

ORACLE®
SPARC Servers

オラクルのSPARC T8サーバーと SPARC M8サーバーのアーキテクチャ

Software in Silicon : リアルタイム・エンタープライズ向け
セキュア・インフラストラクチャの実現

Oracle ホワイト・ペーパー | 2017年9月



ORACLE®

目次

目次	1
概要	2
機能の比較	4
SPARC M8 プロセッサ	6
SPARC M8 プロセッサのアーキテクチャ	7
プロセッサ コアとキャッシュのアーキテクチャ	8
Software in Silicon テクノロジー	9
インメモリ・クエリ・アクセラレーション	10
インライン圧縮解凍	11
Java ストリーム・アクセラレーション	12
Oracle Number アクセラレーション	13
暗号化アクセラレーション	13
Silicon Secured Memory	14
SPARC M8 プロセッサベースのサーバー・ファミリーの概要	17
メモリ・サブシステム	17
I/O サブシステム	18
I/O コントローラ ASIC	18
NVM Express テクノロジー	19
内蔵 USB ストレージおよび Oracle Solaris ブート・プール	19
PCIe アダプタ・カード	20
SPARC T8-1、T8-2、および T8-4 サーバー	20
SPARC T8-1 サーバー	21
SPARC T8-2 サーバー	23
SPARC T8-4 サーバー	25
SPARC M8-8 サーバー	29
サーバー・コンポーネント	29
CPU、メモリ、I/O ユニット・シャーシ	29
CPU、メモリ、I/O ユニット・ボード	30
インターコネクト・アセンブリ	31
サービス・プロセッサとサービス・プロセッサ・モジュール	31
システム・ラックとパワー・ディストリビューション・ユニット	32
SPARC M8-8 サーバーのアーキテクチャ	32
SPARC M8-8 サーバー（シングル物理ドメイン）	33
SPARC M8-8 サーバー（デュアル物理ドメイン）	34
Oracle Solaris	36
仮想化	39
システム管理	41
Oracle ILOM とサービス・プロセッサ	41



電力管理	42
Oracle Enterprise Manager Ops Center	43
信頼性、可用性、保守性	44
高度な信頼性機能	44
エラー検出、診断、およびリカバリ	45
ホットサービス対応冗長コンポーネント	46
結論	47
追加情報	48



概要

最新のテクノロジーによって、IT インフラストラクチャは新しい方向に進んでいます。急速に進化するデジタル市場で要求されるパフォーマンスを実現するには、ビッグ・データ、ソーシャル・ビジネス、モバイル・アプリケーション、クラウド、リアルタイム分析のすべてで、先進的なソリューションと十分な処理能力が必要です。顧客はますますビジネスのスピードを追求するようになっており、組織は顧客の条件に合わせて顧客と関わる必要があります。高レベルのセキュリティを確保して機密情報を管理するとともに、毎日毎時間大量のデータを取得、分析、処理することが必要不可欠となってきました。

これらの課題のため、IT システムの設計、投資、運用の方法は、過去数十年に比べて劇的に変化することになります。データベース (SQL) および Java は、最新のクラウド対応アプリケーションの開発における事実上の標準言語となってきました。データの量、種類、速度が爆発的に増大した結果、組織がよりの確に素早く決定を下すために、セキュアで効果的な分析方法の必要性が高まっています。組織が、コストの削減、運用効率の向上、新しい収益源となる革新的なテクノロジーの実用化を迫られているまさにその時に、複雑な IT インフラストラクチャによるメンテナンスの難しさや高コストが達成の妨げになっています。コンピュータへの侵入によって毎年何十億ドルが失われており、データ・セキュリティを後付けの機能のように考えることはできなくなっています。

オラクルの新しい SPARC M8 プロセッサベースのサーバーは、オラクルの第 2 世代 Software in Silicon テクノロジーを導入することにより、サーバー・テクノロジーが新しいレベルに引き上げられています。これらのサーバーでは、圧縮データのインメモリ分析を行うための Data Analytics Acceleration (DAX) ユニット、Silicon Secured Memory テクノロジー、そして高速なワイドキー暗号化を使用することにより、データベースと Java アプリケーション両方の高速化とセキュリティを実現しています。SPARC M8 プロセッサは、わずか 2U のスペースで最大 256 のスレッドをサポートするアーキテクチャによって効率をさらに向上させています。また、スレッドあたりのパフォーマンス、信頼性、可用性、保守性 (RAS) と電力効率も改善されており、加えて他社の設計よりも大きなメモリ容量ならびに広い I/O 帯域幅を持っています。

オラクルの SPARC M8 プロセッサベースのサーバーでは、世界でもっとも効率的なプラットフォームによってリアルタイム・エンタープライズの競争力を高め、これにより組織が、今日のデジタル市場で競争を続け、コストと時間を節約し、収益の拡大を可能にします。同時に、オラクルの技術革新によって価値を創出し、組織におけるコストの低減と ROI の向上を可能にします。SPARC M8 プロセッサの先進技術に基づいたオラクルの新しい SPARC サーバー・ファミリー (図 1) では、SPARC M8 プロセッサを 1 基から 8 基まで拡張できます。これらのサーバーは、非常に高いレベルで統合することによってセキュリティを強化し、コストを低減し、信頼性を向上させており、柔軟に拡張可能な製品ファミリーを構成しています。最適化されたシステム設計により、あらゆるエンタープライズ・サービスとアプリケーション・タイプに対応できます。管理インターフェースの統一性と標準技術の採用によって管理コストの削減を促進するとともに、革新的なシャーシ設計により、最新のデータセンターを高密度化し、効率性、経済性を向上させています。

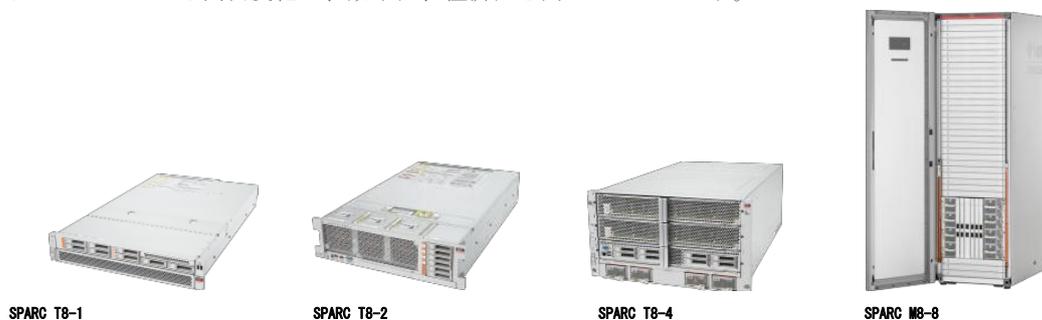


図1：オラクルのSPARC M8プロセッササーバー製品ファミリー

機能の比較

表1に、SPARC T8-1、T8-2、T8-4、およびM8-8サーバーの機能の比較を示します。

表1：SPARC M8プロセッサ・サーバーの機能比較

機能	SPARC T8-1	SPARC T8-2	SPARC T8-4	SPARC M8-8
フォーム・ファクタ	2U、 奥行き737mm / 29インチ	3U、 奥行き753mm / 29.6インチ	6U、 奥行き835mm / 32.9インチ	システム・ラック： 幅600mm、 奥行き1,200mm、 高さ2m / 78.7インチ スタンドアロン：10U、 奥行き813mm / 32インチ
物理ドメイン	1 32コア 5.0GHz SPARC M8 プロセッサ、64MB 共有レベル3 キャッシュ 1または2（静的）			
プロセッサ	プロセッサあたり最大256スレッド Silicon Secured Memory 32のDAXエンジンで、インメモリ・クエリ・高速化、インライン圧縮解凍、Javaストリーム・高速化を実行 次の16種類の業界標準暗号化アルゴリズムと乱数生成を直接サポートする暗号化命令アクセラレータを各コアに搭載：AES、Camellia、CRC32c、DES、3DES、DH、DSA、ECC、MD5、RSA、SHA-1、SHA-3、SHA-224、SHA-256、SHA-384、SHA-512			
プロセッサ数	1	2	2または4	2~8
最大コア数	32	64	128	256
最大スレッド数	256	512	1,024	2,048
メモリ	16GB、32GB、または64GBの DDR4-2400メモリDIMM		32GBまたは64GBのDDR4-2400メモリDIMM	
メモリ容量1	最大1,024GB 最小128GB	最大2,048GB 最小256GB	最大4,096GB 最小512GB	最大8,192GB 最小512GB
内蔵2.5インチ・ディスク・ドライブ・ベイ	8	6	8	N/A
SASによる内蔵2.5インチ・ディスク・ドライブ・ベイのサポート	RAID 0/1/10/1E対応の1個の統合型SAS3コントローラにより、最大8台の2.5インチSASハードディスク・ドライブ（HDD）またはソリッド・ステート・ドライブ（SSD）をサポート	RAID 0/1/10/1E対応の2個の統合型SAS3コントローラにより、最大6（2+4）台のSAS HDDまたはSSDをサポート	RAID 0/1/10/1E対応の2個の統合型SAS3コントローラにより、最大8（4+4）台のSAS HDDまたはSSDをサポート	N/A
NVMeによる内蔵2.5インチ・ディスク・ドライブ・ベイのサポート	工場出荷時構成のPCIeスイッチ（オプション）1枚で最大4台の2.5インチNVMe SSDをサポート	工場出荷時構成のPCIeスイッチ（オプション）1枚または2枚で最大4台の2.5インチNVMe SSDをサポート	PCIeスイッチ（オプション）2枚で最大8（4+4）台の2.5インチNVMe SSDをサポート	N/A

1. 物理メモリ容量。DIMMSペアリングは、メモリがフル搭載された状態で有効になり、メモリ容量の1/16が予約。DIMMSペアリングにより、システムの動作の中断、メモリ容量の損失、エラー保護機能の変更を引き起こすことなく、DIMM全体を自動的に構成から外すことが可能。

表1：SPARC M8プロセッサベースのサーバーの機能比較（続き）

機能	SPARC T8-1	SPARC T8-2	SPARC T8-4	SPARC M8-8
Oracle Flash Accelerator F640 PCIeカード (NVMe) の最大数	6	6	8	12
リムーバブル・メディア	USB経由で接続された外部DVD			DVDなし（USBおよびiKVM経由でアクセス）
管理ポート	イーサネット1000Base-Tポートx1 ² シリアルRJ45ポートx1			イーサネット1000Base-Tポートx2 ² （アクティブ/スタンバイ） シリアルRJ45ポートx2（アクティブ/スタンバイ）
ビデオ・ポート	HD-15 VGAビデオ・ポートx1	HD-15 VGAビデオ・ポートx2		N/A
USBポート	USB 2.0（フロント）ポートx2およびUSB 3.0（リア）ポートx2		USB 3.0ポート x4	N/A
イーサネット	統合型10GBase-Tポート x4 ³ 統合型イーサネット・コントローラx1			PCIeアダプタ・カード経由
PCIe 3.0 ロープロファイル・スロット	スロットx6 x8スロットx6、または x16スロットx2と x8スロットx2 4つのPCIeルート・コンプレックスによりサポート	スロットx8 x8スロットx4と x16スロットx4 8つのPCIeルート・コンプレックスによりサポート	ホットプラグ対応 スロットx16 x8スロットx8および x16スロットx8 12のPCIeルート・コンプレックスによりサポート	ホットプラグ対応 スロット最大24個 プロセッサあたり x16スロットx3 スロットあたり1つのPCIeルート・コンプレックス
PCIeルート・コンプレックスの総数	5	10	20	最大32
冗長電源	冗長（1+1）ホットスワップ対応AC 1,200Wの冗長電源x2	冗長（1+1）ホットスワップ対応AC 2,000Wの冗長電源x2	冗長（N+N）ホットスワップ対応AC 3,000Wの冗長電源x4	冗長（N+N）ホットスワップ対応AC 3,000Wの冗長電源x6
冗長ホットスワップ対応ファン（N+1）	デュアルファン・モジュールx4、上部取り付け	デュアルファン・モジュールx3、上部取り付け	デュアルファン・モジュールx5、リア取り付け	デュアルファン・モジュールx8、フロント取り付け
オペレーティング・システム	パフォーマンスおよび機能（Software in Siliconテクノロジーにより有効化される機能を含む）を強化するため、Oracle Solaris 11.3以降の使用を推奨。 制御ドメイン、ルート・ドメイン、I/Oドメイン：Oracle Solaris 11.3 SRU 24以降 ⁴ 次のバージョンがゲスト・ドメイン内でサポート。 » Oracle Solaris 11.3 SRU 24以降 ⁴ » Oracle Solaris 10 1/13 ⁵ Oracle Solaris 8またはOracle Solaris 9においてのみ動作が保証されているアプリケーションは、Oracle Solaris 10ゲスト・ドメイン内で稼働しているOracle Solaris 8またはOracle Solaris 9のブランド・ゾーンで実行可能。			

2. 1000Base-Tは、全二重の場合にのみ、10MB/秒、100MB/秒、1GB/秒に自動ネゴシエート。
3. 10GBase-Tは、全二重の場合にのみ、100MB/秒、1GB/秒、10GB/秒に自動ネゴシエート。最大15,500バイトまでのジャンボ・フレームをサポート。
4. Oracle Solaris 11.3 SRU 23より前のバージョンのOracle Solaris 11は、SPARC M8プロセッサベースのサーバーでは非サポート。
5. 必須のパッチ要。

SPARC M8プロセッサ

第2世代の Software in Silicon 機能と新設計のコアが組み込まれたオラクルの SPARC M8 プロセッサは、世界最高の処理速度を実現し、マルウェアやソフトウェア・エラーから画期的な方法でデータを保護します。

SPARC M8 プロセッサには、特定のソフトウェア機能またはプリミティブを高速化するハードウェア・ユニットが組み込まれています。これらオンチップのデータ・アナリティクス・アクセラレータ (DAX) ユニットのデータベース・クエリ処理をオフロードし、リアルタイムでデータ圧縮解凍を実行し、Java ストリームを加速化します。インメモリ・クエリ・アクセラレーションは、他のプロセッサに比べて最大 7 倍の性能を実現します。インライン圧縮解凍機能により、パフォーマンスを低下させることなく、同じメモリ・フットプリントで最大 2 倍の量のデータを保存できます。SPARC M8 プロセッサの新しいコアには、Oracle Database での実数処理を高速化する Oracle Number ユニットも組み込まれています。

SPARC M8 プロセッサの Silicon Secured Memory 機能 (図 2) では、データの整合性をリアルタイムでチェックして、ポインタ関連のソフトウェア・エラーとマルウェアからデータを保護します。これにより、ソフトウェアで実装した場合に非常に負荷のかかる複雑な処理を、ハードウェアによる低オーバーヘッドの監視機能で実行することができます。Silicon Secured Memory を活用することで、不正なメモリ・アクセスや誤アクセスの検出から原因の診断、そして適切なりカバリ・アクションまで、一連の処理をアプリケーションで実行できるようになります。SPARC M8 プロセッサでは、各プロセッサ・コアに直接統合されている暗号化命令アクセラレータが強化されています。このアクセラレータにより、16 種類の業界標準暗号化とハッシュを実現して、セキュア・コンピューティングに付きもののパフォーマンスとコスト面の問題をなくすことができます。

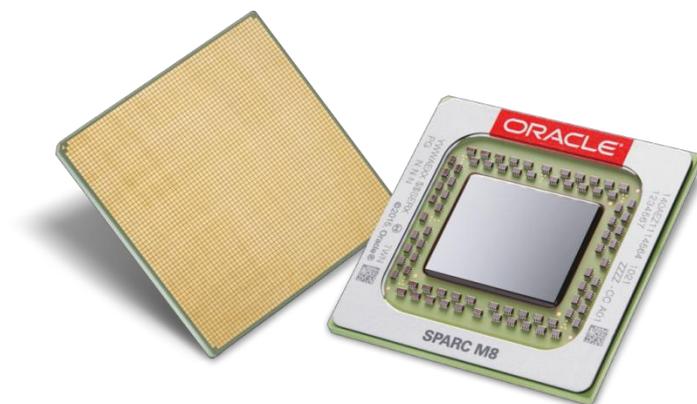


図 2: SPARC M8 プロセッサでは、32 個のコアと第 2 世代の Software in Silicon 機能を組み合わせて、アプリケーションとデータベースの処理速度を加速します。

新しいコア、オンチップ L2 および L3 キャッシュ設計の強化、プロセッサ周波数の増加によって、スレッドあたりのパフォーマンスが向上しています。64MB の L3 キャッシュを完全に共有し、もっとも近いパーティションにホット・キャッシュ・ラインを移設することで、待機時間の最小化とパフォーマンスの最大化が図られています。コアとキャッシュによるアーキテクチャは、サーバーの仮想化とプラグラブル・データベースに最適です。この設計により論理ドメイン間やデータベース間の通信が最小限になるため、システム管理とパフォーマンスの調整が容易になりました。このプロセッサではプロセッサの動作を動的に変えることができます。具体的には、最大 256 個のスレッドを使用してスループットを最大化したり、実行するスレッド数を減らして各スレッドのリソース割

当てを増やし、スレッドあたりのパフォーマンスを高めたりすることができます。この柔軟性により、システムとして最適な結果が得られるように、全体的なスループットとスレッドあたりのパフォーマンスのバランスを取ることができます。

32 コア SPARC M8 プロセッサは、従来の SPARC プロセッサとバイナリ互換性があります。仮想化されたクラウド・コンピューティング環境に最適であり、多数の仮想マシンをサポートし、卓越したマルチスレッド・パフォーマンスを発揮します。このプロセッサにより、最大限の効率と予測可能性で、新しいネットワーク・サービスを迅速に拡大できます。

