



# PRIMEQUEST サーバ:Linux と Windows によるミッションクリティカルサーバ

2006年3月

## 提供先

富士通

## 目次

概要	1
はじめに	1
業界標準でありながら他に類のない アプローチ	3
PRIMEQUEST ファミリーの概要	4
PRIMEQUEST の構成	6
データのミラーリング	8
柔軟なパーティション機能	9
IDEAS の結論	11

## 概要

2005年4月、富士通は PRIMEQUEST の販売開始により、エンタープライズサーバの製品ラインを拡充しました。Linux および Windows のオペレーティング環境や、Intel プロセッサを搭載したサーバに魅力を感じるユーザが増加していることに富士通は以前から着目していました。Itanium 2 搭載の PRIMEQUEST は、ミッションクリティカルな Linux 環境や Windows 環境で高パフォーマンスサーバを必要とするユーザのニーズを満たすものです。

富士通の PRIMEQUEST は、8 プロセッサモデルの PRIMEQUEST 420 が追加されたことにより、中規模からハイエンドの範囲に相当する Itanium ベースのサーバを提供しています。PRIMEQUEST は、ミッションクリティカル環境で SPARC/Solaris PRIMEPOWER システムと相互に補完しあうものです。現在リリースされている Madison 9M チップを搭載したモデルとして、新たに加わった中規模機の PRIMEQUEST 420、16 プロセッサの PRIMEQUEST 440、そして 32 プロセッサまで拡張可能な PRIMEQUEST 480 という製品ラインナップになります。

PRIMEQUEST は標準の Intel プロセッサを採用していますが、だからといって PRIMEQUEST が競合他社の Itanium ベースのプラットフォームと同等であるということではありません。メインフレームや RISC プラットフォームを設計してきたエンジニアのノウハウを活用して、ミッションクリティカルな運用を保証する、独自性のあるサーバを開発したのです。このホワイトペーパーでは、他の競合製品とは一線を画す PRIMEQUEST の先進的な設計に焦点を当てます。

## はじめに

長年メインフレームやエンタープライズクラスの UNIX サーバを開発してきた歴史から、富士通はビジネスに不可欠なコンピューティングのニーズを満たす堅牢なシステムをユーザに提供する業界リーダーとしての地位を確立しています。

本文書は Ideas International, Inc. (IDEAS) に著作権があり、米国法および国際著作権法と条例によって保護されています。本文書は、書面による IDEAS の同意なしに、コピー、複製、検索システムへの保存、別の形式への変換、公共および個人のウェブサイトまたは掲示板への掲載、または第三者へのライセンス再譲渡を行うことは禁じられています。本文書の著作権表示を隠すことや削除することは禁じられています。本文書に記載されているすべての製品の商標と登録商標および社名は保護されています。

本文書は信頼性があると思われる情報、およびソースに基づいて作成されています。本文書は改変しないで使用してください。IDEAS は、本文書に含まれるデータ、題材、品質の正確性および内容の適時性について保証するものではなく、またそれらの責を負うものではありません。本文書に含まれるデータは、変更される可能性があります。IDEAS にはデータの変更について、読者に情報を提供する責任はありません。さらに、IDEAS は本文書に記載された製品、サービス、および会社に関する所見を変更することがあります。

IDEAS は本文書に含まれる情報に基づいて下された決定や、また性能結果および他の成果に対する読者による再現の試みについても、責を負うものではありません。本文書は、将来価値や業績水準を予測するために使用しないでください。本文書において取り上げた製品またはサービス、あるいは本文書に記載されているベンダーが提供した他の製品やサービスを保証するために使用しないでください。

**INTEL は ITANIUM への関与を継続**

Itanium 2 を採用していない一部のベンダーは、Itanium をニッチ製品であると位置付けようとしています。このようなベンダーは、Itanium の出荷量は Xeon に対抗できないとし、Intel はハイエンドプロセッサをエンハンスすることに興味を失ったと主張します。

明らかに、これは真実ではありません。Intel は、ハイボリュームチップとエンタープライズチップの両方のファミリーの拡大を図ることを繰り返し強調しています。Intel は、エンタープライズサーバや業界標準サーバには、要件の異なるチップ設計が必要であるとしています。確かに、小規模な業界標準サーバはハイボリュームで出荷されることとなりますが、エンタープライズサーバは高価であるため、Intel にとっては非常に利益の上がる市場になる可能性があります。

Intel は業界共同体の「Itanium Solutions Alliance」を熱心にサポートしており、Itanium 2 とその後続チップの資金として数十億ドルを投じ続けています。富士通も、Itanium Solutions Alliance の創立メンバーであり、Itanium はビジネスクリティカルな集中コンピューティング環境に最適であることを実証しています。

