



オラクル ホワイトペーパー
2010 年 10 月

オラクルの SPARC T3-1、SPARC T3-2、 SPARC T3-4、SPARC T3-1B サーバー によるサービス稼働時間の最大化

はじめに	1
IT サービスの重要な役割	3
信頼できる IT サービスの必要条件	3
信頼性、可用性、保守性への広がる需要	3
SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバー	5
RAS を見据えた設計	7
可用性の改善: 冗長性、保守性、自己診断機能	10
オラクル仮想化テクノロジーによる障害分離	13
Oracle VM Server for SPARC	13
Oracle Solaris コンテナ	16
IT サービス稼働時間についての広範な焦点	17
障害管理アーキテクチャによる診断機能の高速化	17
Oracle Solaris での停止時間の削減	19
オラクルのソフトウェアによる効率的なシステム管理	21
サーバーの RAS の計測	22
可用性に関するベンチマークのフレームワーク	22
R-Cubed のメトリックス	23
まとめ	24

はじめに

IT サービスは、急速に変化する今日の世界的市場で企業が効果的に競争するための重要な役割を果たしています。どのような規模や業種の企業でも、日常業務や顧客対応、さらに収益計上など、さまざまな面で IT に依存しています。無停止の業務運営が現在の傾向となっており、24 時間休みなしで業務を行うため IT 部門に重責がかかります。その結果、システムやソフトウェアアプリケーションの信頼性、可用性、保守性 (RAS) 機能の重要性がますます高まっています。

信頼性と可用性と保守性は複雑に関係しており、IT サービスの稼働時間の最大化にはどれも欠かすことのできない重要な要素です。システムの信頼性機能は、障害発生 の 頻度を最小化してデータの整合性を確保します。可用性機能は、システム障害やエラーイベントの発生時における IT サービスの継続的なアクセスをサポートします。保守性メカニズムは、コンポーネントのアップグレードや修理のサービスサイクルの短縮を促進します。これら 3 つの RAS 要素を同時に実現させるような設計は、時として難題をもたらします。たとえば冗長性を最大化すると、システムの可用性は高まりますが、コンポーネント数を増やすことになるため潜在的な信頼性レベルが低下します。これからも RAS 機能を改善し続けるために、オラクルはエンジニアードメトリックスを注意深く使用して、新世代のあらゆるサーバーの RAS レベルのバランスをとり最適化するための設計に役立ちます。

オラクルの SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバーの優れた RAS 機能は、ビジネスクリティカルな IT サービスの稼働時間を最大化するのに最適です。信頼性の非常に高い部品と、比較的少ない数のコンポーネントにより、システムエラーの危険性を最小化します。複数の PCI Express (PCIe) ルートコンプレックス、プロセッサ、メモリー DD3 DIMM、および I/O カードを構成できるため、回復力が高まっています。さらに、これらのサーバーには、冗長なホットスワップディスク、電源装置、ファンのほかに、メモリーページリタイアメント (MPR)、コアとスレッドのオフライン化機能、プロセッサとキャッシュの一貫性、統合ディスク RAID 機能、広範な ECC ハードウェア保護機能が含まれています。SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバーはまた発熱を最小に抑えるため、環境に関連した障害を避けることができます。オンボードサービスプロセッサと Integrated Lights Out Management (ILOM) ソフトウェアは、管理を簡単にします。

オラクルはあらゆるプラットフォーム層で RAS に注力しています。オラクルは、プロセッサ、サーバー、仮想化、OS、システム管理などのあらゆる層のテクノロジーに精通しており、緊密に統合され信頼性の極めて高いプラットフォームを作り出しています。エネルギー効率のよい SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバーのプラットフォームと Oracle VM Server for SPARC (旧 Sun Logical Domains)、Oracle Solaris 10、およびオラクルの管理ツールを組み合わせることにより、プラットフォームの安定性がさらに向上しています。オラクルのテクノロジーを利用することで、組織や企業は総所有コスト (TCO) を最少化し、資産活用を最適化して、IT サービスの稼働時間レベルを最大限に引き上げるような、優れた IT ソリューションを作成することができます。

IT サービスの重要な役割

IT サービスは、たとえば政府機関、Fortune 500 にランキングされている企業、小さな地方自治体、起業したてのベンチャー企業など、あらゆる種類の組織や企業において重要な役割を担っています。中には、その存在自体が IT や Web 接続に依存している組織や企業もあります。電子店舗、ソーシャルネットワークサイト、ボイスオーバー IP (VoIP) がその一例です。従来のビジネスも IT にかかり依存しています。たとえば製造業者の中には、サプライチェーンのパートナーとの連絡や、効率的に操業、製造、および流通を行う企業間電子商取引などに IT サービスを利用している企業があります。実際に、数多くの組織や企業が、IT サービスに依存して、ビジネスチャンスの拡大、重要なビジネス機能の行使、操業コストの削減を実現しています。

ネットワーク中心のテクノロジーが進化するにつれ、IT の重要性はますます高まっています。新規ユーザーやデバイス接続の急速な増加に伴い、強いネットワークサービスの必要性和 Web 2.0 の創発が新しい IT サービスの需要を促進しています。さらに、エンドユーザーはこれらの新しい IT サービスに対して安定した可用性、アクセシビリティ、および応答性を期待します。その結果、サービスプロバイダや IT 部門には、サービスの割り込みを最小限に抑えなければならないという重責が常にかかります。

信頼できる IT サービスの必要条件

今日、電子的な業務や通信への依存がほとんどの組織や企業に広がっており、IT サービスは重要な収益経路に位置づけられています。システムの停止時間は、分単位で数百ドルもしくは数千ドルを損失するような財政結果を生み出すこととなります。また、ネットワーク側の事業運営にとって停止時間は収益が損失するだけでなく、ブランドイメージが傷つき、顧客満足度が低下し、潜在的なセキュリティの危険性や規制違反が高まることも意味します。このように IT への依存度はとても高く、利害関係にも深く関係するため、多くの環境において停止時間は受け入れがたいものになっています。

これらの責務がサービスレベルの契約書に明確に記載されようと、単に暗黙的な合意であろうと、IT 部門や CIO が可用性の期待に応える責任を負わされていることがよくあります。IT サービスの真の可用性は、プロセッサ、システムのプラットフォーム、OS、管理ツールなど、可用性に寄与するあらゆる要素で決まります。あらゆる分野の業務で信頼性、可用性、保守性の高い IT 製品やソリューションへの需要が高まっているため、組織や企業は計画的および計画外停止時間を最小限に抑える努力をしています。

信頼性、可用性、保守性への広がる需要

IT システムの RAS 機能は、さまざまな組織や企業にとってさらに重要になっています。RAS は多くのプロジェクトにとって大切ですが、予算、ワークロード量、およびそのアーキテクチャーのアプローチはさまざまです。大規模なスケールアップ (垂直スケール) 可能なシステムは広範な RAS 機能として知られており、一部の IT サービスの展開に適しています。しかし、この手法は、すべてのプロジェクトに適していたり、コスト効率が良いとは限りません。たとえば、ステートレスネットワーク中心の IT サービスでは、スケールアウト (水平スケール) アーキテクチャーを採用し、複数のシステムやアプリケーションインスタンスにワークロードを分散させることによって最大のパフォーマンスを得ています。

複数のシステムにネットワーク中心のアプリケーションを分散させる技法では、個々のサーバーに対する可用性へのプレッシャーは軽くなりますが、各システムの継続した操作が重要になります。

水平スケールアーキテクチャーの仮想化テクノロジーを採用する傾向が高まり、その環境下での RAS 機能の重要性が増しています。仮想化テクノロジーは、複数のアプリケーションを 1 台のサーバー上で実行可能にするため、組織や企業では IT 資産を最大限活用することができます。ますます強力になるマルチスレッドおよびマルチコアプロセッサと粒度の高い仮想化機能により、組織や企業で 1 台のサーバー上にホストするアプリケーションやサービスの数が増え続けています。強力なプロセッサと仮想化機能を組み合わせることにより、組織や企業は水平スケール可能な複数のアプリケーションを最少の数のサーバーに統合させることができます (図 1)。仮想化および統合された水平スケールアーキテクチャーのアプローチでは、パフォーマンスと効率性は最大になり、管理作業と取得原価は減少します。どの統合戦略でも、1 台のサーバーに配置するアプリケーションの数が増えれば、システムの RAS 機能の重要性が増します。

