

社会基盤を支える高信頼基幹IAサーバ PRIMEQUEST 1000シリーズ

PRIMEQUEST 1000 Series: High-Reliability, Mission-Critical IA Server Supporting Social Infrastructure

あらまし

クラウド時代を迎えるにあたり、サーバプラットフォームでは高信頼性・高可用性だけでなく、ICTリソースの有効活用、効率運用のために仮想化技術が基本技術として要求されている。

富士通は2005年、オープンスタンダードであるIA技術とLinux/Windowsをベースに、メインフレーム開発で長年蓄えてきたノウハウと経験を結集し、基幹IAサーバPRIMEQUESTを開発した。市場投入以来、オープンミッションクリティカル市場を開拓し、様々な社会基盤システムに採用されている。そして更なる進化を目指し、従来シリーズの特長に加えて仮想化対応の強化、長期サポートの提供、グリーンICT対応の強化などを備える新モデルPRIMEQUEST 1000シリーズを開発し、2010年3月から販売を開始した。

本稿では、新モデルである1000シリーズのねらいと特長、採用技術について説明する。

Abstract

As we enter the age of Cloud computing, server platforms are required to incorporate virtualization as basic technology for effective use and efficient operation of ICT resources, in addition to high reliability and availability. In 2005, as a result of bringing together the know-how and experience acquired over many years of developing mainframes based on Linux/Windows and open standard IA technology, Fujitsu developed the mission-critical IA server PRIMEQUEST. This server has opened up a new market for open mission-critical systems and been used for various social infrastructure systems ever since it was brought to market. For further evolution, we developed PRIMEQUEST 1000 Series, containing new models provided with enhanced virtualization, long-term support and improved green ICT capability in addition to the features of the existing series. It went on sale in March 2010. This paper describes the aim and features of the 1000 Series and the technologies adopted.



廣瀬元義
(ひろせ もとよし)

IAサーバ事業本部IAサーバ事業部 所属
現在、基幹IAサーバの技術支援業務に従事。



佐藤啓治
(さとう けいじ)

IAサーバ事業本部プロダクト開発統括部 所属
現在、基幹IAサーバの開発業務に従事。



栗田真広
(くりた まさひろ)

IAサーバ事業本部システム技術開発統括部 所属
現在、サーバ管理ソフトの開発業務に従事。



本田育史
(ほんだ やすふみ)

プラットフォーム技術本部クラウドインフラセンター 所属
現在、クラウドインフラモデルの開発に従事。

ま え が き

富士通のサーバ戦略は、「高信頼・高品質、高性能、最先端技術を適用し、社会基盤システムや企業内データセンター最適化を可能とする、ミッションクリティカルサーバを継続的に提供する」ことである。基幹IAサーバPRIMEQUESTは、「メインフレーム/UNIXサーバの信頼性・堅ろう性と、Linux/Windowsサーバの汎用性・経済性の両者の利点を兼ね備えたサーバ」として、「オープンミッションクリティカルの世界を切り拓く」ことを目的として開発された。発売以降、“IT Japan Award 2010” 経済産業大臣賞（グランプリ）を受賞した、東京証券取引所様の“arrowhead”をはじめとして、いくつもの社会基盤システムや、企業内の大規模・基幹システムに採用されている。

その後、2010年3月31日のインテル社のXeon プロセッサ7500番台の発表に合わせて新モデルPRIMEQUEST 1000シリーズの販売を開始した。

本稿では、PRIMEQUEST 1000シリーズについて、開発のねらいと特長、高信頼・高可用性を実現する機能、グリーンICTへの機能、運用管理機能、長期サポート、仮想化対応、OSIV/XSP動作機構を概説する。

PRIMEQUEST 1000シリーズの開発のねらいと特長

従来のPRIMEQUESTは、金融システムや社会基盤システムを中心とする、大規模・基幹システムを主な適用分野としていた⁽¹⁾

近年の仮想化技術の進歩により、基幹システムにも仮想化技術を適用し、計算機資源の効率化が望まれるようになった。これは、社内システムにクラウド技術を適用する、プライベートクラウドへの流れとなって現れている。

仮想化以前は一般のPCサーバで構築していた領域のシステムも、10数台～数10台ものサーバを仮想化技術によって統合すると、1台の物理サーバのダウンによる影響範囲が拡大するため、仮想化基盤を支えるために信頼性・可用性の高いサーバが必要とされるようになった。

● ねらい

PRIMEQUEST 1000シリーズの開発に当たっては、従来の大規模・基幹システムとしての利用に加

えて、仮想化基盤およびプライベートクラウド基盤としての利用に適した性能・機能・可用性の要件を満たすことをねらいとした。

同時に、競争力のある強い商品づくりを実現するため、商品企画から開発・拡販戦略に至るビジネス全体を、製販一体の観点からマネジメントする体制を設け、営業/SEなどの意見を基に、市場ニーズをタイムリに取り入れながら開発を行った。

その一つとして、従来機種で採用していたItaniumプロセッサに比べて、OS・仮想化ソフトウェアだけでなく、データベースなどのミドルウェア製品や業務/業種向けパッケージソフトウェア製品、運用/管理ソフトウェア製品の品揃えや、性能面でも優位となってきたXeon プロセッサの採用の要求があった。

開発部門では、インテル社と協議を行い開発中であった最新Xeon プロセッサに実装される機能をあらゆる面から検討した。その結果、性能・仮想化における優位性に加えて、従来のPRIMEQUESTの機能と組み合わせることで、従来機種に遜色のない信頼性・可用性を有する、インテル社のXeon プロセッサ7500番台を搭載したPRIMEQUEST 1000シリーズを実現した。