表 2 に、オラクルの SPARC M8、SPARC M7、SPARC M6、および SPARC T5 プロセッサの比較を示します。

表 2 : SPARC M8、SPARC M7、SPARC M6、および SPARC T5 プロセッサの機能比較

機能	SPARC M8 プロセッサ	SPARC M7 プロセッサ	SPARC M6 プロセッサ	SPARC T5 プロセッサ
CPU周波数	5.06GHz	4.13GHz	3.6GHz	3.6GHz
アウト・オブ・オーダー実行	あり	あり	あり	あり
命令発行幅	4	2	2	2
データ/命令プリフェッチ	あり	あり	あり	あり
SPARCコア	第5世代	第4世代	第3世代	第3世代
プロセッサあたりのコア数	32	32	12	16
コアあたりのスレッド数	8	8	8	8
プロセッサあたりのスレッド数	256	256	96	128
システムのソケット数	最大8	最大16	最大32	最大8
プロセッサあたりのメモリ	最大16枚のDDR4 DIMM	最大16枚のDDR4 DIMM	最大32枚のDDR3 DIMM	最大16枚のDDR3 DIMM
キャッシュ	32KB L1命令キャッシュ (4ウェイ) 16KB L1データ・キャッシュ (4ウェイ) 共有型256KB L2命令キャッシュ (4ウェイ、クアドコアあたり) 128KB L2データキャッシュ (8ウェイ、コアあたり) 共有型64MB (L3) キャッシュ	16KB L1命令キャッシュ (4ウェイ) 16KB L1データ・キャッシュ (4ウェイ) 共有型256KB L2命令キャッシュ (4ウェイ、クアドコアあたり) 共有型256KB L2データキャッシュ (8ウェイ、コアあたり) 共有型64MB (L3) キャッシュ	16KB L1命令キャッシュ (4ウェイ) 16KB L1データ・キャッシュ (4ウェイ) 128KB L2キャッシュ (8ウェイ) 共有型48MB L3キャッシュ (12ウェイ)	16KB L1命令キャッシュ (4ウェイ) 16KB L1データ・キャッシュ (4ウェイ) 128KB L2キャッシュ (8ウェイ) 共有型8MB L3キャッシュ (16ウェイ)
ラージ・ページのサポート ¹	16GB	16GB	2GB	2GB
電力管理の粒度	チップの1/2	チップの1/4	チップ全体	チップ全体
テクノロジ	20nmテクノロジ	20nmテクノロジ	28nmテクノロジ	28nmテクノロジ

1.Oracle Solaris 11.3でのラージ・ページのサポート

SPARC M8 プロセッサのアーキテクチャ

商用ワークロードを適切なレベルのスループットで処理するため、SPARC M8 プロセッサには新たなコアとキャッシュの階層がその他の改良点とともに実装されており、他社のプロセッサより最大で2倍の処理速度を実現できます。電力管理機能は、システム内のパフォーマンスを高める上で引き続き大きな役割を果たしており、動的電圧周波数スケールリング (DVFS) 機能も提供されます。

図 3 に、SPARC M8 プロセッサのアーキテクチャを示します。SPARC M8 プロセッサには、32 個のコアがあり、L3 キャッシュが完全に共有されます。4 つのメモリ・コントローラ・ユニット (MCU) を搭載し、それぞれが高速リンク経由で Buffer-on-Board (BoB) ASIC に接続されます。BoB には DDR4 チャンネルが 2 つのあり、各チャンネルが 1 つのメモリ DIMM に接続されています。サポートされる SPARC M8 プロセッサあたりの DDR4 DIMM の数は最大で合計 16 枚です。

SPARC M8 プロセッサでは、インクルーシブ・ディレクトリを使用して 8 ウェイのグルーレス・スケーリングに対応できます。他の SPARC M8 プロセッサとの接続とコヒーレンスを保つため、8 つのコヒーレンシ・リンク (CL) が用意されています。SPARC M8 プロセッサベースのサーバーでは、2 つの I/O リンク (IL) が I/O コントローラ ASIC に接続されます。ロジックの追加なしで、最大 8 基の SPARC M8 プロセッサを単一の対称型マルチプロセッシング (SMP) システムに接続可能です。

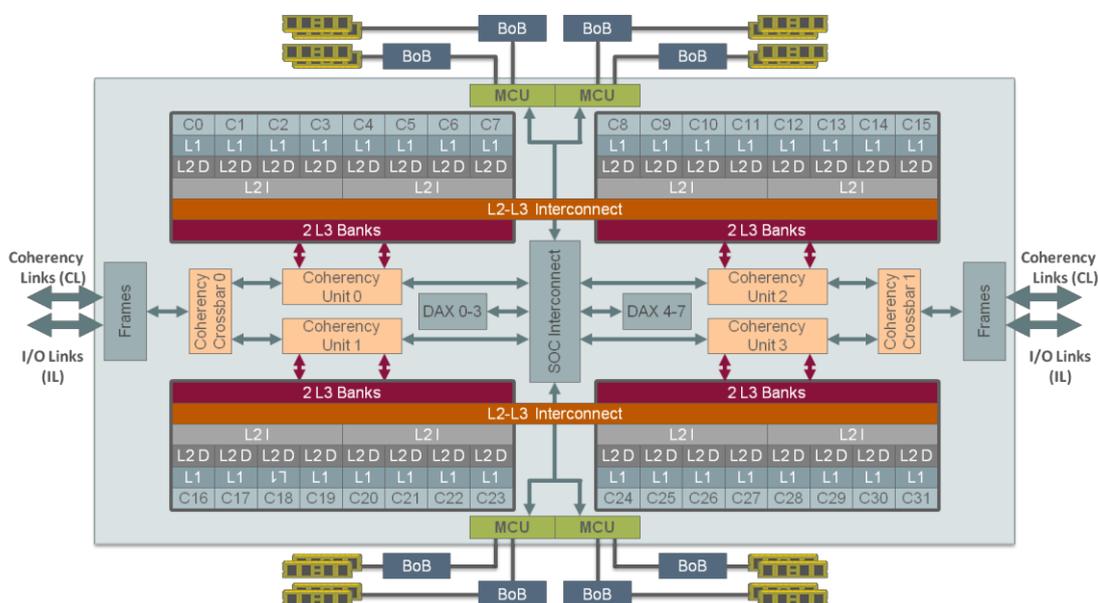


図3：SPARC M8プロセッサの特長：2つのパーティションにグループ化された32個のコア、4つのSPARCコア・クラスタ、4つのメモリ・コントローラ・ユニット (MCU)、8つのデータ分析アクセラレータ (DAX) ユニット

プロセッサのコアとキャッシュのアーキテクチャ

SPARC M8 プロセッサのコアは、イシュー幅が 4 で、最大 192 の実行中の命令をアウト・オブ・オーダーで処理し、最大 8 つのハードウェア・スレッドをサポートします。コアには、スレッドあたりのパフォーマンスが最大限になるように最適化するダイナミック・スレディング機能があります。ソフトウェアを使用し、クリティカル・スレッドの最適化により各コアで最大 8 つのハードウェア・スレッド (ストランド) をアクティブ化できます。コアのリソースは、プロセッサ・ハードウェアにより、アクティブなストランドに動的かつシームレスに割り当てられます。

プロセッサのコアと最終レベルのキャッシュが 2 つのパーティションとして編成されており、各パーティションに 16 個のコアと 32MB の L3 キャッシュがあります。各コアには、専用の 32KB の L1 命令キャッシュ、16KB の L1 データ・キャッシュ、128KB の L2 データ・キャッシュが組み込まれています。4 つのコアで 256KB の L2 命令キャッシュを共有します。プロセッサあたりの L3 キャッシュは、容量の合計が 64MB (すべての内部キャッシュを含みます) で、完全に共有され、16 ウェイ・セット・アソシエイティブ方式が採用されています。すべての L3 パーティションで、SPARC M8 プロセッサの 32 個あるどのコアからのリクエストであっても対応できます。ホット・キャッ

シュ・ラインは、パフォーマンスの最適化を図ってもっとも近い L3 キャッシュ・パーティションに移されています。

コア内のメモリ管理ユニット (MMU) では、ハードウェア・テーブル・ウォーク (HWTW) が行われ、8KB、64KB、4MB、256MB、16GB のページがサポートされています。

Software in Siliconテクノロジー

ほとんどのプロセッサ・チップ開発では、汎用処理能力を向上させて高速化することに力を入れています。数年前オラクルは、インメモリ・データベース機能をチップに直接搭載することによってメモリ内のデータをハードワイヤード保護する画期的なプロジェクトを開始しました。オラクルは、プロセッサ、システム、アプリケーションのレベルでこの革新的な手法を取り入れてアプリケーション・パフォーマンスを最適化する唯一のベンダーです。SPARC M7 は、Software in Silicon 機能をプロセッサ自体に組み込んでこの機能を利用するように設計された最初のプロセッサでした。SPARC M8 プロセッサは、SPARC M7 を基盤にして設計されており、第 2 世代の Software in Silicon テクノロジーを実現します。

SPARC M8 プロセッサでは、オンチップのアクセラレータを組み込むことにより、メモリ内のデータベース・クエリ処理をオフロードし、データ圧縮解凍をリアルタイムで実行し、Java ストリームを高速化します。強化された暗号化命令アクセラレータと新しい Oracle Number ユニットの、各プロセッサ・コアに直接組み込まれています。以下に示す Software in Silicon の機能はこれらを総合して、セキュリティ、パフォーマンス、効率の面で大きなメリットを実現します。

- » DAX ユニットの**インメモリ・クエリ・アクセラレーション**は、他のプロセッサと比較して最大で 7 倍のパフォーマンスを発揮します。
- » **インライン・データ圧縮解凍**を使用すると、パフォーマンスを低下させることなく、同じメモリ・フットプリント内に最大 2 倍の量のデータを保存可能です。
- » **Java ストリーム・アクセラレーション**により、Java 8 アプリケーションでは新しいストリーム・ライブラリを使用して DAX ユニットを利用することができ、これにより合理化分析能力が高まり、Java ストリーム操作のパフォーマンスが著しく向上します。
- » **Oracle Number ユニット**では、独自のオンチップ・アクセラレータを利用して、Oracle Database 独自のプリミティブ・データ型である Oracle Number データ型で実行される算術処理のパフォーマンスを向上しています。
- » **Silicon Secured Memory** では、リアルタイムのデータ整合性チェック機能によってポインタ関連のソフトウェア・エラーやマルウェアからデータを保護し、高コストのソフトウェア・インスツルメント処理をハードウェアによる低オーバーヘッドの監視機能で置き換えます。Silicon Secured Memory によりアプリケーションでは、メモリへの不正なアクセスや誤ったアクセスを検出し、原因を診断して、適切なリカバリ操作を実行できます。
- » **暗号化アクセラレータ**により、セキュア・コンピューティングに付きもののパフォーマンス面とコスト面の問題を排除することができます。障害を取り除くことは、現代の事業運営においてますます必要不可欠な課題となっています。

すべてのコアに搭載されている暗号化命令アクセラレータと Oracle Number ユニットに加えて、SPARC M8 プロセッサには 8 個の第 2 世代 DAX ユニットが組み込まれており、それぞれの DAX ユニットにパイ

プライン（エンジン）が4つあります。これらのエンジンでは、32の独立したデータ・ストリームを処理し、プロセッサ・コアをオフロードして他の処理に振り向けることができます。DAX エンジンでは、圧縮解凍、スキャン、フィルタ、ジョインなどのクエリ関数を処理できます。

DAX ユニットでは、オーバーヘッドが非常に小さいプロセス間通信ときわめて高速なアトミック操作を使用します。たとえば、別々のプロセッサに存在する DAX ユニットが、メッセージを交換し、リモート・メモリの場所にアクセスして、CPU を介さずにロックを交換できます。この機能を利用するには、Oracle Database In-Memory オプションが有効な Oracle Database 12c、および

Oracle Solaris 11.3 以上が必要です。以降のセクションでは、オンチップ・アクセラレータによって実現される Software in Silicon 機能について説明します。

既存のアプリケーションは、適切な Oracle Solaris ライブラリを使用し、テスト環境で検証されることにより、リコンパイルしなくても Silicon Secured Memory で実行可能です。ソフトウェア開発者は、Oracle Solaris のオープンな API を使用して、Silicon Secured Memory と DAX テクノロジーを利用できます。

インメモリ・クエリ・アクセラレーション

インメモリ・クエリ・アクセラレーションは、高速分析応答をその主要な設計方針として設計された Oracle Database In-Memory と連携するように設計されました。データをデータベースに保存してアクセスする従来の方法では、行形式が使用されます。この方法は、挿入と更新が頻繁に行われるトランザクション・ワークロードとともに、レポート形式のクエリの場合に適しています。しかし、分析の実行にもっとも適しているのは列形式です。Oracle Database In-Memory を使用すれば、オンライン・トランザクション処理（OLTP）用の行形式と分析処理用の列形式の両方を備えた二重形式のアーキテクチャにすることが可能です。¹

Oracle Database In-Memory では、データがインメモリのカラムストア（列形式）に保存されます。インメモリ列に保存されているデータに対して、一連の圧縮アルゴリズムが自動的に実行されるため、ストレージの使用量が節約されます。さらに、クエリが実行されると、圧縮形式のままデータをスキャンおよびフィルタリングするため、データを圧縮解凍する必要がありません。図4に示すように、インメモリのカラムストアには In-Memory Compression Unit（IMCU）が作成されています。インメモリ列形式データは、このようなより小さな IMCU に断片化されるため、データ全体を対象とするクエリが実行される場合には並列化が可能です。



図4：インメモリ列形式データがより小さなIMCUに断片化されるため、並列処理が可能

¹ インメモリ・クエリ・アクセラレーションは、以下を前提として SPARC M8 プロセッサベースのサーバーでサポートされます。Oracle Solaris 11.3 以上、Oracle Database 12c 12.1.0.2 バンドル・パッチ、および Oracle Database In-Memory オプション。

SPARC M8 プロセッサのコアがデータベースのクエリを受信すると、そのクエリはオンチップ・アクセラレータにオフロードされます。高速化されたデータベースの処理は次のように行われます。

- » 選択：フィルタリングにより列を減らす
- » スキャン：検索 (“where”句)
- » 抽出：圧縮解凍
- » 変換：参照により大/小の結合を高速化する

クエリがオフロードされると、解放されたコアでは、より高レベルの SQL 関数などの他のジョブを再開できます。一方、アクセラレータではクエリを実行して結果を L3 キャッシュに格納し、コアによる高速アクセスを可能にします。関連コアでは、クエリが完了したことが通知されると、その結果を取得します。

このクエリのオフロード・メカニズムによる処理の高速化以外の利点は、各 SPARC M8 プロセッサ内の 32 のアクセラレータ・エンジンによって促進される大規模並列化があります。プロセッサ内の 32 の各コアには、これらすべてのアクセラレータ・エンジンへのアクセス権があり、それらのアクセス権を同時に使用して 1 つのクエリを完全に並列実行できます。

このパラレル化の仕組みはプロセッサによって実現され、追加の処理を実行するためにアプリケーション・コードやデータベースは必要ありません。アクセラレータでは、プロセッサの超高帯域幅のインタフェースにより、メモリ・サブシステムからデータ・ストリームを直接取り込むことができます。その結果、インメモリ・データに対するクエリは、プロセッサ・コアに接続するキャッシュ・アーキテクチャによって制御されるのではなく、メモリ・インタフェースによって決まる速度で実行できます。

インライン圧縮解凍

圧縮は、より多くのデータをメモリおよびストレージに保存する上で非常に重要なカギとなります。圧縮解凍速度は、書込みよりも読取りの方が多い一般的なデータベース・アプリケーションにおいてもっとも重要です。今日のプロセッサの圧縮解凍速度はディスク・アクセスの場合には十分ですが、フラッシュ・メモリの場合には遅く、インメモリ・データベース・アプリケーションにとって非常に大きなボトルネックとなっています。

この課題を解決するため、クエリ処理に不可欠なステップとしてインライン圧縮解凍機能が DAX ユニットに実装されています。アクセラレータはデータの圧縮解凍とクエリ機能を 1 つのステップで実行するため、読取りと書き込みを複数回行う必要がなくなります。その結果、ペナルティは発生せず、インライン圧縮解凍を 120GB/秒を超えるメモリ速度で実行できます。圧縮解凍は次の順序で実行されます。

- » プロセッサ・コアがクエリ処理をアクセラレータにオフロードし、アクセラレータは圧縮データをそのまま読み取ります (OZIP 圧縮を使用)。
- » アクセラレータは、追加の読取りまたは書き込み処理を行うことなく、1 つのステップでデータをそのまま圧縮解凍してクエリを評価します。
- » 次にプロセッサ・コアは、最終的な結果を圧縮解凍済みデータとして書き出します。

Javaストリーム・アクセラレーション

Java プログラミング言語は常に、オブジェクト指向プログラミングとともに、分析アプリケーションとフレームワークの実行における先駆者となってきました。今日、多くの Java アプリケーションではコレクションという Java クラスを利用しています。コレクションにより、事前定義されたデータ構造体とメソッドを利用して、プログラマー・フレンドリーな方法で大量のデータを効率よくグループ化し、処理することができます。ただし、コレクション・クラスのメソッドの多くでは計算処理が非常に集中し、特に大量のインメモリ・データを分析する場合に顕著です。これは、コレクションに追加できるようにするために、まずコレクションのすべての要素を計算する必要があります。

Java 8 では、コレクションを使用する代わりに、Java ストリームの概念を導入しました。概念上、Java ストリームは固定データ構造体であり、その中の要素はオンデマンドで計算され、ソース・データは変更されません。Java ストリームを使用すると、アプリケーションでは SQL 文のように宣言的にデータを処理できます。

Java ストリームを使用することには多くの利点があります。Java ストリームを使用すると、ストリームでは必要とときにだけデータについての計算が実行されて、そのため不要なメモリのコピーが行われないので、コーディングとともに処理の実行が簡素化されます。さらに、その基盤となる設計のおかげで、Java ストリームではマルチスレッドとマルチコアのアーキテクチャの利点を活かすことができます。

ストリームの特性を以下に示します。

- » 要素のシーケンス：ストリームは特定のタイプの一連の要素を順次送り、オンデマンドで要素を取得/計算します。要素を保存することはありません。
- » ソース：ストリームの入力ソースはコレクション、配列、または I/O リソースです。
- » 集計演算：ストリームでは、フィルタリング、マッピング、制限、削減、検索、照合などの集計演算をサポートしています。
- » パイプライン処理：ほとんどのストリーム操作ではストリームが返されるため、パイプライン処理を行うことができます。これらの操作は中間操作といい、入力を受け取って処理し、出力を次のストリーム操作に返すのがその機能です。

- » 自動反復：ストリーム操作は、明示的な反復を必要とするコレクションとは対照的に、提供されたソース要素に対して内部で反復を実行します。

SPARC プロセッサの DAX ユニットは、Scan、Select、Extract、Fill、Translate などの特殊化された関数を超高速に実行するように設計されています。Java ストリームの実装と機能性に基づくストリーム操作は、DAX ユニット独自の機能を利用する場合に最適です。

オラクルは、DAX ユニットを利用する新しいストリーム API をリリースしました。Oracle Solaris パッケージの一部として入手可能なライブラリにより、Java プログラマーは、標準ストリーム API のものと同じインタフェースを利用しながら SPARC M8 プロセッサの DAX ユニットを使用できます。必要なのは、ソース・ファイルの import 文にわずかな変更を加えることだけです。さらにライブラリは、有利な場合のみ整数ストリーム関数を DAX ユニットにオフロードするように設計されており、有利でない場合は通常のストリーム操作を実行します。

Java ストリームを処理する場合、DAX ユニットには大きなメリットがあり、データソースのサイズとデータに対して実行される操作のタイプに応じて、操作によってはパフォーマンスが最大 20 倍アップします。