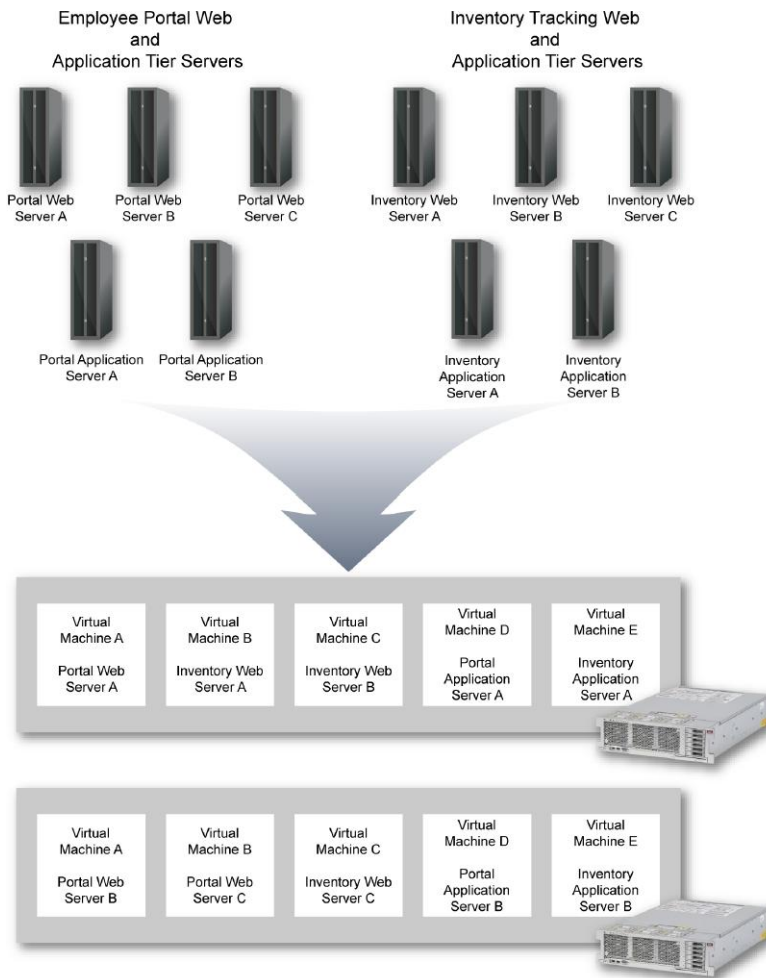


図 1. 仮想化テクノロジーによりサーバーの完全な統合が促進され、プラットフォームの RAS 機能の重要性が大幅に増加します。

SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバー

エネルギー効率のよい高性能 SPARC T3-1、SPARC T3-2、SPARC T3-4、SPARC T3-1B サーバーは、水平と垂直の両方のスケーラビリティを独自に融合し、従来の中規模からハイエンドのシステムと同等のパフォーマンスと信頼性を組織や企業に提供しています (図 2)。このようなサーバーは、第 4 世代のチップマルチスレッド化テクノロジー (CMT) に基づいた高効率性システムの新しい動きを表しています。



図 2. 最新の SPARC T シリーズファミリー (オラクルの SPARC T3-1B ブレードサーバー、SPARC T3-1 サーバー、SPARC T3-2 サーバー、SPARC T3-4 サーバー) には高度な RAS 機能があり、IP サービス稼働時間を最大化します。

SPARC T3-1、T3-2、T3-4、T3-1B サーバーはオラクルの SPARC T3 プロセッサを使用しています。このプロセッサには、優れた UltraSPARC T2 プロセッサや T2 Plus プロセッサの性能も含まれています。そのため、これらのサーバーは、Web スケールでのビジネスの需要を満たし、仮想化されて環境効率の高いデータセンターを作成し、スループットに影響を及ぼさずに企業のアプリケーションをセキュリティ保護するのに最適です。これらのサーバーには、組織や企業が IT サービスのディペンダビリティ (包括的な信頼性) を改善するのに役立つ、次のような主要な設計要素が含まれています。

- **部品数の削減** - プラットフォーム全体の安定性と信頼性の向上につながります。
- **プロセッサのスレッドとコアのオフライン化** - 障害発生時におけるシステム運用の継続をサポートします。
- **冗長性およびホットスワップコンポーネント** - システムの回復力と保守性の向上の基盤となります。
- **パリティ保護およびエラー訂正能力** - システム全体を通してエラーを検出および訂正し、データの整合性を確保します。
- **組み込み RAID 機能** - システムの信頼性を高め、データの整合性を強化します。
- **システム監視** - 物理センサーを使用した温度、電圧、電流、ファンの回転速度の計測と、論理センサーを使用したパフォーマンスとシステム管理パラメータの測定を行います。
- **オンボードサービスプロセッサと ILOM ソフトウェア** - 遠隔管理を簡素化し、管理作業をかなり柔軟にします。

- **優れたエネルギー効率** – データセンターでの放熱を削減し、熱による障害の危険性を最小化します。
- **堅牢な仮想化テクノロジー** – アプリケーション間の障害分離を促進し、IT サービスの稼働時間をさらに改善する RAS 機能の向上に貢献します。
- **幅広い障害管理** – システム、仮想化テクノロジー、OS の各層を含む主要な要素の障害とエラー状態の予防保守管理を行います。

表 1 に、オラクルの SPARC T3-1、T3-2、T3-4、T3-1B サーバーの機能を示します。

表 1. SPARC T3-1、T3-2、T3-4、T3-1B サーバーの機能

機能	SPARC T3-1 サーバー	SPARC T3-2 サーバー	SPARC T3-4 サーバー	SPARC T3-1B ブレードサーバー
CPU	• 16 コア 1.65 GHz SPARC T3 プロセッサ	• デュアル 16 コア 1.65 GHz SPARC T3 プロセッサ	• デュアルまたはクアッド 16 コア 1.65 GHz SPARC T3 プロセッサ	• 8 または 16 コア 1.65 GHz SPARC T3 プロセッサ
スレッド	• 128 以内	• 256 以内	• 512 以内	• 128 以内
メモリー容量	• 128G バイト以内 (8G バイト DDR3 DIMM)	• 256G バイト以内 (8G バイト DDR3 DIMM)	• 512G バイト以内 (8G バイト DDR3 DIMM)	• 128G バイト以内 (8G バイト DDR3 DIMM)
最大内蔵ディスクドライブ	• 16 HDD 以内 (2.5 インチ SAS2 300 G バイトディスクドライブ)、RAID 0/1、(5/5 + BBWC)	• 6 HDD 以内 (2.5 インチ SAS2S 300 G バイトディスクドライブ)、RAID 0/1	• 8 HDD 以内 (2.5 インチ SAS2 300 G バイトディスクドライブ)、RAID 0/1	• 4 HDD 以内 (2.5 インチ SAS2 300 G バイトディスクドライブ)、RAID 拡張モジュール (オプション)
ビデオ	• 1 VGA ポート	• 1 VGA ポート	• 1 VGA ポート	• 1 VGA ポート (ドングル経由)
取り外しとプラグイン可能な I/O	• Slimline DVD-R/W • 5 個の USB 2.0 ポート	• Slimline DVD-R/W • 5 個の USB 2.0 ポート	• DVD なし (RKVMS 経由) • 4 個の USB 2.0 ポート	• DVD なし • 3 個の USB 2.0 ポート
PCI	• 6 x8 PCIe Gen2 スロット	• 8 x8 PCIe Gen2 スロット • 2 x4 PCIe Gen2 スロット	• 16 EM x8 PCIe Gen2	• RAID 0 と 1 または 5/6 向けのファブリック拡張モジュール (オプション) • PCIe Gen2 をサポート
Ethernet	• 4 個のオンボード ギガビット Ethernet ポート (10/100/1000) • 2 個の 10 ギガビット Ethernet ポート (XAUI コンボスロット (共有 w/PCIe) を経由)	• 4 個のオンボード ギガビット Ethernet ポート (10/100/1000) • 2 個の 10 ギガビット Ethernet ポート (XAUI コンボスロット (共有 w/PCIe) を経由)	• 4 個のオンボード ギガビット Ethernet ポート (10/100/1000) • 8 個の 10 ギガビット Ethernet ポート (XAUI 2 QSFP Quad コネクタ経由)	• 2 個のオンボード ギガビット Ethernet ポート (10/100/1000) • 2 個の 10 ギガビット XAUI Ethernet ポート 拡張モジュール (オプション)
電源装置	• 2 個のホットスワップ可能 AC 1200 W 電源ユニット (N+1 の冗長性)	• 2 個のホットスワップ可能 AC 2000 W 電源ユニット (N+1 の冗長性)	• 4 個のホットスワップ可能 AC 2060 W 電源ユニット (N+N の冗長性)	• SunBlade 6000 モジュール システムシャーシに内蔵
ファン	• 6 個のホットスワップ可能 ファンモジュールと、モジュールごとに複数の反転ファン。 (N+1 の冗長性)	• 6 個のホットスワップ可能ファントレイと、モジュールごとに複数の反転ファン。 (N+1 の冗長性)	• 5 個のホットスワップ可能 ファンモジュールと、モジュールごとに複数の反転ファン。 (N+1 の冗長性)	• SunBlade 6000 モジュール システムシャーシに内蔵

オペレーティングシステム	• Oracle Solaris 10 Update 8 + MU9、Solaris 10 Update 9	• Oracle Solaris 10 Update 8 + MU9、Solaris 10 Update 9	• Oracle Solaris 10 Update 8 + MU9、Solaris 10 Update 9	• Oracle Solaris 10 Update 8 + MU9、Solaris 10 Update 9
--------------	--	--	--	--

RAS を見据えた設計

SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、組織や企業の IT サービスの稼働時間を最大化するのに役立ちます。データ整合、障害分離、エラー訂正を自動的に行うコンポーネントや機能を冗長に含めると、SPARC T3-1、T3-2、T3-4、T3-1B サーバーの堅牢性が向上します。オンラインによる保守機能や簡素化された保守手順により、企業は計画的な停止の必要性を回避でき、また広範な障害管理と自己修復機能により、計画外の停止や復旧時間を削減することができます。