● 特長

従来機種の特長を継承・強化したPRIMEQUEST 1000シリーズは、下記の特長を備えている。

- (1) 最大8CPU（64コア）のスケラビリティにより、性能を従来機種の2倍程度、コストパフォーマンスを3倍以上に向上した。
- (2) 仮想化基盤・プライベートクラウド基盤に必要なメモリとI/Oの拡張を図った。
- (3) Xeon プロセッサ7500番台のQPI（Quick Path Interconnect）技術とメモリミラーによるシステムミラー⁽²⁾相当の信頼性と、Itaniumプロセッサ相当のRAS（Reliability, Availability, Serviceability）機能を搭載した。
- (4) 従来機種からの特長であるパーティショニング機能、フレキシブルI/O機能、予備システムボード機能、OSIV/XSP動作機構を提供した。
- (5) 電源・ファン・HDDなど、ユニットレベルでの冗長性と保守性を維持している。
- (6) 部品レベルでの信頼性の確保に加え、部品点

- 数/種類を削減し、合わせてコストダウンも図った。
- (7) 対応するOS（仮想化ソフトウェアを含む）、ミドルウェア、パッケージの種類を大幅に増やした。
- (8) ハードウェアを長期サポート（最長10年）している。

また、環境負荷低減にも配慮した製品を提供することはICTベンダの責務でもあり、下記のようなグリーンICT機能も強化し、富士通スーパーグリーン製品として、お客様に提供している。

- ・大幅な省電力/省スペース/軽量化
- ・消費電力・温度・風量の見える化
- ・低騒音化、リサイクル率の向上
- ・各国の環境規制に適合

高信頼・高可用性を実現する機能

ミッションクリティカルシステムには24時間365日稼働する高い信頼性と可用性が要求される。ここではPRIMEQUEST 1000シリーズのハードウェア概要および高信頼・高可用性を実現する機能について述べる。

● パーティショニング機能とフレキシブルI/O

パーティショニング機能は従来のPRIMEQUESTから継承している重要機能の一つであり、筐体内のハードウェア資源を複数のシステムに分割し、分割した単位で独立したシステムを稼働させる機能である。パーティショニング機能はハードウェアによる分割のためシステム同士の障害隔離性に優れ、用途や規模、求められる信頼性が異なる複数のシステムを、PRIMEQUESTの高処理能力を最大限に生かしながら効率良く搭載することができる。

パーティショニング機能には以下の特長がある。

- (1) 同一筐体内に複数の業務を構築でき、柔軟なシステム運用が可能になる。
- (2) 任意の分割されたシステムの障害を、ほかの分割されたシステムに影響を与えないようハードウェアによって保護している。

パーティショニング機能は、PRIMEQUESTシリーズの特長であるフレキシブルI/O機能と組み合わせることにより、より柔軟なシステム構成を実現している。

フレキシブルI/O機能とは、図-1に示した

PRIMEQUEST 1000シリーズの構成要素であるCPU、メモリを搭載したシステムボード（SB）と、PCI Express（PCIe）スロットおよびオンボードI/Oを搭載するI/Oボード（IOB）およびGiga-LAN/SAS/PCI-BOX インタフェースボード（GSPB）を、その搭載場所に左右されず、任意に組み合わせるパーティションを構成できる機能のことである。フレキシブルI/O機能は、SBからのPCIe信号をスイッチで接続変更してI/Oリソースに結線することで実現している。

パーティショニングの粒度は、SBは1ユニット単位、I/Oでは図-1に示すLIOB、LGSPB単位である。LIOBはIOBを、LGSPBはGSPBをそれぞれ論理的に2分割可能なボードである。これらを組み合わせることにより、筐体内に四つまでのパーティションを構成することができる。例えば、筐体内のすべてのリソースを使い、8CPUとすべてのオンボードI/Oおよび40本のPCIeスロットを使った大きなパーティションを構成する、またはSBをすべて分割して筐体内に四つのパーティションを作り、各パーティションのI/Oリソース量を業務に合わせて変更することも可能である。当然ながら、その中間である2パーティション、3パーティション構成も可能

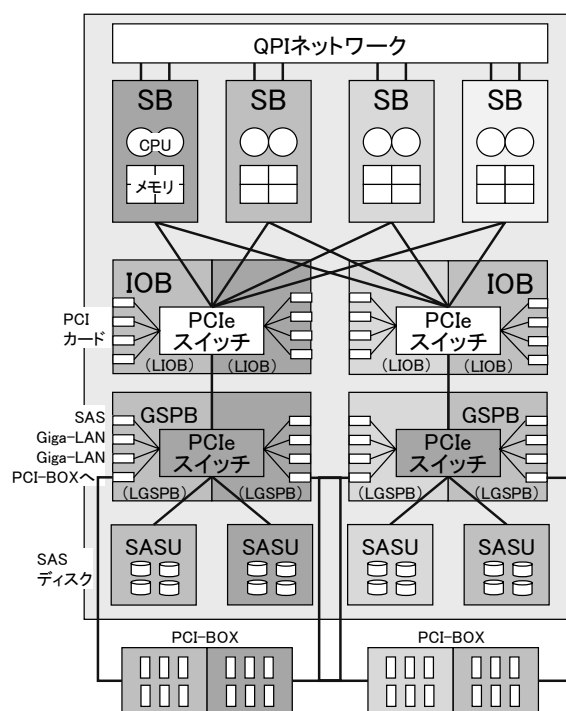


図-1 PRIMEQUESTの構成（概念）
Fig.1-Block diagram of PRIMEQUEST.

であり、その際のSB数、I/O数は前述の粒度単位で設定できる（図-2）。

● リッチI/O

PRIMEQUESTはI/Oの拡張性を重視しており、PCI-BOXを接続することにより、最大40本までPCIeスロットを拡張できる。この多数のI/Oを使ってマルチパスを構成し、システムの信頼性を高めることができる。

また、40本のPCIeスロットはすべてPHP（PCI Hot Plug）に対応しており、カセット形態で実装している。これにより、PCIカードに故障が発生した場合でも、運用を継続しながらのカードの交換が可能である。

● Reserved SB機能

Reserved SB機能とは、あらかじめ筐体内に予備のSBを実装しておき、運用中のSB障害が発生した

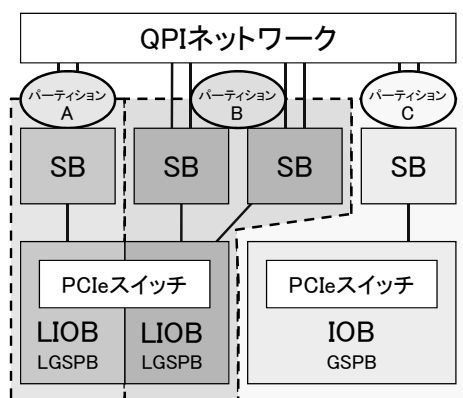


図-2 パーティションの概念
Fig.2-Concept of partitioning.