Oracle Number アクセラレーション

Oracle Database には、プリミティブ・データ型が独自の方法で実装されています。SQL 文の各列の値と定数には、特定の保存形式、制約、値の有効範囲が関連付けられているデータ型があります。Oracle Database の表の列ごとに、必ずデータ型を指定する必要があります。

Oracle Database には、特に固定小数点数と浮動小数点数用の Oracle Number という独自のデータ型があり、このデータ型の場合には固定小数点数と浮動小数点数を格納できます。実質的に何桁の数値でも格納することができ、Oracle Database が稼働しているさまざまなシステム間において最大 38 桁の精度で移植可能なことが保証されています。

SPARC M8 プロセッサの 32 のコアそれぞれに、Oracle Number の算術演算速度を加速するように特別設計された Oracle Number ユニットが組み込まれています。Oracle Number データ型では、新しい 4 つの命令 (ONadd、ONsub、ONmul、ONdiv) を使用することができ、最大 22 バイトの固定小数点数と浮動小数点数のすべての長さをネイティブにサポートします。

Oracle Number データの算術演算処理は計算集中タスクです。これは、それぞれの演算には複数の命令が必要であり、データベースの算術処理にはタスクごとに何百万回の算術演算が必要な場合があるからです。SPARC プロセッサの新しい Oracle Number ユニットでは、演算あたり命令を 1 つだけ受け入れることによって処理速度を高めます。その結果、パフォーマンスが大きく向上します。Oracle Number のビット長が長い場合 (16 バイト超) にはパフォーマンスが 10 倍以上になるため、演算処理に必要な時間が短縮され、コアが解放されて他の処理を行うことができます。

レポートやデータウェアハウスなどの算術演算集中型のワークロードの場合は、この機能によりリソース使用率を大幅に節約し、デプロイメントに必要なコンピュータ・ノードの総数を減らすことができます。

暗号化アクセラレーション

セキュリティの強化はかつてないほど重要となっており、SPARC プロセッサとシステムには、長年にわたってプロセッサベースの暗号化アクセラレーション機能が組み込まれてきました。SPARC M8 プロセッサの 32 のコアそれぞれに、暗号化方式、ハッシュ、キー操作、チェックサムで構成される業界最多の暗号化スイートを含む暗号化命令アクセラレータが組み込まれています。コアは 16

種類の業界標準暗号化アルゴリズムと乱数生成をサポートします。暗号化の加速は、Oracle Solaris の暗号化フレームワークでサポートされます。

SPARC M8 プロセッサでは、AES、Camellia、CRC32c、DES、3DES、DH、DSA、ECC、MD5、RSA、SHA-1、SHA-3、SHA-224、SHA-256、SHA-384、SHA-512 などのサポート対象アルゴリズムを使用した、暗号化方式のハードウェア実装へのアクセスが許可されます。暗号化方式は、コプロセッサとしてではなく、適切なパイプライン自体に実装されています。この方法では、より効率的なハードウェアベースの暗号化を実装し、特権レベルの変更もないため、暗号化アルゴリズムの計算効率が大幅に高まります。

また、データベースの操作により、命令パイプライン自体に実装された各種暗号化方式をはるかに効率的に利用できます。SPARC M8 プロセッサに組み込まれた暗号化機能を Oracle スタックのすべての階層で使用することにより、ほとんどパフォーマンスを低下させることなく、データ・セキュリティを大幅に強化することができます。

Silicon Secured Memory

SPARC M7、SPARC S7、および SPARC M8 プロセッサの Silicon Secured Memory は、ハードウェア内に動的なポインタ・チェック機能を配置することによってハードウェアベースでメモリを保護する機能です。この機能では、メモリ参照エラーを検出および報告し、メモリ内のデータへの意図しない、または悪意のあるアクセスを阻止します。

C や C++ などの一部のプログラミング言語には、ソフトウェア・エラーによって発生するメモリ破損に対する脆弱性が残っています。この種のメモリ参照のバグは検出がきわめて難しく、通常は、データが破損してかなり時間が経過してからデータの破損に気付きます。さらに厄介なことに、データベースとアプリケーションには、何千万もの命令行があり、開発者の数が数千人に達することもあります。重要なのは、バッファ・オーバーフローなどのエラーが組織をリスクにさらしかねないセキュリティ侵害の主な要因になるということです。

現在のアプリケーションは、大きな共有メモリ・セグメント上で動作する多くのスレッドを使用します。これらのアプリケーションのバグまたはポインタの問題が原因で、きわめて予測不可能な動作が行われる可能性があり、アプリケーション開発者は問題のトラブルシューティングと診断に大量の時間を費やすこととなります。サイレント・データ破損とバッファ・オーバーランは、これら診断が難しい問題のうちの2つです。

Silicon Secured Memory により、どちらの問題の場合であっても、アプリケーション開発者がメモリ参照バグのトラブルシューティングに費やす時間を劇的に短縮できます。サイレント・データ破損の場合は、Silicon Secured Memory によってアプリケーションが直ちに対処するよう促し、コストのかかるリカバリ作業を回避することができます。

図 5 に、サイレント・データ破損の問題を示します。図では、2 つのアプリケーション・スレッド (A と B) が同じメモリ位置に誤ってアクセスしています。それぞれのスレッドがアクセスすべきメモリ領域が色分けされていますが、ソフトウェアのプログラミング・エラーによって、スレッド A がスレッド B の赤で囲んだ領域に誤って書き込むという事態が発生します。このエラーは通常すぐには見つからず、そのメモリがスレッド B によって読み取られないと検出されない可能性があります。この場合、スレッド B のデータは気付かれることなくスレッド A によって破損されてきたことになり、多くの場合、破損の原因を追跡するのはきわめて困難です。この問題は非常に診断が難しく、通常、重大な結果を招く可能性のあるソフトウェア・バグとして現れます。

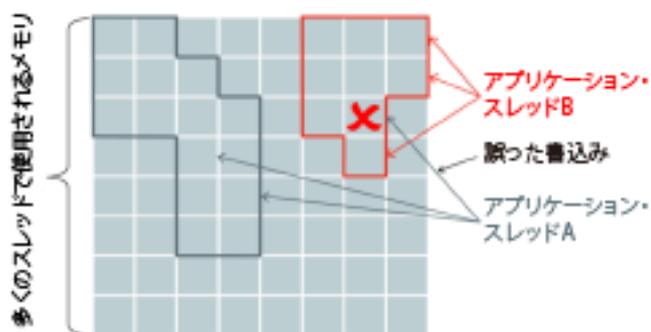


図5：サイレント・データ破損は、2つのスレッドが同じメモリ位置に誤って書き込みを行うときに発生

バッファ・オーバーランは、アプリケーションで発生する可能性がある別の問題です。簡単にいえば、バッファ・オーバーランとは、アプリケーションが割当て領域以外の場所に誤ってデータの書き込みを開始したことを示します（図 6）。このエラーにより、機密データが他のメモリ位置に漏れ、アプリケーションではそのことを認識していない状況が起こり得ます。その結果、悪意のあるアプリケーションはこれらすべての機密情報を読み取ることができます。バッファ・オーバーランは甚大なセキュリティ上の被害を引き起こしかねず、今日の世界では、多くの場合に悪意のあるウイルス攻撃という形で見られます。

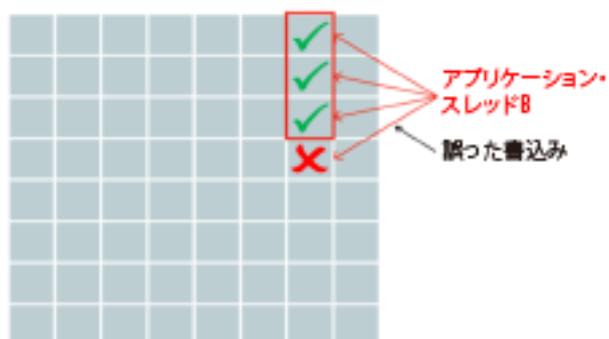


図6：重大なセキュリティ・リスクになりかねないバッファ・オーバーラン

Silicon Secured Memory では、（メモリ）ポインタごとにメモリ・バージョンとしてのキーを使用することによってこれらの問題に対処しています。メモリ割当て処理の間に、対応コードがメモリ・バージョンとしてメモリに書き込まれます。ポインタがこのメモリにアクセスすると、キーとコードがハードウェアで比較されます。一致する場合、そのアクセスは許可されます。一致しない場合は、メモリ参照エラーが発生しており、即座に検出されます。

SPARC M8プロセッサベースのサーバー・ファミリーの概要

SPARC M8 プロセッサベースのサーバーは、高いレベルのセキュリティ、パフォーマンス、効率性を必要とするクラウド・インフラストラクチャ用に設計されています。これらの SPARC サーバーは、データベース、Java、ミドルウェア、エンタープライズ・アプリケーションに最適であり、比類のないスループット・パフォーマンスとメモリ帯域幅を実現します。このサーバー製品ファミリーは 1~8 基の SPARC M8 プロセッサをサポートし、非常に多様なアプリケーション、機能、容量に対応できます。

このサーバー・ファミリーに共通するハードウェア機能には、以下が含まれます。

- ▶ SPARC M8 32 コア 5.06GHz プロセッサ (第 2 世代 Software in Silicon テクノロジーを実装)
- ▶ 32GB および 64GB の DDR4-2400 メモリ DIMM (SPARC T8-1 および T8-2 サーバーでは 16GB の DIMM も使用可能)
- ▶ PCIe 3.0 x16 対応拡張スロット
- ▶ NVMe Express (NVMe) フラッシュ・デバイスをサポート
- ▶ オンボード 12GB/秒 SAS3 I/O コントローラ (SPARC T8-1 サーバー、T8-2 サーバー、T8-4 サーバー)
- ▶ InfiniBand ネットワーク経由の起動をサポートする内蔵 USB (eUSB) ストレージ・デバイス

メモリ・サブシステム

各 SPARC M8 プロセッサは、8 つの Buffer-on-Board (BoB) ASIC を介して最大 16 枚の DDR4 メモリ DIMM をサポートします。プロセッサあたり最大 1TB のメモリがサポートされ、16 枚の 64GB DIMM それぞれが専用のメモリ・チャネルを使用します。SPARC M8 プロセッサのメモリ・リンクには 374GB/秒の物理帯域幅があり、これは、16 の DDR4-2400 チャネルが必要とする物理集約帯域幅 307GB/秒を十二分に上回ります。ハーフおよびフル・メモリ構成がサポートされています。詳細な構成ポリシーについては、このホワイト・ペーパーのモデル別のセクションで後述します。

メモリ DIMM によって提供され、個々の SPARC M8 プロセッサによって制御される物理アドレス空間は、パフォーマンスを最大化するためにインターリーブされます。ハーフ・メモリ構成は 8 ウェイでインターリーブされ、プロセッサあたり 16 枚の DIMM を装着したフル・メモリ構成は 16 ウェイでインターリーブされます。SPARC M8 プロセッサでは、15 ウェイのインターリーブ構成もサポートされています。16 ウェイから 15 ウェイ構成への切り替えは動的に実行されます。この機能は、DIMM スペアリングという機能の基盤であり、SPARC M7 プロセッサベースのサーバーで初めて導入されました。DIMM スペアリングにより、DIMM 交換サービス作業の必要性が軽減され、システムのアップタイムが増加します。

DIMM スペアリングは障害の発生した DIMM を構成から除外する機能であり、この機能により計画外のシステム中断が回避されます。各 DIMM の容量の 1/16 を未使用のままにすることで、問題のある DIMM を除外してその内容を残りの 15 枚の DIMM に再マッピングすることができます。DIMM が故障したと判断されると、アプリケーション・サービスを中断することなく、自動的に DIMM スペアリングが実行されます。このプロセスの間、システム・メモリ容量は変わらず、エラー保護は DIMM スペアリングの実行後もそのまま維持されます。システムは、容量を失ったり障害の発生数が増えたりすることなく稼働し続けます。その結果、システムを停止してハードウェアの保守を行う必要がなくなります。実際の DIMM の交換プロセスは、同じメモリ・バンク内の 2 枚目の DIMM を交換す

る必要が生じるまで待つことができます。

DIMM スペアリングは、フル・メモリ構成（プロセッサあたり 16 枚の DIMM）の SPARC M8 プロセッサベースのサーバーで有効にすることができます。DIMM スペアリングは、ハーフ・メモリ構成ではサポートされません。推奨ではありませんが、フル・メモリ構成で DIMM スペアリングを無効にすることも可能です。

I/Oサブシステム

SPARC T8 プロセッサベースのサーバーは、同じ基本 I/O サブシステム設計を共有しています。各 SPARC M8 プロセッサは、I/O リンク（IL）を介して 1 つまたは 2 つの I/O コントローラ ASIC に接続されます。SPARC M8 プロセッサと I/O コントローラ ASIC には IL が 2 つあります。

SPARC M8 プロセッサと使用中の I/O コントローラ ASIC 間の接続には、2 種類の実装方法があります。SPARC T8-1 サーバーと SPARC M8-8 サーバーでは、各 SPARC M8 プロセッサが両方の IL を使用して 1 つの I/O コントローラ ASIC に接続されています。SPARC T8-2 サーバーと SPARC T8-4 サーバーでは、プロセッサまたは IL の 1 つが使用できない場合でも、2 つの I/O コントローラ ASIC（および PCIe デバイス）への接続を確立できるよう、クロスオーバー接続スキームを利用します。クロスオーバー接続では、プロセッサの一方の IL が 1 つの I/O コントローラに接続され、もう一方の IL が別の I/O コントローラに接続されます。詳細については、このホワイト・ペーパーのモデル別のセクションを参照してください。

I/OコントローラASIC

PCIe インフラストラクチャは、I/O コントローラ ASIC によって実現されます。それぞれの ASIC には、集約帯域幅が 72GB/秒になる 5 つの PCIe 3.0 ルート・コンプレックスがあります。I/O コントローラ ASIC によって PCIe ファブリック全体がホストされるため、プロセッサが追加または除外されても PCIe ファブリックはそのままの状態を維持されます。結果として、デバイスへの接続が I/O コントローラ ASIC にあるルート・コンプレックス内で固定されるため、PCIe デバイス・パスは変更されません。SPARC M8 プロセッサベースのサーバーで使用される I/O コントローラ ASIC は、次に示すような優れた革新性があります。

- » 2 つの x16 IL が SPARC M8 プロセッサに接続され、各 IL は 2 つの x8 接続で構成されます。
- » 各 x8 IL 接続でシングル・レーン障害に対応できます。
- » 各 IL がハードウェア・キャッシュ・コヒーレンシに関与します。
- » SPARC T8-2 サーバーと T8-4 サーバーでは、デュアルホスト・プロセッサのフェイルオーバーが使用されます。
- » SR-IOV に準拠しています。
- » DMA ストリーム単位でアドレス変換が行われます。
- » DMA ストリーム単位でパケットの緩やかなパケット順序付けが行われます。

- » 4 つの PCIe 3.0 x16 ポートを装備しており、これらのポートは 4 つに分岐可能で、x16 PCIe ポート 1 つ、x4 PCIe ポート 4 つ、または x8 PCIe ポート 2 つとして実装できます。
- » 1 つの PCIe 3.0 x8 ポートを装備しており、このポートは 2 つに分岐可能で、x8 PCIe ポート 1 つまたは x4 PCIe ポート 2 つとして実装できます。
- » 5 つの PCIe 3.0 ポートのそれぞれが、独立したルート・コンプレックスです。

NVM Expressテクノロジー

SPARC M8 プロセッサベース・サーバーは、NVM Express (NVMe) として知られる新しいフラッシュ・ストレージ・テクノロジーに対応しています。NVMe 仕様では、ソリッド・ステート・ドライブ (SSD) 用に最適化した PCIe ベースのインタフェースを定義しています。NVMe ベースの SSD では、不揮発性メモリを利用して、SAS または SATA ベースの SSD よりも短い待機時間と優れたスループット・パフォーマンスの両方を実現しています。NVMe では PCIe 信号を利用しており、ドライブあたり 8 ギガ転送/秒 (GT/秒) の x4 インタフェースを実現して、ドライブへの約 4GB/秒全二重帯域幅を可能にします。

SPARC M8 プロセッサベースのサーバーすべてで、Oracle Flash Accelerator F640 PCIe カード (ロープロファイル PCIe カード上の NVMe ベースの SSD デバイス) に対応しています。SPARC T8-1、T8-2、T8-4 サーバーでは、SAS ベースの HDD と SSD にも対応可能な選択したドライブ・ベイに内蔵 2.5 インチ・スモール・フォーム・ファクタ (SFF) NVMe SSD を搭載できます。SFF NVMe ドライブを使用する場合には、工場出荷時構成の NVMe PCIe スイッチ・カードとケーブルが必要です。このスイッチ・カードでは x8 PCIe 3.0 インタフェースを使用し、最大 4 つの x4 ダウンストリーム・リンク (NVMe ドライブあたり 1 つ) に対してファンアウト機能と電気的なリタイミング機能を提供します。

NVMe デバイスはホットプラグ対応ですが、OS 固有のホットプラグ手順に従う必要があります。管理者は `nvmeadm` コマンドを使用して、ドライブ状態とファームウェア・レベルの表示、温度のチェック、エラー・ログの取得、SMART データへのアクセスを行うとともに、Secure Erase (完全消去) やローレベル・フォーマットを実行できます。

内蔵USBストレージおよびOracle Solarisブート・プール

SPARC M7 および SPARC M8 プロセッサベースのサーバーは、より多様なデバイスから起動可能な新しいブート・プロセスに対応しています。従来のブート・プロセスでは、システム・ファームウェアからブート・デバイスにアクセス可能になっている必要があります。この理由により、たとえば InfiniBand 経由のネットワーク・ブートはこれまでサポートされていませんでした。

新しい Oracle Solaris のブート・プロセスには、ブート・プールという概念が取り入れられています。ブート・プールは、ブート・アーカイブの保存に使用されるブート・デバイスです。SPARC M8 プロセッサベースのサーバーでは、1 つ以上の内蔵 USB (eUSB) ストレージ・デバイスをグループ化してブート・プールを形成し、OpenBoot PROM ファームウェアからブート・プールにアクセスできるようにしています。eUSB ストレージは、工場出荷時にシステムに組み込まれる内蔵 USB フラッシュ・メモリ・デバイスです。SPARC T8-1 サーバー、T8-2 サーバー、T8-4 サーバーには、eUSB デバイスが 1 つ組み込まれています。SPARC M8-8 サーバーには、CPU、メモリ、I/O ユニット (CMIOU) シャーシ・ボードごとに 1 つの eUSB デバイスが組み込まれているため、ブート・プールを複数のストライプ化された eUSB ストレージ・デバイスで構成することができます。