大規模マルチスレッド、マルチソケットのシステムオンチップ (SoC) SPARC T3 プロセッサにより、SPARC T3-1、T3-2、T3-4、T3-1B サーバーのコンポーネント数は少なくなっています。40 nm の製造プロセスに基づいた SPARC T3 プロセッサには、たとえば低遅延データ転送のための 2 個の 8 レーン PCIe Generation 2 インタフェース、2 個の 10 ギガビット Ethernet ポート、ワイヤスピードの暗号化を実現する 1 コアあたり 1 個のストリーム処理ユニット (SPU)、乗算と加算を同時に行う FMADD 命令を含む完全にパイプライン化された浮動小数点ユニットといった高度なサーバー機能が組み込まれています (図 3)。SPARC T3 プロセッサにはまた、マルチソケット機能、キャッシュ一貫性リンク、および FB-DIMM シングルリンクのフェイルオーバー処理機能もあります。

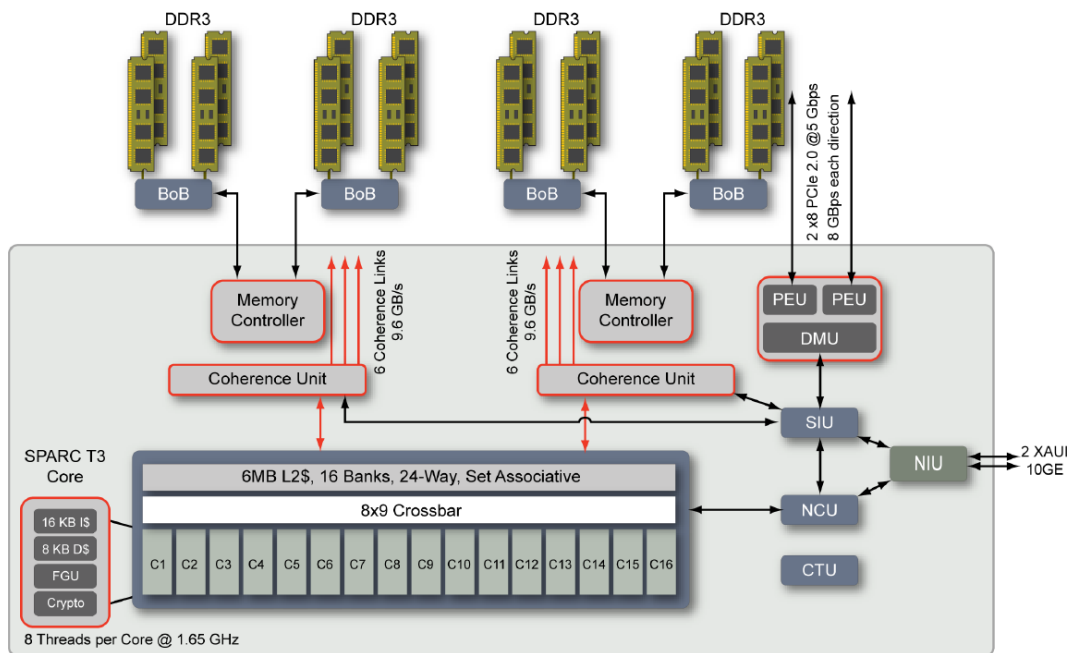


図 3. 多重処理 SPARC T3 プロセッサの SoC 設計は、大規模な計算能力、I/O、ネットワーク、暗号化機能を 1 つのチップに結合しています。

SPARC T3 プロセッサには、次のような設計上の主要な要素が含まれています。

- 2 個もしくは 4 個のプロセッサによるマルチプロセッサの実装をサポート
- 6 個の一貫性リンク
- 1 プロセッサにつき 128 の同時スレッドをサポート
- 1 プロセッサにつき 16 コア。各コアに 2 個の整数処理パイプライン
- 1 コアにつき 8 スレッド
- 16 個の新規に設計され完全にパイプライン化された浮動小数点ユニット (1 コアにつき 1 個)
- 16 個のストリーム処理ユニット (1 コアにつき 1 個。コアと並列して処理する暗号化コプロセッサとして動作)
- オンチップのキャッシュおよびメモリー管理
- 2 個のオンチップ PCIe Generation 2 I/O インタフェース
- 2 個のオンチップ 10 GbE XAUI ポート
- 一貫性リンクと FBDIMM のようなシングルリンクのフェイルオーバー

SPARC T3 プロセッサの SoC 設計では、オンボードコンポーネントの接続のための ASIC を追加する必要が減少します。その結果、SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、複数のシングルスレッドのデュアルコアもしくはクアッドコアプロセッサを搭載した従来の多くのサーバーよりも、障害の起きやすい部品やピンの数が少なくなります。図 4 に、SPARC T3 プロセッサにより劇的に簡素化されたシステム設計を示します。



図 4.16 のコアと 128 のスレッドを持つ SPARC T3 プロセッサは、計算能力を最大化すると同時にシステムコンポーネントの数を最小化します。また、さらに多くのプロセッサ、複数のシステムボード、および非常に複雑な設計を持つシステムよりも、優れた信頼性を提供します。

可用性の改善: 冗長性、保守性、自己診断機能

エラー状態を修正するためのシステム割り込みの必要性を最小にすると、システムとシステムにホストされている IT サービスの可用性は向上します。SPARC T3-1、T3-2、T3-4、T3-1B サーバーには、継続的なシステム運用に影響を及ぼすことなく保守手順を実行できる数多くの設計機能があります。

冗長なホットスワップコンポーネントと RAID 機能

オラクルの SPARC T3 サーバーは、高いレベルの稼働時間と障害からの迅速な復旧を得られるように構築されています。1 プロセッサにつき 2 つの PCIe ルートコンプレックスがあり、さらに複数の CPU (T3-2 と T3-4)、メモリー DDR3 DIMM、および I/O カードを構成できるため、オラクルの SPARC T3-1、T3-2、T3-4、T3-1B サーバーは回復力が向上しています。ホットスワップおよびホットプラグシャーシマウント型ハードドライブ、ファンユニット、および電源装置により、保守性と可用性が改善されています。冗長なコンポーネントがシステムに構成されている場合、管理者はシステム運用を継続しながらソフトウェアコマンドを使用してディスク、電源装置、およびファンユニットの取り外しや交換を行うことができます。

SPARC T3-1、T3-2、T3-4、T3-1B サーバーの組み込み RAID 機能によってデータの冗長性が高まり、追加費用なしでパフォーマンスを向上させることができます。これらのサーバーはオンボードのハードウェア RAID 0/1 をサポートしているため、任意の 2 つの内蔵ドライブ間におけるデータのストライプ化やミラー化が可能です。さらに Oracle Solaris を実行中の SPARC T3-1、T3-2、T3-4、T3-1B サーバーでは、ソフトウェアの RAID 機能を利用できます。たとえば、Oracle Solaris ZFS を内蔵もしくは外付けストレージデバイスに使用することで、シャーシの枠を超えた柔軟性と冗長性を提供できます。

ハードウェア保護—場所を選ばない ECC

ハイエンドシステムの自己診断機能、エラー訂正機能、およびパリティチェック機能により、IT サービスの稼働時間を最大化することができます。SPARC T3-1、T3-2、T3-4、T3-1B サーバーにはメインフレームクラスのプロセッサの RAS 機能があり、オンチップメモリー上のデータ整合性を維持することでシステムの稼働時間を改善します。これは独自の機能であり、他のボリュームマーケットのシステムにはないものです。

SPARC プロセッサの広範なデータ保護機能は、自己診断と訂正処理を行い、エラー状態になってもシステムの継続運用を促進します。SPARC T3 プロセッサを通して、パリティ保護が次の要素とともに提供されます。

- I (命令) キャッシュのタグとデータ
- D (データ) キャッシュのタグとデータ
- ECC データ保護付き内部 L2 キャッシュタグ
- 命令変換ルックアップバッファ (ITLB)
- データ変換ルックアップバッファ (DTLB)
- モジュラー演算メモリーとストアバッファのアドレス
- 一貫性リンク

SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、ソフトウェアとハードウェアの修正フローの組み合わせを利用して、データの整合性を最大化します。エラー発生時には、I キャッシュや D キャッシュのシングルエラー訂正 (SEC) にハードウェアの再取得 (refetch) が使用されます。整数レジスタファイル (RF)、浮動小数点 RF、ストアバッファデータ、トラップスタック、およびその他の内部配列には ECC 保護が提供されます。

SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、連想メモリー (CAM) テクノロジを使用して、ハードウェアベースの仮想アドレス (VA) ルックアップを実装します。物理アドレス変換は RAM に実装されます。パリティ保護は VA ルックアップと物理アドレス変換の両方に提供されます。物理アドレス変換機能のパリティ保護が一般的になってきていますが、可変ページサイズのサポートによってハードウェアベースの VA ルックアップのパリティ保護がさらに困難になっています。ハードウェアベースの VA ルックアップのパリティ保護に対するオラクルのアプローチでは、ほかのプラットフォームでは得られない高い信頼性を提供しています。