場合でも、自律的に故障したSBを切り離し、予備のSBを短時間でリブートして組み込み、SBを交換することができる機能である（図-3）。Reserved SB機能を利用すると、SB上でハードウェア故障が発生しても、パーティション内のSB資源を減少させることなく、早期の復旧が可能となる。

Reserved SB機能によりSBの入替えが発生しても、システムから見えるPCIバスツリー構造を維持しなければならない。このため、複数SBでパーティションを構成する場合は代表SBからのPCIeのみI/Oに接続し、いずれのSBが切り替わってもバスツリーが変化しないように構成する。

また、Reserved SB機能では、筐体内の別パーティションで動作しているSBも予備SBとして指定できる。例えば、同一筐体内に運用系のパーティションとテスト系のパーティションがあったとき、運用系パーティションの予備SBとしてテスト系のSBを指定しておくことができる。この状態で運用系のSBに故障が発生すると、テスト系のパーティションに対し自動的にシャットダウン指示を行い、シャットダウン完了後に予備指定されているSBを運用系に組み込んで自動的に再起動する。これにより、可用性を高めることができる。

● ハードウェアの冗長化

PRIMEQUESTでは電源、ファン、MMB（サーバ管理専用ユニット）などほとんどすべてのコンポーネントの冗長構成を可能としている。これらハードウェアを冗長構成にすることで、構成品の一部が故障しても残りの構成部品で連続運転が可能にな

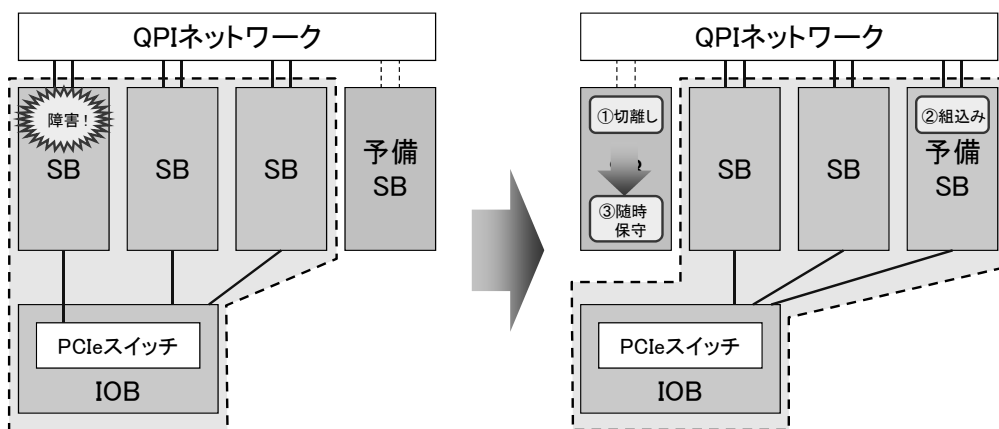


図-3 Reserved SBの動作概念
Fig.3-Operation of Reserved SB.

る。以下にPRIMEQUEST 1000シリーズのハードウェア構成における主な冗長化対策を述べる。

(1) メモリミラーリング

メモリを二重化し、二重化した両方のメモリにデータを保持する。片方のメモリで訂正不可能なエラーが発生した場合でも、もう一方のメモリのデータを使用することで、システムの継続運転と安定した動作を実現する。データの完全な二重化であるため、Double Device Data Correctionなどのエラー訂正機構よりも高い信頼性を有する。

(2) 電源供給系

一系統AC受電の場合は、PSU（電源ユニット）を $n+1$ 冗長にすることができる。つまり、PSUが1台故障した場合でも業務の運用は継続され、運用中の交換も可能である。

二系統AC受電の場合は、系統ごとに必要数のPSUを搭載し $n+n$ 冗長にすることができる。これにより、AC電源供給元から二重化することが可能である。

(3) 冷却ファン

冷却ファンは標準で $n+1$ 冗長しており、冷却ファンが1台故障した場合でも残りのファンで冷却可能な設計をしている。冷却ファンが故障しても業務を停止することなくそのファンを交換可能である。

(4) ハードウェアRAID

PRIMEQUEST 1000シリーズでは、RAID 0, 1, 5, 6, 10, 1Eをサポートしている。

(5) MMBの二重化

PRIMEQUESTには装置全体を管理するMMBが搭載されている。MMBは専用のCPUやメモリを持ち、パーティションとは独立したファームウェアが動作している。MMB単体でネットワークポート、ファームコードを二重化している。さらにMMBを二重化して、MMBハードウェアで故障が発生した場合は自動的に待機系のMMBで運用を継続する。詳細は「運用管理機能」の章で記載する。

● 縮退機能

(1) QPIの再送と縮退、迂回

最新のインテル Xeon プロセッサでは、CPU間およびCPUとチップセット間の通信にQPIと呼ばれる高速システムバスを使用している。QPIはCRC (Cyclic Redundancy Check) によりチェックされており、エラー発生時には再送を行う。バス

の固定故障時には故障箇所を避けて1/2および1/4までのバス幅の縮退を行う。この再送およびバス幅の縮退は運用を継続したまま実施できる。QPIの特定バスが故障で使えなくなった場合はそのバスを迂回するルートを設定し、自動的に再起動する。構成条件により迂回ルートが設定できない場合はSBを縮退して再起動する。