OpenBoot PROM ファームウェアでは、ローカル・ブート・プールが存在するためにブート・アーカイブをロードし、それに続いて iSCSI over IP over InfiniBand (IPoIB) を使用してルート・プールに root ファイル・システムをマウントできます。SPARC M8 プロセッサベースのサーバーには、新しいブート・プロセス用のフォールバック・メカニズムも採用されており、eUSB ベースのブート・アーカイブが使用できない場合に使用可能です。ブート・アーカイブはシステム・サービス・プロセッサ (SP) のフラッシュ・メモリ内にあります。このブート・アーカイブは工場出荷時に SP にロードされ、他のブート方法を使用できない場合にのみ使用することを意図されています。

PCIeアダプタ・カード

SPARC M8 プロセッサベースのサーバーは、PCIe 3.0 x8 と x16 の両方の拡張カード・スロットを搭載しています。サポートされているオプションと要件はサーバー・モデルによって異なります。リリース時にオラクルから入手可能なアダプタ・カードは次のとおりです。

- » Oracle Flash Accelerator F640 PCIe カード : 6.4TB、NVMe PCIe 3.0
- » Oracle Storage Dual Port 16GB または 32GB Fibre Channel PCIe HBA
- » Sun Storage 16GB FC PCIe Universal HBA
- » Oracle Quad Port 10GBase-T Adapter
- » Sun Quad Port GbE PCIe 2.0 ロープロファイル・アダプタ、UTP
- » Oracle Quad 10GB または Dual 40GB Ethernet Adapter
- » Sun Dual 10GbE SFP+ PCIe 2.0 ロープロファイル・アダプタ
- » Oracle Storage 12 GB SAS PCIe HBA、外部 : 8 ポート
- » Oracle Dual Port QDR InfiniBand Adapter M3

SPARC T8-1、T8-2、および T8-4サーバー

SPARC T8-1、T8-2、および T8-4 サーバーは、信頼性を最大限に高め、電力消費と複雑さを最小限に抑えつつ、画期的なセキュリティとパフォーマンスを発揮するように設計されています。これらのシステムは、システム間のフェイルオーバーに対応するアプリケーションによってシステムを複製することで高可用性を実現するスケールアウト・アプリケーションに最適です。共通の共有コンポーネントとサブシステムを使用することにより、SPARC T8-1、SPARC T8-2、SPARC T8-4 の各サーバーでは、個別の設計を採用してシステムをその特定の設計ポイントと機能に応じて最適化しています。これらのサーバーの特長は、シャーシ、コンポーネント、サブアセンブリの堅牢な設計、機能強化されたシステムとコンポーネントの保守性、最小限のケーブル配線によるエアフローの最大化です。

- » SPARC T8-1 サーバーは、シングル・プロセッサのエントリー・モデルですが、256 のハードウェア・スレッドと 1TB のメモリにより、従来のデュアルプロセッサ搭載サーバーよりも優れた性能を発揮します。SPARC T8-1 サーバーは、このサーバー・ファミリーの全モデルと同様に、Silicon Secured Memory によるセキュアな処理、インメモリ・データベース・クエリの高速化とインライン圧縮解凍による高パフォーマンスなど、完全な Software in Silicon 機能を備えています。

- » SPARC T8-2 サーバーでは、SPARC T8-1 サーバーに比べて、2 倍の容量と多くのリソースを使用できます。デュアルプロセッサ SPARC T8-2 サーバーには、プロセッサと I/O コントローラ ASIC 間の革新的なクロス接続、統合された SAS デュアル・コントローラ、オプションのデュアル NVMe PCIe スイッチ（内部 NVMe SSD 用）などの可用性を高める機能も追加されています。
- » SPARC T8-4 サーバーには、フロントアクセス可能なプロセッサ・モジュールに最大 4 基のプロセッサを搭載できます。2 基のプロセッサによるエントリ構成は、必要に応じて容易に 4 基のプロセッサに拡張できます。SPARC T8-4 サーバーには、多数の専用 PCIe ルート・コンプレックスでサポートされる個別にホットプラグ可能な PCIe カードも搭載されており、仮想化プライベート・クラウドにワークロードを統合するのに最適なサーバーとなっています。

SPARC T8-1 サーバー

エントリ・モデルの SPARC T8-1 サーバーの特長は、2U エンクロージャ内に 1 基の SPARC M8 プロセッサを搭載していることです。標準装備として 8 個のオンボード・メモリ DIMM スロットが搭載されており、オプションのデュアル・メザニン・カード（メモリ・ライザー）を使用して合計 16 個まで拡張すれば、システム・メモリの容量を最大 1TB まで増やすことができます。図 7 に、SPARC T8-1 サーバーの前面図と背面図を示します。SPARC T8-1 サーバーは、システム背面からアクセス可能なロープロファイル PCIe 3.0 拡張スロットを 6 個搭載しています。

8 つの 2.5 インチ・スモール・フォーム・ファクタ（SFF）フロントローディング・ドライブ・ベイすべてが 1 つのオンボード 12GB/秒 SAS HBA に搭載されており、これにより RAID 0/1/10/1E による保護を実現できます。Oracle Solaris ZFS では、より高いレベルの RAID を使用できます。最大 4 台の 2.5 インチ SFF NVMe SSD をサポートする工場出荷時構成オプションも用意されています。このオプションには、3 番の PCIe スロットに取り付けられる NVMe PCIe スイッチ・カードとドライブ・ケージへの内部配線が含まれます。SAS ドライブと NVMe ドライブは混在させることができます。SPARC T8-1 サーバーには他にも以下の標準機能があります。

- » 4 つの 10GBase-T Ethernet ポート（RJ45）をシステム背面に装備（オンボード・ネットワーク・インタフェース・コントロール（NIC）経由）全二重の場合にのみ、100MB/秒、1GB/秒、10GB/秒の速度に自動ネゴシエート最大 15,500 バイトのジャンボ・フレームに対応可能
- » 2 台のホットスワップ対応 1+1 電源装置（PSU）（200～240VAC で 1200W）をシステム背面から挿入
- » それぞれ二重反転ファンを備えた 4 個のトップローディング・ホットスワップ対応ファン・モジュールを搭載
- » 前面に USB 2.0 ポート 2 つと、背面に USB 3.0 ポート 2 つと VGA ビデオ・ポート（HD15）1 つを装備
- » USB ポート経由で接続する外付け DVD ドライブを使用可能

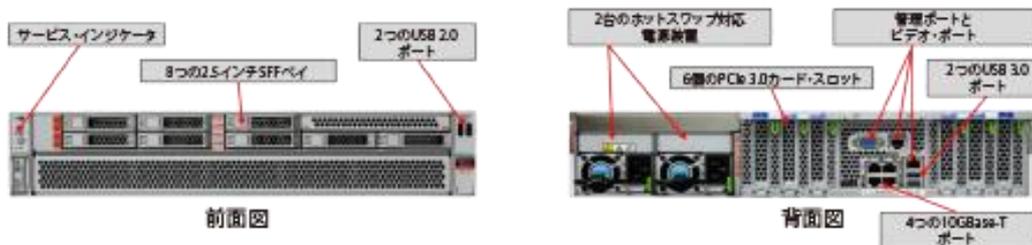


図7：SPARC T8-1 サーバーの前面図と背面図

図 8 に、SPARC T8-1 サーバーのブロック図を示します。SPARC M8 プロセッサは、オンボード・メモリ・スロットとともにオプションのメモリ・ライザーに接続して、最大 16 個の DDR4 DIMM スロットをサポートします。SPARC M8 プロセッサは、PCIe ルート・コンプレックスを実装した I/O コントローラ ASIC に直接接続します。オプションの NVMe PCIe スイッチ・カードにより、最大 4 台の 2.5 インチ NVMe デバイスを中央の 4 つの SFF ドライブ・ベイに配置できます。

サービス・プロセッサ (SP) 上で動作する Oracle Integrated Lights Out Manager (Oracle ILOM) リモート・コンソールにより、リモート・キーボード、ビデオ、マウス (rKVM) の機能が提供されます。SP と通信するためのインターフェースとして、別個のシリアル (RJ45) および Ethernet (1000Base-T、RJ45) 管理ポートを装備しています。また、オンボード 10GbE-T ネットワーク・ポートを (サイドバンド機能を有効にした状態で) 使用して、最大 10GB/秒の速度で SP に接続することもできます。6 個の PCIe スロットのうち 2 個 (3 番と 4 番) は、隣接するスロット (それぞれ 2 番と 5 番) が空のときには x16 接続が可能です。

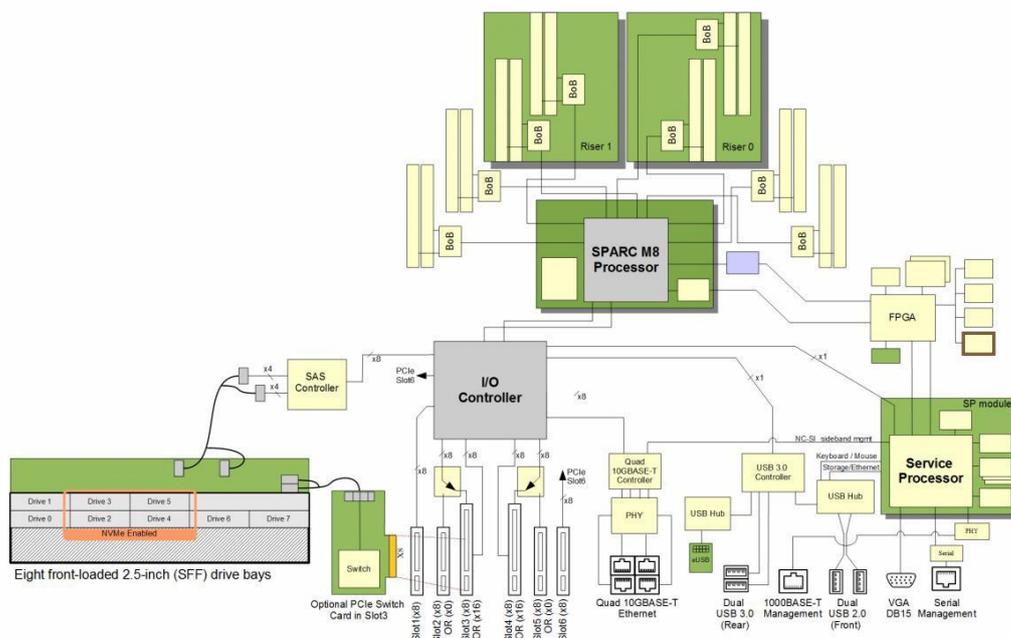


図8：SPARC T8-1 サーバーの特長であるシングル 32コア SPARC M8プロセッサおよび I/OコントローラASIC

表3に、PCIeデバイスが共有する、SPARC T8-1 サーバーの I/Oコントローラ ASIC上の5つのルート・コンプレックスの対応関係を示します。

表3：SPARC T8-1 サーバーのルート・コンプレックス対応表

ルート・コンプレックス	ターゲット	速度
ルート・コンプレックス0	PCIeスロット6	x8
	SASコントローラ	x8
ルート・コンプレックス1	PCIeスロット4	x8またはx16
	PCIeスロット5	x8またはx0
ルート・コンプレックス2	PCIeスロット1	x8
	クアド10GbEコントローラ	x4
ルート・コンプレックス3	PCIeスロット2	x8またはx0
	PCIeスロット3	x8またはx16
ルート・コンプレックス4	USBコントローラ	x4ポートのx1
	サービス・プロセッサ (オンボード・グラフィックス)	x4ポートのx1

SPARC T8-2 サーバー

SPARC T8-2 サーバーの特長は、3U ラックマウント型エンクロージャ内に 2 基の SPARC M8 プロセッサを搭載していることです。それぞれのライザー・カードに 2 枚または 4 枚のメモリ DIMM が装着されており、8 枚のライザー・カードに最大で 2TB のシステム・メモリをマウントできます。図 9 に、SPARC T8-2 サーバーの前面図と背面図を示します。SPARC T8-2 サーバーは、システム背面からアクセス可能なロープロファイル PCIe 3.0 拡張スロットを 8 個搭載しています。

6 つの 2.5 インチ SFF フロントローディング・ドライブ・ベイが 2 つのオンボード 12GB/秒 SAS HBA に搭載されており (2 つの HBA で 2 台と 4 台のドライブを分担)、これにより RAID 0/1 で保護できます。ドライブ・ベイを 4 つ搭載した SAS HBA は、RAID 10/1E にも対応できます。Oracle Solaris ZFS では、より高いレベルの RAID を使用できます。最大 4 台の 2.5 インチ SFF NVMe ドライブを搭載可能な工場出荷時構成オプションも用意されています。それらのオプションには、ドライブ・ケージに内部配線された NVMe PCIe スイッチ・カードが含まれます。次のオプションを選択すると、SAS ドライブと NVMe ドライブを混在させることができます。

- » **シングル NVMe PCIe スイッチ・カード。** このスイッチ・カードは 1 番の PCIe スロットに装着され、4 つすべての NVMe 対応ドライブ・ベイ (上部 4 つのドライブ・ベイ) をサポートします。
- » **デュアル NVMe PCIe スイッチ・カード。** スイッチ・カードが 1 番と 2 番の PCIe スロット (別個のルート・コンプレックス) に取り付けられ、各スイッチで 2 つの NVMe 対応ドライブ・ベイをサポートします。

SPARC T8-2 サーバーには他にも以下の標準機能があります。

- » 4 つの 10GBase-T Ethernet ポート (RJ45) をシステム背面に搭載 (オンボード NIC 経由) 全二重の場合にのみ、100MB/秒、1GB/秒、10GB/秒の速度に自動ネゴシエート最大 15,500 バイトのジャンボ・フレームに対応可能
- » 2 台のホットスワップ対応 1+1 電源装置 (PSU) (200~240VAC で 2000W) をシステム背面から挿入

- » それぞれ二重反転ファンを備えたシャーシ上面から取り付ける 3 個のホットスワップ対応ファン・モジュール
- » 前面に USB 2.0 ポート 2 つと、背面に USB 3.0 ポート 2 つ装備
- » 1 つの VGA ビデオ・ポート (HD15) をシャーシ背面に装備
- » USB ポート経由で接続する外付け DVD ドライブを使用可能

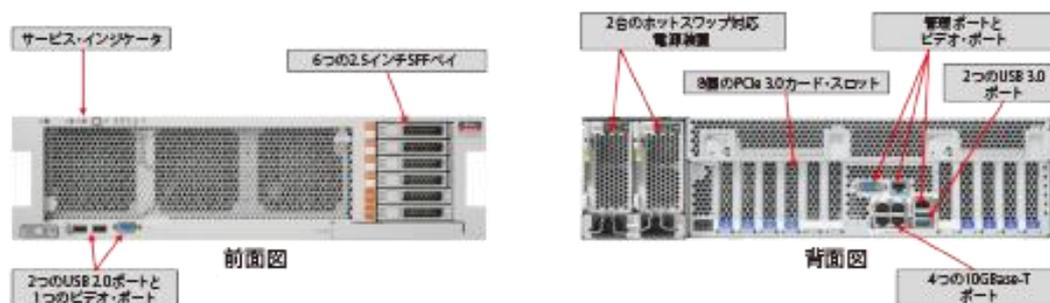


図9：SPARC T8-2 サーバーの前面図と背面図

図 10 に、SPARC T8-2 サーバーのブロック図を示します。2 つのソケットの SPARC M8 プロセッサがメモリ・ライザーに接続し、最大で 32 個の DDR4 DIMM スロット (プロセッサ・ソケットあたり 16 個) をサポートします。プロセッサ同士は冗長コヒーレンシ・リンクによって連結されており、2 つの I/O コントローラが両方の SPARC M8 プロセッサに接続されています。このクロス接続設計により、プロセッサで障害が発生した場合でも、すべての I/O デバイスへのアクセスが維持されます。システムがプロセッサ 1 基のみで起動した場合でも、PCIe デバイス・パスと 10 個すべての PCIe ルート・コンプレックスは影響を受けません。6 個の PCIe スロットのうち 4 個は x16 対応であり、4 個は x8 で配線されています。x16 対応 PCIe スロットには、高スループットのデバイスに必要な不可欠な専用のルート・コンプレックスがあり、他のデバイスがそれらのルート・コンプレックスを共有することはありません。

サービス・プロセッサ (SP) 上で動作する Oracle ILOM リモート・コンソールにより、リモート・キーボード、ビデオ、マウス (rKVM) の機能が提供されます。SP と通信するためのインタフェースとして、別個のシリアル (RJ45) および Ethernet (1000Base-T, RJ45) 管理ポートを装備しています。また、オンボード 10GBase-T ネットワーク・ポートを (サイドバンド機能を有効にした状態で) 使用して、最大 10GB/秒の速度で SP に接続することもできます。

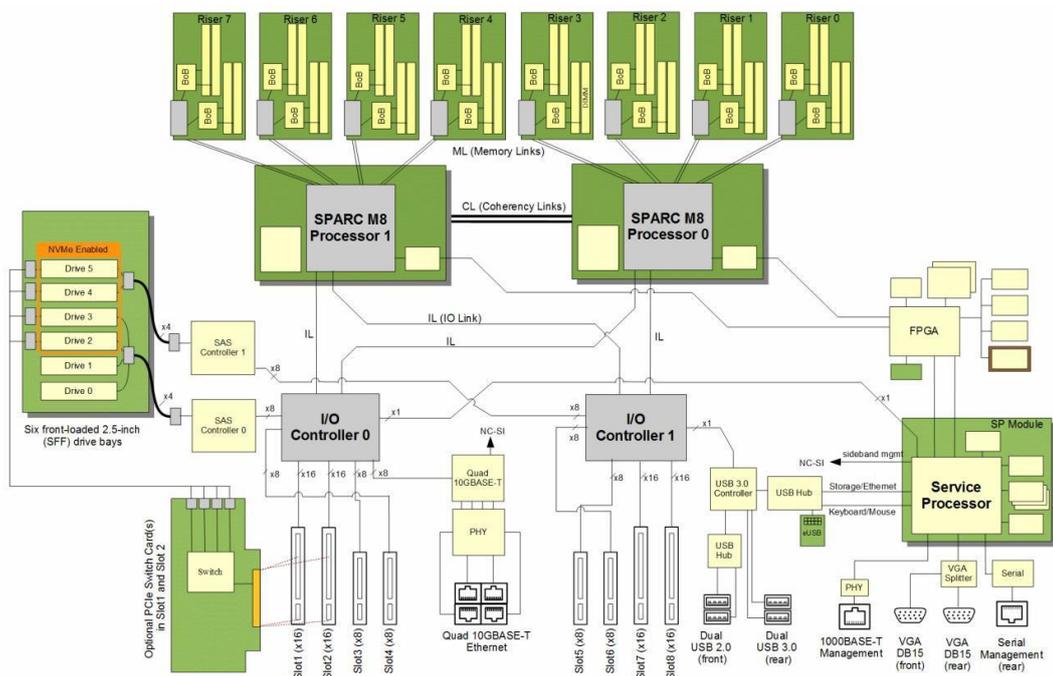


図10：2つの I/OコントローラASICにクロス接続して高い可用性を実現する SPARC T8-2 サーバーのデュアル SPARC M8プロセッサ

表 4 に、PCIe デバイスが共有する各 I/O コントローラ上のルート・コンプレックスの対応関係を示します。

表4：SPARC T8-2サーバーのI/Oコントローラとルート・コンプレックスの対応表

I/Oコントローラ	ルート・コンプレックス	ターゲット	PCIe速度
I/Oコントローラ0	ルート・コンプレックス0	PCIeスロット4 SASコントローラ0	x8 x8
	ルート・コンプレックス1	PCIeスロット3 クアド10GbEコントローラ	x8 x8
	ルート・コンプレックス2	PCIeスロット2	x16
	ルート・コンプレックス3	PCIeスロット1	x16
	ルート・コンプレックス4	サービス・プロセッサ (オンボード・グラフィックス)	x4ポートの x1
I/Oコントローラ1	ルート・コンプレックス0	PCIeスロット6 SASコントローラ1	x8 x8
	ルート・コンプレックス1	PCIeスロット8	x16
	ルート・コンプレックス2	PCIeスロット7	x16
	ルート・コンプレックス3	PCIeスロット5	x8
	ルート・コンプレックス4	USBコントローラ	x4ポートの x1