富士通は PRIMEPOWER の継続的なエンハンスと SPARC64 へのコミットにより、ユーザをサポートし続けます。事実、Sun Microsystems は今後共同で販売する SPARC ベースの Solaris サーバファミリーの開発については、富士通の助力を仰いでいます。

また、64 ビット化が始まって以来、ハイボリュームの業界標準ソリューションに対する市場要求が強くなっていると富士通は感じています。PRIMERGY は、Intel Xeon と AMD Opteron プロセッサ上で Windows および Linux が稼動します。PRIMERGY は、先進的なブレードソリューションからスケーラブルな中規模システムまでに対応しています。

富士通は、Linux 上でのエンタープライズアプリケーションの利用をユーザが望んでいるということを把握しています。オープンソースの Linux コードが徹底的に調査されているという事実から、これらのユーザは Linux がセキュアな環境を提供できるという安心感を持っています。例えば、政府機関は特定のベンダーによってロックインされることよりも、Linux が提供するオープンでコラボレイティブな環境を選択しています。さらに、Linux や Windows は複数のプラットフォーム上で稼動し、広範な開発者コミュニティがあり、拡張アプリケーションの種類も豊富です。

PRIMEQUEST は業界標準に対応したソリューションをミッションクリティカルエンタープライズ領域へ広げます。この Itanium 2 ベースのファミリーは、富士通のノウハウを活用して設計された、可用性に富む強力なサーバです。PRIMEQUEST を導入することで、エントリや中規模モデルで運用されることの多かった Linux または Windows 上で信頼性を求められるエンタープライズアプリケーションを実行できます。

2005 年 4 月、富士通は PRIMEQUEST ファミリーの先駆けとして、ハイエンドの PRIMEQUEST 440 および 480 モデルを発表しました。現在、新規追加された PRIMEQUEST 420 により、富士通のミッションクリティカルな Itanium サーバは中規模サーバにまで拡大しました。最大で 8 プロセッサまで構成できる PRIMEQUEST 420 モデルは、1~8 個の Itanium Madison プロセッサをサポートしており、ユーザに可用性の高い中規模サーバを提供します。PRIMEQUEST 製品ラインは、420、440、および 480 モデルで構成されており、Itanium プロセッサを最大 32 個まで搭載できます。富士通によれば、デュアルコアの Montecito チップが入手可能になった際には、プロセッサ能力も 64 プロセッサコアまで拡張されるということです。

PRIMEQUEST 440 と 480 は、2005 年夏から出荷されています。すでに PRIMEQUEST を導入している富士通ユーザからは、機能に満足していると報告されています。しかしながら、PRIMEQUEST の機能をまだよく知らない潜在的なユーザが多くいます。

富士通のコンピューティングシステムは、世界的に高い評価を受けています。富士通は、日本で 50 年以上もコンピュータを供給してきており、コンピュータシステムの主導的な供給者であり続けています。富士通と

シーメンスのパートナーシップは、ヨーロッパにおいてより高い評価を受けています。北米ではメインフレームサプライヤとして重要な役割を担っていたことがあります。しかし、メインフレームが特化された役割に進化したため、ユーザはハイボリュームのプラットフォームのサプライヤに注目していきました。現在、ユーザは再び信頼性の高いシステムを探しています。PRIMEQUEST はビジネスに不可欠なコンピューティングのニーズに対応しています。このホワイトペーパーでは、PRIMEQUEST をまだよく知らないユーザのために、富士通の Itanium サーバファミリーの利点を取り上げます。

## 業界標準でありながら他に類のないアプローチ

ユーザの視点からすると、業界標準のプラットフォームを選択することによって、複数のベンダーの提案を検討できるという自由が得られます。同時に、ベンダーは業界標準に準拠しながら、自社システムの差別化を図るという課題に取り組んでいます。

富士通の場合、PRIMEQUEST はミッションクリティカル領域に適した堅牢なサーバを提供することに焦点を当てて差別化を図っています。事実、ハイエンドサーバは、ハイエンドプロセッサによって約束されたパフォーマンスを提供できなければなりません。富士通では、メインフレーム、スーパーコンピュータ、および 128-way の SPARC SMP を開発したエンジニアを活用して、遅延時間が低く、高帯域幅のシステムバスを PRIMEQUEST に組み込みました。エンジニアたちは、パフォーマンスの設計だけでなく、冗長性を通じて高い可用性を実現しました。富士通が設計したチップセットは、アドレスバスやデータバスだけでなく、メモリまでもミラー化して構成できます。