(2) PCIeの再送と縮退

SBからIOB、さらに個々のI/Oに至るPCIeはそれぞれCRCによりチェックされており、エラー発生時には再送を行う。バスはLaneという単位で構成されており、バスの固定故障時には故障箇所を避けて必須Lane0側に向かって1/2, 1/4, 1/8までのバス幅の縮退を行う。再送およびバス幅の縮退は運用を継続したまま実施できる。特定バスが故障により使えなくなった場合は、配下のデバイスを切り離して自動的に再起動する。同じ手法でPCIeスイッチやI/Oデバイスが故障した場合も故障影響のある部分を切り離して再起動する。

● 活性保守

筐体内で複数パーティションが動作している状態で、一つのパーティションで故障が発生して部品交換が必要になった場合でも、ユニットを共用していない限りにおいて、ほかの動作中のパーティションには影響を与えずに部品を交換することができる。

PRIMEQUESTはケーブルレス思想で設計しており、装置の引き出しや内部ケーブルの取外しをすることなく各ユニットを容易に交換することができる。

各ユニットを受け止め、ユニット間の信号を結線しているミッドプレーンは、装置中央に位置し、仮に故障すると装置を全停止した上で大がかりな交換作業が必要になる。しかし、PRIMEQUESTのミッドプレーンに電子部品は存在せず、コネクタおよびプリント基板での結線機能のみであるため、故障の確率は皆無に等しい。

● 故障解析とログ

PRIMEQUESTではBIOS、BMC、MMBの大きく3種類のファームウェアが動作している。ハードウェアに異常が発生した場合、各ファームウェアが収集する情報をMMBに集約させ、故障被疑箇所を総合的に判断する。操作、環境変化によるイベント情報や上記エラー情報などはMMBのシステムイベ

ントログとして記録され、異常発生時の解析に用いられる。

● 最新インテル Xeon プロセッサの採用

最新のインテル社のXeon プロセッサでは、前述のQPI障害回復機能のほか、Recoverable MCA (Machine Check Architecture) など、RAS機能が大幅に強化され、従来のPRIMEQUESTで採用したItaniumプロセッサに遜色のない信頼性レベルに達している。インテル社とはItaniumプロセッサの開発時にRAS機能の強化、評価試験などで協業しており、その資産は現在も受け継がれている。次世代プロセッサにおいても信頼性向上に向け協業を図っていく。

PRIMEQUEST 1000シリーズでは、以上のような機能、構成によりミッションクリティカルサーバとしての高信頼・高可用性を実現している。

グリーンICTへ機能実装

PRIMEQUEST 1000シリーズは、従来機種に比べてグリーンICT機能を大幅に強化し、地球環境に配慮した製品である。主な改善点を以下に示す。

- (1) CPUの消費電力を最大1040 Wに削減（従来機3328 W）。
- (2) 高スイッチング周波数および高効率ダイオードを採用し、80Plus Gold認証を取得した電源ユニットを同等クラスのサーバでは初めて採用。また、高変換効率DC-DCコンバータも採用（図-4）。
- (3) ヒートパイプを利用したCPUヒートシンク、高温部品に吸気を直接導くダクト構造、冷却風

の流れを活性保守時にも乱さないシャッター機構などを採用して筐体内の冷却を効率化し、冷却ファンの負荷を軽減（図-5）。

- (4) 1%刻みのファン回転数のダイナミックな制御により冷却を最適化し、ファン消費電力を削減するとともに稼働時の騒音も軽減。
 - (5) CPUコアの高密度化、構成部品の省電力、電源の高効率化によって、省スペース（高さ 約1/3・体積 約1/7）、軽量化（約1/5）を実現。また、全モデルをラックマウント型に統一することで、従来上位機種では本体と別ラックに収容する必要があった周辺装置群を同一ラックに収容可能。
 - (6) 筐体の消費電力や温度だけでなく、ファシリティの最適化のため、風量の見える化も実現。
 - (7) 低騒音化（66 dB→59 dB）。
 - (8) 再資源化可能率（リサイクル）99%以上。
 - (9) 欧州RoHS指令ほか、各国の環境規制に適合。
 - (10) マニュアルのDVD化による省資源化。
 - (11) ビデオリダイレクション機能による遠隔操作により、管理者の移動に伴うエネルギーの削減。
- これらの改善により、サーバ単体で年間、杉の木に換算して2000本以上のCO₂削減を実現した。

PRIMEQUEST 1000シリーズによって、サーバを統合し、さらに富士通製ミドルウェアの“ServerView Resource Coordinator VE”を利用して、サーバの利用効率を最適化することで、さらなる省エネルギー効果を生むことができる。

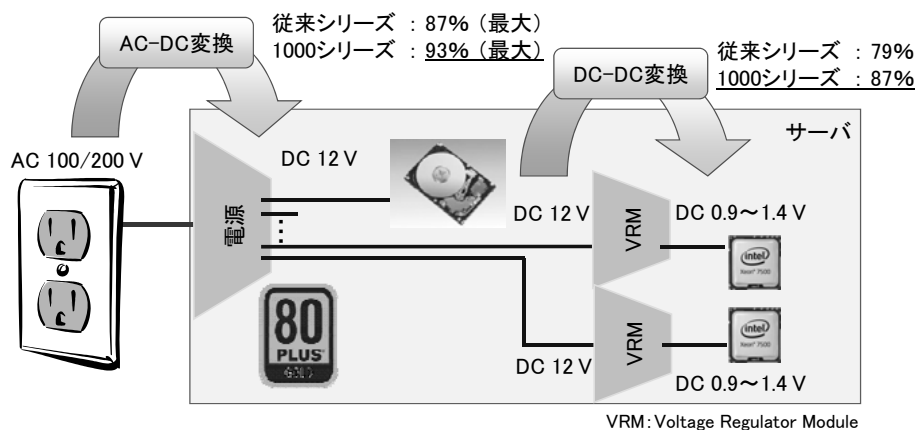
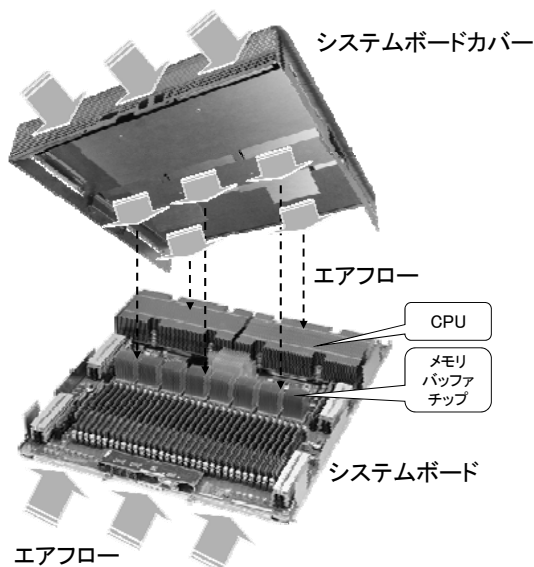