堅牢な SPARC T3 プロセッサ

SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、UltraSPARC T2 Plus プロセッサの障害管理機能を活用します。SPARC T3 プロセッサは、コア、スレッド、統合 I/O ルートコンプレックス、プロセッサ一貫性リンク、専用の暗号化および浮動小数点ユニット (FPU) といった主要な要素に対して、障害監視と復旧のメカニズムを提供します。保守作業をスピードアップさせるために、障害メッセージは注意を要する現場交換可能ユニット (FRU)、システムへの影響、修正作業の提案などを明確に提示します。

- スレッドおよびコアのオフライン化。** SPARC T3 プロセッサは、CPU のエラーを診断し、スレッドおよびコアレベルでオフライン操作を実行することができます。その際、隣接するスレッドやコアに影響を及ぼしたり、IT サービスの実行に割り込んだりすることはありません。すべてのスレッドで共有しているプロセッサリソース内で障害が発生した場合は、コア全体がオフライン化され、適切な障害分離が行われます。
- 一貫性リンク。** ECC およびパリティ保護された一貫性リンクは SPARC T3 プロセッサ内のソケットを接続し、キャッシュデータの整合性を確保します。
- メモリー。** SPARC T3 プロセッサのメモリーサブシステムは、メモリーバンク、DDR3 DIMM、およびページレベルでの診断とページリタイアメント機能を提供します。また、各メモリーチャンネルには複数の高速シリアルレーンが各方向に含まれており、シングルレーンの自動フェイルオーバーをサポートして、過度のエラーによりレーンに障害が発生した場合でもシステム運用の続行を可能にします。
- I/O サブシステム。** SPARC T3 プロセッサは I/O サブシステムおよび PCIe ファブリックの障害管理と診断を行います。ルートコンプレックス自体の診断に加えて、プロセッサは PCIe ファブリック診断に対する Oracle Solaris の追加サポートも利用できます。
- 暗号化装置。** ある暗号化装置 (SPU) で障害が診断されると、オフライン化機能が働いて、暗号化ドライバが障害の発生したデバイスを使用するのを防ぎます。そのとき、別のプロセッサコア上にある残りの暗号化装置は使用可能であり、プロセッサのスレッドはアクティブのままです。

Integrated Lights Out Manager (ILOM) による簡潔化された遠隔保守性

エラーの通知が遅れると、システム運用の再開に要する時間が長くなります。また、厄介なシステム保守作業によってシステムの停止時間が長くなり、システムの新たな構成上の問題を引き起こしかねません。ILOM 3.0 ソフトウェアと SPARC T3-1、T3-2、T3-4、T3-1B サーバーに埋め込まれたサービスプロセッサにより、遠隔のシステム管理は簡単になり、管理作業は簡素化され、保守タスクは早くなります。

サービスプロセッサは、サーバー上でサーバーの待機電力を使いながら独立して実行します。したがって、サーバーの OS 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、ILOM ファームウェアとソフトウェアは機能し続けます。Integrated Lights Out Manager は、SPARC T3-1、T3-2、T3-4、T3-1B サーバーの次の状態を監視します。

- CPU の温度状態
- ハードドライブのステータス
- 格納装置の温度状態
- ファンの回転速度とステータス
- 電源装置のステータス
- 電圧状態
- Oracle Solaris ウォッチドッグ、ブートのタイムアウト、自動サーバー再起動イベント

Integrated Lights Out Manager は、専用 Ethernet 接続を介した、管理者による SPARC T3-1、T3-2、T3-4、T3-1B サーバーの監視と制御を可能にし、Secure Shell (SSH)、Web、および Integrated Platform Management Interface (IPMI) アクセスをサポートします。Integrated Lights Out Manager 機能は、端末または端末サーバーへの接続用の専用シリアルポートからでもアクセス可能です。ILOM コマンド行およびブラウザベースのインタフェースは、地理的に分散していたり物理的にアクセス不可能なマシンの遠隔管理を簡素化します。また、Integrated Lights Out Manager は、サーバーのシリアルポートとの物理的な距離の近さを要する診断の遠隔実行も可能です。Integrated Lights Out Manager は、サーバーに関連する別のイベントと同様に、ハードウェア障害についての警告を、電子メールを通じて配布するように設定できます。

環境の監視

システム運用の安定性に、環境の要素は重要な役割を果たします。SPARC T3-1、T3-2、T3-4、T3-1B サーバーの未然保守アプローチでは、過度の温度、システム内の通気の不足、電源装置の故障、およびハードウェアの故障などから内部のコンポーネントを保護し、システムの電源消費について報告します。さらに、SPARC T3-1、T3-2、T3-4 サーバーの Intelligent Fan Control (IFC) 機能により、不要な振動を引き起こすファンの回転速度や頻度を回避することができます。振動を最小化すると、コンポーネントの摩耗を削減できます。

温度センサーは SPARC T3-1、T3-2、T3-4、T3-1B サーバー全体にわたって配置されていて、システムや内部コンポーネントの周囲を取り巻く温度を監視します。ファンの回転速度は温度状態を補正するために自動的に調節されますが、実際には消費電力を抑えるため可能な限り下げられます。

センサーが監視する温度が下限のしきい値を下回るか上限のしきい値を上回ると、監視サブシステムソフトウェアが温度の警告を行います。その温度の状態が変わらずに続き、危険なしきい値に到達すると、監視サブシステムは前と後ろのパネルにあるオレンジ色の保守要求 (Service Required) インジケータを点灯し、システムのシャットダウンを始めます。電力サブシステムも同様で、電源装置は常に監視されており、障害が発生すると前と後ろのパネルにあるインジケータが点灯します。システムコントローラに障害が発生すると、深刻なダメージからシステムを守るため、ハードウェアのシャットダウンが強制的に行われます。自動的にシステムシャットダウンが行われたあともインジケータは点灯したままになっているため、問題の診断に役立ちます。

オラクル仮想化テクノロジーによる障害分離

数多くのアプリケーションを 1 つのサーバーに統合することは、膨大な数の IT サービスのホスティングや管理にかかるコストを削減したい企業にとって必要不可欠です。Oracle VM Server for SPARC と Oracle Solaris コンテナの 2 つのテクノロジーにより、SPARC T3-1、T3-2、T3-4、T3-1B サーバーのかなりのシステムリソースを仮想化することができます。オラクルの仮想化テクノロジーは、統合の方針の範囲内でソフトウェアアプリケーションの障害を分離できます。そのため、その障害が同じプラットフォームにホストされている別の IT サービスに影響を及ぼすことはありません。つまり、オラクルの仮想化テクノロジーによって、停止時間を削減しながらリソースの利用率を改善することができます。

Oracle VM Server for SPARC

どの仮想化ソリューションも、主要なアプリケーションやサービスの RAS への貢献度を注意深く検討する必要があります。ファームウェアとハードウェアの両方に実装すると、Oracle VM Server for SPARC はソフトウェアベースの仮想化ソリューションよりも高い信頼性を実現します。堅牢な障害分離に加え、Oracle VM Server for SPARC には構成内に冗長性を生成するためのオプションがいくつか用意されており、予防保守的な障害管理機能を提供し、ゲストドメインの復旧時間を短縮する独自の I/O アーキテクチャーを実装しています。

Oracle VM Server for SPARC は、ファームウェアとハードウェアの構造全体で広範なリソースの分離を行います。各論理ドメインの作成、破壊、再構成、およびリブートは独立して実行できるため、サーバーの電源を再投入する必要はありません。論理ドメインはまた、完全に独立したマシンとして管理され、次のようなりソース制御はローカライズされます。

- カーネル、パッチ、およびチューニングのパラメータ
- ユーザーアカウント
- ディスク
- ネットワークインタフェース
- MAC アドレス
- IP アドレス

Oracle VM Server for SPARC は SPARC ハイパーバイザ (マザーボードにロードされたシステムファームウェアスタックに含まれるファームウェア層) を使用してマシンハードウェアを仮想化し、OS とハードウェア間のリンクを切り離します (図 5)。つまり、作成可能な仮想マシンの数は、システムにインストールされている物理的なハードウェアデバイスの数ではなく、ハイパーバイザの機能に依存しています。SPARC T3-1、T3-2、T3-4、T3-1B サーバーでは、128 の論理ドメインを確立できます¹。各論理ドメインは完全な仮想マシンとして独立した OS インスタンスや実行環境を実行し、仮想化 CPU、メモリー、ストレージ、コンソールおよび暗号化デバイスを含んでいます。

すべての論理ドメインのインスタンスは、前述のような基盤となるテクノロジー構造に依存していますが、ドメインによっては別の役割も存在します。コンテキストや用途に応じて、1 つの論理ドメインが次の役割のうちの 1 つまたは複数の機能を果たすこともできます。

- **制御ドメイン**は、ドメインの作成と物理リソースの割り当てを管理します。
- **サービスドメイン**は、CPU、メモリー、ネットワーク、ディスク、コンソール、および暗号化装置といったハードウェアリソースへのアクセスをゲストドメインが管理できるように、ハイパーバイザをインターフェースでつなぎます。
- **I/O ドメイン**は、PCIe カード、ストレージ装置、およびネットワークデバイスといった入出力デバイスへの物理的なアクセスを直接制御します。
- **ゲストドメイン**は、サービスドメインや I/O ドメインが提供する仮想デバイスを使用し、制御ドメインの管理のもとで動作します。