SPARC T8-4 サーバー

SPARC T8-4 サーバーの特長は、6U ラックマウント型エンクロージャ内に 4 基の SPARC M8 プロセッサと最大 4TB のメモリを搭載していることです。図 11 に、SPARC T8-4 サーバーの前面図と背面図を示します。このサーバーのホットプラグ対応 PCIe キャリアには、x8 および x16 の PCIe 3.0 拡張スロットがそれぞれ 8 個ずつ搭載されており、これらのスロットにはシステムの背面からアクセス可能です。

8つの2.5インチSFFフロントローディング・ドライブ・ベイは2つのオンボード12GB/秒SAS HBAで均等に分担され、RAID 0/1/10/1Eで保護できます。Oracle Solaris ZFSでは、より高いレベルのRAIDを使用できます。SAS HDDとSAS SSDに加えて、最大8台の2.5インチSFF NVMe SDDドライブを搭載可能な工場出荷時構成オプションが用意されています。それらのオプションには、ドライブ・ケージに内部配線された最大2枚のNVMe PCIe スイッチ・カードが含まれます。各NVMe PCIe スイッチ・カードにより、4つのドライブ・ベイにNVMeドライブを搭載できます。SASドライブとNVMeドライブは混在させることができます。

SPARC T8-4サーバーには他にも以下の標準機能があります。

- » 4つの10GBase-T Ethernetポート(RJ45)をシステム背面に搭載(オンボードNIC経由)全二重の場合のみ、100MB/秒、1GB/秒、10GB/秒の速度に自動ネゴシエート最大15,500バイトのジャンボ・フレームに対応可能
- » 以下のPCIe ロープロファイル・ホットプラグ対応キャリア I/O スロット16個を4つのI/Oコントローラ ASICでサポート
 - » PCIe 3.0 x8 スロット8個
 - » PCIe 3.0 x16 スロット8個
- » 4台のホットスワップ対応N+N冗長PSU(200~240VACで3000W)をシステム正面から挿入
- » シャーシ上面から5個のホットスワップ対応ファン・モジュールを搭載
- » ローカル・キーボード(前面USB、背面USB、ビデオ)、リモート・キーボード、ビデオ、マウス(KVM)の機能を提供
- » オンボード・サービス・プロセッサと通信するためのインターフェースとして、別々のシリアル・ポートとネットワーク管理ポートを装備
- » USBポート経由で接続する外付けDVDドライブを使用可能

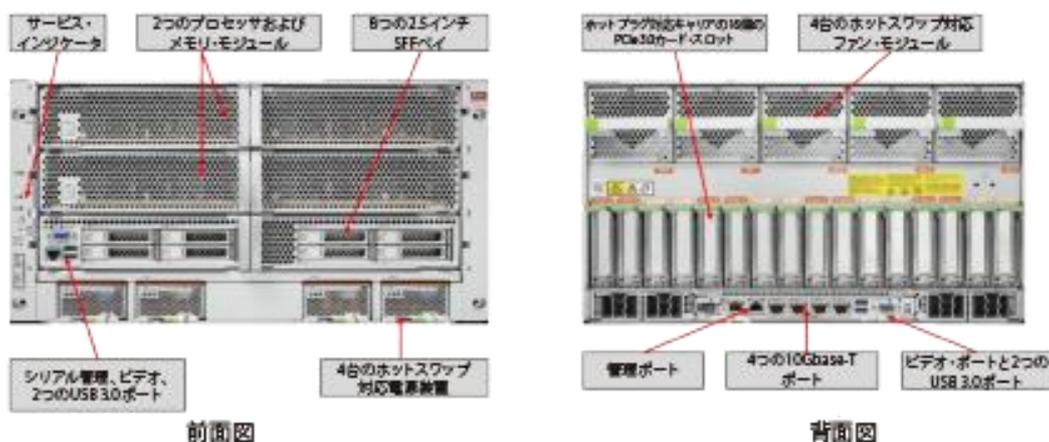


図11 : SPARC T8-4 サーバーの前面図と背面図

図 12 に、SPARC T8-4 サーバーのブロックレベル図を示します。最大 2 つの専用プロセッサ/メモリ・モジュールそれぞれに、2 つの SPARC M8 ソケットと 32 個の DDR4 DIMM スロット（プロセッサ・ソケットあたり 16 個の DIMM スロット、サーバーあたり 64 個の DIMM スロット）が搭載されています。各プロセッサ・ソケットのコヒーレンシ・リンクは、他方のプロセッサ/メモリ・モジュールのプロセッサ・ソケットにミッドプレーンで冗長接続されています。冗長 IL により、各プロセッサ・ソケットが 4 つある I/O コントローラの 2 つに接続されています。4 つの I/O コントローラには、システム内のすべての PCIe スロットの PCIe ルート・コンプレックスが実装されており、大きな I/O 能力を発揮します。SAS HDD、SAS SSD、または NVMe SSD は、すべてのドライブ・ベイに取付け可能です。オプションの NVMe PCIe スイッチ・カードがシステムのシャーシ背面のホットプラグ対応 PCIe キャリア・スロットを使用することはないため、追加の I/O デバイスのためにより多くのスロットを残しておくことができます。

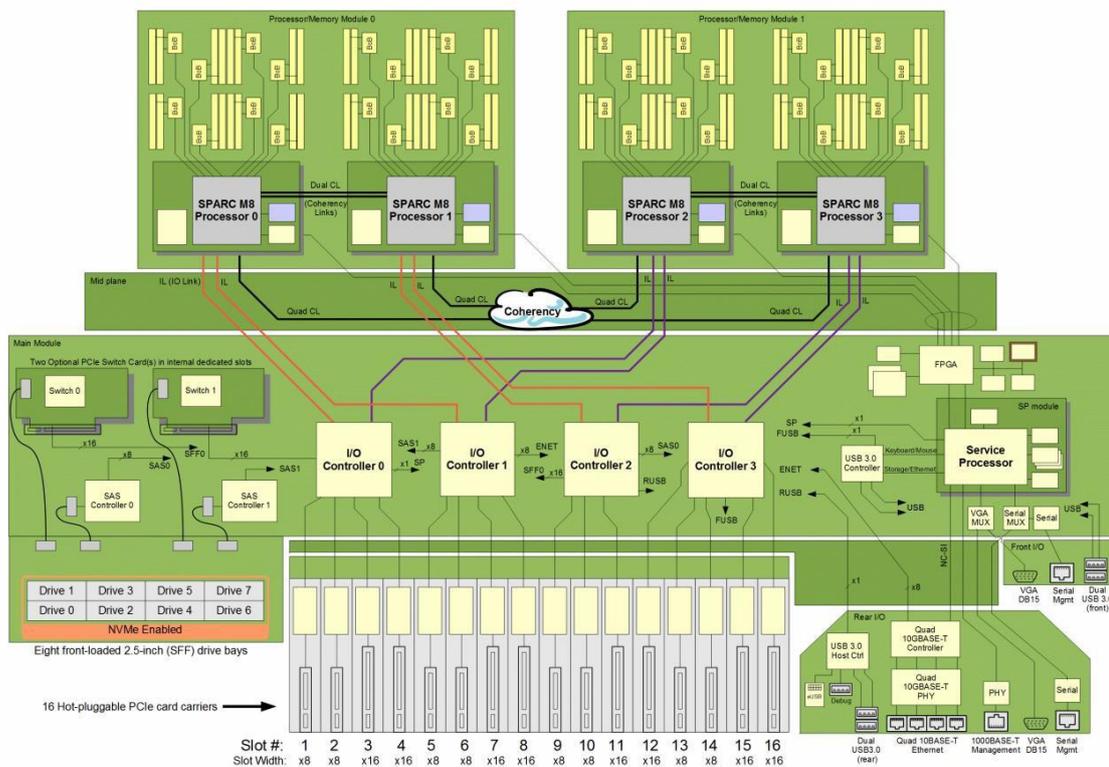


図12：SPARC T8-4 サーバーの特長である 2つのデュアルプロセッサ/メモリ・モジュールと 4つの I/OコントローラASIC

表 5 に、PCIe デバイスで共有する 4 つの I/O コントローラ上のルート・コンプレックスの対応関係を示します。

表5：SPARC T8-4 サーバーのI/Oコントローラとルート・コンプレックスの対応表

I/Oコントローラ	ルート・コンプレックス	ターゲット	PCIe速度
I/Oコントローラ0	ルート・コンプレックス0	PCIeスロット1 PCIeスロット2	x8 x8
	ルート・コンプレックス1	NVMe PCIeスイッチ・カード1	x16
	ルート・コンプレックス2	PCIeスロット4	x16
	ルート・コンプレックス3	PCIeスロット3	x16
	ルート・コンプレックス4	サービス・プロセッサ (オンボード・グラフィックス)	x4ポートの x1
I/Oコントローラ1	ルート・コンプレックス0	PCIeスロット5 PCIeスロット6	x8 x8
	ルート・コンプレックス1	SASコントローラ1	x16
	ルート・コンプレックス2	PCIeスロット8	x16
	ルート・コンプレックス3	PCIeスロット7	x16
	ルート・コンプレックス4	クアド10GbEコントローラ	x8
I/Oコントローラ2	ルート・コンプレックス0	PCIeスロット9 PCIeスロット10	x8 x8
	ルート・コンプレックス1	SASコントローラ0	x8
	ルート・コンプレックス2	NVMe PCIeスイッチ・カード0	x16
	ルート・コンプレックス3	PCIeスロット11	x16
	ルート・コンプレックス4	背面USB 3.0コントローラ	x4ポートの x1
I/Oコントローラ3	ルート・コンプレックス0	PCIeスロット13 PCIeスロット14	x8 x8
	ルート・コンプレックス1	PCIeスロット12	x16
	ルート・コンプレックス2	PCIeスロット16	x16
	ルート・コンプレックス3	PCIeスロット15	x16
	ルート・コンプレックス4	前面USB 3.0コントローラ	x4ポートの x1

SPARC M8-8 サーバー

SPARC M8-8 サーバーは、最新のクラウド・インフラストラクチャ用に設計されています。ミッションクリティカルな大規模コンピューティング環境にふさわしい運用効率、信頼性、スケーラビリティを必要とするデータベースや商用アプリケーションに最適です。SPARC M8-8 サーバーは、メモリ・フットプリントが大きくてメモリ帯域幅が際立って広いため、データベースとエンタープライズ・アプリケーションをメモリ内にすべてデプロイして卓越したパフォーマンスと容量を達成しようとする場合に最適です。

SPARC M8-8 サーバーは、革新的な CPU、メモリ、および I/O ユニット (CMIOU) シャーシとシステム・インターコネクトをベースにした、より高いスケーラビリティ、データセンター機能、および機能性を提供することで、要求の厳しいビジネスおよびデータセンター・アプリケーションに最適なサーバーとなっています。このサーバーでは、信頼性と可用性を確保するための幅広い機能を提供しながら、効果的に拡張して高性能なプロセッサと大容量のメモリをサポートすることができます。SPARC M8-8 サーバーは、2 つの別々の物理ドメイン (PDom) にパーティション化するオプションを選択できるように特別に設計されています。このサーバーは 2 種類の工場出荷時構成で提供され、1 つまたは 2 つの PDom のいずれかの構成で注文できます。

サーバー・コンポーネント

以降のセクションでは、SPARC M8-8 サーバーの主要コンポーネントの概要を示します。

CPU、メモリ、I/Oユニット・シャーシ

SPARC M8-8 サーバーは、1 つの CMIOU シャーシ内に収容されています。このエンクロージャには、プロセッサ・メモリ・ボード、サーバー・プロセッサ (SP)、およびインターコネクト・アセンブリが収容されています。また、電源と冷却ファンも内蔵しています。図 13 に、CMIOU シャーシとその主要コンポーネントの前面図と背面図を示します。

シャーシ前面には、ホットスワップ対応ファン・モジュールと電源装置、および CMIOU ボードをまとめてシステムに接続するインターコネクト・アセンブリが配置されています。シャーシ前面の機能は次のとおりです。

- » ホットスワップ対応ファン・モジュール 8 個
- » ホットスワップ対応 N+N 冗長 PSU (200~240VAC で 3,000W) 6 台

CMIOU シャーシの背面には以下が搭載されています。

- » 最大 8 枚の CMIOU ボード
- » 冗長 SP 2 基

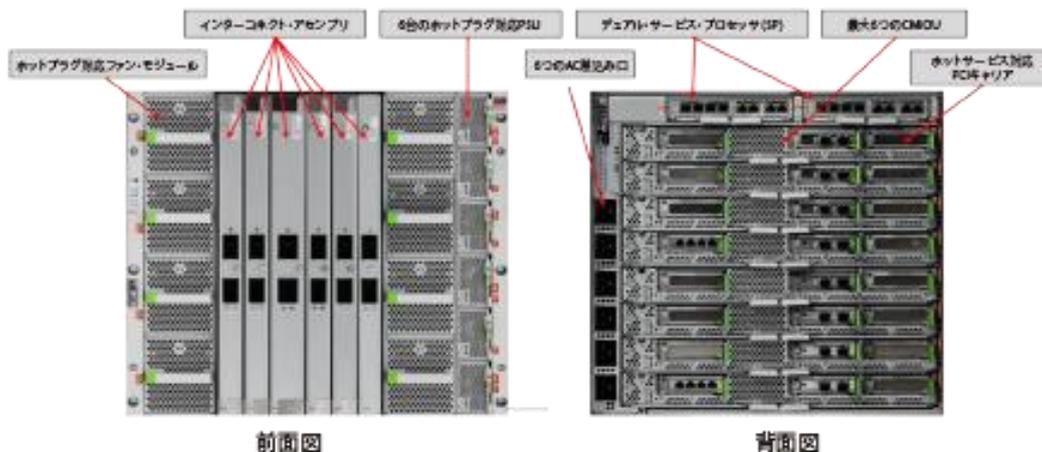


図13：SPARC M8-8 の CMIOUシャーシの前面図と背面図

GPU、メモリ、I/Oユニット・ボード

各 CMIOU ボード・アセンブリには、メザニン・ボードに搭載されている 1 基の SPARC M8 プロセッサと関連メモリ、および I/O コンポーネントが搭載されています。16 枚のメモリ DIMM スロットは、すべてボード上に配置されています。I/O コントローラ ASIC には、3 個の PCIe 3.0 (x16) スロット用の専用ルート・コンプレックスが実装されています。PCIe ホットプラグ対応キャリアはボードに組み込まれています。図 14 に CMIOU ボードのブロック図を、図 15 に上から見た写真を示します。CMIOU シャーシに挿入されると、CMIOU ボードがインターコネクティブ・アセンブリと接続され、これにより CMIOU ボードと SP 間の接続が確立されます。

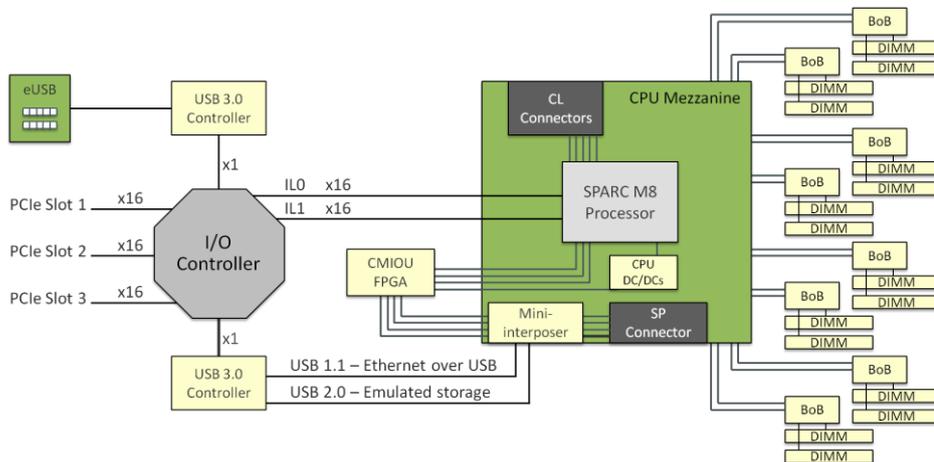


図14：単一 SPARC M8-8 CMIOUボードのアーキテクチャ

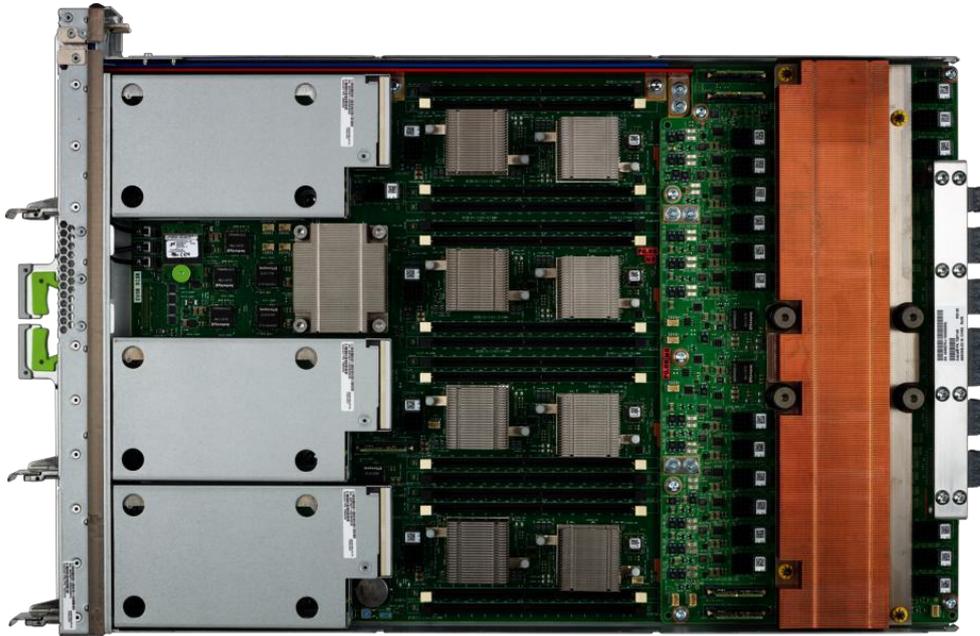


図15：各CMIUボードには、1基の SPARC M8プロセッサ、16個の DDR4 DIMMスロット、1基の I/OコントローラASIC、3つの PCIe x16ホットプラグ対応キャリアが搭載される

