています。

企業は、冗長なインフラストラクチャを統合し、管理を簡素化し、あまり利用されていないシステムを活用する努力をしており、仮想化技術は非常に重要なツールとして知られています。データ損失や破損によって生じる様々な損害はますます増加しており、その影響は財政面のみならず、企業の信頼や評価にもダメージ与えることとなり、セキュリティは非常に重要なものとなっています。その中で最も重要なのが恐らく統合プラットフォームとして機能するサーバであり、計算能力、メモリ容量、I/O スループットなど、様々な方向に速やかにスケールできる必要があります。企業に、このような問題に対処するための専用のインフラストラクチャを用意する余裕はないため、最も要求の高いミッション・クリティカルなアプリケーションに制限を課する結果となります。

業界初の大量スレッド処理のシステム・オン・チップ（SoC: System on Chip）である UltraSPARC T2 Plus プロセッサを搭載する SPARC Enterprise T5440 サーバは、Web 2.0 インフラストラクチャを推進し、その他の厳しいデータセンターの要求に応えるための画期的なパフォーマンスとエネルギー効率を提供します。マルチスレッド・テクノロジーをサポートする業界初のクアッド・ソケット・システムである SPARC Enterprise T5440 サーバは、わずか 4RU（ラック・ユニット）で最大 256 スレッドをサポートし、電力と冷却の厳しい制限に対応しながら、大容量の計算密度を提供します。計算能力のみならず、このサーバは大容量のメモリ・サポート（最大 512GB）と非常に高レベルの I/O スループット、さらに、初期の UltraSPARC T2 の単一ソケット実装時の 4 倍の I/O キャパシティを提供します。

SPARC Enterprise T5440 サーバは、真の高レベルな統合とバランスの取れたシステム設計を採用しており、レイテンシを縮小し、コストを削減し、セキュリティと信頼性を向上させます。このような利点と組み込み型（無償）の仮想化ソフトウェアにより、多層アプリケーションおよびサービス全体が 1 台の SPARC Enterprise T5440 サーバに統合されます。さらに、プロセッサと Solaris オペレーティングシステム（Solaris OS）がともにオープン・ソース・ライセンスにより利用可能であることで、より一層の利点が企業にもたらされます。企業は自由に技術革新でき、ワールドワイドな技術コミュニティに参加できます。

第1章 マルチスレッド・テクノロジーの進化

マルチスレッドテクノロジーを用いた UltraSPARC T1 プロセッサに基づく SPARC Enterprise T1000 および T2000 サーバは、熱狂的な称賛を勝ち取り、市場投入後、急速に浸透していきました。これらのシステムは、1/4 の設置面積と電力で最大 5 倍のスループットを提供し、さらに電力会社^{*1} から初のリベートプログラムが提供されることとなりました。こういったリベートプログラムは世界中で行われるようになりました。今、マルチスレッド・テクノロジーは、Web およびその他アプリケーションの広範囲にわたり絶えず変化している要求を満たすために急速に発展しています。

^{*1} 2006年8月、Pacific Gas and Electric (PG&E)社 は、SPARC Enterprise T1000 / T2000サーバを購入、配備した場合の大幅なエネルギー割引を開始

Web2.0 のビジネスの課題

Web サービスとサービス指向アーキテクチャ (SOA) の普及によって特徴づけられる新しい「参加の時代」は、豊富な新しいコンテンツと広帯域サービスがかつてない大量のユーザに提供する能力が期待されています。さまざまな業界の企業はこの移行を通じて、より大きい市場を相手にし、コストを引き下げ、それぞれの顧客を深く理解することを望んでいます。同時に、有線、無線を問わず、ますます幅広い種類のクライアント・デバイスが何百万もの人々の日常生活にネットワーク・コンピューティングをもたらしています。この流れは、データセンターのスケラビリティと容量に関する要件の見直しを迫り、設置スペース／電力／冷却の制限と衝突する事態をも招いています。

• Webシステムにおけるスケールアップ型の拡張

Webアプリケーションはインフラの配置に新たなペースや必要性を生み出しています。企業は、開発時間やサービス開始までの時間を短縮しなくてはならない一方で、冗長性のある高品質で高性能なアプリケーションやサービスを提供しなければなりません。企業の多くは迅速にスケール変更をできるような余力を持ち、小規模からスタートしなければなりません。というのも、新たな顧客や革新的なWebサービスには数ヶ月で容量が倍増してしまうものもあるからです。

同時に、企業は現在のデータセンターで利用可能な電力、冷却、スペース中で、環境負荷を低減しなければなりません。運用コストがIT予算の最大40%を占める中で、システム管理コストについても見直しが必要となっています。簡易化とスピードが最重要事項であり、これが、ダイナミックに景気に反応する能力を与えます。企業はまた、ベンダーのロックイン状態を無くすことを目指しています。それは企業が過去、現在、そして将来の投資を守ろうと考えているためです。オープン・スタンダードで構築されたオープンプラットフォームは、参入/退出コストを削減すると同時に、最大限の柔軟性を提供する手助けとなります。

● データセンター運用における仮想化と環境効率化の推進

多くのデータセンターは、サービスの規模と共に、業務的および技術的な処理を並行して行なうため、標準プラットフォームの配置を少なくすることに重点を置いています。このプロセスには、使用中の広範囲にわたるサーバ・インフラを、ビジネスの俊敏性を高め、災害復旧を改良し、運用コストを削減するための効果的な仮想化ソリューションとの統合が含まれます。このように着目すると、データセンターが消費する電力のワットあたりのパフォーマンス量を改善することで、エネルギー削減や、データセンターの容量制限の解決を促進することができるのです。

環境効率は具体的なメリットをもたらします。二酸化炭素排出量を削減することでエコロジー改善となり、法的な面で企業の社会的責任を達成し、企業が支払う電気代も改善します。システムはより密度の高く能力のあるコンピューティング・インフラに統合されるので、データセンター用面積の需要が少なくなります。注意深く計画を進めると、この方法は過度の熱負荷が原因のハードウェア障害を削減することとなり、サービスの使用可能時間や信頼性を改善することができます。高レベルの標準RAS（信頼性・可用性・保守性）を装備したサーバは非常に需要があると考えられます。

● 企業におけるスピードと安全性

企業は顧客やパートナーとの全てのコミュニケーションを保護することに対して非常に関心があります。リスクに応じて、エンド・ツー・エンドの暗号化がセキュリティや、守秘義務による信用への高まりから必要となっています。暗号はまた、ストレージにとっても重要であり、保存・蓄積されたデータの安全を守り、改ざんやデータの不整合を検出するメカニズムを作り出しています。

しかし、増加する暗号化の費用は、コンピュータのリソースへの負担増となります。セキュリティもまた、顧客の経験や低速度のトランザクションに対しボトルネックを与えないような回線速度が必要となります。パフォーマンスへの影響やコストの増加が無く、クライアントのセキュリティやプライバシーを確保するのに役立ち、企業にはコンプライアンスに役立つものがソリューションとして提供されなければなりません。

マルチスレッド・テクノロジーの革新

既存のプロセッサやシステムの機能を追い抜くためには、新しいアプローチが必要でした。

ムーアの法則と従来のプロセッサ設計の限界

たびたび引用されるムーアの法則では、最小部品コストに関連する集積回路におけるトランジスタの集積密度は、18~24ヶ月ごとに倍になると述べられています。ムーアの法則が提唱されてから30年以上経ち、プロセッサのパフォーマンスは新たな段階に向

かっています。プロセッサの設計者は、目標である命令レベルの並列性(instruction-level parallelism：ILP)と共に、複雑なプロセッサを構築するためにチップ面積を利用しました。現在、この従来のプロセッサは、シングル・インストラクション・パイプラインを高速化するため、様々な高度な方式で、非常に高い周波数を実現しています。

- 大容量キャッシュ
- スーパースcalar設計
- アウト・オブ・オーダー
- 超高速クロック周波数
- パイプライン
- 投機実行、プリフェッチ

これらの技術は、重要なマルチギガヘルツの周波数を持った高速プロセッサを誕生させた一方で、複雑でパワー重視のプロセッサを生み出す結果となり、現在のデータセンターで見られるようなタイプの業務に適さない場合が出てきています。実際、多くのデータセンターの業務は、これらのプロセッサに提供され苦勞して得たILPを活用することができていません。高度に共有化されたメモリや高度な同時ユーザ、計算業務などのアプリケーションは、単一スレッドでの処理(ILP)よりも一般的に多数の同時スレッド処理(thread-level parallelism, TLP)に集中しています。

事態をますます悪化させているのは、既存のアプリケーションにあるILPのほとんどがすでに抽出され、メリットが少なくなっていることです。さらに、マイクロプロセッサの周波数スケールリングは、マイクロプロセッサの電力問題のために横ばい状態です。より高速なクロックで、一連のプロセッサの生成は、以前より一層パワーのあるものが要求されているようであり、実際は、マイクロプロセッサの周波数スケールリングは、2-3GHz内で横ばい状態です。パイプラインのスーパーScalar・プロセッサを配置することはより能力が必要であり、プロセッサを冷却する根本的な機能によってこの方法を制限しています。

マルチコア・プロセッサにおけるマルチプロセッシング

これらの問題を解決するために、マルチプロセッサ業界の大半はムーアの法則で提供されたトランジスタを使用して、単一の物理的なダイ上に2もしくは4つの標準的なプロセッサ・コアをグループ分けし、マルチコア・プロセッサ（もしくはマルチスレッド・プロセッサ）を設計してきました。多くのマルチスレッド設計で作り出された個別のプロセッサ・コアには、以前の単一プロセッサ・チップほどのパフォーマンスは認められず、事実、単一スレッド・アプリケーションを実行させる場合には単一コア・プロセッサより遅いと認識されていました。しかし、複数のプログラム(または、複数のスレッド)が並列に処理(スレッドレベルでの並列処理)出来るため、チップ全体として性能は向上しています。

残念ながら、大半の現在使用可能な（もしくはまもなく使用可能になる予定の）チップ・マルチプロセッサは既存の（シングル・スレッド）プロセッサ設計からコアを単

純に複製しています。これでは、総パフォーマンスの中でわずかな改善しか得られません。それは、メモリ速度やハードウェア・スレッド・コンテキスト・スイッチングといったキー・パフォーマンス問題については着手できていないためです。その結果、これらの設計がいくつかの付加的なスループットやスケーラビリティを生み出す一方で、相当量の電力を消費し、著しい熱を発生させ、全体のパフォーマンスで釣り合いのとれた高速化にはならないのです。

最先端マルチスレッド・テクノロジー

我々は、早くからプロセッサ速度とメモリ・アクセス速度の間の不均衡を認識していました。プロセッサ速度が2年ごとに2倍になっていくのに対して、メモリのアクセス性能は6年ごとにしか倍増しません。その結果、メモリ待ち時間がアプリケーションのパフォーマンスにおいて大きな影響力を持ち、クロック速度におけるメリットを大きく消失しています。この大きな不釣り合いは、メモリがアクセス性能ではなく、密度やコストを重視した結果です。

残念ながら、プロセッサとメモリ速度の間に関連したギャップは超高速プロセッサに85%のアイドル状態の時間を残し、メモリ処理の完了を待つことになります。皮肉なことに、従来のプロセッサの実行パイプラインはより早くより複雑になるにつれ、メモリ待ち時間の影響は大きくなります。更に悪いことに、アイドル状態のプロセッサは電力が必要であり、熱を発生させます。周波数（ギガヘルツ）が実際のパフォーマンスには追従できていないのが現状です。

マルチスレッディングは、UltraSPARC T1プロセッサと共に最初に導入され、CMPの優位性を利用し、重要な機能も付加しています。それは周波数よりむしろスレッドと共に拡大することです。従来の単一スレッド化したプロセッサや大半のCMPプロセッサと異なり、ハードウェアのマルチスレッド・プロセッサのコアは、アクティブなスレッド間での急速なスイッチングを許容しています。それ以外のスレッドはメモリで失速します。図1はマルチスレッド、Fine-Grained multi threading (FG-MT)、マルチコア・マルチスレッド・テクノロジーの違いを示しています。ここで鍵となるのは、マルチスレッド・プロセッサ内の各コアは各クロック・サイクル上のマルチプル・スレッド間でスイッチするよう設計されていることです。その結果、プロセッサの実行パイプラインはアクティブな状態で実際の作業を行っているのと同時に、失速したスレッドのメモリ動作が継続しています。

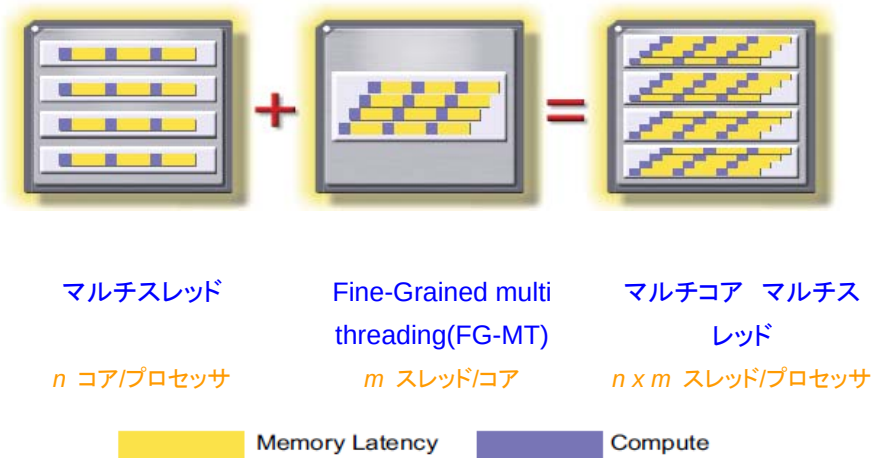


図 1. マルチコア マルチスレッド・テクノロジーは、マルチプロセッシングとFG-MTを結合したものです。

マルチスレッディングはどんなクロック・サイクルでも実業務を行うために実行パイプラインの能力を上げているので、実質的な価値を提供できます。多くの実行スレッドがリソースを共有しているので、実行パイプラインが大きく強化されています。プロセッサとメモリ・サブシステムはプロセッサ実行パイプラインに並んでアクティブなままなので、メモリの待ち時間というネガティブな要素が効率的に解消します。それぞれのコアは単純なパイプラインを実装し、そのパイプラインは周波数よりもスレッドの拡張性を重視しているため（ILP上のTLPを重視）、より冷却でき、起動の際のエネルギーの省エネルギーを実現しています。この革新的な方式がマルチスレッド・プロセッサ技術です。（コアごとのマルチプル・アクティブ・スレッド・コンテキストとマルチプル・物理インストラクション・実行パイプライン（各コアに1つ））。さらに、UltraSPARC T2/UltraSPARC T2 Pulsプロセッサでは更なるスケーラビリティの向上のため、コアあたり2つの実行パイプラインを実装しています。