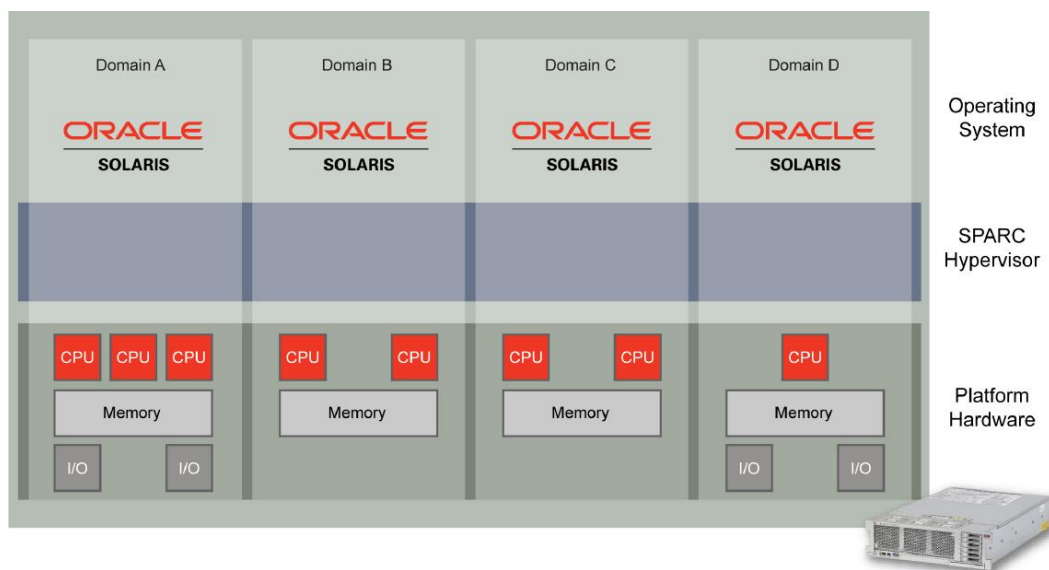


図 5. SPARC T3 ハイパーバイザは、仮想マシンの OS インスタンスと割り当てられたハードウェアリソースを、Oracle VM Server for SPARC を通じて分離します。

¹ 特定のシステムに適した論理ドメインの最適な数は、使用可能なリソースやアプリケーションの集合によって変わります。

論理ドメインの分離

組織や企業は、仮想化テクノロジーを利用して、複数のビジネスクリティカルなアプリケーションを同じサーバーにホストすることがあります。そのため、ある仮想マシンのシステム障害や保守イベントを確実に分離して、別のパーティションにある IT サービスの可用性に影響を及ぼさないようにする必要があります。

多くの仮想化ソリューションでは、制御ドメインやサービスドメインのリブートにシステム全体のリセットが必要です。しかし Oracle VM Server for SPARC では、ほかのすべての論理ドメイン (物理的なハードウェアリソースから直接制御されている論理ドメインを含む) から独立して、論理ドメインのリブートとリセットを行うことができます。マシンの電源を再投入したり、別のドメインの継続実行に影響を及ぼしたりすることなく、ゲストドメインを独立して構成、開始、および停止することができます。さらに仮想 I/O インタフェースは、同じプラットフォーム上の別のドメインに影響を及ぼすことなく、必要に応じて接続や切断を行うことができます。管理者は、OS インスタンスが実行中でも論理ドメインに仮想 CPU を直接追加および削除することもできるため、計画的な停止時間の必要性を回避できます。

堅牢な仮想化 I/O

I/O デバイスとの継続的な通信は、数多くの IT サービスを配信する上で重要です。Oracle VM Server for SPARC では、I/O ドメインは I/O デバイスを物理的に接続する論理ドメインです。I/O ドメインは通常、サービスドメインとしても機能し、仮想デバイス内の別の論理ドメインへの I/O アクセスを共有します。Oracle VM Server for SPARC には I/O ドメインやサービスドメインの障害状態の影響を軽減させるための機能やアーキテクチャーの要素があり、I/O デバイスとの連続したアクセスを可能にします。

- **I/O 障害イベントのマスク。** Oracle VM Server for SPARC には、I/O 接続の一時的な切断による影響を軽減させる機能があります。I/O デバイスにおける障害や保守操作に応じて、Oracle VM Server for SPARC は Oracle Solaris とインタフェースで連結して、影響を受けたすべての仮想 I/O デバイスへの I/O 操作を正常に中断します。物理 I/O デバイスが復旧したあと、仮想 I/O デバイスは再接続し、中断した位置から I/O トランザクションを再開します。
- **冗長な I/O パス。** I/O の中断は障害をマスクする手段の 1 つですが、ほんの短い時間でも I/O サービスを欠かすことができないアプリケーションは数多く存在します。冗長な I/O パスを構成し、Oracle Solaris の I/O マルチパスソフトウェアを使用することにより、1 つの I/O ルートが失われた場合でも、ディスクやネットサービスへの継続的なアクセスが確保されます。SPARC T3-1、T3-2、T3-4、T3-1B サーバーには複数の PCIe ルートコンプレックスが装備されているため、2 つ以上のサービスドメインを構成することができます。各サービスドメインは、追加のホストバスアダプタ (HBA) がインストールされている場合に物理的なディスクサブディスクへの独自のパスを提供することができます。このシナリオは、従来の I/O 接続モデルで 2 つの HBA を採用するのと似ています。一度ゲストドメイン内で初期化されると、Oracle Solaris I/O マルチパスソフトウェアは、障害発生時に、2 つのパス間における自動的なフェイルオーバーのメカニズムを提供します。また、Oracle Solaris I/O マルチパスソフトウェアには手動のパス切り替え機能が含まれているため、管理者はプライマリ I/O ドメインの再構成時やリブート時に I/O を代替パスにリダイレクトすることができます。

- **I/O ドメインのクラスタ化。** クラスタ化テクノロジーを採用することで、仮想化 I/O サービスの可用性をさらに向上させることができます。Oracle VM Server for SPARC テクノロジーを Oracle Solaris クラスタとともに使用することで、I/O ドメインのフェイルオーバーに備えます。I/O ドメインをクラスタ化すると、I/O デバイスサービスの自動フェイルオーバーに備え、I/O 接続の復旧に要する時間を最小限にし、IT サービス全体の可用性を最大限向上させることができます。

ブート時間の短縮

障害もしくは保守のイベントが発生した場合、システムをオンラインに戻すのに要する時間を短縮することは、IT サービスの可用性を最大にするために重要です。論理ドメインに IT サービスをホストすると、実際に復旧時間を短縮することができます。Oracle VM Server for SPARC では、論理ドメインの計画的もしくは計画外のリブート時に、OS インスタンスは時間のかかる I/O 起動手順を省略することができます。ほとんどのドメインでは仮想 I/O デバイスを使用して I/O バスの所有権の負担を委任し、デバイスに対するバスを検索して、I/O ドメインにデバイスドライバをロードします。仮想 I/O デバイスを使用するドメインには、検索するための I/O バストポロジと I/O デバイスへの物理的な接続は含まれていません。OS の起動時に、時間のかかる I/O 初期化ステップの必要性をなくすことで、復旧時間をスピードアップしています。

Oracle Solaris コンテナ

Oracle Solaris コンテナは、Oracle Solaris のシステムイメージとドメインのどちらで使用されても、柔軟なソフトウェア定義の境界を利用してソフトウェアアプリケーションとサービスを分離させることができます。仮想化とソフトウェアのパーティション化における画期的なアプローチであるコンテナテクノロジーにより、Oracle Solaris の 1 つのインスタンス内に数多くのプライベートな実行環境を作成することができます (図 6)。

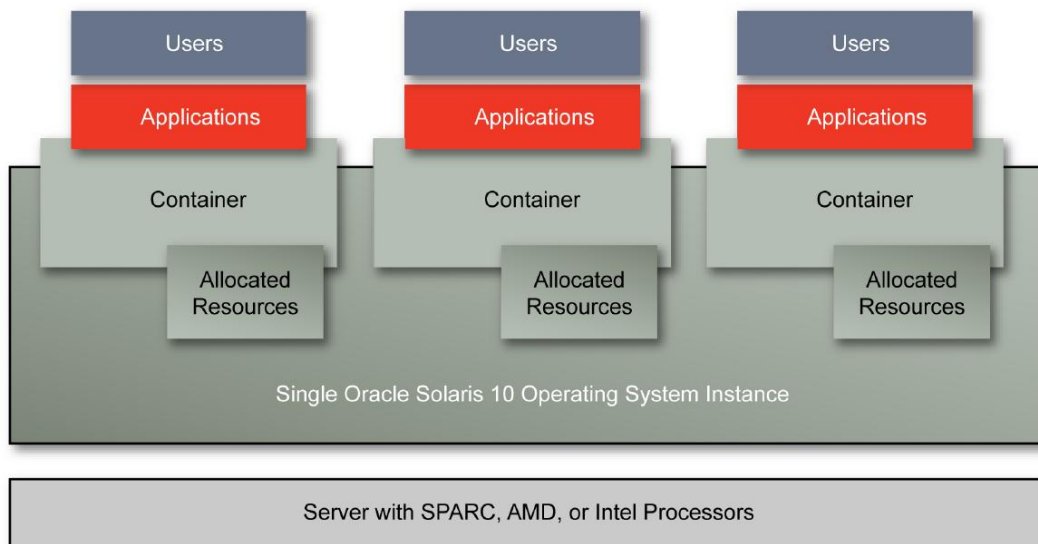


図 6. Oracle Solaris コンテナは、柔軟なソフトウェアメカニズムによってアプリケーションを分離します。

Oracle Solaris コンテナは、完全に分離されセキュリティー保護されたアプリケーションの実行環境を提供し、システムリソースのきめ細かい管理を可能にします。各コンテナは、ユーザー、デバイスパス、CPU とメモリーのリソース、およびネットワークに関して別々に管理できます。動的なリソースの再割り当て機能により、未使用のシステムを必要に応じてコンテナ間でシフトすることで、ホストされているアプリケーションへのサービスの質が向上します。また、リソースを重要なサービスに割り当てたり、予約したりすることで、優先度の低いワークロードとの計算能力の競合を削減できます。コンテナが提供するきめ細かいリソース管理により、高いレベルのサービスの可用性を保ちながら計算能力を効率的に使用することができます。