PRIMEQUEST システムのミラーアーキテクチャは、Standard Mirror Mode と Extended Mirror Mode の 2 つのモードを提供します。これらのいずれも、ロックステップで稼動する冗長ハードウェアを構成します。その結果、これらのコンポーネントのいずれかに障害が発生しても、PRIMEQUEST のパーティションには障害が発生しません。Standard Mirror Mode は、冗長システムバスを提供します。Extended Mirror Mode は、複数のメモリビットエラーまたはメモリモジュール全体の障害からの保護として、メモリまでも二重化します。

また、PRIMEQUEST はホットプラグもサポートします。例えば I/O ユニット、ファンおよびファントレイ、電源装置、System Management Board、オペレータパネル、およびキーボード/ビデオ/マウス(KVM)ユニットなど、ほとんどのコンポーネントがホットプラグ可能となっています。

また、PRIMEQUEST はクラスタ化してさらに可用性を高めることができます。従来の物理的に別個のサーバを 2 台以上で構築するクラスタだけ

ではなく、PRIMEQUEST 440 と 480 は、同一サーバ内のパーティション間クラスタもサポートしています。

PRIMEQUEST の高可用性機能は、エンタープライズ市場の他の競合サーバとは一線を画しています。このホワイトペーパーでは、これらの競合サーバより優れた利点である PRIMEQUEST の特徴のいくつかを取り上げます。<sup>1</sup>

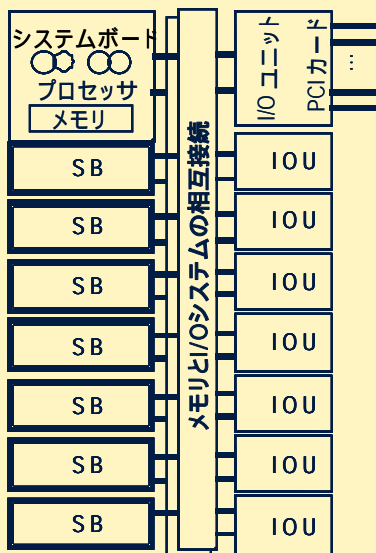
## PRIMEQUEST ファミリーの概要

PRIMEQUEST ファミリーは、拡張性、共有メモリ、シンメトリックマルチプロセッサ(SMP)を有する 3 つのモデルから構成されます。このシステムは、システムボード(SB)、I/O ユニット(IOU)、クロスバー、および System Management Module というモジュール型の要素によって構成されています。各システムボードには、4 つのプロセッサチップと 32 のメモリ DIMM スロットがあります。このシステムボードはクロスバースイッチに接続されているため、すべてのメモリと I/O はすべてのプロセッサからアクセスできます。図 1(次ページ)に、PRIMEQUEST ファミリーのブロック図を示します。PRIMEQUEST 480 は、8 つのシステムボードすべてを装着でき(図に示したとおり)、32 個の Itanium 2 プロセッサチップを使用できます。PRIMEQUEST 440 は 4 つのシステムボードと 4 つの IOU から構成されており、最大で 16 プロセッサのシステムとして構成できます。最近発表されたモデル 420 は、最大で 2 つのシステムボードと 2 つの IOU を装着でき、1~8 個の Itanium 2 プロセッサチップを持つ中規模システムを構成できます。

一見したところ、図 1 は他のモジュール式のスケラブルな SMP のハイレベルのブロック図と同様です。しかし、富士通がスーパーコンピューティングやメインフレームのノウハウを生かして、競合製品より低い遅延時間でより高い帯域幅を得られるクロスバースイッチを開発したことは、ブロック図だけではわかりません。例えば、32 プロセッサの PRIMEQUEST 480 の場合、I/O 帯域幅のピークが 75 GB/s であるのに対して、32 プロセッサコアの IBM p5-590 では I/O 帯域幅のピークは 48 GB/s になっています。8 プロセッサの PRIMEQUEST 420 の相互接続チップセットは、25.6 GB/s クロスバー帯域幅を提供するのに対して、HP の sx1000 チップセットは、8 プロセッサの HP rx7620 に 8 GB/s の帯域幅を提供します。もちろん、ユーザは必ずしもすべてのチップセットの仕様を知っている必要はありません。富士通の熟練したエンジニアリングチームが設計した PRIMEQUEST は、同様のプロセッサチップを使用している他のモジュラ