UltraSPARC T2 Plus プロセッサ

複雑なシングル・スレッド・プロセッサとは異なり、マルチスレッド・プロセッサは利用可能なトランジスタ・バジェットを利用して、複数のハードウェア・マルチスレッド・プロセッサ・コアを1つのチップ・ダイに実装しています。UltraSPARC T2プロセッサは、CMTモデルのレベルを引き上げ、プロセッサあたり最大8つのコアを提供します。各コアは、それぞれ2つの独立したパイプラインを介して最大8つのスレッドをサポートし、周波数を上げずに、UltraSPARC T1プロセッサのスループットを効果的に2倍に上げます。新しいUltraSPARC T2 Plusプロセッサは、実証済みのマルチスレッドアーキテクチャのメリットをマルチソケットの実装まで拡張して、ミッション・クリティカルなOLTPデータベースのワークロードに対応します。多くのシングル・スレッド・プ

ロセッサ設計とは異なり、UltraSPARC T2 Plusプロセッサは利用可能なトランジスタ・バジェットを使用して、単一プロセッサ・ダイのホスティングで、大量スレッド処理のシステム・オン・チップ (SoC) を実装しています。

- プロセッサあたり最大64スレッド (8コアx,8スレッド)
- オン・チップLevel1、オン・チップLevel2のキャッシュ
- コアごとの浮動小数点機能
- コアごとの暗号化アクセラレータ
- オン・チップPCI Expressインタフェース
- オンチップのキャッシュ・コヒーレンス・ユニット

SoC設計により、UltraSPARC T2 PlusプロセッサはCPUの汎用性を大幅に向上させ、8つの浮動小数点演算ユニット (コアあたり1個) を構築します。拡張された浮動小数点演算能力は、UltraSPARC T2 Plusプロセッサを従来のマルチスレッディングに配慮したデータセンター・スループット・アプリケーションのみならず、数値計算アプリケーションの世界にまで広げました。無償のセキュリティと暗号化処理アクセラレーションは、オンチップでコア単位での暗号化処理アクセラレータによって提供されます。さらに、UltraSPARC T2 Plusプロセッサからのデータ入出力機能は、統合されたオンチップPCI Expressインタフェース、メモリ管理ユニット、キャッシュ・コヒーレンス・ロジック、およびマルチソケット・システム設計を促進するコヒーレンス・リンクによって効果的に支援されます。

SPARC Enterprise T5440 サーバ

SPARC Enterprise T5440サーバは、UltraSPARC T2 Plusプロセッサの多量のリソースをコスト効率の良いマルチソケット・コンピューティング・プラットフォームに導入するように設計されています。クアッド・ソケットのSPARC Enterprise T5440サーバは、デュアル・ソケットのSPARC Enterprise T5240サーバの最大2倍のスループットを提供し、ワット当たりのパフォーマンスに関して競合をリードしています。エンタープライズ機能とデータセンターのデザイン・フォーカスを、統合された暗号化処理アクセラレーションとオンチップI/Oテクノロジーに結び付け、従来のシステム設計からの真の脱却を実現しています。



図 2. 4RUで最大4つのUltraSPARC T2 Plusプロセッサを提供するSPARC Enterprise T5440 サーバ

概要

統合プラットフォームである SPARC Enterprise T5440 サーバは、ロジカル・ドメイン (LDoms) および Solaris コンテナなどの組み込み型の無償の仮想化技術により、拡張性とバランスのとれたシステム設計を実現する優れたサーバです。UltraSPARC T2 Plus プロセッサの最大 4 つのソケット、大容量のメモリ・サポート、および相当量の I/O スループットを持つ SPARC Enterprise T5440 サーバは、4RU(ラック・ユニット)だけで、従来は全ラックを必要としたインフラストラクチャを提供できます。コア単位の浮動小数点演算能力により、マルチスレッド・テクノロジーのメリットは、マルチスレッドのコマーシャル・ワークロードに加え、技術的ワークロードにも利用可能です。SPARC Enterprise T5440 サーバは、SPARC Enterprise T5120/T5220 および T5140/T5240 サーバを補完するように設計されており、マルチスレッドテクノロジーを利用したシステムの能力はミッドレンジ・ワークロードの取り組みにも使用されます。

- **効率的かつ予測可能なスケーラビリティ**

SPARC Enterprise T5440サーバはそれぞれ、最大256スレッド、最大512GBメモリ、大容量I/O帯域幅をわずか4RUでサポートします。つまり、このようなサーバは統合および仮想化に理想的なプラットフォームといえます。

SPARC Enterprise T5440サーバは、優れた計算能力とI/Oスループットを提供しつつ、消費電力(実効値)は1,800ワット未満であるため、厳しい電力、冷却、および設置面積の制約を持つ組織において、成長するITおよびWebインフラストラクチャの要求に応える上で役立ちます。

- **“Time to Market”への加速**

Solaris OSを実行するSPARC Enterprise T5440サーバは、初期のUltraSPARCシステムとの完全なバイナリ・コードの互換性を提供するため、投資を保護し、開発時間を短縮できます。

- **簡易化された管理**

SPARC Enterprise T5440サーバは、SPARC Enterprise T5120/T5220/T5140/T5240サーバと互換性を持つILOM (Integrated Lights Out Management) サービス・プロセッサを提供します。ILOMは、遠隔地からの監視および管理を支援するためのコマンド・ライン・インタフェース (CLI)、Webベースのグラフィカル・ユーザ・インタフェース (GUI)、インテリジェント・プラットフォーム管理インタフェース (IPMI) 機能を提供します。これらのシステムのILOMは、SPARC Enterprise T1000およびT200サーバに精通している管理者に対して、ALOM (Advanced Lights Out Management) 下位互換性モードを提供します。

- **業界で最もオープンなプラットフォーム**

SPARC Enterpriseサーバは、GPL(GNU General Public License)のもと供給される唯一の主流プロセッサおよびハイパーバイザーという点で、業界で最もオープンなプラット

フォームです。Solaris OSは、フリーでオープンであり、完全なバイナリ・コード互換性と企業規模の機能を提供します。

- **仮想化と統合用の業界最高のツール**

マルチスレッド・テクノロジーは統合に適しており、各階層の技術スタックで仮想化のための低レベルのマルチスレッド・サポートを提供します。論理ドメインはUltraSPARC T2 Plusプロセッサのソケットあたり64スレッドを有効利用して、マルチゲストのオペレーティング・システム・インスタンスをサポートする一方で、Solaris コンテナはSolaris OSインスタンスの中で仮想化を提供します。先進のZFSファイル・システムは、ストレージの仮想化により大幅な拡張性を提供します。

- **”eco-responsible”を継承**

SPARC Enterprise T1000/T2000は業界初の”eco-responsible”サーバです。SPARC Enterprise T5440サーバもこの考え方を踏襲し、広範囲に渡る業務用、技術用の処理で、ワットあたりのパフォーマンスにおいて最高の成果を出しています。さらに、UltraSPARC T2 Plusプロセッサはコアとメモリの両レベルで独自の電力管理機能を内蔵しています。

- **システムとデータセンターの信頼性**

信頼性はアプリケーションの可用性を維持し、コストを抑えるうえできわめて重要です。SoC 設計によって統合レベルが上がると同時に、SPARC Enterprise T5440 サーバが部品数の大幅な削減をもたらし、それが信頼性／可用性／保守性 (RAS) の水準向上につながっています。消費電力を削減し、ワット当りの性能を引き上げることによって、発熱量とそれに付随する問題が大幅に減少します。インテリジェントなファン制御と振動監視機能によって問題を事前に検知できます。Solaris 予測的セルフヒーリングなどのテクノロジーがハードウェアに統合されており、システムの可用性維持に役立ちます。

- **ゼロ・コスト・セキュリティ**

現在、電子的な侵入や窃盗が多発しており、今日ほど安全な通信やデータ保護が重要視されたことはありません。UltraSPARC T2 Plus プロセッサには、最大8つの暗号化アクセラレータが搭載されるため、ネットワーク上の通常のテキストを送信したり、ストレージ・システムに通常のテキストを保存しておいたりする必要はありません。SPARC Enterprise T5440 サーバは、競合他社のプロセッサや専用暗号化アクセラレータ・カードと比べて、1秒あたりより複雑な暗号化機能をサポートして、システムへのオーバーヘッドを最小化しています。

革新的なシステム設計

データセンターのシステム設計者において、個々のシステムの機能ではカバー出来ない、特殊で差し迫ったニーズがあります。

密度、パフォーマンス、スケーラビリティは、すべて必要な検討項目ですが、システム側には実用性も必要であり、電力、冷却、保守性を考慮した現在のデータセンターの戦略に適合したものでなければなりません。

SPARC Enterprise T5440 サーバは、革新的なデザイン哲学を共有しています。そのデザイン哲学には以下のものが含まれます。

- **最大演算密度：**SPARC Enterprise T5440 サーバは、CPU コアやメモリ、ストレージ、I/O に関して優れた密度を提供しています。
この密度への重点的な取り組みにより、4RU(ラック・ユニット)を実現しました。
SPARC Enterprise T5440 サーバは、4RU で以前のサーバや競合のサーバの全ラックに取って代わることが出来ます。
- **共通の共有管理：**SPARC Enterprise T5440 サーバは、SPARC Enterprise T5120/T5220/T5140/T5240 サーバと共有するサービス・プロセッサにより、管理の容易性と保守性を提供するように設計されています。システムとコンポーネントは簡単に識別できるように設計されており、また、ホットスワップ・コンポーネントによりオンラインでの交換が可能です。
- **継続的な投資の保護：**SPARC Enterprise T5440 サーバは投資の保護を最大限にするように設計されています。

マルチスレッド・テクノロジーのような革新的なテクノロジーと同様に、Solaris のバイナリ互換性はアプリケーションが修正を加えることなく動作することを意味しています。

表 1 は SPARC Enterprise T5140/T5240/T5440 サーバの比較です。

表 1. SPARC Enterprise T5140 / T5240 / T5440 サーバ 特長

特長	SPARC Enterprise T5140 サーバ	SPARC Enterprise T5240 サーバ	SPARC Enterprise T5440 サーバ
CPU	1.2GHz UltraSPARC T2 Plus プロセッサ (4, 6, 8 コア) ×2	1.2 GHz (6, 8 コア) または 1.4 GHz (8 コア) の UltraSPARC T2 Plus プロ セッサ ^c ×2	1.2 GHz または 1.4 GHz (8 コア) UltraSPARC T2 Plus プロセッサ×2~4
スレッド数	最大 128	最大 128	最大 256
メモリ容量	最大 64GB (1, 2, 4 または 8GB FB-DIMM)	最大 128GB (1, 2, 4 または 8GB FB-DIMM) ^d	最大 512GB (2, 4 または 8GB FB-DIMM)
内蔵ディスク・ ドライブ	最大 4 台の 2.5 インチ SAS 73 または 146 GB ディス ク・ドライブ	最大 8 または 16 台の 2.5 インチ SAS 73 または 146 GB ディス ク・ドライブ	最大 4 台の 2.5 インチ SAS 73 または 146 GB ディス ク・ドライブ
記憶装置	CD-RW/DVD-RW	CD-RW/DVD-RW	CD-RW/DVD-RW
USB	USB 2.0 ポート×4	USB 2.0 ポート×4	USB 2.0 ポート×4
PCI	PCI Express x8 : 1 スロット PCI Express x8 or XAUI : 2 スロット	PCI Express x8 : 4 スロット PCI Express x8 or XAUI : 2 スロット	PCI Express x8 : 6 スロット PCI Express x8 or XAUI : 2 スロット
Ethernet	オンボード Gigabit Ethernet (10/100/1000) : 4 ポート	オンボード Gigabit Ethernet (10/100/1000) : 4 ポート	オンボード Gigabit Ethernet (10/100/1000) : 4 ポート
	10 Gb Ethernet ポート (XAUI スロット):2 ポート	10 Gb Ethernet ポート (XAUI スロット):2 ポート	10 Gb Ethernet ポート (XAUI スロット):2 ポート
電源	AC 電源×2 (N+1 冗長)	AC 電源×2 (N+1 冗長)	AC 電源×4(2+2 冗長)
ファン	ファン・トレイ×6(活性交 換、2 ファン/1 トレイ、 N+1 冗長)	ファン・トレイ×5(活性交 換、2 ファン/1 トレイ、 N+1 冗長)	ファン・ユニット×4(活性 交換、N+1 冗長)
フォーム・ファクタ	1RU(ラック・ユニット)	2RU(ラック・ユニット)	4RU(ラック・ユニット)
オペレーティング・ システム	Solaris 10 8/07 OS 以降、 必須パッチ	Solaris 10 8/07 OS 以降、 必須パッチ	Solaris 10 5/08 OS 以降、 必須パッチ

a: 10Gigabit Ethernet ポートにアクセスするためにはオプションの XAUI アダプタカードが必要。

XAUI アダプタカードは PCI Express スロットを占有

b: XAUI アダプタカードを使用すると、1 つの RJ45 Gigabit Ethernet ポートが使用できなくなります。

2 つの XAUI ポートを使用した場合には、Gigabit Ethernet ポートは 2 ポートのみ利用可能。

c: 1.4GHz プロセッサ且つディスクドライブ 16 台搭載可能装置を使用時には、200-240V の電源が必要

d: 256GB の最大メモリ構成には、オプションのメモリメザニンキットが必要

優れた信頼性、可用性、保守性(RAS)