図-4 電源変換効率の向上
Fig.4 Efficiency of voltage conversion.

■ ダクト構造で、冷却風を高発熱部品へ集中



■ 冷却風の回り込み防止構造

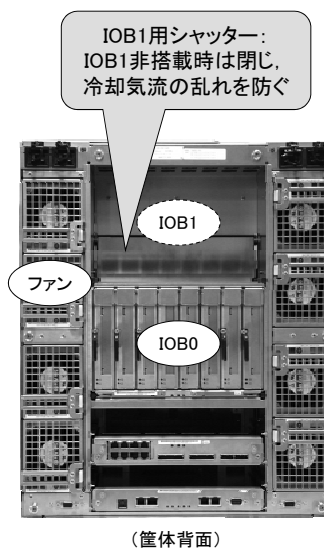


図-5 冷却効率の向上
Fig.5-Efficiency of cooling.

運用管理機能

サーバシステムの安定稼働のために、運用管理機能は重要な要素である。PRIMEQUEST 1000シリーズでは、従来のモデルから継承して筐体全体の監視を行うMMBと、各パーティション上にPRIMERGYを含めた富士通IAサーバ共通のサーバ管理ソフトウェア（SVS：ServerView Suite）を搭載し、さらにPCIカードの異常監視やクラスタの高速切替えなどのミッションクリティカル要件に対応した機能を従来のモデルから継承してPRIMEQUEST専用エージェントソフトウェア（PSA：PRIMEQUEST Server Agent）で提供する（図-6）。

● 運用管理機能の概要

(1) MMB

MMBはPRIMEQUEST筐体全体のハードウェア監視を行う。SB、IOB、SASU（SASハードディスクユニット）/SASRU（内蔵ディスクユニット）、GSPB、電源、ファン、PCI_BOXなどが搭載されている各種ハードウェアコンポーネントやパーティション管理を行う。

MMBは各ハードウェアコンポーネントと内部LANで接続されており、各種センサとも内部バスで接続されており、システム全体を常時監視・管理

している。MMBはオプションで冗長構成をとることができ、MMB自身やネットワークの障害時は、自動的にMMBが切り替わり、システムの継続した監視を行う。MMBが異常を検出した場合は各種通報アクション（REMCSセンタへの通知、snmpトラップ、メール通報）を実現する。

また、MMBはBMCと連携しパーティションの遠隔操作を行う機能（ビデオリダイレクション/リモートストレージ、テキストコンソールリダイレクション）を提供する。ビデオリダイレクション機能を使用する場合、ユーザの端末にはWeb経由でVGAの出力が表示される。

また、ユーザの端末に接続されているマウスとキーボードの入力は、LAN経由でパーティション側に接続される。テキストコンソールリダイレクションを起動すると、ユーザの端末にはWeb経由でパーティションからのシリアル出力が表示される。リモートストレージ機能は、ビデオリダイレクション機能を使用する場合に使用できる。ユーザ端末のCD/DVDドライブ、ISOイメージ（CD/DVD）、フロッピーディスク、USBデバイスをストレージデバイスとしてパーティションと共有できる。ISOの場合は、端末のISOがパーティション側にエミュレートされるドライブとして見える。

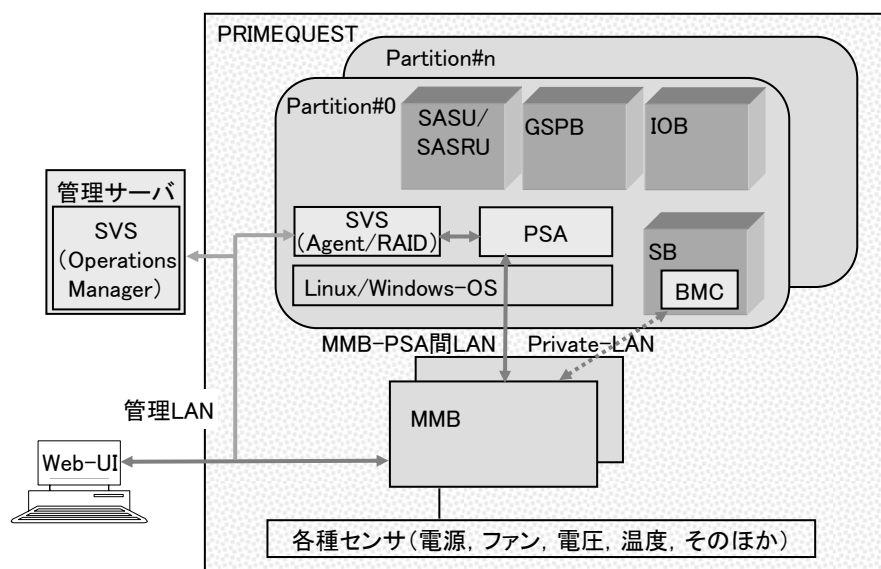


図-6 PRIMEQUESTの運用管理の構成
Fig.6-Configuration of PRIMEQUEST server management.