<sup>1</sup> 本ドキュメントは、2006年3月に執筆されており、ここで取り上げた PRIMEQUEST と競合モデルは、その時点で入手可能なものです。いずれ Montecito などの新しいプロセッサチップが販売されると、これらのアップグレードされたチップが新しいモデルに組み込まれます。アップグレードされたモデルの詳細は、本ドキュメントで説明している内容とは異なる場合がありますが、各ベンダーの全体的な製品ファミリーは、本ドキュメントで取り上げたものの一貫性があると思われます。

図 1

PRIMEQUEST アーキテクチャの  
ブロック図

—実装より高いシステムパフォーマンスを得られるということユーザーは高く評価しています。

この図では、システムボードと I/O ユニットが互いに物理的に分離していることに注目してください。これについては後述しますが、システムをパーティション化すると、システムボードと I/O ユニートを独立して別個のパーティションに割り当てることができます。これは、他の競合製品にはない柔軟性です。

図 1 を詳細に見ると、PRIMEQUEST の相互接続は、実際には複数のクロスバーファブリックとして実装されています。実際、複数の相互接続ファブリックは、ユーザーが PRIMEQUEST を高可用性のミッションクリティカルなコンピューティングとして信頼する重要な理由の 1 つです。

図 1 の相互接続ブロックは、複数のアドレスクロスバーとデータクロスバーによって構成されています。Standard Mirror Mode では、アドレスクロスバーは、ロックステップ二重化方式で実行されます。

この概念をおおまかに説明すると、アドレスクロスバーは要求されたデータのアドレスをすべてのメモリ、I/O、およびシステムボードにブロードキャストします。エラーが検出されなかった場合、不正な場所にデータが書き込まれることとなりますが、これは後で深刻なシステムの障害やデータの破損を引き起こす可能性があります。PRIMEQUEST には 2 つのアドレスバスが実装されています。各システムボードには、別個のチップが搭載されており、それぞれアドレスバスのインターフェースとして機能します。Standard Mirror Mode 操作は、これら 2 つのアドレスバスを同一のバストランザクションを二重化したコピーとして扱います。1 つのバス(またはこれに関連付けられているチップセット)の断続的なエラー、または障害は、システムの運用には影響を与えません。これは、二重化されたミラーバスが正しいアドレス情報の独立したコピーを持っているからです。<sup>2</sup>

他のベンダーでは、アドレスの相互接続に各種のエラー検出メカニズムを実装しています。しかし富士通は、標準の運用モードでミラー化されたアドレスバスの冗長性を提供するという点で、他社より一歩先んじています。

富士通のハードウェアデザインは、アドレスバスだけではなくデータバスも二重化されることに注目してください。Extended Mirror Mode は、アドレスバスのミラー化と同じ方法で、このデータ相互接続を冗長的に構

<sup>2</sup> 優れたメモリ性能と I/O トラフィックのある環境では、ミラー化されていない動作を選択することで、両方のアドレスバスを使用してスループットを最大化できます。ただし、富士通の見積もりによれば、ほとんどの作業負荷については、2 つのバスの予想される総負荷が非常に低いため、1 つしか使用されず、もう 1 つはミラー化されたコピーとして使用され、全体のパフォーマンスが 1%未満低下します。したがって、PRIMEQUEST サーバでは、デフォルトのミラー化オプションは Standard Mirror Mode になっています。

成します。しかし、ここでは Extended Mirror Mode の他の側面を取り上げる前に、PRIMEQUEST サーバのメモリ容量について説明しましょう。

## PRIMEQUEST の構成

要件を最も満たすプラットフォームを選択する前に、ニーズを満たすために適した機能を提供するソリューションの候補（ショートリスト）を作成して、選択肢を絞り込んでおく必要があります。エンタープライズコンピューティングでは、アプリケーションの負荷を処理するために十分なプロセッサ、メモリ、および I/O を構成できるサーバを見つけることとなります。表 1 に示すように、PRIMEQUEST モデル 420、440、および 480 では、中規模からハイエンドまでの構成を提供します。どれもすべて強力なプロセッサ、大量の高速 DDR2 メモリ、および必要な I/O を装着するための十分な PCI スロットと内部ディスクベイを提供します。

表 1. PRIMEQUEST の仕様

PRIMEQUEST モデル	プロセッサ GHz	プロセッサ	最大メモリ (DDR2)	最大 PCI スロット (拡張した場合)	内部 ディスク
PRIMEQUEST 420	Itanium 2 1.5, 1.6 GHz	1 ~ 8	256 GB	4 PCI-X 2 PCI-Express (合計 18PCI-X/E)	8 drive 584 GB
PRIMEQUEST 440	Itanium 2 1.5, 1.6 GHz	1 ~ 16	512 GB	16 PCI-X (64 PCI-X)	16 drive 2.35 TB
PRIMEQUEST 480	Itanium 2 1.5, 1.6 GHz	1 ~ 32	1 TB	32 PCI-X (128 PCI-X)	32 drive 4.7 TB