コンテナの中のアプリケーションは分離されているため、あるコンテナでの動作は、別のコンテナで実行中のプロセスからの監視や干渉から保護されています。スーパーユーザーのプロセスであっても、別のコンテナ内の動作を監視したり、影響を及ぼしたりすることはできません。Oracle Solaris コンテナのソフトウェア障害とセキュリティーを分離する機能はまた、使用頻度の低いアプリケーションが別のコンテナに影響を及ぼすことを防ぎます。Oracle Solaris コンテナと Oracle VM Server for SPARC を別々もしくは組み合わせて使用することにより、組織や企業はアプリケーションとリソースの優先度を動的に制御して、サービスレベルをよりよく定義し、そのレベルを満たすことができます。

IT サービス稼働時間についての広範な焦点

OS や管理ツールの選択は、IT サービスの稼働時間にかなり影響を及ぼします。信頼性の高いソリューションの構築に向けて、オラクルはプラットフォームのあらゆるレベルで RAS を統合しています。つまり、Oracle Solaris 10 には Oracle Solaris の広範な障害管理アーキテクチャーである予測的自己修復機能が含まれており、SPARC T3-1、T3-2、T3-4、T3-1B サーバーのハードウェア、仮想化、および OS 層全体のエラー処理を担っています。これらのサーバー上で Oracle Solaris とオラクルの管理ツールを一緒に使用すれば、ホストされている IT サービスの可用性をさらに高めることができます。

障害管理アーキテクチャーによる診断機能の高速化

障害が発生して保守が必要な場合、IT サービスが運用可能状態に戻るまでに要する時間は重要です。当初 Oracle Solaris に導入された障害管理フレームワークを拡張した自己修復機能が、SPARC T3-1、T3-2、T3-4、T3-1B サーバーのプラットフォームや個々のハードウェアコンポーネント、および Oracle VM Server for SPARC のすべてのテクノロジ層のエラーイベントを管理します。ハードウェアやソフトウェアの障害を予防保守的に診断、分離、および復旧させることによって、予測的自己修復テクノロジは障害に関する重要情報を提供し、IT サービスの再開を高速化して、特定のシステム障害を防止するのに役立ちます。

予測的自己修復

Oracle Solaris の予測的自己修復機能は予防的にシステムコンポーネントの監視と管理を行うため、IT サービスの可用性を最大限高めることができます。予測的自己修復は革新的な機能であり、多くのハードウェアやアプリケーションの障害を自動的に診断、分離、および復旧させます。その結果、ビジネスクリティカルなアプリケーションや必要不可欠なシステムサービスは、ソフトウェア障害やハードウェアコンポーネントの大きな故障、およびソフトウェア構成の問題などが発生した場合でも処理を継続して行うことができます。

Oracle Solaris の障害マネージャーとサービス管理機能は、予測的自己修復機能の主要なコンポーネントです。障害マネージャーはハードウェアやソフトウェアのエラーに関するデータを受け取り、自動的にその潜在的な問題を診断します。問題が診断されると、障害マネージャーは自動的に反応し、障害の発生したコンポーネントをオフライン化して管理者にエラー状況を示す信号を送ります。信号は、コンソールメッセージ、およびシャーシの前と後ろにあるシステムとコンポーネントのステータスインジケータを介して伝えられます。サービス管理機能は、プロセスではなくサービス自体に自動的な自己修復を許可します。Oracle Solaris のベースサービスの記述には、開始、停止、再開についてのすべての依存性情報が含まれています。サービス管理機能のもとで実行するようにユーザーアプリケーションを構成することは比較的簡単であり、これを行うことによって SPARC T3-1、T3-2、T3-4、T3-1B サーバーにホストされた IT サービスの障害を効率的に管理することができます。

SPARC T3-1、T3-2、T3-4、T3-1B サーバーの予測的自己修復

SPARC T3-1、T3-2、T3-4、T3-1B サーバーは、たとえば syslog、CPU オフライン、MPR、および I/O リタイアエージェントといった、Oracle Solaris にすでに含まれている複数の予測的自己修復応答エージェントを活用します。サービスプロセッサ上の ILOM ソフトウェアによって応答エージェントがさらに追加されます。これらのエージェントには、エラーイベントの発生した障害 FRU を更新する動的現場交換可能ユニット ID (FRUID) のエージェントと、特定の障害に対応したシステムおよびコンポーネント (FRU) のステータスインジケータを点灯する LED エージェントが含まれています。障害の診断と復旧のステータスは、論理ドメインとサービスプロセッサの間で同期をとられます。SPARC T3-1、T3-2、T3-4、T3-1B サーバーの障害イベントには、該当する FRU 部品と障害コンポーネントのシリアルナンバーが含まれています。

次の機能は、SPARC T3-1、T3-2、T3-4、T3-1B サーバーに実装された予測的自己修復機能の粒度と性能の一例です。

- SPARC T3 プロセッサには、レベル 1 (L1) とレベル 2 (L2) のキャッシュ、スレッド毎のレジスタ、TLB、および暗号化装置にエラーレポートメカニズムが含まれており、予測的自己修復テクノロジーによって診断可能です。
- 修正可能なメモリーエラーが SPARC T3-1、T3-2、T3-4、T3-1B サーバーで発生した場合、そのエラーは予測的自己修復診断の待ち行列に入ります。必要に応じたメモリーページの自動リタイアメントと DDR3 DRAM エラーの予測的分析や自己修復により、潜在的なシステムの割り込みが回避されます。
- 予測的自己修復機能によって管理されている SPARC T3-1、T3-2、T3-4、T3-1B サーバー内の障害には、ハイパーバイザによって仮想化されたデバイスと同様に、強化および非強化ドライバを持つ動的 I/O デバイスが含まれます。

- 一貫性リンクの単一レーンのフェイルオーバーが SPARC T3-1、T3-2、T3-4、T3-1B サーバー内で引き起こされた場合、その障害の内容は記録され、インジケータのライトは点灯し、システム運用は継続されます。
- PCIe スイッチと SAS/SATA コントローラは、強化されたデバイスドライバを使用して、エラーハンドラがエラーを検出した時にレポートを作成します。エラーは完全な分析と診断に転送されます。

ドメインの予測的自己修復

仮想環境の中でエラー処理を正しく行うためには、メッセージや警告を仮想マシンインスタンスにまで到達させる必要があります。予測的自己修復機能は、SPARC T3-1、T3-2、T3-4、T3-1B サーバーのハードウェアを密接に統合するだけでなく、ドメイン内における障害処理を管理します。システム全体を見渡して、プロセッサ、メモリー、コンソール、暗号化装置といったハイパーバイザが抽出したデバイスのエラー状況をハイパーバイザに到達し、標準エラーレポートを作成して host-to-service-processor メールボックスチャンネルを通じてサービスプロセッサに送ります。診断エンジンは各エラーレポートの完全な分析を行い、潜在的に影響を受ける OS インスタンスに警告を送信します。Oracle Solaris では予測的自己修復にナレッジベースが役立ちます。ナレッジベースを使用して、Oracle VM Server for SPARC の実行に関連するエラーメッセージを文書化された事前定義済みの訂正動作に関連付けることで、システムの復旧時間を短縮できます。

Oracle Solaris での停止時間の削減

最も信頼性の高い OS の 1 つである Oracle Solaris の開発に、20 年以上もの歳月を費やしてきました。主要な計算要素、すなわち OS、ネットワーク、ユーザー環境は Oracle Solaris 内で合体して安定した高品質の IT サービスの基盤を提供しています。実際に、多くの組織や企業では、Oracle Solaris を実行しているシステムが数ヶ月もしくは数年にわたり再起動を必要とせず、継続して実行していると述べています。

Oracle Solaris は可用性を重視して設計されています。構築されているカーネルは小さくてコンパクトなため、OS の障害やそれに付随して起こるプラットフォームの停止時間の潜在性を制限することができます。さらに Oracle Solaris はカーネル、共有ライブラリ、アプリケーションを明確に区別することで、アプリケーション障害による影響を制限しています。また、ほとんどのパッチや増加しているソフトウェア更新を Oracle Solaris に適用する際にシステムをオフライン化する必要がないため、稼働時間が増えて保守性が向上します。Web ベースのインストールやグラフィカルなプロセスマネージャーなどの使いやすい機能により、オペレータによるエラーの危険性を減らし、保守時間を最小限に抑えることができるため、可用性を最大限に引き上げることができます。非常に高い信頼性、持続的な可用性、簡素化された保守性を伴って IT サービスを開発、展開、および管理するのに役立つ Oracle Solaris の主要な機能について、以降の節で説明します。