SPARC Enterprise T5440 サーバは優れた RAS 機能 (信頼性、可用性、保守性)を提供します。信頼性の高い部品と、コンポーネント数が減少したことによって、システム・エラーの発生を最小限に抑えることができます。2つの PCI Express コントローラ、および複数プロセッサ構成が可能なことから、システム障害からの回復力を高めています。さらにこのサーバには、冗長構成が可能な、活性交換可能なディスクドライブ、電源装置、ファンと共に、プロセッサコアやスレッドのオフライン機能などの機能を持っています。以下に示す SPARC Enterprise T5440 サーバの重要な設計要素は、IT サービスの信頼性の向上に貢献しています。

- 部品点数の削減
- プロセッサ・コアとスレッドのオフライン機能
- 冗長構成かつ、活性交換可能なコンポーネント
- パリティによるエラー訂正機能
- システム監視
- ILOM サービスプロセッサ
- 優れたエネルギー効率
- 強力な仮想化技術
- 包括的な障害管理

第 2 章

第 2 章 マルチスレッド・テクノロジーを搭載する UltraSPARC T2 plus プロセッサ

UltraSPARC T2/UltraSPARC T2 Plusは業界初のシステム・オン・チップ（SoC : System on Chip）と呼べるプロセッサです。今日までシステム・レベルで提供されてきた多くの機能と、コア/スレッドを1つのプロセッサ上に統合した革新的なプロセッサです。

UltraSPARC T2 Plus: 世界初の大量スレッド処理 SoC

UltraSPARC T2 Plusプロセッサは、コンピューティング、セキュリティ、そしてI/Oを1つのチップに統合することにより、高価なカスタム・ハードウェアとソフトウェアの開発を不要としました。従来のUltraSPARCプロセッサとのバイナリ互換性を維持した上で、これほど小さいスペースで、少ない電力要件で済むプロセッサは他にありません。企業は、最大限の効率と予測可能性を提供する新しいネットワーク・サービスを迅速に拡張できます。図3に、UltraSPARC T2 Plusプロセッサを示します。



図 3. マルチスレッドテクノロジーを搭載するUltraSPARC T2 Plus プロセッサ

表2 に、UltraSPARC T2/UltraSPARC T2 PlusプロセッサとUltraSPARC T1プロセッサの比較を示します。

表 2. UltraSPARC T1 および UltraSPARC T2/UltraSPARC T2 Plus プロセッサの特長

特長	UltraSPARC T1	UltraSPARC T2	UltraSPARC T2 Plus
コア数／プロセッサ	最大 8	最大 8	最大 8
スレッド数／コア	4	8	8
スレッド数／プロセッサ	32	64	64
ハイパバイザー	あり	あり	あり
ソケット数	1	1	2 または 4 ^a
メモリ	メモリ・コントローラ ×4、 最大 16DIMM	メモリ・コントローラ ×4、 最大 16 FB-DIMM	メモリ・コントローラ ×2、 最大 16 FB-DIMM
キャッシュ	16KB L1 命令キャッシュ、 8KB L1 データ・キャッシュ、 3MB L2 キャッシュ (4 バンク、12 ウェイ・アソ シアティブ)	16KB L1 命令キャッシュ、 8KB L1 データ・キャッシュ、 4MB L2 キャッシュ (8 バンク、16 ウェイ・アソ シアティブ)	16 KB L1 命令キャッシュ、 8 KB L1 データ・ キャッシュ、 4 MB L2 キャッシュ (8 バンク、16 ウェイ・ アソシアティブ)
テクノロジー	90nm テクノロジー	65 nm テクノロジー	65 nm テクノロジー
浮動小数点演算ユニット	1 FPU／チップ	1 FPU／コア、8 FPU／ チップ	1 FPU／コア、8 FPU ／チップ
整数演算ユニット	実行ユニット×1	実行ユニット×2／コ ア	実行ユニット×2／コ ア
暗号化処理機能	高速モジュール算術 演算 (RSA)	ストリーム処理ユ ニット／コア、市場で 広く使用されている 10 種類の暗号化アル ゴリズムに対応	ストリーム処理ユ ニット／コア、市場 で広く使用されてい る 10 種類の暗号化 アルゴリズムに対応
ネットワーク・ インタフェース	—	10Gb Ethernet×2	—
I/O インタフェース	—	PCI Express(x8)	PCI Express(x 8)、 コヒーレンシ・ロ ジックとリンク (4.8Gb/sec)

a. External Coherency Hub を4ソケット実装に使用

次世代へ進化したマルチスレッド設計

開発チームが次世代 マルチスレッド・プロセッサの設計を始めたとき、主に次の目標を想定しました。

- UltraSPARC T1 プロセッサのスループット性能を 2 倍に引き上げ、増え続ける Web アプリケーションの需要を満たす
- 浮動小数点性能を向上させ、これまで以上に大きく、多様性のあるワークロードをサポートする
- ネットワーク集中型コンテンツに対応するため、ネットワーク機能を高速化する
- データセンターの暗号化をエンド・ツー・エンドで行う
- サービス水準を引き上げ、停止時間を短縮する
- データセンターのキャパシティを改善しながら、コストを削減する

マルチスレッディングアーキテクチャはきわめて柔軟性が高いため、プロセッサ／コア／統合化コンポーネントをモジュール方式でさまざまな組み合わせることができます。社内のエンジニアリング作業では、上記の考慮事項に基づいて、すぐれた UltraSPARC T1 のアーキテクチャをさらに改良することに関して、各種の方法が比較検討されました。社内のエンジニアリング作業では、上記の考慮事項に基づき、UltraSPARC T1 のアーキテクチャをさらに改良する方向で、様々な比較検討なされました。例えば、単純にコアの数を増やせばスループット性能は向上しますが、使用するダイ面積も広くなり、浮動小数点プロセッサなどの統合コンポーネントの統合が難しくなります。

最終的な UltraSPARC T2 Plusプロセッサの設計において、パフォーマンス向上時の最大のボトルネックはメモリ・レイテンシであると結論付けました。そこで、各コアがサポートするスレッド数を増やすこと、ネットワーク帯域幅を増やすことで、UltraSPARC T2 Plusは UltraSPARC T1プロセッサの倍のスループットを得るというアプローチに落ち着きました。

図4に、8コアのUltraSPARC T2 Plusプロセッサが処理するスレッド・モデルを示します。

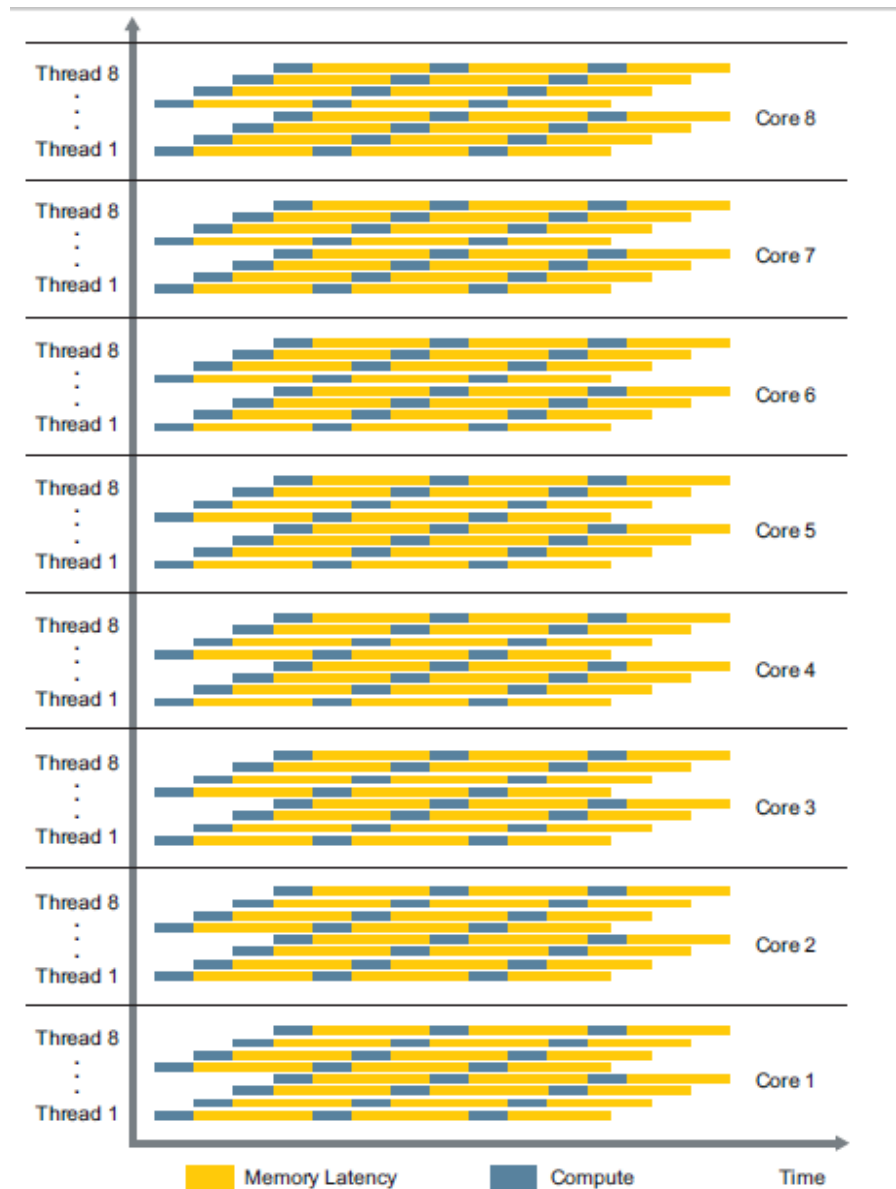


図 4.8 コアの UltraSPARC T2 Plus プロセッサ 1 つで最大 64 スレッドを実行でき、各コアで同時に最大 8 つのスレッドを実行できます。

各UltraSPARC T2/UltraSPARC T2 Plusプロセッサは最大 8 コアからなり、各コアでは最大 8 スレッド (64 スレッド/プロセッサ) のスイッチングが可能です。さらに、各コアには 2 つの整数実行ユニットが搭載されているため、1 つの UltraSPARC コアで同時に 2 つのスレッドを実行することが出来ます。

UltraSPARC T2 Plus プロセッサ・アーキテクチャ

UltraSPARC T2 Plusプロセッサの洗練された堅牢なアーキテクチャによって、スループットコンピューティング・イニシアチブをより一層推進します。そして、アプリケーション性能の最大化を目指します。

図 5 に UltraSPARC T2 Plusプロセッサのブロック・ダイアグラムを示します。

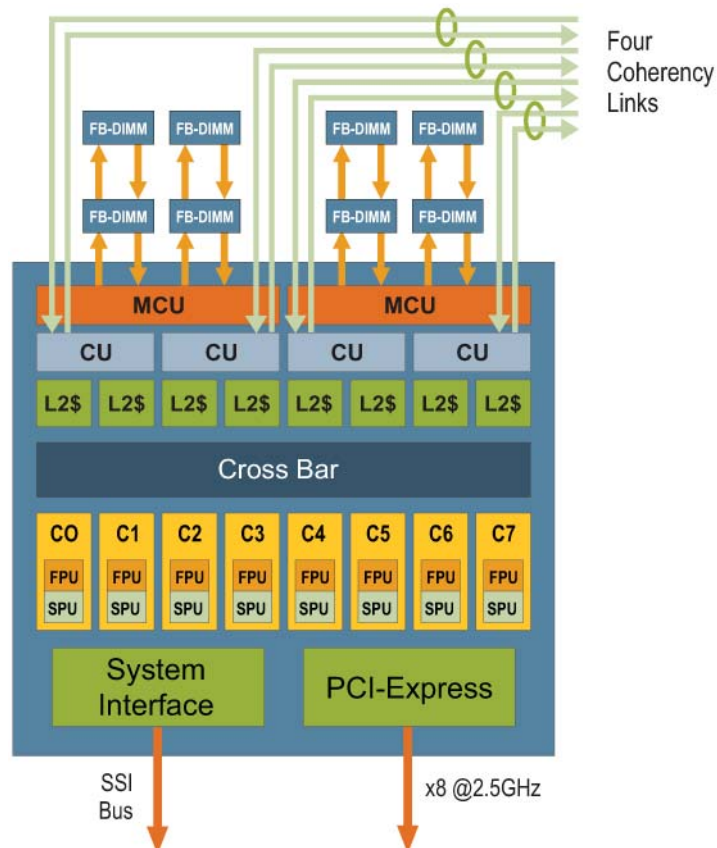


図 5. UltraSPARC T2 Plusプロセッサは、4 本のコヒーレンシ・リンクを用いて、他の最大4つのプロセッサと接続します。

UltraSPARC T2 Plusプロセッサの8つのコアは、フル・オンチップの非ブロッキング型8×9クロスバー・スイッチによって相互接続されます。このクロスバーは、各コアをL2キャッシュの8つのバンクに接続し、さらにI/O用のシステム・インタフェース・ユニットに接続します。クロスバーは約300GB/秒の帯域幅を提供し、コアからバンクへの書き込みでは8byte、バンクからコアへの読み取りでは16byteをサポートします。システム・インタフェース・ユニットは、個別のキャッシュ・バンクによって、ネットワークとI/Oをメモリに直結します。FB-DIMMの採用により、メインメモリ、キャッシュ間で上り/下り専用レーンをサポートし、パフォーマンスを向上すると同時にレイテ