(2) SVS

PRIMERGYと共通のシステム統合管理ツール群である。システム導入から稼働、そして運用までを、Deployment（構築）、Operation（運用）、Analysis（診断・分析）、Integration（統合）というカテゴリに定義し、簡単・確実な導入、24時間365日の安定稼働や容易なシステム管理などを提供する。

初期導入、再構築作業の支援についてはServerView Installation Managerを提供する。

導入作業の簡素化、推奨ドライバ、セキュリティパッチの確実なインストールの実施を支援することで安定稼働を実現する。

運用監視、診断・分析・統合の支援については、ServerView Operations Managerを提供する。各パーティション上に搭載するServerView Agentでパーティションごとに検出できるCPU、メモリ、HDDなどの構成管理、システムリソース（CPU使用率など）を監視し、パフォーマンスの状況把握、システム異常の通知などを実現する。管理サーバまたは管理パーティションにインストールするServerView Operations Managerで各パーティションや、ほかのPRIMERGYを含めた複数のサーバをGUIにより統合管理することが可能である。さらに各パーティションから通知されるイベントを受領した場合、各種のアクション（メール通知、

snmpトラップ通知、警告灯通知、ポップアップ通知など）により、システム管理者に通知することが可能である。

アレイコントローラの監視支援についてはServerView RAID Managerで提供する。各パーティション上に搭載し、アレイコントローラの構成、HDDの状態、バッテリーの状態などを監視し、異常を検出すると通知する。また、ディスクアレイの構築、バッテリーのキャリブレーションの機能を実現する。

(3) PSA

各パーティション上にPRIMEQUEST専用の監視エージェントであるPSAが搭載される。MMBとは内部LANにより密に連携している。Linux/Windows上で動作するソフトウェアであり、PCIカード・HDDの構成と状態を監視する。また、マルチパスドライバ（GDS、GLS、MPDなど）の異常監視も行う。異常を検出した場合はMMBに通知し、MMBより各種通報アクションを実現する。また、ミッションクリティカルシステムに向けたクラスタリングソフトウェアPRIMECLUSTERに向けて、MMBと連携してOS状態の変化時に即時通知を行ったり、OSを強制停止させたりする機能を提供し、迅速かつ確実な切替えを実現している。

(4) SVS/PSAの連携

SVSが通知する異常について、PSAが検出し

MMB経由でREMCSセンタに通報を行うことを実現している。

● 今後の取組み

パーティション上に搭載する管理ソフトウェアは現在、SVSとPSAが存在するがIAサーバで一元化するためにPSAで提供している機能をSVSに統合する予定である。具体的には、PSAのPCIカード、HDDについてOSのドライバが検出した異常の監視、マルチパスドライバの異常監視、およびPRIMECLUSTER連携機能である。これにより、SVSの機能強化、IAサーバ管理基盤の統一およびミッションクリティカルシステムに対応した運用管理機能の強化を行い、PRIMERGYを含めたIAサーバ全体の管理機能の向上を図っていく。

長期サポートへの取組み

とくに社会基盤や金融関係のシステムでは、従来から24時間365日の可用性とともに、長期間の安定した稼働が強く求められている。

また、それ以外の分野の基幹業務システムでも、

長期間の利用が望まれるようになっており、「ベンダー企業への要望（JUAS 2007年度ハード・ソフトの寿命研究プロジェクト報告より）」の調査結果にも、下記の2点として現れている（図-7）^③

- (1) ハードウェア、ソフトウェアともに保守サポート期間を延長してほしい。
- (2) ハードウェアについては、部品の枯渇に対する対策を検討してほしい。

最近では、これらの要望に応えるため、OS/MW/パッケージなどのソフトウェアも、長期利用を可能とする製品やサービスを提供することが多くなってきた。

PRIMEQUESTシリーズでは、信頼性の向上のため部品レベルから厳しい品質基準を設定して選別を行っているが、冷却ファン・HDD・バッテリー・コンデンサには寿命があり、そのままでは長期にわたって故障率を一定基準内に維持することは難しい（図-8）。

そこで、PRIMEQUEST 1000シリーズでは、長期利用の要望に応えるべく、10年間の保守部品供

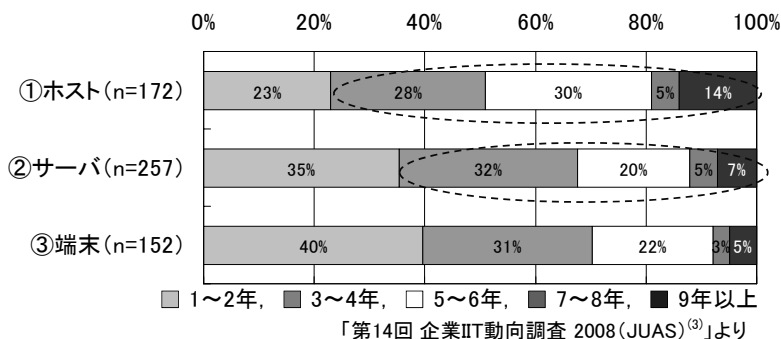


図-7 ユーザ企業が望むハードウェアの保守保証の延長期間
Fig.7-User demand for extended period of hardware maintenance and guarantee.

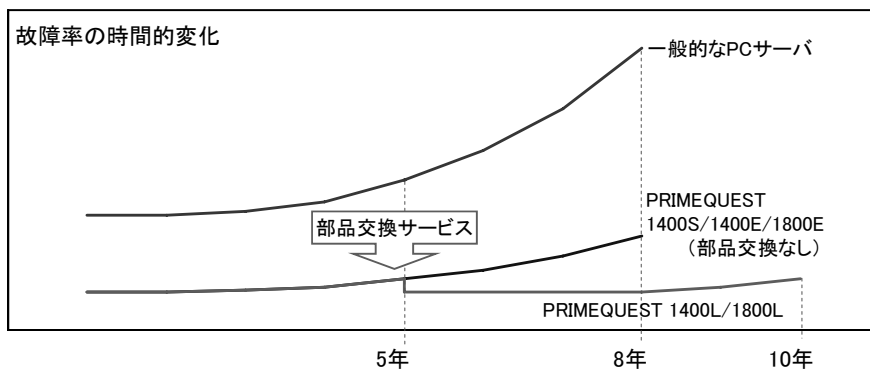


図-8 利用期間と故障率
Fig.8-Running time and failure rate.