メモリーページリタイアメント

多くのシステムでは、修正可能な永続メモリーと修正不可能な永続メモリーのどちらのエラーに対処する場合でも、サーバーの停止時間がある程度必要になります。予測的自己修復テクノロジフレームワークの中には MPR 機能があり、システム割り込みを行わずにメモリーの問題箇所を分離することができます。障害マネージャー機能の診断ソフトウェアは、配下のハードウェアから検出されたメモリーの修正可能なエラーと修正不可能なエラーを継続的に調べます。MPR は、修正可能なエラーを持つメモリーページと、修正不可能なエラーを含む再配置可能なクリーンページを、ユーザーアプリケーションに割り込むことなくリタイアさせます。また MPR は、ユーザープロセスへの影響を最小限に抑えながら修正不可能なエラーを含む再配置可能なダーティーページを分離して、システム全体の機能停止を防ぐことができます。

Oracle Solaris ファイルシステム

信頼性の高いデータサブシステムは、高可用性 IT サービスを作成する上で重要です。組織や企業は、メタデータのロギングのような高い回復力機能を提供したり、システム障害時にデータ破壊を防いだり、迅速に復旧したりする際に、依然として Oracle Solaris 10 の UNIX ファイルシステムに依存しています。Oracle Solaris 10 には、Oracle Solaris ZFS が含まれています。Oracle Solaris ZFS は、データ整合性への革新的なアプローチによってデータ管理が劇的に進化したファイルシステムです。

Oracle Solaris ZFS は管理上のエラーに対して、さらに高度な保護機能と、64 ビットのチェックサム機能や広範なデータ更新といったエンドツーエンドのデータ整合性要素を提供します。ディスク上のデータの自己整合性を常に確保するために、Oracle Solaris ZFS は、コピーオンライト機能やエンドツーエンドのチェックサム機能といった証明済みの最先端テクノロジを提供します。データは常に、データへのポインタが変更されて書き込みがコミットされる前にディスク上の新しいブロックに書き込まれます。ファイルシステムは常に一貫しているため、システムが正常ではない方法でシャットダウンされたときでも、ファイルシステムのチェック (fsck) のような時間のかかる復旧手順は必要ありません。コピーオンライト機能を使用すると、管理者は一貫したバックアップを作成したり、任意の時点にデータをロールバックしたりすることができます。

Oracle Solaris ZFS の組み込まれた Oracle Solaris 10 は、すべてのデータに対してエンドツーエンドのチェックサムを行う唯一の OS です。Oracle Solaris ZFS は定期的にデータを読み込んでチェックすることで整合性を確認し、ミラープールにエラーが存在すれば破壊されたデータを自動的に修復します。可用性を保つためのこの絶え間ない警戒により、費用も時間も消費するデータ損失や、以前は検出できなかった暗黙のデータ破壊を防ぎます。

Oracle Solaris フラッシュアーカイブと Oracle Solaris Live Upgrade

計画的な停止時間が一年間のシステム割り込みの大半を占めることがよくあります。Oracle Solaris と Oracle Solaris Live Upgrade のフラッシュアーカイブ機能を使用して効率よくインストールやアップグレードを行うことで計画的停止時間の必要性を削減できます。フラッシュアーカイブと Oracle Solaris Live Upgrade により、SPARC T3-1、T3-2、T3-4、T3-1B サーバーの展開やアップグレードをほんの数分で完了させることができます。

- フラッシュアーカイブ機能により、IT 関連の組織や企業では、ニーズに応じて調整された OS 構成を使用して、システムのインストールおよび更新を迅速に行うことができます。フラッシュアーカイブテクノロジーのツールにより、システム管理者はインストール可能なカスタムイメージ (アプリケーション、パッチ、パラメータなど) を迅速に構築し、ハードウェアのフルスピードに近いデータ転送率でインストールすることができます。
- Oracle Solaris Live Upgrade ソフトウェアは、Oracle Solaris の複数のオンディスクインスタンスをアップグレードおよび管理するためのメカニズムを提供します。このテクノロジーにより、サーバーをオフライン化することなく実行中の本稼働システムに新しい OS やパッチをインストールすることができます。停止時間が必要なのは、新しい構成をリブートするときだけです。Oracle Solaris Live Upgrade には、管理者が必要に応じてシステムを初期状態に迅速にロールバックすることができる機能もあります。

オラクルのソフトウェアによる効率的なシステム管理

管理手順を自動化するテクノロジーは、障害を防ぎ、システム管理者が膨大な数のサーバーを効率的に管理するのに役立ちます。その際、ILOM ソフトウェアに加えて、強力な Oracle Enterprise Manager Ops Center を利用することができます。この洗練されたツールは監視機能と管理機能を自動化し、管理の負担と一般的なエラーの発生率を軽減します。Oracle Enterprise Manager Ops Center は、複雑なソフトウェアのインストールと構成を自動化するため、運用が簡単になります。

Oracle Enterprise Manager Ops Center

急速に変化する IT インフラストラクチャーで高いレベルの可用性を提供するには、プロセスエラーの可能性を最小限に抑えながら、サーバーのプロビジョニングと再プロビジョニングを効率的に行う能力が必要です。Oracle Enterprise Manager Ops Center は異種ネットワークインフラストラクチャーでの複雑なシステムインストールおよび構成の作業を自動化し、IT サービスの展開を高速化および簡素化します。Oracle Enterprise Manager Ops Center は、システムインフラストラクチャーのライフサイクル管理を簡素化するための広範なソリューションを提供します。Oracle Enterprise Manager Ops Center を使用することにより、オラクルの Sun SPARC アーキテクチャーに基づいた x64 ベースの数百ものサーバーを 1 つのコンソールから検出、プロビジョニング、監視、更新、および管理することができます。Oracle Enterprise Manager Ops Center ソフトウェアではまた、企業の需要の変化に応じたリソースの割り当てや割り当て解除を柔軟に行うことができます。

データセンターの管理タスクを簡素化する目的で設計された Oracle Enterprise Manager Ops Center には、電力の遠隔制御、OS の展開とパッチ適用、システム BIOS およびファームウェアの更新、イベントのロギングと通知、およびハードウェアと OS の監視の機能が含まれています。さらに、このソフトウェアでは、システムの論理グループを作成し、複数グループ間での処理実行を、1 つのノード上で実行しているのと同じくらい簡単に行うことができます。Oracle Enterprise Manager Ops Center はシステムの監視と保守を簡単かつ迅速に実行できるため、運用の効率性を高め、プロセスエラーによる停止時間を最小限に抑えます。

サーバーの RAS の計測

オラクルは、システム運用を全体的に見渡し、信頼性と可用性と保守性 (RAS) の相互関係を研究して、3 つの要素をすべて同時に最適化する努力を続けています。新しいシステムの設計に繰り返しメトリックスを使用することにより、オラクルは、現世代や前世代のオラクルサーバーや競合システムを超えた継続的な改善を達成しています。システムの可用性の重要度が高まり、オラクルは計測の業界標準を確立するための投資と企業努力を積極的に行なっています。

可用性に関するベンチマークのフレームワーク

ベンチマークは、コンピュータシステムのパフォーマンスを向上させるためのリサーチと開発に役立ちます。システムの RAS 機能の重要性が増しているため、エンジニアが可用性レベルのベンチマークを計測したいと望むのは自然な流れです。しかし、この重要な分野において業界が認める標準は、依然として完成していません。システムのディペンダビリティと計測に関する業界標準の欠如を改善するための組織的なアプローチを取り入れる目的は、R-Cubed² と呼ばれる可用性ベンチマークのフレームワークを作成することにあります。オラクルの R-Cubed 可用性ベンチマークのフレームワークはさまざまなシステムに適用でき、方法論についての情報は、産業、学術機関、技術的なコミュニティー全体が再利用できるように、広く公開されています。

サーバーが停止した時間を 1 年あたりの分単位で測るという従来の計測に加えて、R-Cubed の可用性ベンチマークのフレームワークでは、ディペンダビリティ全般、つまりシステムが使用可能状態を続け、サービスの割り込みを防ぐ能力にも焦点を当てています。したがって、R-Cubed フレームワークは次の 3 つの主要な属性を計測することで、システムの RAS レベルを示します。

- **障害および保守率**は、特定の時間内におけるシステム障害の頻度を計測します。
- **堅牢性**は、障害によるシステム運用の機能低下の範囲と、オンラインで復旧を行う可能性について示します。
- **復旧**は、システムの特性を調べ、障害や保守イベントの後で運用状態に戻すのに必要な作業の量を計測します。

R-Cubed フレームワークの 3 つの要素を最適化すると、システムのディペンダビリティは高くなりますが、この 3 つの R-Cubed の要素をすべて同時に向上させるのには注意が必要です。たとえば、サーバーを少ないコンポーネント数で設計すると、障害率や保守率が下がり信頼性は高まりますが、冗長な部品による堅牢性が低下することになりかねません。冗長性がないと、1 つのコンポーネントエラーや障害でもシステム停止を引き起こすことがあり、可用性が高まる代わりに信頼性が低下してしまいます。R-Cubed フレームワークを使用することにより、オラクルはバランスのとれたアプローチによってシステム全体の RAS を最大限引き上げます。