インターコネクト・アセンブリ

SPARC M8-8 サーバーのシステム・インターコネクトは、インターコネクト・アセンブリに実装されます。インターコネクト・アセンブリは、CMIU シャーシの 10U スペース内に収容されています。5 つのインターコネクト・アセンブリにより、CMIU ボード間がコヒーレンシ・リンク (CL) で接続されます (グルーレス・システム・インターコネクト)。各 CMIU ボードとシャーシ上部のデュアル SP 間の通信には、1 つのインターコネクト・アセンブリが使用されます。電気的に分離された 1 つまたは 2 つの PDom に対応するため、インターコネクト・アセンブリの配線方法は 2 種類の SPARC M8-8 サーバーで異なります。

サービス・プロセッサとサービス・プロセッサ・モジュール

SPARC M8-8 サーバーの特長は、ホットプラグ対応の冗長サービス・プロセッサ (SP) を搭載していることです。CMIU シャーシには、SP が 2 基配置されています。各 SP には 1 つまたは 2 つのシリアル (RJ45) ポートと 1 つの 1000Base-T 管理ポートがあります (10/100/1000MB/秒に自動ネゴシエート)。SP は、SP インターコネクト・アセンブリ経由で CMIU ボードと通信します。

各 SP には 1 つまたは 2 つのサービス・プロセッサ・モジュール (SPM) が組み込まれています。SPM は、Oracle ILOM ソフトウェアを実行するコンポーネントであり、これによりサーバー・システムで SP 機能を使用できます。冗長 SP 機能とフェイルオーバー機能を常時利用できるようにするため、一部の構成では各 SP にデュアル SPM が搭載されています。以下に、サーバー・モデルに応じた各種構成を示します。

- » 1つの PDom を使用する SPARC M8-8 サーバーでは、2つの SP それぞれに SPM が1つだけ組み込まれています。これは、冗長化するには十分です。一方の SP はプラットフォームを管理するアクティブ SP として機能し、もう一方の SP は障害発生時にアクティブ SP の役割を引き継ぐスタンバイ SP として機能します。
- » 2つの PDom を使用する SPARC M8-8 サーバーでは、SP それぞれに2つの SPM が組み込まれ、サーバーは2組の SPM を使用してシステム内の2つの PDom を管理します。2つの SPM はアクティブな SP として動作し、残りの2つはスタンバイ SP として動作します。アクティブ SPM とスタンバイ SPM は別々の SP 内に配置されています。

システム・ラックと配電ユニット

SPARC M8-8 サーバーは通常、工場出荷時に Oracle Rack Cabinet 1242 に収容されます。ラックには三相デュアルグリッド電源用の配電ユニット (PDU) が2台搭載されています。この PDU の電源は、CMIU シャーシと同じラック内の他の機器に電力を供給する単相出力です。PDU は、シリアル (RJ45) ポートと 100Base-T (RJ45) 管理ポートを1つずつ搭載した電力監視モジュールを内蔵しています。

オラクルでは、工場出荷時にラックに収容された状態の SPARC M8-8 サーバーを注文することをお勧めします。ただし、このシステムをスタンドアロン・システムとして注文するオプションもあり、その場合は現場でラックに収容します。スタンドアロン型 SPARC M8-8 サーバー・エンクロージャの高さは 10U です。さらに、ラック下部または上部の 3U ユニットが三相電源ケーブル用に取り分けられています。PDU はキャビネット内に縦に取り付けられ、ユニット・スペースを使用しません。1台の SPARC M8-8 サーバーを Oracle Rack Cabinet 1242 に収容した場合は、29のユニット・スペースを他のデバイス用に使用できます。

SPARC M8-8サーバーのアーキテクチャ

両方のタイプの SPARC M8-8 サーバーに、合計で最大 8 基のプロセッサを搭載できます。シングル PDom の SPARC M8-8 サーバーは、1つの対称型マルチプロセッシング (SMP) サーバーとして構成されます。デュアル PDom 構成の SPARC M8-8 サーバーは、電気的に分離した2つのハードウェア・パーティションで構成されます。SPARC M8-8 サーバー・ハードウェアはすべて1つの CMIU シャーシ内に収容されるため、デュアル PDU とともに工場出荷時にラックに収容することをお勧めします。2種類の SPARC M8-8 サーバーのほとんどのハードウェア・コンポーネントは同一です。ただし、インターコネクト・アセンブリの配線は、最大 8 ウェイの1つの PDom または分離した 4 ウェイの2つの PDom のどちらをサポートするかによって異なります。この配線は工場出荷時に設定され、後で変更することはできません。また、サービス・プロセッサのハードウェアは、冗長化された PDom の数によって異なります。

表 6 に、SPARC M8-8 サーバーのおもな特性を示します。

表6：SPARC M8-8サーバーの機能

機能	説明
最大プロセッサ数	8 または 2x4
最大スレッド数	2,048 または 2x1,024
最大メモリ容量	8TB または 2x4TB (64GB DIMMIに基づく)
最大PCIe 3.0 (x16) スロット数	24 または 2x12
サービス・プロセッサ (SP)	シングルPDom：ホットプラグ対応SPx2。それぞれに SPモジュール (SPM) を1つ搭載して、1つの PDomで冗長化 デュアルPDom：ホットプラグ対応SPx2。それぞれにデュアルSPMを搭載して、2つの分離した PDomで冗長化
ファン	ホットスワップ対応N+1ファン・モジュールx8、二重反転ファン搭載、前面搭載
電源	ホットスワップ対応N+N冗長電源装置x6 (それぞれ3,000W、200~240VAC)、前面搭載
PDU	システム・ラック内に2台のPDU。デュアルグリッド構成の場合、PDUごとに3本の三相電源コード。PDUからシステム・シャーシに接続する単相電源コード6本

SPARC M8-8 サーバー (シングル物理ドメイン)

シングル PDom が設定された SPARC M8-8 サーバーは、1つの CMIU シャーシに収容され、最大 8 基の SPARC M8 プロセッサを搭載したシングル PDom として構成されます。SPARC M8-8 サーバー内の各 SP には、1組の SPM だけで SP 機能を冗長化できるため、サービス・プロセッサ・モジュール (SPM) が 1 つ組み込まれています。デフォルトでは、最初の CMIU ボード (スロット番号 0 に搭載) が SP0/SPMO に接続されてアクティブになります。スロット番号 1 の CMIU ボードは、スタンバイ状態で SP1/SPMO に接続されます。図 16 に、I/O コントローラ ASIC の専用ルート・コンプレックスでサポートされる合計 24 個の PCIe 3.0 (x16) スロットを含む、SPARC M8-8 サーバー I/O デバイスの概略図を示します。

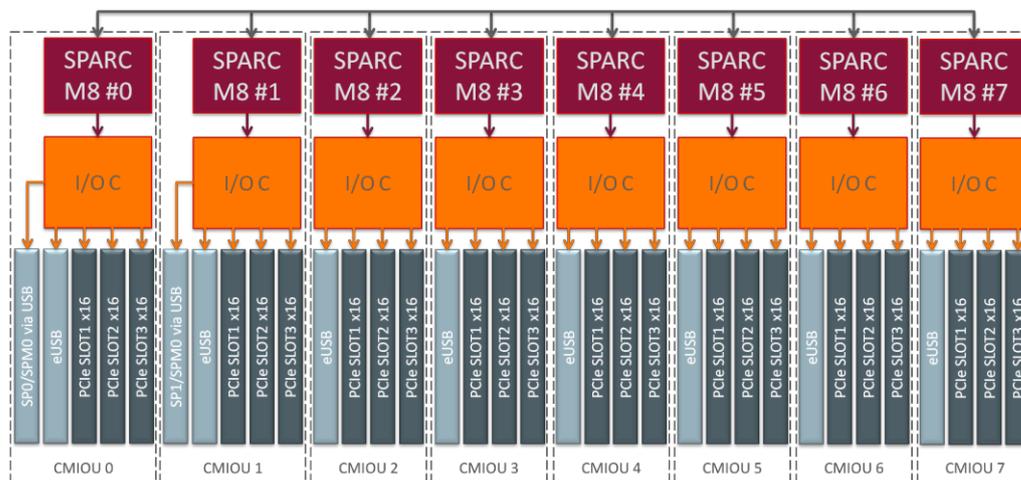


図16：この SPARC M8-8 サーバーには最大8基の SPARC M8プロセッサ を搭載。工場出荷時に シングルPDom で構成

8 ウェイのグルーレス CPU 間ファブリックは、各プロセッサから他の 7 基のプロセッサへの 7 つの x16 コヒーレンシ・リンク (CL) で構成されています (図 17)。

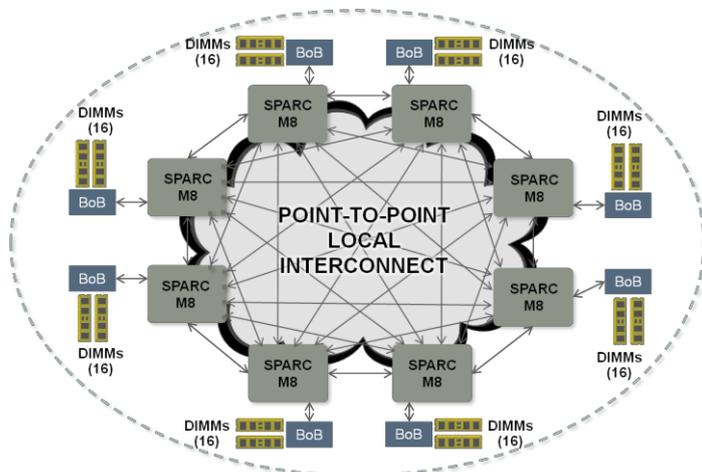


図17: SPARC M8-8 サーバーの 8ウェイ全対全インターコネク

SPARC M8-8 サーバー (デュアル物理ドメイン)

デュアル PDom 構成の SPARC M8-8 サーバーも 1 台の CMIU シャーシに収容されますが、工場出荷時に電氣的に 2 つの PDom にパーティション化されています。インターコネク・アセンブリの配線は、PDom が 1 つだけのタイプの SPARC M8-8 サーバーで使用される配線とは異なります。

2 基の SP それぞれにデュアル SPM が組み込まれて、2 つの PDom で冗長サービス・プロセッサ機能が提供されます。PDom は各サービス・プロセッサ内の 1 つの SPM に関連付けられており、アクティブ/パッシブ・ペアとして動作します。SPM への接続は、各 PDom の最初の 2 つの CMIU ボードから行われます。アクティブな SPM または SP で障害が発生すると、SP のアクティブ/パッシブ機能がスタンバイ接続に切り替えられます。

図 18 に、デュアル PDom 構成の SPARC M8-8 サーバーの PCIe デバイスの概略図を示します。この図は、PDom と SP 間の接続を示しています。デフォルトでは、PDom 0 の場合、スロット番号 0 の CMIU ボードから SP0/SPM0 への接続がアクティブになります。スロット番号 1 の CMIU ボードは、スタンバイ状態で SP1/SPM1 に接続されます。PDom 1 の場合も同様に、スロット番号 4 と 5 の CMIU ボードのそれぞれに、SP0/SPM1 と SP1/SPM1 へのアクティブ/スタンバイ状態の接続があります。このデバイス図には、I/O コントローラ ASIC の専用ルート・コンプレックスでサポートされる PCIe 3.0 スロットへの x16 接続も示されています。

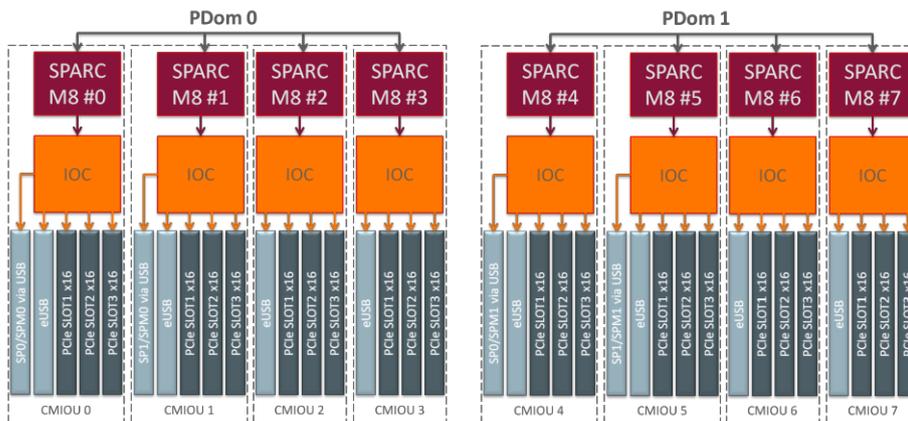


図18：SPARC M8-8 サーバーは、電気的に分離した 2つのPDom が工場出荷時に構成された製品として注文可能

図 19 に、デュアル PDom 構成の SPARC M8-8 サーバーのグルーレス・システム・インターコネクトを示します。このシステム構成には、2つの独立した 4 ウェイ、全対全インターコネクトがあります。4 ウェイの CPU 間ファブリックは、各プロセッサから他のすべてのプロセッサ間を連結する 6 つの x16 コヒーレンシ・リンク (CL) で構成されます。

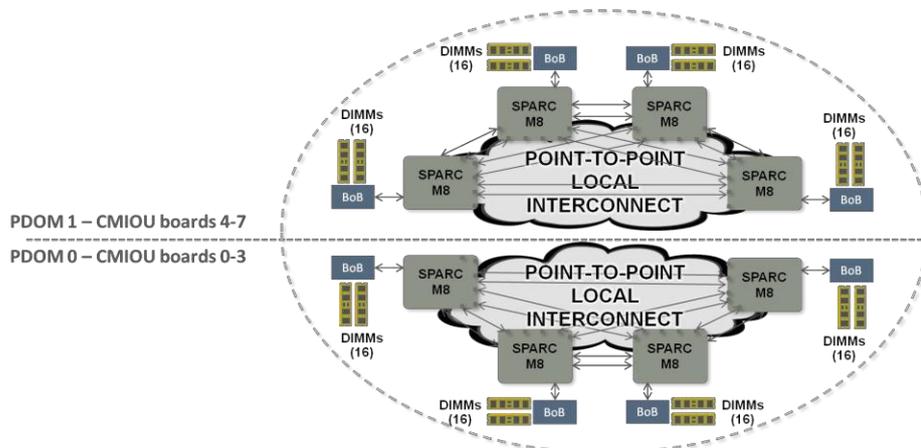


図19：2つの全対全インターコネクトで構成される デュアルPDom 構成の SPARC M8-8 サーバー

Oracle Solaris

Oracle Solaris には、垂直/水平方向に拡張された環境で、仮想化、最適な使用、高可用性、比類ないセキュリティ、および優れたパフォーマンスを実現するための主要な機能が搭載されています。Oracle Solaris は、広範な SPARC および x86 ベースのシステムで動作し、既存のアプリケーションとの互換性が保証されています。Oracle Solaris 11.3 リリースは、先進的な Software in Silicon 機能一式を含む、SPARC M7 および SPARC M8 プロセッサベース・サーバーの大量のリソースを十分に活用するよう特別に設計されています。Oracle Solaris 10 サポートは、サービス・ドメイン（制御ドメイン、ルート・ドメイン、I/O ドメインなど）が提供する仮想ディスク、コンソール、ネットワーク・デバイスだけを使用するゲスト・ドメインでのみ利用できます。ゲスト・ドメインには物理 I/O デバイスがなく、サービス・ドメインに依存しています。表 7 に、SPARC M8 プロセッサベース・サーバーの各種ドメインでサポートされている Oracle Solaris 11 と 10 のリリースを示します。

表7：SPARC M8プロセッサベース・サーバーでサポートされているOracle Solarisリリース

Oracle Solarisのバージョン	制御ドメイン	ルート・ドメイン	I/Oドメイン	ゲスト・ドメイン
Oracle Solaris 11.3 SRU 24 ¹	✓	✓	✓	✓
Oracle Solaris 10 1/13 ²				✓

1. Oracle Solaris 11.3 SRU 23 より前の Oracle Solaris 11 のバージョンは、SPARC M8プロセッサベース・サーバーではサポートされません。
2. 必須パッチを追加で適用してください。

SPARC M8 プロセッサ・ベース・システムのもっとも魅力的な特長の 1 つは、このシステムが、Oracle Solaris およびそのサポート対象のアプリケーションでは良く知られている SMP システムとして見えることです。さらに、Oracle Solaris には、オラクルのマルチコア/マルチスレッド・アーキテクチャ上のアプリケーションのパフォーマンスを改善する、次のような数多くの機能が搭載されています。

» Oracle Solaris のソフトウェア定義型ネットワーク (SDN)。Oracle Solaris 11 は、Oracle Solaris の既存の統合ソフトウェア定義型ネットワーク・テクノロジーを拡張し、高価なネットワーク・ハードウェアによるオーバーヘッドを加えることなくアプリケーションの俊敏性を大幅に強化します。完全に分散されたシステム・セット上でのアプリケーション主導のマルチテナント・クラウド仮想ネットワーク、物理ネットワーク・インフラストラクチャからの分離、およびアプリケーションレベルのネットワーク・サービスレベル・アグリーメント (SLA) を実現できるようになりました。これらすべての機能が、プラットフォームの一部として組み込まれています。

- » 拡張機能と新機能は次のとおりです。
 - » 仮想 NIC (VNIC) と仮想スイッチによるネットワーク仮想化
 - » ネットワーク・リソース管理と統合サービス品質 (QoS) により、VNIC およびトラフィック・フローに帯域幅制限を適用
- » Oracle Solaris 11 に付属の OpenStack ディストリビューションのコア機能が実現するクラウド対応
- » Oracle Solaris Elastic Virtual Switch および Virtual Extensible LAN (VXLAN) によるアプリケーション駆動型のマルチテナント・クラウド仮想ネットワーク