給を保証するロングライフモデルを、同等クラスのサーバでは初めて提供した。ロングライフモデルのハードウェア保守サポートサービスには、保守部品の供給保証以外に、5年目にファンやHDDなど、時間経過とともに損耗を避けられない部品の交換サービスも含まれている。これにより、10年間の安定したハードウェア利用を可能とした。

仮想化対応の方針

この章ではPRIMEQUEST 1000シリーズの仮想化対応の方針について説明する。

ITリソースの有効活用、効率運用のためには仮想化技術がますます重要となっている。とくにITリソースをプール化し、リソースを柔軟に払い出して使用するクラウドコンピューティングでは、サーバ・ネットワーク・ストレージを含めた仮想化が前提の技術となっている。

PRIMEQUEST 1000シリーズでは、ハードウェアによる障害隔離性の高いパーティションとソフトウェアによる仮想マシン（VM）機能の二つの機能を組み合わせることで、より効率的なサーバ仮想化環境を提供する。中でも仮想化環境を支える技術として以下の特長を備えている。

● 大容量メモリ

サーバ上でVMを多数動作させるためには、CPUよりも先にメモリやI/Oがボトルネックとなってくる。このため、高密度に仮想集約を考えた場合には搭載できるメモリ容量がキーとなってくる。

PRIMEQUEST 1000シリーズではCPUあたり16スロットのメモリを装備し、最大で128スロット1Tバイトまで搭載できる。これはVMあたり平均メモリ容量2～4Gバイトとした場合、256台ものサーバ集約に相当する。

● 豊富なI/Oスロット

メモリの次にボトルネックとなってくるものとして、I/Oスループットが挙げられる。PRIMEQUEST 1000シリーズでは、オンボードの1Gbps LAN 16ポートに加え、PCIe拡張スロットも最大で40スロット（標準16、拡張24）まで装備可能である。

高密度に仮想集約した場合においてもVMあたりのLAN性能確保が容易となる。サーバのLAN構成をVMに単純に引き継がせることで、複雑なVLAN

設定を不要にできるメリットもある。

● 高信頼性・高可用性

つぎに、高密度仮想集約の際に課題となるのは、サーバに故障が発生した場合の影響範囲である。最悪の場合は、サーバ故障が全VMに及ぶことである。PRIMEQUEST 1000シリーズでは、この課題に対し、以下の取組みで対応している。

- ・部品レベルの信頼性確保
- ・パーティション機能による影響範囲の限定
- ・ReservedSBとFlexible I/Oによる短時間（1時間程度）リカバリ

このように、大規模仮想集約時でも安心して使用できる信頼性と可用性を提供する。

● ソフトウェアライセンス節約

VMwareなど一般的な仮想化ソフトウェアはCPUコア数に応じて課金されるケースが多い。したがって、コアに仮想マシンを多数集約できることはソフトウェアライセンスの節約にもつながる。

● ソフトウェアの選択性

仮想化運用を実現可能な以下のソフトウェアに対応している。

(1) 仮想化ソフトウェア

- ・VMware/HyperV/RH Xenをサポート

(2) クラウド運用管理ソフトウェア

- ・ServerView Resource Orchestrator
- ・Systemwalker Service Catalog Manager
- ・Systemwalker Software Configuration Manager
- ・Systemwalker Runbook Automation

● シンプルなサーバリソース追加

2CPU機などのスケールアウト型サーバと異なるPRIMEQUEST 1000シリーズは、8CPUまでのスケールアップサーバである。このため、拡張時にも設計変更が必要なくなり、単なるCPUやメモリの増設相当で済む。スケールアウト型サーバの増設時に必要なシステム再構築や、運用テストが不要となり、サーバリソース追加が容易で増設コストが安価に済む。

● 適用事例

豊島区役所様ではメインフレームベースの従来基幹システムをPRIMEQUEST 1000シリーズとVMwareで仮想集約するなど導入実績を上げている⁽⁴⁾

また、SAP社、Microsoft社と3社共同検証プロジェクトであるPRIME3⁽⁵⁾では、PRIMEQUEST

1000シリーズとETERNUSディスクアレイ、HyperVを使用した仮想集約で次世代SAPシステムの安定動作の検証実績を上げている（図-9）。

● 今後の取組み

お客様社内システムにクラウド技術を適用した、プライベートクラウドシステムを支えるサーバ製品群の一つとして、検証済モデルの提供などを通じて仮想化対応を強化していく。

OSIV/XSP動作機構と導入のメリット

富士通製メインフレーム（以下、グローバルサーバ）⁽⁶⁾のGS・PRIMEFORCE（プライムフォース）用OS“OSIV/XSP”と、同OSに対応した“AIM”などのミドルウェア、業務アプリケーションをPRIMEQUESTシリーズ上でそのまま稼働できる。その際に必要となるハードウェアオプションがOSIV/XSP動作機構⁽⁷⁾である。

OSIV/XSP動作機構は、CPU-OS間のファームウェアレイヤに位置し、インテル Xeon プロセッサをOSIV/XSPで扱えるグローバルサーバのCPUとして仮想化し、OSIV/XSPの命令コードをインテル Xeon プロセッサの命令コードに変換して実行する（図-10）。