表 2 に示すように、IBM System p5 サーバは、同クラスの PRIMEQUEST サーバと同等のメモリ容量を提供できません。たとえば、IBM の p5-550 と p5-550Q は 64 GB の DDR2 メモリしか使用できませんが、PRIMEQUEST では、この 4 倍を使用できます。また、p5-590 がサポートする最大メモリは、より低速な DDR1 であり、p5-590 の DDR2 の容量は 8 分の 1 になってしまいます。

表 2. IBM System p5 の仕様

IBM System p5 モデル	プロセッサ、GHz	チップ/コア の最大数	最大メモリ (注記があるもの 以外は DDR2)
p5-550	POWER5+: 1.65、1.9 GHz	2 / 4	64 GB
p5-550Q	POWER5+: 1.5 GHz	4 / 8	64 GB
p5-560Q	POWER5+: 1.5 GHz	8 / 16	128 GB
p5-570	POWER5+: 1.9、2.2 GHz	8 / 16	512 GB (2.2 GHz の場合) 256 GB (1.9 GHz の場合)
p5-590	POWER5: 1.65 GHz	16 / 32	1 TB (DDR1) 128 GB (DDR2)

## MONTECITO については?

現在「Madison 9M」プロセッサチップは、PRIMEQUEST に使用されており、すべての RISC 代替チップに対して高い競争力を持っています。近いうちに、劇的に新しい Itanium 実装が入手可能になります。現在のコードネームは「Montecito」です。Montecito の各チップには 2 つのプロセッサコアが搭載されており、各コアは 2 つの命令ストリーム「スレッド」を実行します。

PRIMEQUEST モデルは、増分アップグレードを行えるように設計されており、Madison と Montecito システムボードの両方を搭載したシステムをサポートすることになっています。

Montecito が販売されると、PRIMEQUEST のパフォーマンスは劇的に増大します。PRIMEQUEST を導入しているユーザは、PRIMEQUEST を他のエンタープライズサーバと差別化しているミッションクリティカルな高可用性の属性を享受しながら、今後大幅にパフォーマンスを増大させる可能性があることを待望できます。

表 3 では、HP Integrity サーバのメモリも、PRIMEQUEST が提供するメモリ容量には及ばないことが示されています。HP が出荷している Integrity プラットフォーム(sx1000 チップセット使用)は、より低速な PC133 SDRAM のみをサポートしています。最近、HP は DDR2 をサポートする新しいモデル(sx2000 チップセット使用)を発表しましたが、これらの新しいモデルも PRIMEQUEST のメモリ容量には及びません。同様に、表 4 に示すように、PRIMEQUEST は Unisys ES7000 のモデルと比較した場合でも、プロセッサコアごとにより多いメモリ容量を構成できます。また、Unisys Itanium プラットフォームの ES7000 モデル 400 の場合、パーティションごとに最大で 16 の Itanium チップに拡張できますが、これも、PRIMEQUEST 480 が提供するシングルシステムイメージの最大量の半分に過ぎません。

表 3. HP Integrity の仕様

HP Integrity モデル	プロセッサ、GHz	プロセッサの 最大数	最大メモリ
rx7640 (sx2000)	Itanium 2: 1.6 GHz	8	128 GB (DDR2)
rx8640 (sx2000)	Itanium 2: 1.6 GHz	16	256 GB (DDR2)
Superdome (sx2000)	Itanium 2: 1.6 GHz	64	1 TB (DDR2)
rx7620-16 (sx1000)	Itanium 2: 1.5, 1.6 GHz Itanium 2: 1.1 GHz (mx2)	8 16 (mx2)	128 GB (SDRAM)
rx8620-32 (sx1000)	Itanium 2: 1.5, 1.6 GHz Itanium 2: 1.1 GHz (mx2)	16 32 (mx2)	256 GB (SDRAM)
Superdome (sx1000)	Itanium 2: 1.6 GHz Itanium 2: 1.1 GHz (mx2)	64 128 (mx2)	1 TB (SDRAM)

表 4. Unisys ES7000 の仕様

Unisys モデル	プロセッサ、GHz	チップ/コアの最大数	最大 メモリ
ES7000 モデル 400	Itanium 2: 1.3, 1.5, 1.6 GHz Xeon: 2.8, 3.0 GHz	すべてのパーティション で 32 チップ パーティションごとに 最大 16 チップ	パーティシ ョンごとに 256 GB
ES7000 モデル 600	Xeon EM64T: 2.8, 3.3 GHz	32 / 64	512 GB