ンシを低減させています。このアプローチによって、DDR2 以上の帯域幅が得られます。

各コアにフル・パイプラインの浮動小数点演算ユニットとグラフィック・ユニット (FGU : Floating point / Graphics Unit) を実装するとともに、ストリーム処理ユニット (SPU : Stream Processing Unit) も実装しています。FGU は UltraSPARC T1 をはるかに上回る浮動小数点演算性能を発揮し、SPUは、市場で広く使用されている10 種類の暗号化アルゴリズム(DES, 3DES, AES, RC4, SHA1, SHA256, MD5, 鍵長 2048 bit の RSA, ECC, CRC32)で暗号化を高速に処理します。これら暗号化処理ユニットのオンチップ化により、CPU性能の劣化を気にすることなく、追加コストを気にすることなく、高速なエンド・ツー・エンドの暗号処理が可能となります。

UltraSPARC T2 Plusアーキテクチャは、マルチソケットのシステム実装をサポートするために、4つのコヒーレンシ・ユニット (CU) を提供します。4つのコヒーレンシ・チャンネル (コヒーレンシ・リンク) が提供され、1つのチャンネルがそれぞれのコヒーレンシ・ユニットに関連付けられます。これらのリンクは、FB-DIMMインタフェースを介してキャッシュ・コヒーレンシ (スヌーピー) プロトコルを実行し、各方向で200Gbpsを超える最高4.8G転送/ポートを行います。

UltraSPARC T2 Plusプロセッサは、2ソケットと4ソケットの実装をともにサポートできます。UltraSPARC T2 Plusのデュアル・ソケット実装 (SPARC Enterprise T5140および T5240によって提供される) は、プロセッサの4つのコヒーレンシ・リンクを相互接続するため、回路を追加する必要はありません。4ソケットの実装では、追加のロジックを4つのExternal Coherency Hub (ECH) の形態で追加する必要があります。図6に、一般的な4ソケット実装を示します。External Coherency Hubの詳細については、後で説明します。

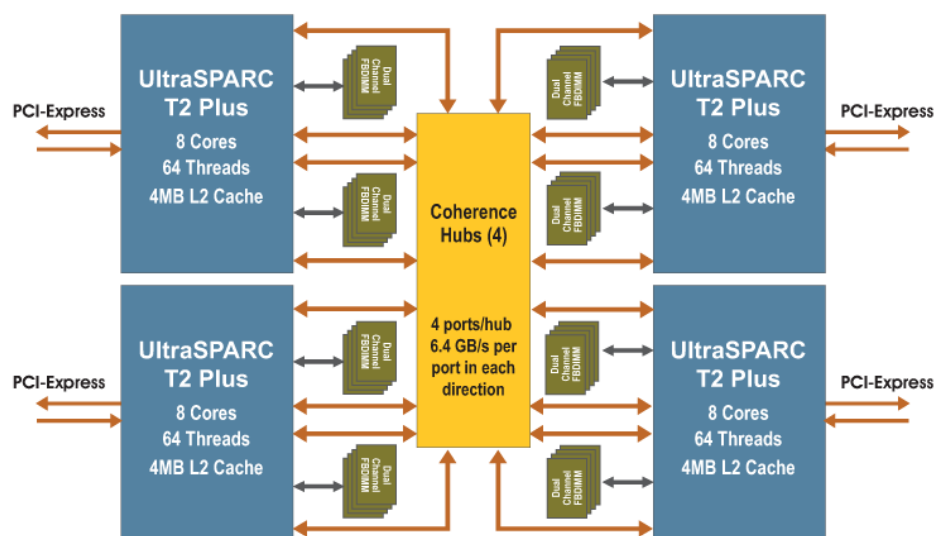


図 6. SPARC Enterprise T5440 サーバの UltraSPARC T2 Plus のクアド・ソケット構成

コア・アーキテクチャとパイプライン

UltraSPARC T2プロセッサとUltraSPARC T2 Plusプロセッサのコア・デザインは同じです。図 7 に、UltraSPARC T2プロセッサのシングルUltraSPARCコア（プロセッサあたり最大8コアをサポート）のブロック・ダイアグラムを示します。

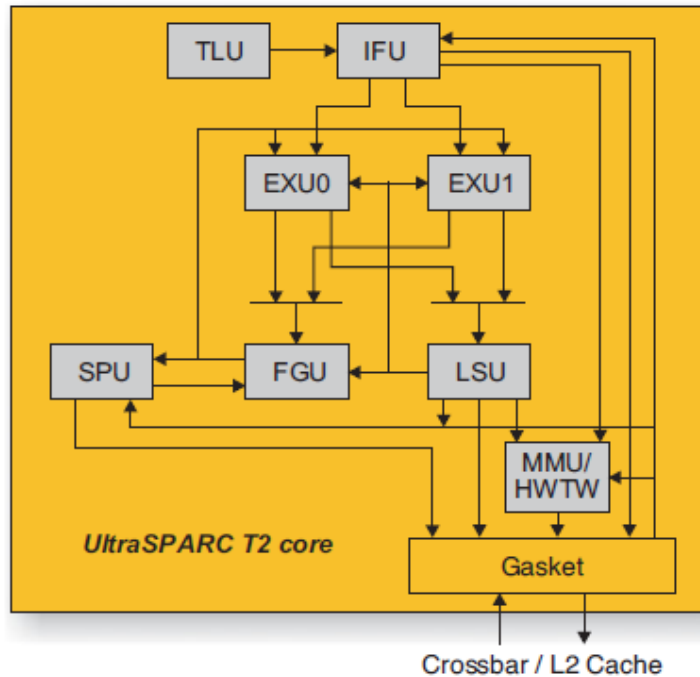


図 7. UltraSPARC T2 のコアのブロック・ダイアグラム

各コアに実装されているコンポーネントは、次のとおりです。

- **トラップ・ロジック・ユニット (TLU : Trap Logic Unit)** — TLUはマシン・ステータスを更新し、例外と割り込みを処理します。
- **命令フェッチ・ユニット (IFU : Instruction Fetch Unit)** — IFUは 16 KB の命令キャッシュ (32byte ライン、8 ウェイ・セット・アソシアティブ) と 64 エントリのフル・アソシアティブ命令変換ルックアップ・バッファ (ITLB : Instruction Translation Lookup Buffer) からなります。
- **整数実行ユニット (EXU : Integer Execution Unit)** — コアごとに2つのEXUが提供され、4 スレッドで 1 つのユニットを共有します。スレッドごとに 8 つのレジスタ・ウィンドウが用意されており、整数レジスタ・ファイル (IRF : Integer Register File) のエントリ数はスレッドごとに 160 です。
- **浮動小数点／グラフィック・ユニット (FGU : Floating Point/Graphics Unit)** — FGUは各コア内部にあり、コアに割り当てられた 8 スレッドすべてで共有されます。浮動小数点レジスタ・ファイルのエントリ数は、スレッドごとに 32 です。
- **ストリーム処理ユニット (SPU : Stream Processing Unit)** — コアごとに 1 つずつ、

暗号コプロセッサを搭載するSPUが組み込まれています。

- **メモリ管理ユニット(MMU : Memory Management Unit)** – MMUは、ハードウェア・テーブル・ウォーク (HWTW : Hardware Table Walk) を実行し、8 KB/64 KB/4MB/256MBのページングをサポートします。

各 UltraSPARC T2/UltraSPARC T2 Plusプロセッサのコアは、8 ステージの整数パイプラインと 12 ステージの浮動小数点パイプラインから構成されます (図 7)。各サイクルで実行する 2 つのスレッドを選択するため、「ピック (Pick)」という新しいパイプライン・ステージが追加されています。



図 8. UltraSPARC T2 のコアごとの整数パイプラインと浮動小数点パイプライン

デュアル・パイプラインの機能を説明するために、図 9 にロード・ストア・ユニット (LSU : Load Store Unit) を使用する整数パイプラインを示します。インストラクション・キャッシュは、コア内の 8 つすべてのスレッドによって共有されます。

least-recently-fetched (フェッチが最も古い) アルゴリズムを使用して、次にフェッチするスレッドが選択されます。各スレッドはスレッド固有のインストラクション・バッファ (IB : Instruction Buffer) に書き込まれ、8 スレッドのそれぞれがコア内の 2 つのスレッド・グループのうちのどちらか一方に静的に割り当てられます。

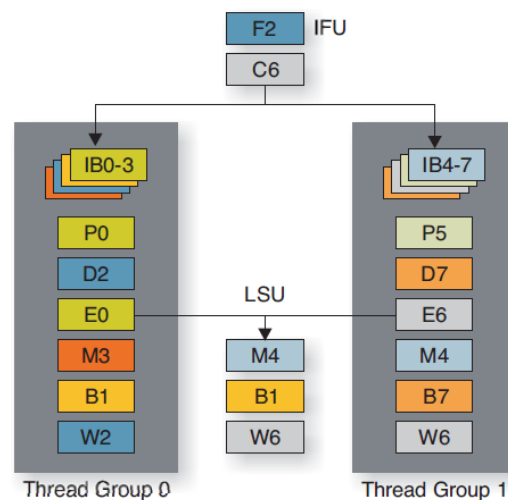


図 9. スレッドはほとんど制限を受けずに、パイプラインのステージ間でインターリーブされます (整数パイプラインの場合、文字はパイプライン・ステージを表し、数字はスケジュールされているスレッドを示します)。

「ピック」ステージでは、各スレッド・グループ内のサイクルごとにスレッドが 1 つ選択されます。各スレッド・グループ内でのピックアップ（選択）は、相互に独立しており、least-recently-picked（ピックが最も古い）アルゴリズムを使用して、次に実行するスレッドが選択されます。デコード・ステージは、ピック・ステージでは処理されないリソースの衝突を解決します。図に示すように、スレッドはほとんど制約を受けずに、パイプライン・ステージ間でインターリーブされます。2 つのスレッド・グループのどちらかに分割する前に、どのスレッドがフェッチ・ステージまたはキャッシュ・ステージにあってもかまいません。ロード／ストアと浮動小数点ユニットは、8 つすべてのスレッド間で共有されます。これらの共有ユニット上でスケジューリングされるのは、どちらかのスレッド・グループから 1 つのスレッドだけです。

ストリーム処理ユニット（SPU）

UltraSPARC T2 Plus コア上の各 SPU は、コアと同じ周波数で動作します。独立した 2 つのサブ・ユニットがコアのクロスバー・ポートを共有する DMA エンジンとともに提供されています。

- Modular Arithmetic Unit（MAU）は、FGU 乗算器を共有し、RSA 暗号化／復号、バイナリおよび整数多項式関数、および楕円曲線暗号（ECC : Elliptic Curve Cryptography）^{*2}を実行します。*2 楕円曲線暗号は将来Solaris™OSリリースでサポート予定
- 暗号／ハッシュ・ユニットは、RC4, DES, DES3, AES-128/192/256, MD5, SHA1, SHA-256 のサポートを提供します。

SPU は、UltraSPARC T2 プロセッサ上 10Gigabit Ethernet ポートで、暗号化／復号処理を実行するよう設計されています。

PCI Express インタフェースをチップに内蔵

UltraSPARC T2 Plus プロセッサは、ポイント・ツー・ポイントのデュアル・シンプレックス・チップ・インターコネクトによって、レーンごとに双方向に 4 GB/秒で動作する PCI Express インタフェースを搭載しています。統合 IOMMU は、PCI Express BDF 番号を使用し、I/O の仮想化とプロセス・デバイスの分離をサポートします。総 I/O 帯域幅は 3 ～ 4 GB/秒で、最大ペイロード・サイズは 128 ～ 512 byte です。オフチップ PCI Express スイッチとの統合用に 1 つの x8 SerDes インタフェースが用意されています。

電源管理

UltraSPARC T2 Plus プロセッサには、プロセッサ上のコアとメモリの双方を制御する電源管理機能が実装されています。この管理機能には、実行命令数の制御、アイドル状態にあるスレッド／コアのパーキング機能、コア／メモリのクロックオフが含まれます。以下は、想定する利用例です。

- 実行されない条件分岐などに対する制限

- データ・パス／制御ブロック／アレイにおける広範なクロック・ゲーティング
- デコード・ステージに余分なストール・サイクルを注入できるパワー・スロットリング

第3章 サーバ・アーキテクチャ

SPARC Enterprise T5440 サーバは、SPARC Enterprise T5140 および T5240 サーバ機能を基盤として、画期的なパフォーマンスとスケーラビリティを提供するように設計されています。SPARC Enterprise T5440 サーバは、UltraSPARC T2 Plus プロセッサの強みをさらに拡大し、革新的な筐体設計の特長を生かしています。それによって、優れた CPU、メモリ、I/O 拡張能力を備えた、高スケーラブルな UltraSPARC T2 Plus ベースのサーバを実現しています。

システムレベルのアーキテクチャ

SPARC Enterprise T5440サーバの設計努力は、最小限のレイテンシで、最大4つの UltraSPARC T2 Plusプロセッサのためのキャッシュ・コヒーレンシを提供することに集中しました。また、実績のある筐体設計が、予測可能な熱および電力で計算密度を提供します。SPARC Enterpriseサーバのマザーボードのブロック・ダイアグラムを図10に示します。