OSIV/XSP動作機構は、システムボード・CPU・メモリ・ディスクなどのハードウェアと合わせて

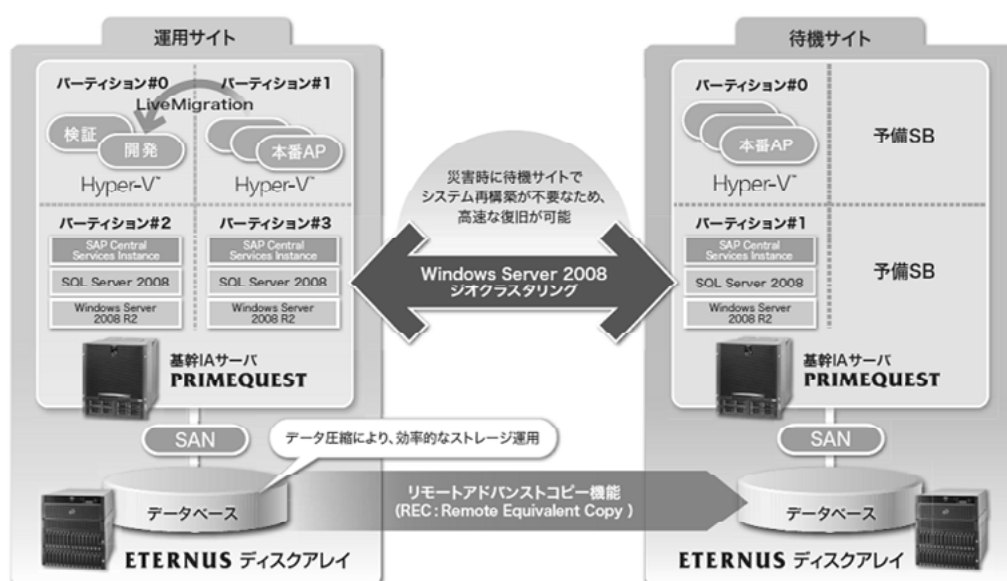
セット化し、「OSIV/XSP動作機構キット」として提供している。

また、グローバルサーバからの移行を容易にするために、OSIV/XSP動作機構を搭載したPRIMEQUESTは、グローバルサーバとの高い互換性を保っている。以下に、その一部を紹介する。

- (1) グローバルサーバ用仮想化ミドルウェアAVM/EXSが動作可能。
- (2) BMC（Block Multiplexer Channel）/ OCLINK（Optical Channel Link）チャンネルを装備し、グローバルサーバのI/O装置をPRIMEQUESTに直接接続し使用することが可能。



図-10 OSIV/XSP動作機構の概念
Fig.10-OSIV/XSP translation facility.



■ <http://jp.fujitsu.com/solutions/sap/brochures/pdf/prime3project3.pdf>

図-9 PRIME3での検証
Fig.9-Verification by PRIME3 project.

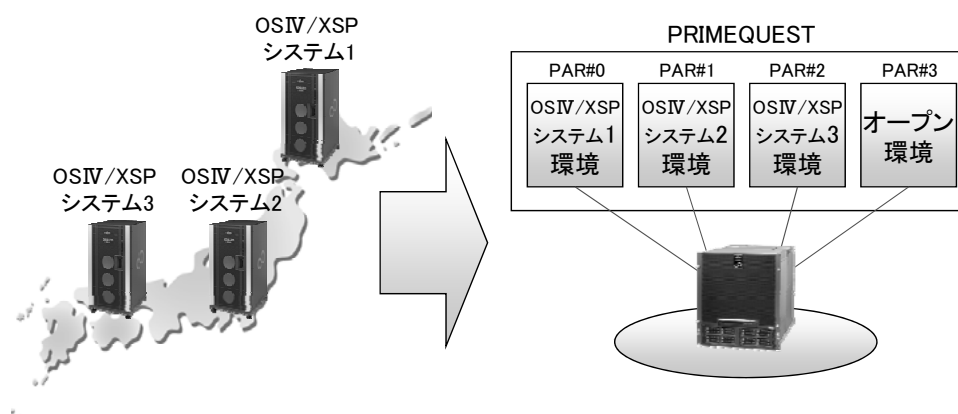


図-11 OSIV/XSPとオープンシステムの統合
Fig.11-Integration of OSIV/XSP and open systems.

(3) 周辺I/O装置・空調設備などの電源連動、異常発生時のアラームなど、グローバルサーバと同一の運用が可能。

これら高い互換性により、ユーザの現行アプリケーション資産を手直することなく、またシステム運用を変更することなく、グローバルサーバと同様の運用が可能である。

● 導入のメリット

さらに、PRIMEQUESTと融合したことにより、これまでに得られなかったメリットを享受することが可能となる。

PRIMEQUESTのパーティショニング機能により、独立したシステムを複数構築することで、個々に導入していた複数のOSIV/XSPシステムを1台のPRIMEQUESTに統合することが可能になる。さらには、オープンシステムも同一のPRIMEQUESTに統合できる(図-11)。

また、メインフレーム用OSだけでなく、オフィスコンピュータPRIMERGY 6000シリーズのOSであるASPにも対応を計画中である。

む す び

本稿では、富士通の基幹IAサーバPRIMEQUESTの新モデルであるPRIMEQUEST 1000シリーズのねらいと特長、採用技術について紹介してきた。今後も最先端技術の採用および主要

パートナーとの強力な連携により、基幹IAサーバの開発とオープンミッションクリティカル市場の拡大を牽引していく。

参 考 文 献

- (1) 黒川明佳ほか：基幹IAサーバ“PRIMEQUEST”のねらい. *FUJITSU*, Vol.56, No.3, p.188-193 (2005).
- (2) 浜田王才：基幹IAサーバ“PRIMEQUEST”の高信頼システムミラー機構. *FUJITSU*, Vol.56, No.3, p.207-210 (2005).
- (3) (社)日本情報システム・ユーザー協会：第14回企業IIT動向調査2008.
<http://www.juas.or.jp/servey/it08/index.html>
- (4) 富士通：豊島区役所様 導入事例.
<http://jp.fujitsu.com/solutions/theme/infrastructure/virtual/casestudies/toshimaku/>
- (5) 富士通, マイクロソフト, SAPジャパンによる3社共同検証プロジェクト「PRIME3」.
<http://jp.fujitsu.com/solutions/sap/brochures/pdf/prime3project3.pdf>
- (6) 江川博之ほか：グローバルサーバGS21. *FUJITSU*, Vol.53, No.6, p.471-476 (2002).
- (7) 斉藤友嗣ほか：メインフレーム資産を活用したオープン化事例. *FUJITSU*, Vol.60, No.5, p.451-457 (2010).