では、富士通はなぜ他の競合製品より多いメモリをサポートするように PRIMEQUEST を設計したのでしょうか。これらの追加メモリは、メモリ消費量が多いアプリケーションや、ミッションクリティカルな環境のメモリのミラー化に使用できます。

多くのアプリケーションでは、メモリを大量に消費します。例えば、Business Intelligence アプリケーションでは、トレンドやデータの関係を見つけるために大量のデータを検索します。他のアプリケーションでは、

多数の同時ユーザへのサービスを提供し、各ユーザは独自にデータを必要とします。メモリを追加することは、パフォーマンスを向上するための最も簡単な方法になることがよくあります。ユーザは、アプリケーションを調査して、よりメモリ容量が大きい PRIMEQUEST にはどのような利点があるかを考慮する必要があります。さらに、PRIMEQUEST と他のエンタープライズサーバを比較し、DDR2 と富士通独自の高パフォーマンスインターフェースチップによってもたらされる利点を検証するとよいでしょう。プロセッサコアごとに大量のメモリが必要な負荷がある場合、PRIMEQUEST は競合機種では不可能な構成で使用できます。

メモリ容量による制約を受けることのない他のユーザの環境でも、一般的なシングルエラー訂正/ダブルエラー検出 (SECDED; single-error-correct double-error-detect) 誤り訂正符号 (ECC; Error Correcting Code) やシングルメモリチップ障害 (チップキル耐性と呼ばれることもある) の対処等を越えた高い可用性が必要とされることがあります。Extended Mirror Mode は、このようなミッションクリティカルなエンタープライズ環境に、メモリとデータバスの冗長性を提供します。

## データのミラーリング

業界で一般的であるように、富士通は ECC とチップキルを組み込んで、メモリモジュールの障害がシステムのクラッシュを発生させる可能性を最小限に留めています。それでもなおミッションクリティカルなユーザの環境では、さらに高い可用性が求められています。富士通は、標準でアドレスバスのミラーリングに加えて、データクロスバーやメモリ自体のミラーリングまでもオプションで提供しています。

ミラーリングするように構成した場合、リソースはペア化され、同一の機能が平行して実行されます。ペアの片方に障害が発生しても、もう片方のリソースが有効な結果を提供し、システムは中断されることなく動作を続けます。表 5 に示すように、Standard Mirror Mode は、すべてのシステムボードと I/O ユニットの間のアドレス相互接続として動作するアドレスクロスバーをミラーリングします。Extended Mirror Mode は、4 つのデータクロスバーを 2 つの二重化構成された半分ずつとして配置してメモリを構成するため、各メモリのロケーションは 2 つの別個のメモリモジュール内で二重化されます。

表 5. PRIMEQUEST が提供するシステムミラー機構

	Standard Mirror Mode	Extended Mirror Mode
アドレスバス	ミラー化	ミラー化
データバス	独立	ミラー化
システムボード帯域幅	12.8 GB/sec.	6.4 GB/sec.
IOU 帯域幅	6.4 GB/sec.	3.2 GB/sec.
メモリ	独立	ミラー化



	Standard Mirror Mode	Extended Mirror Mode
メモリ容量	100%	50%
パフォーマンス	>99%	>95%

PRIMEQUEST のデータクロスバーは、4 つの別個のデータバスで構成されます。それぞれのデータバスは、シングルビット ECC によって保護され、システムと I/O ボードにはインターフェース専用チップがあります。より高度な保護を必要とするユーザは、Extended Mirror Mod を使用すると、4 つのデータバスを 2 つのミラー化されたペアとして構成できます。データバスがミラー化されると、システムボードインターフェースの回路もミラー化された構成で実行されるため、インターフェースの回路で発生した障害や、バス自体の障害からシステムを保護できます。

さらに重要なことは、Extended Mirror Mod ではメインメモリに格納されているデータが二重化されることです。前述したように、メモリ ECC はシングルビットエラーを修正するだけでなく、障害が発生したメモリチップからすべての 4 ビットを再構成します。ただし、複数のエラーが発生した場合、そのメモリロケーションの内容は、復元できません。メモリのミラーリングは、データを 2 つの別個のメモリチップのセットに二重化します。複数のビットに障害が発生した場合、1 つのメモリロケーションにエラーが発生しますが、同時にミラー化されたロケーションにエラーが発生する可能性は非常に低くなっています。さらに、インターフェース回路のミラーリングにより、回路障害によるシステム障害が回避されるため、高度にミッションクリティカルなアプリケーションに必要なとされる、堅固なデータ整合性が得られます。