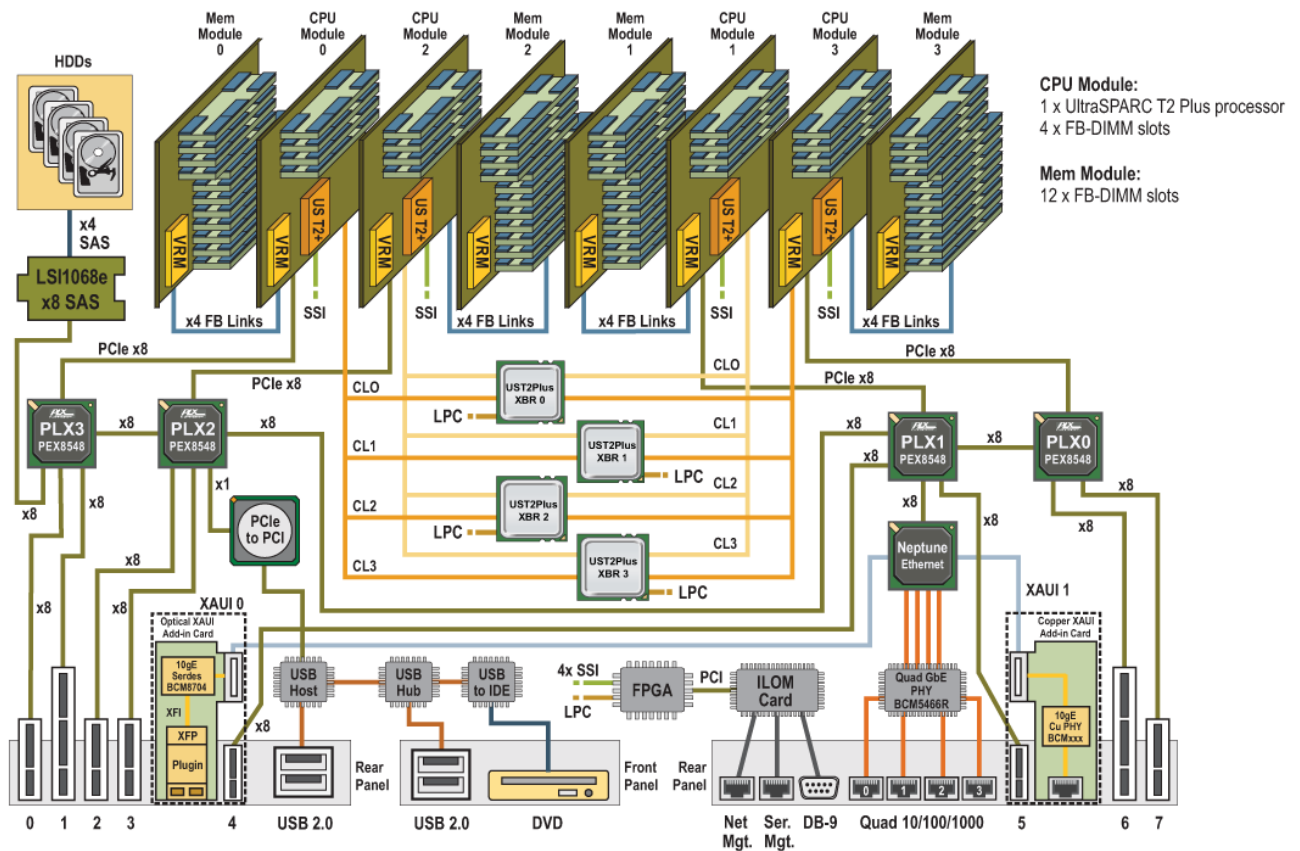


図 10. SPARC Enterprise T5440 サーバサーバのブロック・ダイアグラム

SPARC Enterprise T5440サーバのマザーボードは、プロセッサ、メモリ（それぞれCPUとメモリ・モジュール上）、ディスク・コントローラ、I/Oサブシステムを含むすべてのシステム機能をサポートするプリント基板です。ILOMサービス・プロセッサは、マザーボードに接続されるドーター・カード上で提供されます。I/Oオプションには、USB、DVDコントロール、Quad Gigabit Ethernet、および様々なサードパーティ製のPCI Express拡張にソケットを提供する4つの独立したPCI Expressバスがあります。

SPARC Enterprise T5440サーバのマザーボードには、主に次のような特長があります。

- CPUモジュール
各CPUモジュールはUltraSPARC T2 Plusプロセッサと4つのFB-DIMMソケットで構成。各CPUモジュールに搭載されるメモリは、UltraSPARC T2 Plusプロセッサをサポートするための最小メモリ構成となる。
- メモリ・モジュール
各メモリ・モジュールは最大12個のFB-DIMMソケットを使用することで、対応するCPUモジュールに増設メモリを提供。各メモリ・モジュールは、マザーボードを介して物理FB-DIMMインタフェースによってCPUに接続。
- 4つのExternal Coherency Hub ASIC
External Coherency Hubはそれぞれ、各UltraSPARC T2 Plusプロセッサ上のコヒーレンシ・リンクに接続。
- 4つのPCI Express拡張ハブ
各ハブは、CPUモジュール・ソケットによって各UltraSPARC T2 Plusプロセッサ上のPCI Expressインタフェースに接続。PCI Express拡張ハブは、筐体背面にある8つのPCI Expressスロットに加え、様々なシステム周辺機器をサポート。また、任意のCPUモジュールから周辺機器に接続するための相互接続も提供。
- Neptuneチップとの統合により、10Gigabit Ethernet機能に加え、標準で4つのGigabit Ethernetポート（10/100/1000-BASE-T）を提供

SPARC Enterprise T5440サーバのマザーボードは、前世代のシステムから大幅に簡素化されました。配電ボード（PDB：Power Distribution Board）に接続された1対のメタル・バス・バーを介して、12Vの電力がマザーボードに給電されます。単一のフレックス回路コネクタが、PDBへ重要なあらゆるパワー・コントロール、DVDドライブとSASディスクの信号を送ります。また、ディスク・ドライブのバックプレーンとマザーボードを接続し、内蔵ディスク・ドライブへのデータ・パスを提供します。

External Coherency Hub (UltraSPARC T2 Plus Crossbar)

SPARC Enterprise T5140およびT5240サーバなどのデュアル・ソケット・システムは、2つのUltraSPARC T2 Plusプロセッサのコヒーレンシ・リンクに直接接続するように設計されています。SPARC Enterprise T5440サーバのような大規模システムには、External Coherency Hubが必要になります。External Coherency Hubは、単一のUltraSPARC T2 Plusプロセッサのキャッシュ階層を最大4つのソケットおよびプロセッサ（ノード）にわ

たって拡張します。

External Coherency Hubは、カスタムASICに実装される4ポート・アービター／スイッチで、UltraSPARC T2 Plusプロセッサのコヒーレンシ・コントロール部分に接続します。External Coherency Hubのブロック・ダイアグラムを図11に示します。External Coherency Hubは、次のような機能を実行します。

- 同じアドレスに対するロード、ストア、ライトバックの要求をシリアルライズする
- スヌープをブロードキャストし、応答を集約する
- プロセッサに対する読み取り／書き込み要求のフロー・コントロールを提供
- ECCまたはパリティ保護を提供

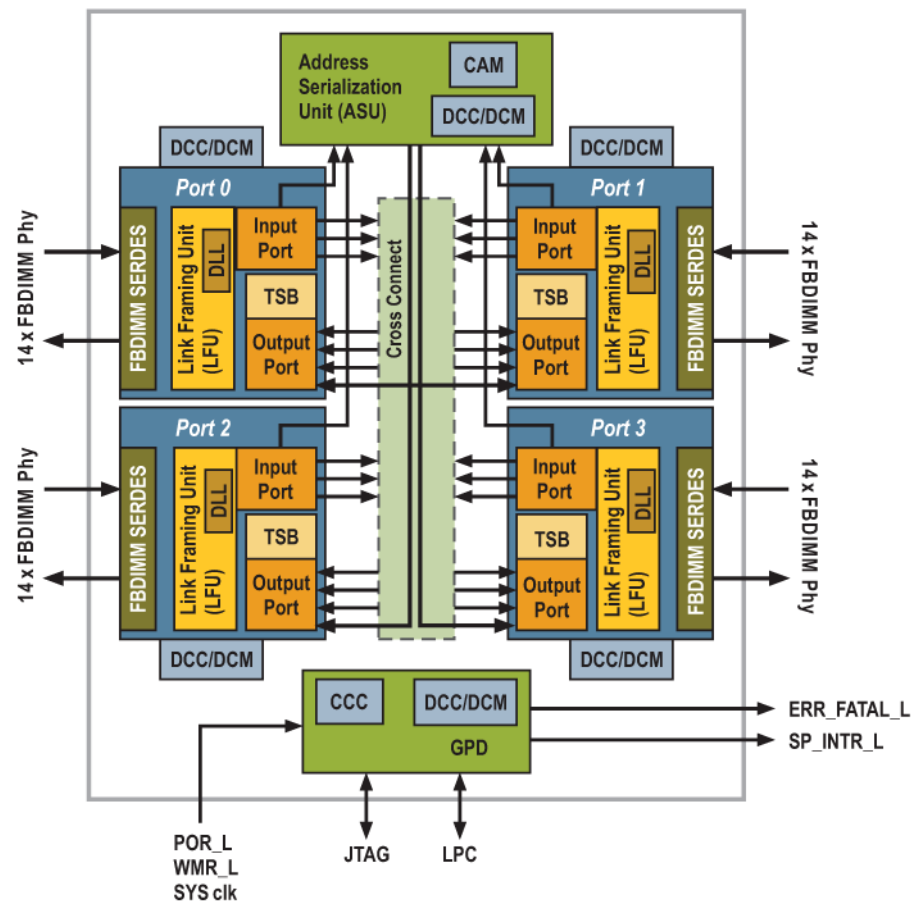


図 11. External Coherency Hubのブロック・レベル・ダイアグラム

メモリ要求がプロセッサのL2キャッシュのヒットに失敗した場合、その要求はExternal Coherency Hubに転送されます。ハブは、システム上の複数のノードに対する一連のメッセージを通してこれらの要求に応えます。ハブは、コヒーレンシ・プレーン内のアドレス・シリアルライズの単一点として機能し、TSO (Total Store Ordering) 一貫性モデルに従うのに必要となるオーダリング・ルールを強制します。External Coherency Hubは、ノードのキャッシュ不可能なスペース、および接続されているPCI Express I/Oファブリック上のI/Oデバイスにマップするロケーションを宛先ノードに転送します。

各SPARC Enterprise T5440サーバは4つのExternal Coherency Hubを提供し、それぞれが4つのCPUモジュール・ソケットすべてに接続されます。External Coherency Hubのコンポーネントは次のとおりです。

- **FB-DIMMインタフェース** : FB-DIMMインタフェース・ブロックにはSerDesコアとシンボル・アライメント・ロジックが含まれており、FB-DIMM Northboundインタフェースに使用されているものと事実上同一です。
- **Link Framing Unit** : 各External Coherency Hub上のLink Framing Unit (LFU) は、FB-DIMM Frame/Cellリンク層の最上位プロトコルを実装し、パケット化、CRCチェック、および受信側でエラーが検出されたパケットの再送信を行います。
- **入出力ポート** : 出力ポートは、メッセージの処理を行う一方で、すべての入力ポートからの送信メッセージを受信します。TSB (Transaction Score Board) は、一貫した読み取りとライトバックを保つために、要求、応答、データの状態をそれぞれ追跡します。
- **アドレス・シリアライゼーション・ユニット (ASU)** : アドレス・シリアライゼーション・ユニットは、同じメモリ・アドレスに対する要求のシリアライズと、特定のアドレスで保留になっている要求リストの維持を行います。ASUは、4つすべてのポートからの要求を扱います。
- **クロス・コネクト (Cross Connect)** : クロス・コネクトは、様々なポート・ツー・ポート・ネットワークを作成する相互接続ワイヤのセットから成ります。

メモリ・サブシステム

SPARC Enterprise T5440サーバでは、UltraSPARC T2 Plusプロセッサ上のオンチップ・メモリ・コントローラが高速シリアル・リンクを介して、FB-DIMMと直接通信します。各UltraSPARC T2 Plusプロセッサには、4つのコヒーレンシ・ユニットのほかに2つのデュアル・チャネルFB-DIMMメモリ・コントローラ・ユニット (MCU) が実装されています。UltraSPARC T2 Plusプロセッサ上の各MCUは、4.8Gb/秒の総レートでデータを転送します。

UltraSPARC T2 Plusプロセッサのほかに、各CPUモジュールは4つのFB-DIMMスロットを提供します。これら4つのスロットで基本のFB-DIMM一式を構成します。これは、単一プロセッサをサポートするための最小条件です。CPUモジュールには常にそれに対応するメモリ・モジュールが付随しており、マザーボード・コネクタを介してFB-DIMMインタフェースによって接続されます。デフォルトでは、メモリ・モジュールは、4つのFB-DIMMが搭載されており、各プロセッサのメモリ容量を2倍にするための十分なDIMMスペースがあります。希望した場合には、メモリを搭載しないメモリ・モジュールを取り付けることができ、正しい気流を確保するためのDIMMフィラーが提供されます。

各メモリ・モジュールは追加で12の FB-DIMMソケットを提供するため、プロセッサあたり合計16メモリ・ソケットが提供されます。メモリ・モジュールのそれぞれの

FB-DIMMが、物理的に1つのCPUモジュールと、1つのプロセッサに接続されることに注意してください。各プロセッサには最大で16個の667 MHz FB-DIMMが関連付けられ、最大システム・メモリ容量は、2GB FB-DIMMを使用した場合で128GB、4GB FB-DIMMを使用した場合で256GB、そして8GB FB-DIMMを使用した場合で512GBになります。

I/O サブシステム

各UltraSPARC T2 Plusプロセッサには、双方向4GB/秒で動作する8レーンのPCI Express (x8) ポートが1つ組み込まれています。SPARC Enterprise T5440サーバでは、このポートはPLXテクノロジーを採用した4つのPCI Expressエクスパンダ・チップによりI/Oデバイスに接続します。すべてのPCI Expressエクスパンダ・チップは、それ自身がPCI Express (x8) インタフェースに相互接続するため、任意のプロセッサが任意のI/Oデバイスにアクセスできることを保証します。PCI Expressエクスパンダ・チップは、PCI Expressカード・スロット、または以下に示すようなPCI Expressとインタフェースするブリッジ・デバイスのいずれかに接続します