² 詳細は “R-Cubed (R3): Rate, Robustness, and Recovery—An Availability Benchmark Framework,” (labs.oracle.com/techrep/2002/abstract-109.html) を参照してください。

R-Cubed のメトリックス

R-Cubed の可用性フレームワークには、次の計測が含まれており、システム設計者がシステム設計のディペンダビリティを測り潜在的な改善を評価するのに役立ちます。

- **障害堅牢性のベンチマーク A (Fault Robustness Benchmark-A, FRB-A)**。障害が起きてもサービスに支障をきたさないシステムに報いるベンチマークです。計測結果は 1 から 100 までの数値で表されます。1 は 1 つの障害がサービスに支障をきたすことを意味し、100 は 1 つの障害では支障をきたさないことを意味します。コストと冗長性のトレードオフを最適化すると、この値は高くなります。つまり、コストを重視する場合には、信頼性の比較的低い部品は冗長にしておき、信頼性の高い方の部品を使用します。
- **保守堅牢性ベンチマーク A (Maintenance Robustness Benchmark-A, MRB-A)**。システムに支障をきたさずに保守できる能力を数値で表します。1 はすべての保守作業がシステム停止を引き起こし、100 はすべての現場交換可能ユニット (FRU) を停止時間なしで交換できることをそれぞれ意味しています。MRB-A の数値は、ホットスワップ可能なシステムに対しては高くなります。一般に、スモールフォームファクタ (SFF) サーバーに対する MRB-A の数値は、オンラインハードウェアサービス向けに設計された大きなサーバーほどは高くなりません。
- **平均サービス間隔 (Mean Time Between Services, MTBS)**。平均故障間隔率 (MTBF) を使用して計算を分離し、サービス例外を招く FRU 故障のみを対象とします。
- **平均システム割り込み間隔 (Mean Time Between System Interrupts, MTBSI)**。システムの停止時間を招くコンポーネント故障と、保守作業のための計画的シャットダウンを数量化します。MTBSI は計画的システム割り込みと計画外システム割り込みを分けて認識します。時として、1 回の計画外割り込みによってシステム運用状態が低下し、故障したコンポーネントを交換するために計画的割り込みが 1 回必要になることがあります。
- **計画外平均システム割り込み間隔 (Unscheduled Mean Time Between System Interrupts, U_MTBSI)**。コンポーネント故障によるシステム割り込み率を計測します。計画外割り込みは、予測的サービス実施に対する最大の脅威となります。
- **可用性**。1 つのシステムの 1 年間の稼働時間率を計測します。

表 2 のデータと図 7 の棒グラフでわかるように、SPARC T3-1、T3-2、T3-4 サーバーは、2 ~ 3 年前にリリースされた Sun Fire V490 サーバーのような旧 Sun システムと比較して、最高の信頼性に向けて格段に進歩を遂げています。

表 2. オラクルの SPARC T3-1、T3-2、T3-4、および Sun Fire V490 サーバーの R-Cubed メトリックス値

システム	障害 堅牢性 ベンチマーク-A	保守 堅牢性 ベンチマーク-A	U_MTBSI	MTBSI	MTBS	可用性
オラクルの SPARC T3-1 サーバー	86.6	53.4	366972	148,526	56220	0.99998
オラクルの SPARC T3-2 サーバー	89.1	62.3	269428	78090	34571	0.99996
オラクルの SPARC T3-4 サーバー	85.1	27.3	143172	40612	27057	0.99993
オラクルの Sun Fire V490 サーバー	58.7	32.77	161,804	83,539.7	7,905	0.99998

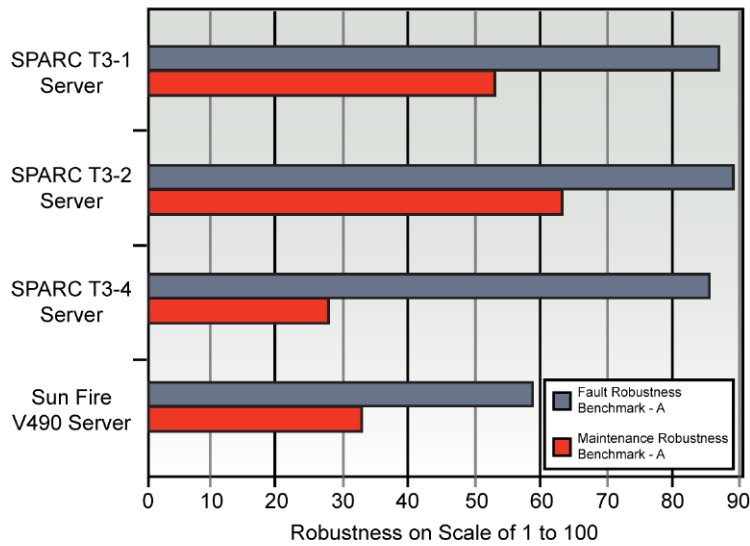
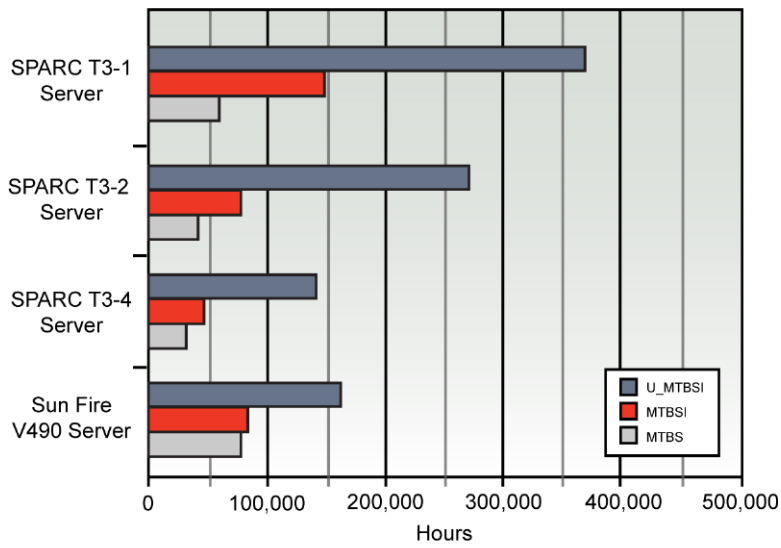


図 7. これらのグラフは、オラクルの T3-1、T3-2、T3-4、および Sun Fire V490 サーバーの R-Cubed ベンチマーク結果を示しています。

まとめ

IT サービスへの高い可用性の必要性はますます高まっています。SPARC T3-1、T3-2、T3-4、および T3-1B サーバーは、IT サービス稼働時間の高い目標を達成するための企業努力をサポートいたします。少ないコンポーネント数、広範なデータ整合性機能、および高いエネルギー効率により、信頼性を高めています。冗長なコンポーネントは高レベルの可用性を促進し、ILOM、ホットスワップ可能コンポーネント、および自己修復機能は保守性を簡素化します。

広範な設計アプローチの一環として、SPARC T3-1、T3-2、T3-4、および T3-1B サーバーのプロセッサ、仮想化テクノロジー、OS、およびソフトウェアツールのすべてが革新的な RAS 機能を提供しています。Oracle VM Server for SPARC と Oracle Solaris コンテナテクノロジーを使用したプラットフォームの仮想化により、IT サービスを障害から分離して資産運用を最適化させることができます。

Oracle Solaris では、以前は大規模システムでのみ使用可能であった MPR や拡張 ECC 保護といった RAS 機能を使用できます。Oracle Solaris の予測的自己修復機能は SPARC T3-1、T3-2、T3-4、および T3-1B プラットフォーム、SPARC T3 プロセッサ、および Oracle VM Server for SPARC に拡張されており、障害発生時におけるシステムの迅速な復旧を促進します。オラクルはまた Oracle Enterprise Manager Ops Center のように管理上のエラーを防ぎ、高い保守性を実現し、復旧時間の短縮に役立つツールも提供しています。

オラクルはソフトウェア製品やハードウェア製品の開発に企業の専門性と革新をもたらします。どの新世代システムも、プラットフォームの信頼性、可用性、保守性を改善しています。したがって、オラクルのシステムは、組織や企業が求める IT サービスの無停止運用のための強力な基盤を提供しています。



オラクルの SPARC T3-1、SPARC T3-2、SPARC T3-4、および SPARC T3-1B サーバーによるサービス稼働時間の最大化
2010 年 10 月

Oracle Corporation
本社所在地
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

米国以外からのお問い合わせ先:
電話: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com



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

Copyright © 2010, 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。本文書は一切間違いがないことを保証するものではなく、さらに、口述による明示または法律による黙示を問わず、特定の目的に対する商品性もしくは適合性についての黙示的な保証を含み、いかなる他の保証や条件を提供するものではありません。オラクル社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。本文書はオラクル社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 は米国 Oracle Corporation 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。

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。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 の商標はライセンスをもとに使用し、SPARC International, Inc の登録商標もしくは商標です。UNIX は X/Open Company, Ltd. からライセンスされている登録商標です。

SOFTWARE. HARDWARE. COMPLETE.