もちろん、メモリのミラーリングによって、メモリの総容量は半分になります。しかし、PRIMEQUEST で構成できるメモリ量の多さを考慮すると、総メモリ容量が少なくなることは、あまり問題にはなりません。確かに、二重化されたメモリロケーションにデータを格納することによって、メモリのコストも倍になります。したがって、Extended Mirror Mode はシステム障害を回避することが非常に重要な環境に最適です。

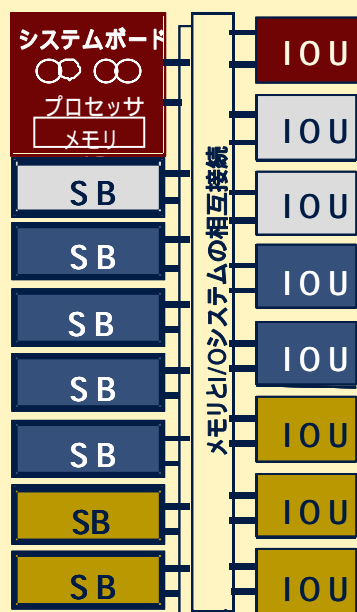
他のベンダーのエンタープライズサーバには、PRIMEQUEST の Extended Mirror Mode に匹敵する冗長性機能を提供するものではありません。メインフレーム水準の高い信頼性と可用性に親しんでいるユーザは、特に PRIMEQUEST のミラーリング機能に価値を見出すでしょう。

## 柔軟なパーティション機能

Itanium 2 プロセッサチップを 32 個まで拡張できる PRIMEQUEST のスケラビリティは、多くの個別アプリケーションにとって、特に Windows や Linux 上で実行している場合は、日常業務の要件を超えたものでしょう。それでもなおユーザは、ピーク要件を満たすために柔軟に構成でき、また日常業務に使用するために、独立した別個のパーティションに分割で

図 2

## PRIMEQUEST の柔軟なパーティション機能



きる大規模なサーバを求めることがよくあります。PRIMEQUEST のパーティション機能は、主要な競合製品より優れています。また、大規模なシステムは、複数のサーバの作業負荷を、管理が容易な中央サーバに統合するためにも使用できます。例えば、異なる作業負荷によって異なるバージョンのミドルウェアが必要になるなど、ソフトウェアスタックの競合が発生することがあります。パーティション化によってエラーを分離できることは、開発中やテスト中のソフトウェアの障害が稼働中の作業不可に影響を与えないという大きな利点があります。

PRIMEQUEST は、モジュール型の要素によって構成されていることを思い出してください。各システムボードには最大で 4 つのプロセッサチップ、128GB のメモリを搭載でき、I/O にはシステムの相互接続クロスバーを通じて接続します。クロスバースイッチに対して、システムボードのセットを互いに分離するように指定し、これらのシステムボードにアクセスするように I/O ユニットの割り当てをすることで、PRIMEQUEST 内に独立したパーティションを作成できます。独立した各パーティションは、オペレーティングシステムの別個のインスタンスを実行しており、ハードウェアまたはソフトウェアに障害が発生しても、他のパーティションとは物理的に分離されています。

システムのパーティション化については、ベンダーによって異なる方法を採用しています。例えば IBM では、仮想パーティションを多く採用しており、パーティションの分離はソフトウェアに依存しています。しかし、IBM の hypervisor などのソフトウェアによるパーティションは、ソフトウェアのシングルポイント障害の影響を受けやすいものです。またソフトウェアのオーバーヘッドは、その作業負荷をネイティブなハードウェアで実行している場合と比較すると、パフォーマンスが低下します。多くのユーザは、ソフトウェアパーティションの役割をよく知っていますが、パーティションを完全に分離する場合は、ハードウェア上で物理的にパーティションを分離する方法のほうが好まれています。

HP が提供するハードウェアパーティションは、大規模なサーバを「セルボード」と HP が呼んでいるハードウェア単位で分割します。しかし、HP のセルボードパーティションユニットには、プロセッサやメモリだけでなく、I/O へのパスも含まれています。このため、各パーティションは、I/O からプロセッサまたはメモリへの比率が固定されています。

PRIMEQUEST では、I/O ユニット(IOU)を柔軟にハードウェアパーティションに割り当てることによって、このような固定化された関係を排除しています。クロスバースイッチを使用することにより、富士通の柔軟な I/O は、I/O ユニットの任意のシステムボードに関連付けることができるため、処理能力と I/O 接続性を適切に混成してパーティションを定義できます。図 2 では、8 つのシステムボードを持つ PRIMEQUEST 480 サーバが 4 つのパーティションに分割されています。各パーティションは、異なるソフトウェアリリースまたは異なるソフトウェアスタックを含めて、独自の