- ディスク・コントローラ —4レーンのPCI Express (x4) ポートに接続されるLSI Logic SAS1068E SAS コントローラ・チップにより制御されます。
- モジュール方式のディスク・バックプレーン —4ディスクのバックプレーンがx4 SAS リンクによりLSI ディスク・コントローラに接続されます。
- 4ポートのGigabit Ethernet —SPARC Enterprise T5440 サーバでは、Neptune チップが2つの10/100/1000 BASE-T ポートと、2つの10/100/1000/10000 BASE-T インタフェースを提供します。これらのインタフェースは、背面パネルの4つのRJ45 コネクタとして実装されています。
- 2ポートの10Gigabit Ethernet —SPARC Enterprise T5440 サーバは、XAUI/PCI Express の共有スロットを通じ、2つの10Gigabit XAUI 接続を提供します。10Gigabit Ethernet インタフェースは、Neptune チップによって提供されます。10Gigabit Ethernet ポートが接続されると、2つのGigabit Ethernet ポートが使用できなくなります。SPARC Enterprise T5120/T5220/T5140/T5240 で使用可能なXAUI 拡張カードをSPARC Enterprise T5440 サーバでもサポートしています。
- USB および DVD —すべてのサーバ上でシングル・レーン PCI Express ポートからPCIブリッジ・デバイスに接続されます。セカンド・ブリッジ・チップが32bit 33MHz PCI バスを複数のUSB 2.0 ポートに提供します。システムのUSB インターコネクトはこれらのポートから制御されます。さらに、DVD はUSB ポートの1つをIDE フォーマットに変換する別のブリッジ・チップにより制御されます。

筐体設計の刷新

SPARC Enterprise T5120/T5220、T5140/T5240、および T5440 サーバはすべて、基本的な筐体設計の技術革新を共有します。そのため、製品群全体で一貫したルック・アンド・フィールを実現できるだけでなく、一貫したコンポーネント配置とコンポーネントの共有によって、管理を簡素化しています。さらに、一貫性のみならず、データセンター用に差別化できる箇所には主要なテクノロジーを配備するという、データセンターを考慮した設計を施しています。

- システムとコンポーネントの保守性の強化

今日のデータセンター環境において、各サーバとコンポーネントを検索し識別することは、課題の 1 つになっています。SPARC Enterprise T5440 サーバは、特定しやすいサーバとモジュールにより、データセンターの自動構成に対応できるよう最適化されています。カラーコード化されたオペレータ・パネルでは分かりやすい診断が行え、システムはホットアイル／コールドアイルのマルチラック配備に対応した設計となっており、前面と背面の診断 LED によって障害のあるコンポーネントを特定することができます。障害コンポーネントの識別には、Fault Remind 機能が使用されます。

電源、ネットワーク接続、および管理のための一貫したコネクタ・レイアウトにより、SPARC Enterprise システム間における移動が簡単に行えます。すべてのホットプラグ・コンポーネントは工具不要で、容易な保守が可能です。たとえば、ヒンジで動く蓋を開けるとデュアル・ファン・モジュールにアクセスできるため、損傷しやすいコンポーネントをさらしたり、不要なダウンタイムを生じさせたりすることなく、保守することが可能です。

- 筐体、コンポーネント、およびサブアセンブリの堅牢な設計

SPARC Enterprise T5440 サーバは、信頼性と安定した冷却運転を維持できるよう綿密に設計された筐体を採用しています。筐体の六角形の通気口などの特長も、高い強度、最大限の気流、そして電気ノイズをバランスよく考慮した結果に基づいて設計されています。

ファン・アセンブリにより、システムの冷却を効果的にする気流が確保されます。ファン・モジュールは筐体から分離され、他のシステム・コンポーネントに回転振動が伝わらないようになっています。さらに、ファンの振動を監視して制御することで、ディスク・ドライブへの共鳴の影響を防ぎます。ファンの速度の監視と制御は、十分な冷却効果を維持するためだけでなく、騒音を縮小し、ファンがシステムの電力を使用するのを最小限にするために採用されています。

計算密度／I/O 密度／ストレージ密度にもかかわらず、SPARC Enterprise T5440 サーバは従来のテクノロジーを用いて適切な温度を維持できます。DC-DC の電力変換を最小限に抑えることによっても、システム全体の効率性に貢献しています。12 V の電力をマザーボードに供給することで、電力変換を不要にしています。こ

のようなアプローチが発熱を抑え、システムの効率をさらに向上しています。

- 気流を最大にするためにケーブルを最小限に
配線を最小限にし、かつ信頼性を向上させるために、各筐体に適した様々な小型ボードが使用されています。これらのボードが SPARC Enterprise T5440 サーバの拡張機能を提供しています。
 - － 配電ボードにより、2つの電源装置からマザーボードに、さらに（コネクタ・ボードを経由して）ディスク・バックプレーンに電源が供給されます。
 - － コネクタ・ボードにより、多数の配線ケーブルが不要となりました。コネクタ・ボードは、制御信号と多くのデータ信号をディスク・バックプレーン／ファン・ボード／PDB に分配するダイレクト・カード・プラグイン・インターコネクトとして機能します。
 - － PCI Express カードは、マザーボードに直接差し込みます。
 - － 2つの XAUI スロットにより、Neptune Ethernet チップ上の 10Gigabit Ethernet インタフェースにアクセスできます。あるいは、これらのスロットを使用して、PCI Express インタフェースにアクセスすることもできます。各スロットには、ファイバーまたは銅線接続用 XAUI カード、または最大で x8 フォーム・ファクタ・エッジ・コネクタを備えた業界標準のロープロファイル PCI Express カードを搭載できます。
 - － ディスク・バックプレーンは、2つの筐体のディスク・ケージに搭載されていて、マザーボードから4チャンネルが独立した mini-SAS ケーブルにより、データを送受信します。SPARC Enterprise T5440 サーバには、4ディスクのバックプレーンが用意されています。
 - － ディスク・バックプレーンによって、CD-RW/DVD-RW ドライブと2つの USB ポートがシステムの前面に配線されます。

SPARC Enterprise T5440 サーバの概要

SPARC Enterprise T5440サーバは、省スペースの4ラックマウント・サーバで、優れた計算能力とスケーラビリティを提供します。マルチスレッド・テクノロジー・アーキテクチャの実証済みのメリットを、Webサービスをサポートする最も重要なOLTPデータベース・ワークロードにまで拡張することで、ミッドレンジ・コンピューティングを変えます。このサーバは、消費電力が大幅に削減され、設置面積もわずかであり、現在、データセンターが抱えている課題に対処できるように設計されています。選択するモデルにもよりますが、SPARC Enterprise T5440サーバには、2～4個の8コアの UltraSPARC T2 Plusプロセッサと最大512GBメモリが搭載されます。

システムマザーボードと筐体

図12は、SPARC Enterprise T5440筐体の上部カバーを外して上から見た図です。取り付けられているすべてのCPUモジュールとメモリ・モジュールが示されています。前述のように、各CPUモジュールは、それぞれ対応するメモリ・モジュールとペアになっています。CPUモジュールは、それと対応するメモリ・モジュールの取付が必須という

訳ではありません。またシステムが動作するために、すべてのCPUモジュールが取り付けられる必要もありません。いずれかのCPUまたはメモリ・モジュールが搭載されていない場合、適切な気流と前面から背面の冷却を保つためにフィルター・モジュールが取り付けられます。

8個すべてのPCI Expressスロットはロープロファイルで、x8 PCI Expressインタフェースに配線されます。ロープロファイル・グラフィック・アクセラレータなどのx16 PCI Expressカードをサポートするために2つのスロットがx16物理コネクタを提供します。2つのPCI Expressスロットは、XAUIカードとスロットを共有しています。XAUIカードは、PCI Expressコネクタの背後にあるコネクタに接続します。そのため、XAUIカードが取り付けられている場合、PCI Expressカードは対応するPCI Expressスロットで使用できません。また、XAUIカードが取り付けられているときは、Neptuneチップに接続するGigabit Ethernetポートの1つが使用できなくなります。

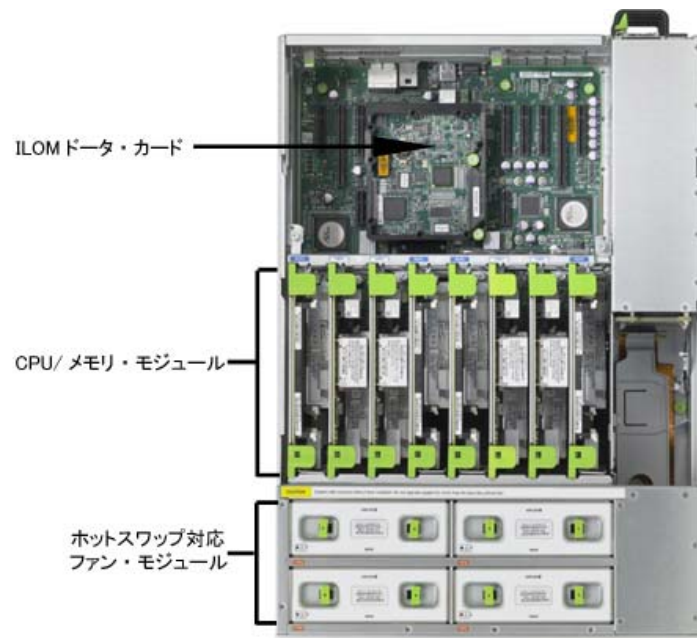


図 12.SPARC Enterprise T5440 サーバを上から見た図

4つのシステム・ファンは筐体の上から取り付けます。また、4つの電源装置では筐体の背面から取り付けます。4つの電源装置で2+2冗長を提供します。システムは4つの電源装置のうち、任意の2つの装置で、フルパワーで動作し続けることができます。したがって、個々の電源装置の障害だけでなく、データセンター全体の回線もカバーできます。

エンクロージャ

4 U のSPARC Enterprise T5440サーバ は、標準 19 inchラックに搭載可能なよう設計されています。(表 3).

表3. SPARC Enterprise T5440 サーバ の寸法と重量

寸法	国際単位
幅	445 mm
奥行き	633 mm
高さ	176 mm
重量 (最大、PCI スロットとラックマウントを除く)	40 kg

SPARC Enterprise T5440サーバ に組み込まれている主要コンポーネントは、次のとおりです。

- 1.2GHzまたは1.4GHz UltraSPARC T2 プロセッサ (8 コア) x2~4
- FB-DIMM (Fully Buffered Dual Inline Memory Module) スロット×16
(2GB/4GB/8GB FB-DIMMをサポート。最大512GB))
- オンボード 10/100/1000 BASE-T Ethernet ポート x4
- PCI Express(x8)スロット x2 (ロープロファイル/MD2, 物理形状はx16)
- PCI Express(x8)スロット x4 (ロープロファイル/MD2, 物理形状はx8)
- PCI Express(x8)スロット x2 (ロープロファイル/MD2, オンチップ10Gigabit Ethernet使用時に用いるXAUIアダプタと兼用)
- USB 2.0 ポート x4 (前面x2、背面x2)
- 2.5" SASディスク・ドライブ x4 (標準2台、最大4台)
- ILOM (Integrated Lights out Management) システム・コントローラ
- 効率性に優れたホットスワップ対応1.120W電源ユニット×4 (2つの独立する電源回路で提供され場合は2+2の冗長構成を提供)
- 環境モニタ/制御下で4台の冷却ファンを搭載 (N+1の冗長構成)。
筐体の上部から直接装脱着可能

前面図と背面図

図 13 に SPARC Enterprise T5440 サーバ の前面パネルと背面パネルを示します。

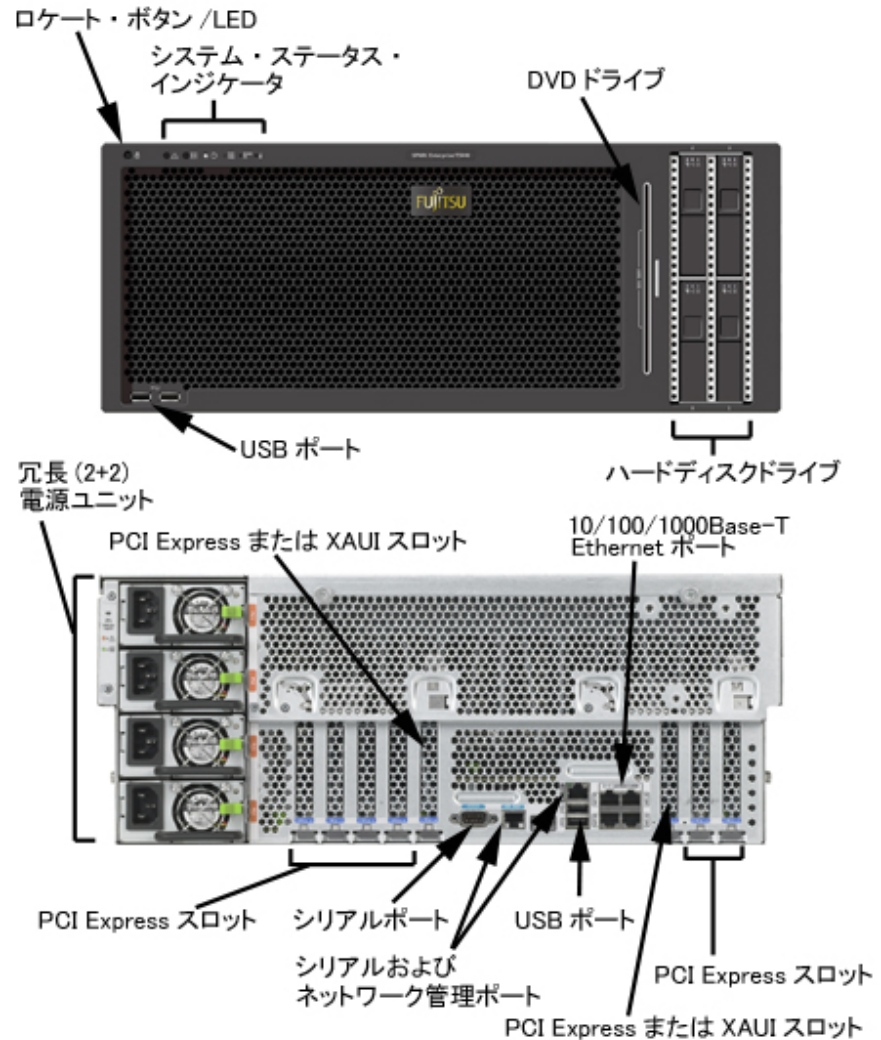


図 13. SPARC Enterprise T5440 サーバ の前面パネルと背面パネル

SPARC Enterprise T5440サーバ の外装部は以下のとおりです。

- システムとコンポーネントのステータス・インジケータ(前面/背面)。
システム状態をロケータ（ホワイト）／保守の必要性（オレンジ）／正常状態（グリーン）として点灯表示
- ホットプラグ対応SAS ディスク・ドライブ x4（前面パネルから装脱着可）
- CD-RW/DVD-RWドライブ x1（前面）
- USB ポート x4（前面パネル x2、背面パネル x2 ）
- ファン内蔵のホットプラグ／ホットスワップ対応の電源装置 x4台（2+2冗長構成）（背面）
- 各電源装置の動作状態を点灯表示する電源装置インジケータ（背面）