- » アプリケーション駆動型 SDN によるアプリケーションレベルの QoS
- » Oracle Solaris Zones と Oracle Solaris 10 Zones の緊密な統合
- » **ライフサイクル管理**。Oracle Solaris 11 には、プラットフォームのソフトウェア・ライフサイクルを管理するための一連のテクノロジーが統合されています。Oracle Solaris 自動インストーラによるセキュアなエンド・ツー・エンドのプロビジョニング、Oracle Solaris Image Packaging System によるフェイルセーフのソフトウェア更新、ZFS ブート環境、Oracle Solaris Unified Archives を使った迅速なアプリケーションのデプロイ、および包括的なコンプライアンス・フレームワークのサポートにより、生産性の向上、ヒューマン・エラーの軽減、IT コストの大幅な削減が可能になります。
- » **高速暗号処理**。Oracle Solaris の暗号化フレームワークおよび SPARC M8 プロセッサにより、高速暗号処理がサポートされます。SPARC M8 プロセッサでは、暗号化ハードウェアの実装にアクセスできます。今回初めて、ユーザーレベルの命令によって、コプロセッサとしてではなく適切なパイプライン自体に暗号が実装されます。つまり、ハードウェアベースの暗号をより効率的に実装でき、特権レベルでの変更が不要になるため、暗号化アルゴリズムの計算が大幅に効率化されます。また、データベースの操作により、命令パイプライン自体に実装された各種暗号化方式をはるかに効率的に利用できます。
- » **クリティカル・スレッドの最適化**。Oracle Solaris 11 および Oracle Solaris 10 では、ユーザーまたはプログラマーが Oracle Solaris スケジューラにクリティカル・スレッドを認識させることができます（それには、コマンドライン・インタフェース (CLI) または関数のシステム・コールを使用してスレッドの優先順位を 60 以上に上げます)。Oracle Solaris スケジューラは通常、システム上の全コアにスレッドを均等に分散します。ただし、プロセスが FX60 モードで実行されている場合、コア全体に対する排他的アクセスをそのスレッドに与えて、残りの実行可能なスレッドを他の使用可能なコアに割り当てようとします。
- » **マルチコア/マルチスレッドの認識**。Oracle Solaris 11 および Oracle Solaris 10 では SPARC M8 プロセッサの階層が認識されるため、Oracle Solaris スケジューラによって、使用可能なすべてのパイプラインで効果的に負荷を分散できます。Oracle Solaris がこれらのプロセッサを 256 基の論理プロセッサとして公開する場合でも、コア・クラスタ、L2 と L3 のキャッシュ階層、サポートするコアとスレッド間の相関関係を把握して、迅速かつ効率的なスレッドの実装を実現します。
- » **きめ細かな管理**。Oracle Solaris 11 および Oracle Solaris 10 には、SPARC M8 プロセッサの個別のコアとスレッド（論理プロセッサ）を有効または無効にする機能が搭載されています。また、プロセッサ・セットなどの Oracle Solaris の標準機能によって、論理プロセッサのグループ定義や、論理プロセッサ上のプロセスまたはスレッドのスケジューリングが可能になります。
- » **インタフェースのバインディング**。Oracle Solaris では、必要または要望に応じて、プロセスや個々のスレッドをプロセッサやプロセッサ・セットに非常に柔軟にバインドできます。
- » **仮想化ネットワークおよび I/O のサポート**。Oracle Solaris には、SPARC M8 プロセッサ上のコンポーネントとサブシステムをサポート、仮想化するテクノロジーが搭載されています。高パフォーマンスのネットワーク・アーキテクチャの一部として、Oracle マルチコア/マルチスレッド対応デバイス・ドライバが提供されているため、仮想化フレームワーク内で実行されるアプリケーションは、I/O デバイスとネットワーク・デバイスを効果的に共有できるようになります。

» **Oracle Solaris における Non-Uniform Memory Access (NUMA) の最適化。**メモリは SPARC M8 プロセッサごとに管理されるため、これらの実装は NUMA アーキテクチャに相当します。NUMA アーキテクチャの場合、プロセッサが自身のメモリへのアクセスに要する時間は、別のプロセッサに管理されているメモリへのアクセス時間よりやや短くなります。Oracle Solaris には、特にアプリケーションへの NUMA の影響を減らし、NUMA アーキテクチャのパフォーマンスを改善することができる次のテクノロジーが搭載されています。

» **メモリ配置最適化 (MPO)。**Oracle Solaris では、MPO を使用してサーバーの物理メモリ全体のメモリ配置を改善することによりパフォーマンスを向上させています。プロセッサのアクセス対象となるメモリは、MPO を介してそのプロセッサにできる限り近い位置に配置される一方で、システム内でのバランスは十分に維持されます。その結果、多数のデータベース・アプリケーションの実行速度が MPO により大幅に向上します。

» **Hierarchical Lgroup Support (HLS)。**複雑なメモリ待機時間の階層を持つシステムのパフォーマンスは HLS によって最適化され、Oracle Solaris の MPO 機能が改善します。Oracle Solaris は HLS を利用して、メモリがどの程度離れているかを判断し、できるだけ待機時間の短いリソースをアプリケーションに割り当てます。特定のアプリケーションがデフォルトでローカル・リソースを使用できない場合、Oracle Solaris では、HLS を使用してもっとも近いリモート・リソースを割り当てます。

» **Oracle Solaris ZFS。**Oracle Solaris ZFS により、データ管理は劇的に進歩します。複雑なストレージ管理の概念を自動化および統合し、世界唯一の 128 ビットファイル・システムで無限のスケラビリティを実現します。Oracle Solaris ZFS は、I/O の発行順序に関する従来の制約をほぼすべて解消するトランザクション・オブジェクト・モデルに基づいており、パフォーマンスを大幅に向上させます。また、サイレント・データ破損の検出と修復を行う 64 ビットのチェックサムですべてのデータが保護され、データの整合性が維持されます。

» **マルチパス・ソフトウェア。**Oracle Solaris のマルチパス・ソフトウェアにより、ストレージ・デバイスやネットワーク・インタフェースなどの I/O デバイスへの冗長物理パスを定義、制御することができます。デバイスへのアクティブ・パスを使用できなくなると、マルチパス・ソフトウェアにより別のパスに自動的に切り替わるか、またはフェイルオーバーして可用性を維持できます。マルチパス機能を利用するには、冗長ネットワーク・インタフェースや 2 枚のホスト・バス・アダプタを同じデュアルポート・ストレージ・アレイに接続するなど、サーバーを冗長ハードウェアで構成する必要があります。

» **OpenStack クラウド管理。**Oracle Solaris 11 には OpenStack デистриビューションがすべて備わっているため、管理者は他のベンダーが提供するインフラストラクチャ製品や仮想化製品などのデータセンター・リソースを 1 つの管理ペインから一元的に共有、管理できます。OpenStack on Oracle Solaris は、Oracle Solaris Zones、ZFS ファイル・システム、Unified Archives、包括的な SDN のような核となるテクノロジー基盤に統合されているため、セルフサービス・コンピューティングが可能です。このため、IT 組織は、エンタープライズ級の信頼性、セキュリティ、およびパフォーマンスを持つサービスを何週間ではなく何分間で提供できます。

» **セキュアで堅牢なエンタープライズ・クラス的环境。**既存の SPARC アプリケーションを変更せずに、引き続き SPARC M8 プロセッサベース・システム上で実行できるため、ソフトウェアへの投資が保護されます。また、マルチレベルの認証セキュリティによって、Oracle Solaris 環境が不正侵入から保護されます。Oracle Solaris の障害管理アーキテクチャによって、Oracle Solaris の予測的自己修復などの要素がハードウェアと直接通信できるため、計画内停止時間と計画外停止

時間の両方が短縮されます。Oracle Solaris DTrace などの効果的なツールにより、システムのリソースを最大限に利用できるようにアプリケーションを調整できます。

仮想化

組織が、使用率を高めながらさまざまなワークロードをより強力で少ない台数のシステムに統合しようとする場合、仮想化は必要不可欠なテクノロジーです。SPARC M8 プロセッサベースのサーバーには、3 つの仮想化機能（複数の PDom（デュアル PDom 構成の SPARC M8-8 サーバー）、Oracle VM Server for SPARC、および Oracle Solaris Zones による OS ベースの仮想化という形での第 3 レイヤー・サポート）が組み込まれています。

- » PDom により、1 つのハードウェア・システムを複数のセキュリティ・サーバーと障害隔離サーバーに、または複数のハードウェア・パーティションに分割します。
- » 論理ドメイン (LDom) は、Oracle VM Server for SPARC を使用して作成され、サーバーまたは PDom を仮想化して複数の仮想マシン (VM) をホストするために使用されます。各仮想マシンでは、それ独自の Oracle Solaris のインスタンスが実行されます。Oracle VM Server for SPARC は、オラクルの SPARC サーバーすべてに付属する無償の機能です。
- » Oracle Solaris Zones で OS を仮想化することで、Oracle Solaris の 1 つのインスタンスによってアプリケーションどうしを安全に分離し、システム・リソースを各ゾーンに割り当てることができます。そのため基本的には、Oracle Solaris オペレーティング・システムの 1 つのインスタンス内に複数の仮想マシンを作成できます。

これらの仮想化テクノロジーは相互に補完されます。実際、多くの場合にベスト・プラクティスにはレイヤー化した仮想化が含まれており、そこでは 2 つまたは 3 つすべてのテクノロジーを導入して、セキュリティ、可用性、パフォーマンス、管理性が最適化されています。さらに、オラクルの仮想化テクノロジーは追加コストなしで組み込まれています。図 20 に、SPARC M8 プロセッサベースのサーバーで利用可能な仮想化テクノロジーを示します。また、図の後にその詳細を説明します。

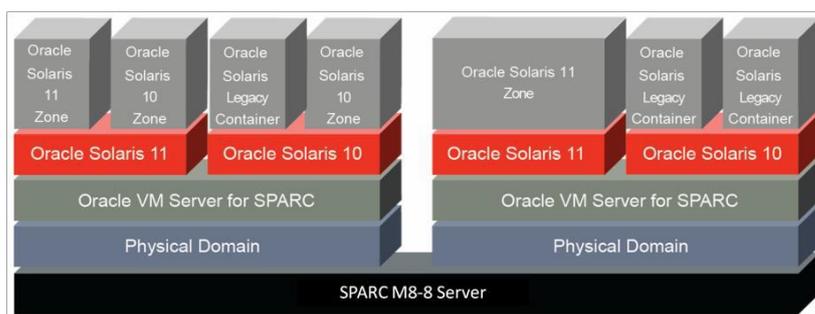


図20 : SPARC M8-8 サーバーでは デュアルPDomをサポートし、一方で SPARC T8-1 サーバー、T8-2 サーバー、T8-4 サーバーは シングルPDom をサポートする

SPARC M8-8 サーバー上の PDom により組織は、1 つのハードウェア・システムを複数のセキュリティ・サーバーと障害分離サーバーに分割できます。SPARC M8-8 サーバーは、工場出荷時デュアル PDom 構成で注文できます。適切に構成することにより、1 つのドメイン内でハードウェアまたはソフトウェアの障害が発生しても、それらの障害は分離されたままになり、他のドメインの動作に影響することはありません。

前世代の SPARC プロセッサと同様に、SPARC M8 プロセッサは Oracle VM Server for SPARC をサポートします。これは、シングル PDom 内で、または SPARC T8-1 サーバー、T8-2 サーバー、T8-4 サーバー、M8-8 サーバーの場合には単一サーバー内で論理ドメイン (LDom) を作成することを可能にするハイパーバイザです。このハイパーバイザは、プロセッサと緊密に統合された安定した仮想マシン・アーキテクチャを提供する小さなファームウェア・レイヤーである点で、従来のハイパーバイザとは異なります。ハイパーバイザは基盤となるマルチコア/マルチスレッド・プロセッサと直接やり取りするため、マルチスレッドが非常に重要です。Oracle VM Server for SPARC は、オラクルのマルチコア/マルチスレッド・テクノロジーを使用するすべてのサーバーでサポートされており、独立したオペレーティング・システム・インスタンスを実行する完全な仮想マシンを実現します。これら完全な仮想マシンは、ルート・ドメインとして構成して、I/O デバイスへの直接アクセス用にこれらの仮想マシンにルート・コンプレックスが割り当てられるようにすることができます。または、ゲスト・ドメインとして構成して、仮想化 I/O デバイスにアクセスするようにすることができます。各オペレーティング・システム・インスタンスには、プロセッサ、メモリ、ストレージ、コンソール、暗号化デバイスが含まれます。

Oracle Solaris 11 には、Oracle Solaris Zones と呼ばれる独自の OS コンテナ・テクノロジーが導入されています。このテクノロジーにより、アプリケーションを実行するための分離されたセキュアな環境を作成できます。Oracle Solaris Zones は、1 つの Oracle Solaris インスタンス内に作成された非常に軽量の仮想オペレーティング・システム環境です。Oracle Solaris Zones を使用すると、アプリケーションとプロセスをシステムの残りの部分から分離できます。この分離により、1 つの Oracle Solaris Zones 内のプロセスと別の Oracle Solaris Zones で実行されているプロセスとの干渉が避けられるため、セキュリティと信頼性が向上します。マルチプロセッサ・システム内の仮想 CPU (または SPARC M8 プロセッサ内のスレッド) は、動的なリソースまたは固定 CPU リソースのいずれかとして割り当てることで、非常に効率的なパーティション環境や強力なパーティション環境、またはその両方を兼ね備えたパーティション環境を実現できます。この共有または専用リソースは、Oracle Solaris Zones 環境の一部として構成したり、ビジネス・ニーズに合わせて動的に再構成したりすることができます。Oracle Solaris Zones はカーネル・ゾーン・モードでも実行できます。その場合、まったくバージョンの異なる Oracle Solaris 11 を実行して、ホスト OS のグローバル・ゾーンとは別の時間にパッチを適用できます。さらに、Oracle Solaris 11 には不変ゾーン機能があり、Oracle Solaris Zones とホスト OS グローバル・ゾーンの両方、および任意の Oracle VM Server for SPARC ドメインで使用可能な読み取り専用ルート・ファイル・システムを実装することにより、ゾーンをロックダウンします。特定の保守作業として実行されない限り、システム・バイナリやシステム構成に対する変更はブロックされます。

オラクルのレイヤー化された仮想化テクノロジーをまとめて使用することにより、回復力のある高可用性システムを構築できます。たとえば、オラクルのエンジニアド・システムと Oracle Maximum Availability Architecture では、これらの仮想化テクノロジーと仮想化ベスト・プラクティスを採用することで、非常に高い信頼性、可用性、保守性 (RAS) を達成しています。仮想化の詳細については、表 9 (39 ページ) の参考資料を参照してください。

システム管理

サーバー・システムのシステム管理を実地でローカルに行うことは、ほとんどの組織にとって現実的ではなくなっています。24 時間のシステム運用、ディザスタ・リカバリのホット・サイトの運用、地理的に分散した組織形態などのため、システムのリモート管理が必要とされます。Oracle サーバーの多くのメリットのうちの 1 つは、Lights-Out データセンターをサポートしていることで、これにより、人件費の高いサポート・スタッフがネットワーク・アクセスの可能な任意の場所で作業することが可能になります。SPARC M8 プロセッサベースのサーバーの設計は、Oracle Integrated Lights Out Manager (Oracle ILOM) ソフトウェアを実行する高性能なサービス・プロセッサ (SP) と組み合わせられています。この仕組みと Oracle Enterprise Manager Ops Center ソフトウェアのおかげで管理者は、ハードウェアに触れることができなくても、ほぼすべてのタスクをリモートで実行および制御することができます。これらの管理ツールとリモート機能により、管理者の負担が軽減され、組織の時間が節約され、運用コストが削減されます。

Oracle ILOMとサービス・プロセッサ

各 SP 上の Oracle ILOM ソフトウェアは、SPARC サーバーのリモート監視/管理機能の中核です。SP は、サーバー・システムから独立した専用プロセッサで構成されており、Oracle ILOM ソフトウェア・パッケージを実行します。より大きな SPARC M8-8 サーバーは、自動フェイルオーバー機能を持つ冗長 SP を特長としており、継続的に稼働するためのホットサービスをサポートできます。サーバーに入力電力が供給されている間、すべてのドメインが非アクティブであっても、SP は常にシステムの監視を続けます。

SP はまた、環境センサーを定期的に監視し、潜在的なエラー状態の事前警告を出し、必要に応じて事前予防的なシステム・メンテナンス・プロシージャを実行します。たとえば、システムに物理的な損傷を与えかねない温度状態に対する応答として、SP でサーバーのシャットダウンが開始されます。管理者は、SP で実行されている Oracle ILOM ソフトウェア・パッケージを使用して、物理ドメインと仮想マシン、およびハードウェア・プラットフォーム自体をリモートで制御および監視できます。

オペレーターは、SP へのネットワーク接続またはシリアル接続を使用して、ネットワーク上のどこからでもサーバーを効率的に管理できます。SP へのリモート接続はオペレーティング・システムとは別に実行され、この接続によりシステム・コンソールを完全に制御し、権限実行できます。Oracle ILOM SP はシステム・コントローラとして動作して、リモートでの運用管理を促進します。この SP はフル機能を備えており、オラクルの他のサーバーで使用される実装方法に似ています。そのため、サーバーは既存の管理インフラストラクチャと簡単に統合できます。Oracle ILOM SP は、効果的なシステム管理に不可欠であり、次の処理を行います。

- » IPMI 2.0 準拠 SP を実装することにより、サーバーのファームウェア、OS、アプリケーション、(Oracle ILOM 4.0 Ethernet 管理インタフェース経由で SP にアクセスする) IPMI ベースの管理ツールに対して IPMI 管理機能を提供します。また、シャーシ内のサーバー・モジュールやシャーシ内のその他の場所にある環境センサーを可視化します。
- » プロセッサ、DIMM、ファン、電源を含むサーバーのインベントリ管理と環境制御を行い、HTTPS、CLI、SNMP によるこのデータへのアクセスを可能にします。

- » テキスト表示のコンソールをリモート操作できます。
- » すべてのシステム・ファームウェアのアップグレードをダウンロードできます。

Oracle ILOM SP によっても管理者は、プラットフォームで稼働するオペレーティング・システムに依存せずに、またシステムの動作に干渉することなく、サーバーをリモート管理できます。Oracle ILOM により、サーバーに関連したハードウェア障害、警告、その他のイベントに関する電子メール・アラートを送信できます。その電子回路は、サーバーのスタンバイ電源を使用して、サーバーから独立して動作します。そのため、Oracle ILOM のファームウェアとソフトウェアは、サーバーのオペレーティング・システムがオフラインになったり、サーバーの電源がオフになったりしても実行され続けます。

Oracle ILOM により、以下のサーバー状態を監視します。

- » CPU の温度状態
- » ハード・ドライブの有無
- » エンクロージャの熱状態
- » ファンの速度とステータス
- » 電源のステータス
- » 電圧状態
- » Oracle Solaris の予測的自己修復、ブート・タイムアウト、サーバーの自動再起動イベント