オペレーティングシステムとして Linux または Windows を実行しています。また図 2 では、I/O ユニットも柔軟に割り当てられています。例えば、あるパーティションには、2 つのシステムボードと 3 つの I/O ユニットが割り当てられているのに対して、別のパーティションには 4 つのシステムボードと 2 つの I/O ユニットが割り当てられています。

Mirror Mode オプションでは、すべてのパーティションを統一する必要がないことに注目してください。つまり、あるパーティションを Standard Mirror Mode で構成し、より大きな冗長性が必要な他のパーティションは Extended Mirror Mode で動作できるということです。

サーバのパーティション化は、同時メンテナンス作業や段階的なアップデートを実行するときにも役立ちます。システムボードまたは I/O ユニットとアクティブなパーティションとを電氣的に分離して、修理を行うための取り外しやコンポーネントのアップグレードを行うことができます。ここでも、PRIMEQUEST のフレキシブル I/O により、システムボードまたは I/O ユニットはサービスやアップグレードのために個別に分離できます。HP のハードパーティションには、このような柔軟性はありません。また IBM はハードウェアパーティションを採用していないため、IBM の p5 サーバでは同時メンテナンスや段階的なアップグレードを実行できません。

## IDEAS の結論

最近まで、ミッションクリティカルなコンピューティングは、メインフレームや少数の UNIX/RISC プラットフォームに委ねられていました。そのころ Linux や Windows は、比較的小規模な構成が中心でした。現在、ユーザはより重要度の高い作業負荷についても、業界標準のハードウェアおよびソフトウェアソリューションを求めようになりました。PRIMEQUEST は、ミッションクリティカルな環境向けに堅牢に設計されたシステムとして、このようなニーズに応えます。

PRIMEQUEST は、中規模モデルの 420 から 32 個の Itanium プロセッサチップモデルの 480 に至るまで、バランスの取れたパフォーマンス、メモリ、および I/O 能力を提供することによって、エンタープライズ環境を対象とした競合サーバに対して優位にあります。PRIMEQUEST のハードウェアベースのパーティション機能は、他のハードウェアベースの実装と比較すると、柔軟な構成ができるという利点があり、またソフトウェアベースのパーティション機能では不可能な同時メンテナンスおよびアップグレードができるという利点があります。

富士通の設計は、他の Itanium ベースの SMP とは明らかに一線を画しており、同社がハードウェアの高可用性を重視していることがわかります。ECC や従来からある他の可用性技術を広く使用するだけでなく、この独自性のあるミラー化されたクロスバーやミラー化されたメモリを使用することで、PRIMEQUEST プラットフォームは、最も要求の厳しいユー

**Americas**

Ideas International, Inc.  
800 Westchester Avenue  
Suite S620  
Rye Brook, NY 10573-1330  
USA  
Tel + 1 914 937 4302  
Fax +1 914 937 2485

**Asia/Pacific and Worldwide Headquarters**

Ideas International Limited  
Level 3  
20 George Street  
Hornsby, NSW, 2077  
Australia  
Tel +61 2 9472 7777  
Fax +61 2 9472 7788

**Europe, Middle East, Africa**

Ideas International Europe  
1 Deanes Close  
Steventon  
Oxon OX13 6SZ  
United Kingdom  
Tel +44 (0) 1235 437 850  
Fax +44 (0) 1235 437 851

[www.ideasinternational.com](http://www.ideasinternational.com)



ザであっても、ミッションクリティカルなビジネスコンピューティングに使用できるでしょう。

エンタープライズコンピューティングを必要とするユーザは、従来からのサプライヤである IBM、HP、Unisys など、各社が提供する各種のシステムを候補として考慮します。しかし、ミッションクリティカルな作業負荷を処理する必要があるユーザは、必ず富士通を候補として検討し、PRIMEQUEST の利点をよく吟味する必要があるでしょう。

**【注意事項】**

This document was translated from the original English language document (PRIMEQUEST Servers: Mission-Critical Servers for Linux and Windows) created by Ideas International Limited. This document was translated into Japanese by Fujitsu Limited, and Ideas International Limited has not reviewed the translated version of the document as the official version.

本資料は Ideas International Limited によって作成された英文原稿資料 (PRIMEQUEST Servers: Mission-Critical Servers for Linux and Windows) から翻訳されたものです。本資料は富士通(株)によって日本語に翻訳されたもので、Ideas International Limited はこの日本語版を正式資料としてレビューしておりません。