- 各ホットプラグ/ホットスワップ対応の電源装置に1つのACプラグ
- 自動検知 10/100/1000 BASE-T Ethernet ポート x4
- シリアル・デバイス用 DB-9 TTYA シリアル・ポート x1 (ILOM システム・コントローラには接続不可)
- PCI Express カード・スロット x8 (2スロットのいずれかがNeptune 10Gb Ethernet チップ接続用XAUIアダプタと兼用)
- ILOM システム・コントローラ用管理ポート x2。
ILOM コントローラへのデフォルト接続用シリアルポート (RJ45)
ネットワーク管理ポート10/100BASE-T (RJ45)

第4章 エンタープライズ・クラスの実運用とソフトウェア

新しいテクノロジーに対応するツールとアプリケーションが市場で揃うには、多くの場合、時間を要するものです。また、新しいテクノロジーを実運用環境中に取り込み、その上で可用性の高いサービスを実現するためには、安定した開発ツール／オペレーティングシステム／ミドルウェア／管理ソフトウェアが必要とされるものです。

UltraSPARC T2 Plus プロセッサは、革新的なテクノロジーでありながら従来のSPARCシステムとバイナリ・レベルで互換性を維持しているため、納入直後からプリインストールされたSolaris OSとプリロードされたツールを利用して、即時機能させることが可能です。さらに、豊富な開発ツールが完備されているため、UltraSPARC T2 Plus プロセッサがもたらす能力を、即実運用環境へ導入し、アプリケーション開発とそのチューニングに着手することができます。

システム管理テクノロジー

どのような企業でもシステム数が増加すれば、ライフサイクル全般にわたってITシステム基盤の管理が複雑になり、困難を増していきます。より効率的なシステム管理には、主要なシステム・コンポーネントの動作を監視するシステム・コントローラとともに、主なシステム管理作業を自動化出来る先進的なフレームワークが必要とされます。

ILOM (Integrated Lights-Out Management) システム・コントローラ

SPARC Enterprise T5440 サーバには、SPARC Enterprise T5120/SPARC Enterprise T5220/SPARC Enterprise T5140/SPARC Enterprise T5240 に標準搭載されてきたシステム・コントローラ、ILOM が搭載されています。したがってこれらのサーバは、既存の管理基盤に容易に統合することができます。SPARC Enterprise T5440 サーバには、システム・マザーボードに接続するドータカード上に、ILOM サービス・プロセッサが搭載されています。

ILOM は、システム管理に求められる次の機能を提供します。

- IPMI 2.0 に準拠したサービス・プロセッサ。IPMI 管理機能をサーバのファームウェア、OS、アプリケーションに提供。さらにイーサネット管理インタフェース経由でのアクセスも可能。環境センサーに可視性を提供 (サーバ・モジュールと、シャーシ内の任意の場所)
- CPU、DIMM、電源装置など、インベントリと環境を管理。HTTPS/CLI/SNMP によるアクセスが可能
- テキスト指向のリモート・コンソール・インタフェースを提供

- システム・ファームウェアのアップグレード

ILOM は、オペレーティングシステムの種類を問いません。ILOM は、システムの動作に影響を与えることなく、サーバのリモート管理を可能とします。ILOM は、ハードウェア障害や警告、その他のイベントを電子メールにて管理者に通知することも出来ます。仮に、オペレーティングシステムがオフラインの場合でも、サーバの電源がオフになっても、ILOM は機能し続けます。ILOM は、これらのサーバの次の条件を監視します。

- CPU の温度
- ハードディスク・ドライブの有無
- エンクロージャ内の温度
- ファンの速度と動作状態
- 電源装置の状態
- 電圧
- Solaris のウォッチドッグ、ブート・タイムアウト、自動サーバ再起動のイベント

サーバ管理ソフトウェア

IT（情報技術）の普及により複雑化したサーバを管理するためには、サーバを統合的に監視・制御できるソフトウェアが必要です。富士通では、Solaris, Linux, Windowsサーバを統合管理できるサーバ管理ソフトウェア Server System Manager を提供しています。また統合運用管理ソフトウェア Systemwalker との連携により業務運用を自律化、ビジネス環境の変化に即対応し、ビジネスの継続を支えます。

Server System Manager (SSM)

Server System Manager は、富士通UNIXサーバ SPARC Enterprise、基幹IAサーバ PRIMEQUEST、PCサーバ PRIMERGYで構築されたヘテロ環境のシステムを統合管理できる、サーバ管理ソフトウェアです。複数のプラットフォームの混在環境において、各サーバのハードウェアの構成表示や状態監視などを行うサーバ管理機能をGUIで提供します。

Server System Managerを利用することにより、1台の管理クライアントから複数のサーバの状態を監視することができます。各サーバのハードウェア構成は、プラットフォーム共通のイメージ（ツリー形式）で表示され、サーバの構成部品の状態を一目で把握でき、統一された操作でサーバの管理を行えます。また、ラックに搭載している装置のレイアウト表示や、CPU、メモリの資源使用状況のリアルタイム表示も可能です。

Server System ManagerはILOMと連携しています。サーバの構成部品に異常が発生すると、ILOMを通じてServer System Managerの管理画面に故障箇所が表示されるため、故障部品を容易に特定できます。さらにこれらの異常発生をメールやSNMP Trapにより通報することができます。また、特定サーバのイベントや特定レベルのイベントのみ通報を行うフィルタリング設定も可能です。

さらに、Solarisコンテナの管理を、管理クライアントから行うことができます。GUIウィザード（対話形式）を用いて、仮想Solaris環境（zone）の作成、設定変更、削除や起動、停止が行えます。

Server System Managerを利用することで、サーバの運用管理作業が容易になり、ハードウェア故障時の対応を迅速に行えます。そのため、運用管理者の作業負担が軽減します。

Enhanced Support Facility (ESF)

Enhanced Support Facility は、ILOMと連携してSPARC Enterprise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表示することができます。他にも、システム情報の一括採取や、/etc/systemsファイルの設定チェックなどが可能です。

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。

マルチスレッド・テクノロジーのスケラビリティとサポート

Solaris 10 OSは、UltraSPARC T2 Plusプロセッサ搭載システムの能力を余すところなく引き出せるよう設計されています。Solaris 10 OS は、垂直／水平方向への拡張に優れているばかりでなく、仮想化／利用率／可用性／セキュリティ／パフォーマンスを高めるため、大きな役割を果たしています。Solaris 10 OS は、SPARC／x64／x86 ベースのシステム上で動作し、既存のアプリケーションとの互換性をも保証しています。

UltraSPARC T2 Plusプロセッサを搭載するシステムの最大の魅力は、Solaris OSとSolaris上で動作するアプリケーションから見た時、それがかつてのSMPシステムの如く扱える点にあります。Solaris 10 OSには、マルチスレッドアーキテクチャ上でアプリケーション・パフォーマンスを向上させるための多数の機能が組み込まれています。

• マルチスレッド・テクノロジー への対応

Solaris 10 OSは、UltraSPARC T2 Plusプロセッサを64の論理プロセッサの固まりとして認識し、使用可能なすべてのパイプライン間で負荷が効果的に分散するよう、スレッ

下間の相関関係に応じ、効果的にスレッドを割り当てます。。

- **きめ細かい管理が可能**

Solaris 10 OSには、個々のプロセッサを有効にしたり無効にしたりする機能が備わっています。UltraSPARC T2 Plusプロセッサの場合、この機能が拡張され、個々のコアおよびスレッド(論理プロセッサ)を有効または無効にすることができます。さらに、プロセッサ・セットなど、Solaris OSの標準機能が論理プロセッサ・グループを定義し、それらのグループに対してプロセスまたはスレッドをスケジューリングする機能を持ち合わせています。

- **インタフェースのバインド**

Solaris OS では、必要に応じて、個々のプロセスとスレッドをプロセッサまたはプロセッサ・セットに自在にバインドできます。

- **仮想化されたネットワーク機能と I/O および高速暗号化処理サポート**

Solaris OSには、UltraSPARC T2 Plusプロセッサ上の拡張機能をサポートし、仮想化するテクノロジーが組み込まれています。これには、オンチップPCI Expressインタフェース/暗号化処理機構が含まれます。ハイパフォーマンス・ネットワーク・アーキテクチャの一部として、マルチスレッディングを認識するデバイス・ドライバが用意されているので、仮想化フレームワーク内で動作するアプリケーションは、I/Oとネットワーク・デバイスを実質的に共有します。また、Solaris Cryptographic Frameworkによって高速な暗号化がサポートされます。

- **Solaris OSのNUMAの最適化**

SPARC Enterprise T5440サーバでは、UltraSPARC T2 Plusプロセッサがメモリを管理し、非均一メモリ・アクセス (NUMA: Non-Uniform Memory Access) アーキテクチャとして実装されています。NUMAアーキテクチャでは、プロセッサが自分のメモリにアクセスするために必要となる速度は、別のプロセッサによって管理されるメモリにアクセスするために必要となる速度と若干異なります。Solaris OSでは、アプリケーションがNUMAアーキテクチャにおけるパフォーマンスを向上できるようなテクノロジーを提供します。

- **メモリ配置の最適化 (MPO: Memory Placement Optimization)**

Solaris 10 OSでは、MPOによって サーバの物理メモリ全体におけるメモリの配置を改善し、パフォーマンスを高めています。MPOを使用することで、Solaris 10 OSは、システム内の均衡を十分に保ちつつ、メモリにアクセスするプロセッサの近くにメモリを配置します。その結果、データベースとHPCアプリケーションの多くが、より迅速に動作できるようになります。

- Hierarchical Lgroup Support (HLS)

HLS は、Solaris OSのMPOの機能を向上させます。HLSを活用すると、Solaris OSは、複雑なメモリ・レイテンシ階層を持つシステムのパフォーマンスを最適化することができます。また、Solaris OSはHLSによりメモリがどの程度離れているかを識別できるため、レイテンシが最小のリソースをアプリケーションに割り当てることができます。あるアプリケーションに対し、デフォルトでローカル・リソースを使用できない場合は、Solaris OSはHLSにより、もっとも近くにあるリモート・リソースを割り当てます。

- **Solaris™ ZFS**

Solaris ZFS (zettabyte file system) により、データ管理／自動化／ストレージ統合が劇的に進歩します。世界初の 128 bit ファイルシステムによる究極のスケーラビリティです。Solaris ZFS は、トランザクション型のオブジェクト・モデルに基づいているため、I/Oの発行順序における従来の制約の大半が解消され、パフォーマンスが飛躍的に向上します。さらに、Solaris ZFS は、エラーとして報告されないデータ破損を検出し、修正する 64 bit チェックサム機能であらゆるデータを保護し、データの整合性を維持します。

- **セキュリティ保護された堅牢なエンタープライズクラス環境**

Solaris が何より優れているのは、バイナリ・コンパティビリティを維持している点です。バイナリ・レベルで互換性が維持されていることから、既存の SPARCプラットフォーム上で動作しているアプリケーションをそのままUltraSPARC T2 Plusプラットフォーム上へと移植出来るのです。

また、Solaris は、この他にも様々な先進的な機能を備えています。マルチレベルのセキュリティ保護機能、障害の自動検出／分離／修復機能(Fault Management ArchitectureおよびSolaris™ Predictive Self Healing) 、アプリケーションのチューニング (Solaris™ Dtrace : Dynamic Tracing) などです。

エンド・ツー・エンドの仮想化テクノロジー

異なるアプリケーションをより強力なシステム上で束ね、使用率を向上させるとともに管理台数を減らし、TCOを低減する、この有力な手段として、今、仮想化テクノロジーに注目が集っています。SPARC Enterprise T5440サーバ は、仮想化テクノロジーが充実したシステムです。プロセッサからメモリ、ネットワーク、I/Oまで、あらゆるレベルできめ細かくリソースを分割することが可能です。しかも、これらの仮想化テクノロジーは、システムの1標準機能として提供されるものであり、利用に際して、追加費用は一切必要ありません。

マルチスレッド対応のハイパーバイザー

UltraSPARC T1/UltraSPARC T2 プロセッサと同様、UltraSPARC T2 Plusプロセッサもマルチスレッド対応のハイパーバイザーを実装しています。このハイパーバイザーは、プロセッサに密に統合され、バーチャル・マシンとして機能する軽量のファームウェアです。ハイパーバイザーは、チップ・マルチスレッドに対応し、UltraSPARC T2 Plusプロセッサと連動し機能します。競合他社とは異なり追加費用が一切必要なく、且つ、オーバーヘッドも小さいハイパーバイザーです。

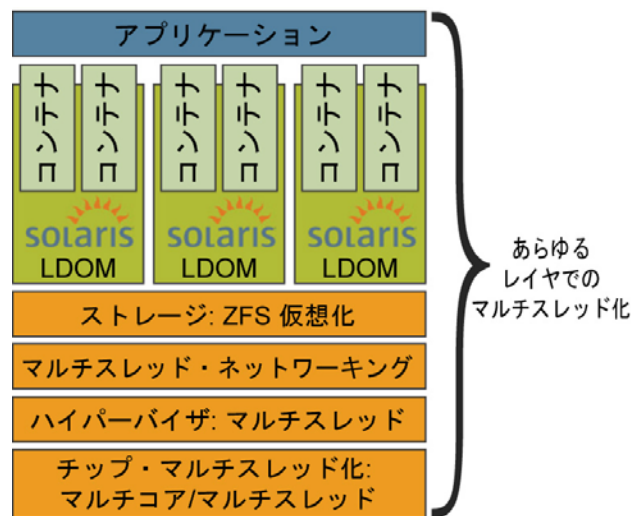


図 14. テクノロジー・スタックのあらゆるレベルで、並列化と仮想化を行います

図 14のようにハイパーバイザーの上部に仮想化テクノロジーに対応する各レイヤが構築されています。SPARC Enterpriseの強みは、プロセッサからアプリケーションまで、アーキテクチャのあらゆるレイヤにおいて完全にマルチスレッド対応していることです。