電力管理

サーバーの電力コストと冷却コストの割合が増大しており、これらのコストを削減することは企業のデータセンターにおける最優先課題となっています。データセンターを拡張するための電力やスペースには限りがあるため、企業はサーバーの電力効率を注視するよう迫られています。所定の消費電力を超えると違約金が発生する契約を電力業者と締結している場合は、企業の管理下でサーバーが自動的に電力消費量を抑えることができるようにする必要があります。電力効率と二酸化炭素排出量も、企業がサーバーを評価する際の評価要素となってきています。

オラクルのマルチコア/マルチスレッド設計本来の効率性に加えて、SPARC M8 プロセッサのコアとメモリの両方のレベルには、電力管理機能が組み込まれています。この機能には、命令速度の抑制、アイドル状態にあるスレッドとコアの停止、コアとメモリの両方でのクロック・オフなどの消費電力を抑えるための機能が含まれます。さらに、以前の SPARC プロセッサではチップ・レベルで電源管理を行うことができたが、SPARC M8 プロセッサではサブチップの電源管理を行うことができます。SPARC M8 プロセッサは、それぞれが 16 個のコアで構成される 2 つのパーティションに分割されています。パーティションごとに、他方のパーティションに関係なく、異なる電力消費レベルを設定できます。これにより、パーティションあたりのパフォーマンス・レベルに差が出るのを可能にし、重要な点として、パーティションごとに LDom をデプロイする場合に異なる SLA を設定できるようにしています。

Oracle ILOM での電源管理サポートのほか、Oracle Solaris 11.3 以降には SPARC M8 プロセッサベースのサーバーをサポートする電源マネージャーが搭載されています。Oracle Solaris では、poweradm 設定に基づいて、有効にする節電機能を決定できます。この設定項目は、システム (Oracle ILOM) ポリシーに基づいてプラットフォームによって設定されますが、Oracle Solaris の管理者はこの設定を無効にすることができます。以下の領域で大きな技術革新が行われています。

- » 実行されない条件分岐などの投機実行の制限
- » データ・パス、制御ブロック、アレイにおける広範囲なクロック・ゲーティング
- » 余分なストール・サイクルをデコード・ステージに注入できるようにする電源調整

Oracle VM Server for SPARC を使用する仮想化環境において電源管理マネージャーは、LDom ゲストを管理する場合に次のタスクを実行します。

- » 電源管理ポリシーに基づいて有効にする節電機能を決定する
- » 電源管理エンジンを呼び出してリソース上で電力状態の変更を開始して、電力の調整/使用率レベルを達成する (Oracle Solaris 11.3 ゲストが所有していないリソースの場合) か、またはハイパーバイザに指示してハイパーバイザ/ハードウェア管理の電力状態を有効または無効にする Oracle Solaris 11.3 ゲストにのみ電源管理ピアがあります。

Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center は、SPARC M8 プロセッサベースのサーバー向けの集約型ハードウェア管理ソリューションで、インフラストラクチャ・スタック全体で管理を統合します。Oracle Enterprise Manager Ops Center の先進の仮想化管理機能とレポート作成機能、アプリケーションとディスク間の管理、インテリジェントな構成管理などの機能により、IT マネージャーは、複雑さを軽減し、インフラストラクチャ管理を合理化および簡素化できます。すべての SPARC サーバーに Oracle Enterprise Manager Ops Center が搭載されているため、データセンターの管理者は、ストレージ、ネットワーク、サーバー、Oracle Solaris、仮想化環境を単一のインタフェースから監視および管理できます。これにより、運用効率が向上して運用コストが低減します。

Oracle Enterprise Manager Ops Center は、オラクルのサーバーとエンジニアド・システム・インフラストラクチャ向けのもっとも包括的な管理ソリューションです。複数のサーバー・アーキテクチャと膨大な数のオペレーティング・システムを管理するためのコンソールを 1 つ備えており、資産検出、ファームウェアとオペレーティング・システムのプロビジョニング、パッチの自動管理、パッチおよび構成管理、仮想化管理、包括的なコンプライアンス・レポート作成機能を使用して、SPARC M8 プロセッサベースのサーバーのコンポーネントを管理できます (図 21)。

Oracle Enterprise Manager Ops Center では、ポリシー・ベースの管理手法でワークフローを自動化し、コンプライアンスを強化します。しかも、それらすべてを直感的な単一のインタフェースを使用して実現しています。IT スタッフは、ビジネス要件に合わせてインフラストラクチャを効率的に導入しながら、データセンターの標準化とベスト・プラクティス、法令遵守、セキュリティ・ポリシーを実装および強化できます。

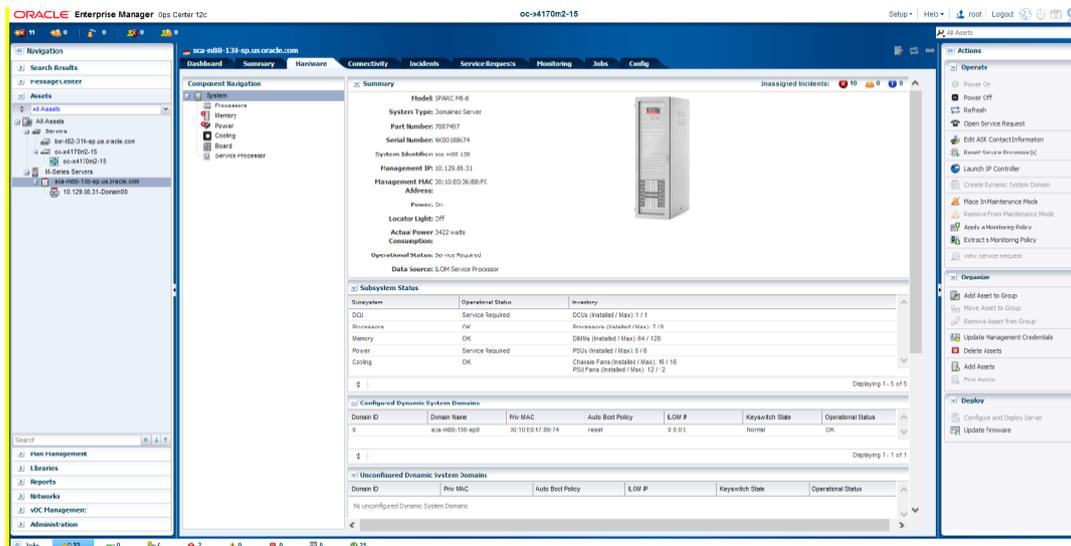


図21 : Oracle Enterprise Manager Ops Center による SPARCサーバーの細部の管理

信頼性、可用性、保守性

計画および計画外のダウンタイムを減らすことは、IT サービスにとってきわめて重要です。システム設計には、主要サービスの可用性に影響を及ぼすことなく、耐障害性、短時間での修復、迅速な拡張を可能にするメカニズムを組み込む必要があります。SPARC M8 プロセッサベースのサーバーは、複雑なネットワーク・コンピューティング・ソリューションと厳しい高可用性（HA）の要件に対応できるように特別に設計されており、ホットサービス可能な冗長システム・コンポーネント、設計全体の診断とエラー・リカバリ機能、組込みリモート管理機能が組み込まれています。これらの信頼性の高いサーバーの高度なアーキテクチャにより、アプリケーションの可用性を高レベルに保ち、さまざまなハードウェア障害からの迅速なリカバリを促進して、企業のシステム運用を簡素化し、コストを低減できます。

高度な信頼性機能

SPARC M8 プロセッサベースのサーバーのコンポーネントに組み込まれている高度な信頼性機能により、サーバーの全体的な安定性が向上しています。サーバー・アーキテクチャ内のコンポーネントの数と複雑さが低減すると、信頼性が高まります。また、高度な CPU 統合と確実なデータ・パスの整合性により、SPARC M8 プロセッサによる自律型のエラー・リカバリが実現し、修正処理の開始にかかる時間が短縮され、結果としてアップタイムが長くなります。主要コンポーネントの冗長性により、フェイルオーバーが可能です。ホットプラグ対応コンポーネントおよびホットスワップ対応コンポーネントによって保守性が向上します。SPARC M8 プロセッサベースのサーバーでは、先進の I/O コントローラによって PCIe ルート・コンプレックスがホストされるため、プロセッサが追加または除外されても、ルート・コンプレックスに変更が加えられることはありません。これらの機能を総合して、非常に堅牢なシステム設計が実現されています。

ファームウェアと Oracle Solaris の予測的自己修復ソフトウェアの両方に実装されている障害管理アーキテクチャにより、SPARC サーバーの信頼性はさらに強化されています。このアーキテクチャでは、プロセッサ、メモリ、I/O デバイスを常に監視します。永続的な CPU ソフト・エラーは、

エラーの性質に応じて、スレッド、コア、またはプロセッサ全体を自動的にオフラインにすることによって解決できます。また、メモリ・ページ・リタイアメント機能では、特定の DIMM メモリのデータに対して複数の修正があった場合にメモリ・ページを先見的にオフラインにすることができます。前述のように、SPARC M8 プロセッサベースのサーバーでは DIMM スペアリングがサポートされており、システムの動作を中断することなく、メモリ容量とエラー保護機能をそのまま維持しながら DIMM 全体をリタイアさせることができます。

エラー検出、診断、およびリカバリ

SPARC M8 プロセッサベースのサーバーには、障害を早期に修正し、周辺コンポーネントがダウンタイムを繰り返さないようにする重要なテクノロジーが導入されています。本質的に信頼性を向上させるアーキテクチャ上の進歩は、サーバーのハードウェア・サブシステム内のエラー検出機能やリカバリ機能によって促進されています。最終的には、次の機能と連携させてアプリケーションの可用性を高めます。

- » エンド・ツー・エンドのデータ保護により、システム全体のエラーが検出、修正され、データの完全な整合性が維持されます。
- » 最先端の障害分離機能により、サーバーはコンポーネント境界内のエラーを分離して、コンポーネント全体ではなく、関連するチップまたはチップの一部だけをオフラインにすることができます。可能な限り小さな単位でエラーを分離することで、安定性が向上し、最大限の処理能力を継続的に発揮できます。この機能は、プロセッサ、メモリ ASIC と DIMM、スイッチ ASIC、接続リンク、I/O コントローラ ASIC、SP に適用されます。
- » 定期的な環境監視により、関連する環境条件とエラー条件の履歴ログが記録されます。
- » ホストのウォッチドッグ機能により、ソフトウェアの動作（ドメインのオペレーティング・システムを含む）が定期的にチェックされます。またこの機能では、SP ファームウェアを使用してエラー通知機能とリカバリ機能がトリガーされます。
- » システムの障害管理アーキテクチャと SPARC M8 プロセッサの動的 CPU リソース構成解除により、確実な障害分離とリカバリが可能です。システムは必要に応じて、実行中のアプリケーションを中断させることなく、プロセッサ・リソース（コアなど）を動的にリタイアさせることができます。
- » コンポーネント・ステータスの定期的なチェックにより、多くのシステム・デバイスのステータスが確認され、差し迫った障害の兆候が検出されます。リカバリ・メカニズムがトリガーされて、システム障害やアプリケーション障害が回避されます。
- » エラー・ロギング、マルチステージ・アラート、現場交換可能ユニット（FRU）の電氣的な識別情報、システム障害の LED インジケータにより、問題を迅速に解決できます。

» ホットサービス対応冗長コンポーネント

今日の IT 組織は、ノンストップで事業を遂行するという難題に直面しています。ネットワーク化されたグローバル経済では、昼夜を問わず収益の機会があるため、計画ダウンタイムを短縮する必要や、場合によっては完全になくす必要に迫られています。SPARC M8 プロセッサベースのサーバーでは、これらの要望に応えるため、ホットプラグおよびホットスワップ対応の組込み冗長ハードウェアを採用して、個々のコンポーネント障害やシステム構成の変更による中断を最小限に抑えます。実際これらのシステムでは、多くの場合、（たいていはユーザーやシステムの機能に影響を及ぼさずに）ハードウェア障害からリカバリできます。

SPARC M8 プロセッサベースのサーバーはホットスワップ対応の冗長電源とファン・ユニットを搭載しており、オプションで、プロセッサ、メモリ DIMM、I/O カードをそれぞれ複数構成することも可能です。管理者は、ホットプラグ対応のディスク・ドライブとディスク・ミラー化ソフトウェアを組み合わせて、冗長な内蔵ストレージを作成できます。SPARC M8-8 サーバーは、ホットプラグ対応の冗長サービス・プロセッサもサポートしています。障害が発生しても、これらの重複コンポーネントによって動作を続行できます。コンポーネントやエラーの種類によっては、システムがデグレード・モードで動作を続行したり、再起動したりする場合があります。その場合、障害が自動的に診断され、関連コンポーネントが自動的にシステム構成から切り離されます。また、これらのサーバー内のホットサービス対応ハードウェアにより保守が高速化され、システムを停止することなくコンポーネントの交換や追加を簡単に実行できます。

交換可能なコンポーネントは、次のカテゴリに分類されます。

- » **ホットサービス対応**。ホットサービス対応コンポーネントは、サーバーの実行中に取り付けと取外しが可能です。
- » **ホットスワップ対応**コンポーネントは、保守前の準備が不要です。
- » **ホットプラグ対応**コンポーネントは、保守前の準備が必要です。
- » **コールドサービス対応**。コールドサービス対応コンポーネントは、保守を行うためにサーバーを停止する必要があります。また、一部の保守手順では、電源装置と電源ソース間の電源ケーブルを取り外すことが必要になります。

表 8 に、SPARC M8 プロセッサベースのサーバーのおもなホットサービス対応コンポーネントを示します。

表8：SPARC M8プロセッサベースのサーバーの顧客側のおもなホットサービス対応コンポーネント

コンポーネント	SPARC T8-1 サーバー	SPARC T8-2 サーバー	SPARC T8-4 サーバー	SPARC M8-8 サーバー
2.5インチSFFドライブ	✓	✓	✓	N/A
電源	✓	✓	✓	✓
ファン/ファン・モジュール	✓	✓	✓	✓
PCIeカード	-	-	✓ ¹	✓
CMIOUボード	N/A	N/A	N/A	✓ ²
SP	-	-	-	✓

1. カード・キャリア内の PCIe カード内部スロット内の NVMe PCIe スイッチ・カードはコールドサービス対応です。
2. CMIOU ボード が含まれている PDom はシャットダウンされている必要があります。その他の PDom は動作可能です。

結論

現代のテクノロジーとビジネスの課題により、急速に進化するデジタル市場に対応するための革新的なソリューション、大きな処理能力、幅広いサーバー・フォーム・ファクタが必要とされています。オラクルの革新的な SPARC M8 プロセッサの利点をベースにしたオラクルの SPARC サーバーという新しいアプローチにより、組織は、柔軟性と高可用性を維持しつつ、もっとも要求の厳しい演算ワークロードを効率良く処理することができます。革新的な Software in Silicon テクノロジーでは、各 SPARC M8 プロセッサの専用高速化エンジンを使用して、Silicon Secured Memory、インメモリ・クエリ・アクセラレーション、インライン圧縮解凍、高速暗号化処理を行ってセキュリティを強化し、アプリケーションに劇的なメリットをもたらします。

SPARC M8 プロセッサベースのサーバーによって実現される広範な処理能力により、組織は、基盤となるインフラストラクチャからコンピューティング・コストをシフトし、ワークロードを生産的な業務で使用することに集中できます。たとえば、新しいレベルのパフォーマンスと効率性により、同じシステム上で OLTP とデータ分析の両方を実行できます。オラクルのレイヤー化した仮想化を Oracle Solaris の機能と組み合わせることで、組織は、卓越した容量とスケーラビリティを備えたもっとも効率的なオンプレミス・クラウドを構築できます。これにより、商品化までの期間を短縮し、収益を高めることができます。

追加情報

詳細については、表 9 に挙げた参考資料を参照してください。

表9：参考資料

Webリソース	URL
オラクルのSPARCサーバー	oracle.com/us/products/servers-storage/servers/sparc/oracle-sparc および oracle.com/technetwork/server-storage/sun-sparc-enterprise
Oracle Solaris	oracle.com/solaris
Oracle Optimized Solutions	oracle.com/optimizedsolutions
オラクルのベンチマーク結果	oracle.com/benchmarks
Oracleヘルプ・センターのドキュメント	docs.oracle.com
セキュリティ	oracle.com/technetwork/topics/security
仮想化	oracle.com/technetwork/topics/virtualization
Oracle Integrated Lights Out Manager	http://www.oracle.com/technetwork/server-storage/servermgmt/tech/integrated-lights-out-manager/index.html
Oracle Enterprise Manager Ops Center	oracle.com/technetwork/oem/ops-center
ホワイト・ペーパーおよび技術記事	URL
『オラクルのSPARC T7サーバーとSPARC M7サーバー：ドメイン構成に関するベスト・プラクティス』	oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/sparc-t7-m7-best-practices-2701865.pdf
『Oracle SPARC仮想化テクノロジーを使用した集約』	oracle.com/technetwork/server-storage/sun-sparc-enterprise/technologies/consolidate-sparc-virtualization-2301718.pdf
『Oracle Solarisカーネル・ゾーンのベスト・プラクティス』	oracle.com/technetwork/articles/servers-storage-admin/solaris-kernel-zones-best-practices-2400370.html
『Oracle VM Server for SPARCのベスト・プラクティス』	oracle.com/technetwork/server-storage/vm/ovmsparc-best-practices-2334546.pdf
『Implementing Root Domains with Oracle VM Server for SPARC』	oracle.com/technetwork/server-storage/vm/ovm-sparc-rootdomains-wp-1914481.pdf
『Oracle SPARCサーバーとOracle Solaris上のOracle Multitenant』	oracle.com/technetwork/articles/servers-storage-admin/multitenant-on-sparc-solaris-2016889.html



Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

海外からのお問い合わせ窓口
電話 : +1.650.506.7000
ファクシミリ : +1.650.506.7200

CONNECT WITH US

-  blogs.oracle.com/oracle
-  facebook.com/oracle
-  twitter.com/oracle
-  oracle.com

Integrated Cloud Applications & Platform Services

Copyright © 2017, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載される内容は予告なく変更されることがあります。本文書は、その内容に誤りがないことを保証するものではなく、また、口頭による明示的保証や法律による黙示的保証を含め、商品性ないし特定目的適合性に関する黙示的保証および条件などのいかなる保証および条件も提供するものではありません。オラクルは本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクルの書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

Oracle および Java は Oracle およびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

Intel および Intel Xeon は Intel Corporation の商標または登録商標です。すべての SPARC 商標はライセンスに基づいて使用される SPARC International, Inc. の商標または登録商標です。AMD、Opteron、AMD ロゴおよび AMD Opteron ロゴは、Advanced Micro Devices の商標または登録商標です。UNIX は、The Open Group の登録商標です。0917

Oracle SPARC T8 と SPARC M8 のサーバー・アーキテクチャ
2017 年 9 月



Oracle is committed to developing practices and products that help protect the environment