Solaris OSは1992 年からマルチスレッド環境に対応しています。このSolaris OSのマルチスレッド化を通して得た知識と経験から、今日の全方位的なエンド・ツー・エンドの並列化と仮想化環境が形作られていると言っても過言ではありません。プロセッサ、ハイパーバイザーのみならずネットワーク、ファイル・システムもマルチスレッド環境に最適化されています。このエンド・ツー・エンドの環境があつてこそ、初めてシステム全体でアプリケーション性能をリニアに高めて行くことができるのです。。

ロジカル・ドメイン (LDoms)

マルチスレッドテクノロジーを利用する SPARC Enterprise T5120/T5220/T5140/T5240/T5440 サーバでサポートされる ロジカル・ドメイン(LDoms) は、CPU/メモリ/ネットワークといったコンピューティング・リソースをハイパーバイザーで抽象化し、分離した動作領域上で異なる Solaris OS を動作させる仮想化テクノロジーの一つです。ロジカル・ドメインのアーキテクチャ内では、Solaris 10 OS が

ハイパーバイザに書き込まれ、ハイパーバイザが各ロジカル・ドメイン内のオペレーティングシステムに対し、使用するサーバ・ハードウェアに関して、安定した、論理的で仮想化可能な方法で表現します。各ロジカル・ドメインは完全に切り離され、1つのプラットフォーム上で作成される仮想マシンの最大数は、システムに組み込まれている物理ハードウェア・デバイスの数ではなく、ハイパーバイザの能力によって決まります。たとえば、SPARC Enterprise T5440 サーバ は、最大 128 のロジカル・ドメインをサポートし、個々のロジカル・ドメインがそれぞれ固有の OS インスタンスを 1 つずつ実行できます。

ロジカル・ドメインを利用することによって、企業は単一プラットフォーム上で複数のオペレーティングシステムを同時に配備する柔軟性が得られます。さらに、管理者は仮想化デバイスの機能を活用して、ある物理マシンから別の物理マシンへ、ロジカル・ドメイン上にホスティングされたソフトウェア・スタック全体を転送できます。ロジカル・ドメインは Solaris コンテナをホスティングし、両方のテクノロジーの分離／柔軟性／管理可能性に関連する機能を獲得できます。ロジカル・ドメインは UltraSPARC T2 Plus プロセッサおよび Solaris 10 OS と緊密に統合されることで、柔軟性が向上し、作業の処理が分離され、サーバを最大限に利用できる可能性が高まります。

LDoms は、ハードウェア／ハイパーバイザー／仮想デバイス／ゲスト・ドメイン／コントロール・ドメイン／サービス・ドメインから構成されます。ハイパーバイザは、各ホスティングされたオペレーティングシステムと、ハードウェア間のインタフェースとして機能します。ハイパーバイザーが制御し、サポートするオペレーティングシステムのインスタンスを「ゲスト・ドメイン」と呼びます。ゲスト・ドメインの作成や、制御を目的としたハイパーバイザー／ハードウェア／その他のドメイン間の通信は、「コントロール・ドメイン」により制御されます。ゲスト・ドメインには、システムとハイパーバイザーをともに制御し、さらに I/O を割り当てる「サービス・ドメイン」から、仮想デバイス・アクセス権が与えられます。

仮想化ネットワークをサポートするため、LDoms は、仮想レイヤ 2 スイッチを実装します。ゲスト・ドメインは、複数の仮想スイッチ（vswitch）に接続可能であり、複数のゲスト・ドメインを一つの仮想スイッチに接続することも可能です。vswitch は、実際の物理ネットワーク・ポートに対応付けることができます。また、対応するポートがない状態で設定することもできます。この場合、vswitch が行うのは、同一サーバ内のドメイン間通信です。この手法により、ゲスト・ドメインはネットワークとの直接通信チャネルも得られます（図 17）。各ゲスト・ドメインは、NIC 全体と NIC が提供する帯域幅を所有していることになりますが、実際に利用可能なのは、総帯域幅の一部分だけです。したがって、各 NIC を必要に応じて設定し、ドメイン毎に帯域幅を設定することができます。vswitch デバイスを専用の物理 Ethernet ポートに接続することによって専用の帯域幅を利用できます。

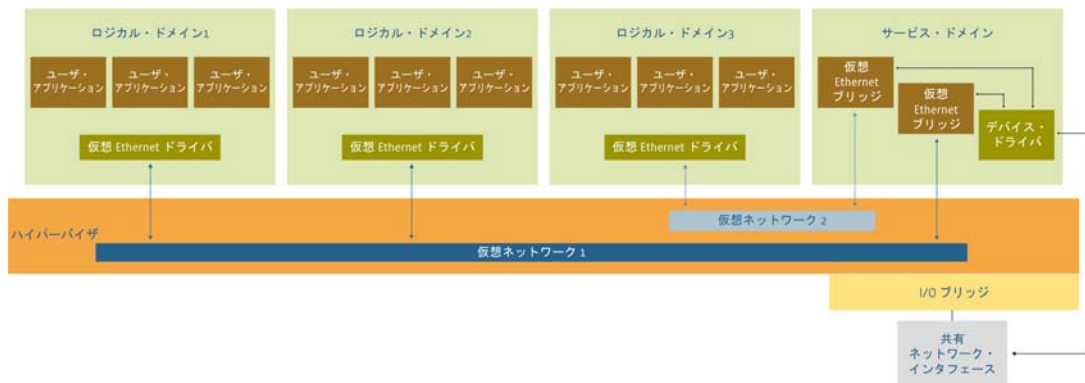


図 13. データはロジカル・ドメインと仮想デバイス間を直接移動します。

Solaris™ コンテナ

OS レベルで仮想化を行う Solaris コンテナは、システム・リソースを効率的に管理し、環境を仮想化し、しかもセキュリティ保護されたアプリケーション実行環境を提供する複数のテクノロジーから構成されるものです。Solaris コンテナには、フェア・シェア・スケジューラと連動する重要なテクノロジーが組み込まれています。

- **Solaris™ Zones**

Solaris 10 OS は、Solaris Zonesという独自のパーティショニング・テクノロジーを提供します。Solaris Zonesにより、アプリケーション実行用に分離したセキュリティで保護された環境を作成することが出来ます。ゾーンは、1 つの Solaris OS インスタンス内で生成される仮想化されたオペレーティングシステム環境です。ゾーンにより、アプリケーションとプロセスをシステムの他の部分から分離することが出来ます。この分離は、セキュリティと信頼性の強化に役立ちます。ある特定ゾーンのプロセスが他のゾーンで実行されているプロセスに影響を与えることはありません。

- **リソース管理**

Solaris OS に用意されているリソース・マネージャを使用すると、特定のアプリケーションにCPU などのリソースを割り当てることが出来ます。マルチプロセッサ・システムにおける CPU (UltraSPARC T2/UltraSPARC T2 Plusプロセッサの場合はスレッド) は、論理上、プロセッサ・セットに分割し、リソース・プールにバインドし、そこから Solaris Zonesに割り当てることが出来ます。リソース・プールによって、CPU リソースの消費が重ならないよう、負荷を分散することが可能です。また、プロセッサ・セットとクラス割り当てのスケジューリングに対応する持続的な構成メカニズムも備わっています。さらに、リソース・プールの動的機能により、需要の変化に応じて、システム・リソースを調整することも可能です。

障害管理と予測的セルフヒーリング

Solaris 10 OS には、障害管理と予測的自己修復を可能とする新しい管理アーキテクチャが盛り込まれています。予測的セルフヒーリングは、さまざまなハードウェアおよびアプリケーション障害を自動的に診断し、分離し、回復する Solaris 10 OS の画期的な機能の一つです。これにより、ソフトウェア障害や主要なハードウェア・コンポーネントの障害、ソフトウェアの誤った設定による問題発生時でも、アプリケーションを停止させずにサービスを継続させることが可能となります。

- **Solaris™ Fault Manager**

Solaris Fault Managerは、ハードウェアおよびソフトウェア・エラーに関連するデータを収集します。Solaris Fault Managerは、自動的かつエラー報告をせずに問題を検出して診断し、さらに、拡張可能な 1 組のエージェントが障害コンポーネントをオフラインにすることによって自動応答を行います。分かりやすい診断メッセージは ナレッジ・ベースの項目にリンクしており、修復作業に人的介入が必要な場合でも、管理者が迷うことはありません。Solaris Fault Managerはオープン設計なので、管理者や現場のスタッフが診断システムの動作を監視することもできます。Solaris Fault Managerによって、障害条件が発生してから自動診断を経て必要な人的介入が完了するまでの、全体の時間が大幅に短縮されるため、アプリケーションの稼動時間が拡大します。

- **Solaris™ Service Manager**

Solaris Service Managerは、アプリケーション・サービスを、管理者が一元的な方法で観察して管理できる最優先オブジェクトに変換することによって、アプリケーション・サービスのための標準制御メカニズムを提供します。これらのサービスは、管理者が誤って終了させた場合、ソフトウェア・プログラミング・エラーの結果として終了した場合、またはハードウェアの問題で停止した場合に、自動的な再起動が可能です。Solaris Service Managerはさらに、サービスをそれぞれの依存関係に基づき並行して開始させるので、システムの起動時間が 75 % も短縮されます。「取り消し (undo) 」機能により、変更を簡単に元に戻せるため、人的エラーに対する安全装置になります。Solaris Service Manager は導入が簡単であることも特長の 1 つです。開発者は各アプリケーションに単純な XML ファイルを追加するだけで、Solaris Service Manager 機能をフル活用するように、大部分の既存アプリケーションを変換できます。

予測的セルフヒーリングと障害管理により制御可能なSPARC Enterprise T5440サーバの主要な機能は以下の通りです。

- CPU オフラインは、動作不安定なコアをオフラインにします。オフライン化されたコアは、リソース・キャッシュに格納され、プロセッサが交換されない限り、

リブート後であってもシステムから分離されたままです。プロセッサの交換後、リソース・キャッシュからクリアされます。

- Memory Page Retirement は、障害フラグの付いたメモリ・ページをリタイアさせるものです。ページは、リソース・キャッシュに格納され、部位特定された DIMM が交換されない限り、リブート後であってもシステムから分離されたままです。DIMM の交換後、リソース・キャッシュからクリアされます。
- I/O Retirement は、エラーと障害を記録します。
- fmlog は、システムが検出した障害を記録します。

第5章 まとめ

Web 2.0 アプリケーションと仮想化、そして効果的なエコシステムを、データセンターが抱える今日の複合的な要求に応じて提供するには、画期的なプロセッサ／システム・プラットフォーム／オペレーティングシステムはもとより、有力なアプリケーション／ミドルウェア／管理テクノロジーまでを含めた、総合的なアプローチが必要と考えています。将来の話としてではなく、現時点で、企業がパフォーマンスとキャパシティのニーズに対処しながら、スペース／電力／発熱を適切に管理できるようにするソリューションがSPARC Enterpriseにはあります。

実績あるUltraSPARC T1とUltraSPARC T2を基礎に、UltraSPARC T2 Plusプロセッサはマルチスレッドアーキテクチャのメリットを、最もミッションクリティカルなアプリケーションやWebサービスをサポートするOLTPデータベースのワークロードにまで拡張します。最大256スレッド、最大512GBメモリ、そして大容量のI/O帯域幅を提供するこのシステムは、本質的なアプリケーションおよびサービスにとって理想的な統合／仮想プラットフォームといえます。プロセッサあたり64スレッド、オンチップ・メモリ管理、PCI Expressインタフェース、オンチップの暗号化処理アクセラレーション、マルチソケット・サポートのためのキャッシュ・コヒーレンスを提供することで、UltraSPARC T2 Plusプロセッサは現代のプロセッサのあるべき姿を再定義しました。このSPARC Enterprise T5440サーバの優れた機能性により、コンパクトな4Uラックマウント筐体で、ワット当たり優れたパフォーマンスを提供する新たなパフォーマンス・レベルで、強力で高スケーラブルなサーバ・プラットフォームを提供します。その結果、低い電力と少ない設置面積で、既存の設備全体に取って代れるデータセンター・インフラストラクチャを実現できます。

SPARC Enterprise T5440サーバは、最も要求の厳しい Web／アプリケーション／データベース／HPC 用途に必要な演算／ネットワーク／I/O リソースを提供し、効果的なサーバ統合を支援します。これらのシステムはエンド・ツー・エンドでマルチスレッド化と仮想化をサポートし、SPARC／Solaris テクノロジーへの投資を保護するとともにオープンソース・ソフトウェア環境のツールを提供することで、ワークロードを統合し、システム・リソースの有効活用を可能にします。企業では、その最重要プロジェクトに、ロジカル・ドメイン、Solarisコンテナといった最新システムを採用し、環境への配慮と収益向上に向けて迅速に対応できます。

詳細情報

富士通製品および SPARC Enterprise T5440サーバに関する詳細は、当社営業または販売パートナーまでお問い合わせいただくか、下記ウェブサイトの関連資料を参照してください。

<http://primeserver.fujitsu.com/sparcenterprise/>

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Limited.

SPARC is a trademark used under license from SPARC International, Inc. of the U.S.A.

Sun, Sun Microsystems, the Sun Logo, Solaris and all Solaris based marks and logos are trademarks or registered trade marks of Sun Microsystems, Inc, in the U.S. and other countries, and are used under license.

SPARC64 is a trademark of Fujitsu Limited under license from SPARC International, Inc. of the U.S.A.

All other company/product names mentioned may be trademarks or registered trademarks of their respective holders and are used for identification purposes only.



FUJITSU