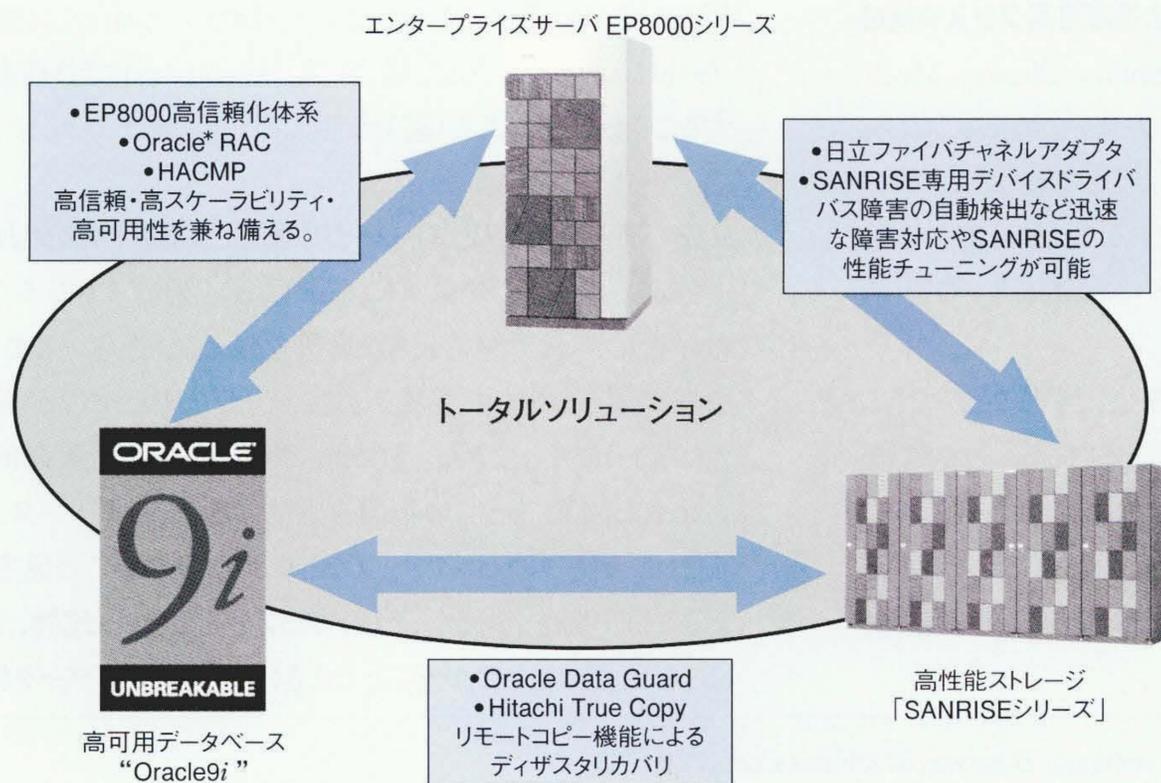


高性能サーバ・データベース・ストレージの連携で実現する高信頼・高可用システムソリューション

Highly Reliable and Highly Available Solutions with the Collaboration of High-Performance Server, Database, and Storage

福田 弘忠 *Hirokata Fukuta* 矢田龍太郎 *Ryûtarô Yada* 新田 英之 *Hideyuki Nitta*



高信頼・高可用トータル システムソリューションの概念

「EP8000+SANRise+Oracle 9i」によって高可用なクラスタ構成が実現でき、一部で障害が発生しても業務の続行が可能である。また、SANRise専用デバイスドライバや「日立ファイバチャネルアダプタ」により、各種障害への対応を強化している。さらに、リモートコピー機能を活用し、メインサイトとリモートサイト間のディザスタリカバリを実現する。

注：略語説明ほか

RAC (Real Application Clusters)
HACMP (High-Availability Cluster Multi-Processing)

* ORACLEは、Oracle Corporationの登録商標である。

企業内で扱うデータが増え続ける一方で、UNIXサーバ^{※1)}やオープンプロダクトを、高い信頼性が要求されるシステムに適用するケースが増加している。このような背景の中で、大容量ストレージ、UNIXサーバ、およびオープンプロダクトを組み合わせることにより、障害を起りにくくするだけでなく、万一障害が起ってもビジネスを停止させない、高信頼・高可用なシステムを構築、運用することが強く求められるようになってきた。

日立製作所は、オープンプロダクトを活用したソリューション事業を強化することを目的として2002年11月に設立した株式会社日立オープンプラットフォームソリュー

ションズと共に、高信頼かつ高可用なシステムを構築するためのソリューションを提案している。具体的には、ミッションクリティカルシステムに適したサーバ「EP8000シリーズ」と、高い可用性とリニアなスケラビリティを両立させたORACLEのデータベース、および高密度実装と高性能を兼ね備えたストレージ「SANRiseシリーズ」を密接に連携させたソリューションである。また、クラスタシステム構成での検証を重ね、そこから得られた性能チューニングや障害発生時のリカバリ手順などのノウハウを、トータルソリューションに反映させている。

1 はじめに

従来、ミッションクリティカル(停止不可)システムは、主に金

融機関のオンラインシステムなど、メインフレーム上で構築されたシステムで運用されてきた。しかし、インターネットの普及やオープンプラットフォームの信頼性向上などにより、IT(情報技術)基盤のライフライン化が進み、EC(Electronic Commerce)

※1) UNIXは、X/Open Company Limitedが独占的にライセンスしている米国ならびに他の国における登録商標である。

サイトをはじめとするオープンな環境でも、24時間365日無停止なシステムが要求されるようになってきた。

ここでは、日立製作所と株式会社日立オープンプラットフォームソリューションズ(以下、日立OPSSと言う。)が提案している、高信頼・高可用システムソリューション、およびサーバとデータベースを連携させたクラスタシステム構成による高可用性、高信頼性の検証について述べる。

2 サーバ・データベース・ストレージの連携

2.1 サーバとデータベースの連携による高可用クラスタ構成

“EP8000”のHACMP(High-Availability Cluster Multi-Processing^{※2)}機能とORACLEデータベースのRAC(Real Application Clusters)機能を組み合わせて使うことにより、従来以上に高可用なシステムを構築することが可能となる。また、ノードを追加することで、パフォーマンスとスループットを向上させることもできる。

HACMPは複数のノードをクラスタ構成にすることができるクラスタソフトウェアであり、RACは、ORACLEデータベースをクラスタ構成にするための拡張機能である。これらの連携によってクラスタシステムを構成し、障害発生時には他のノードへリソースを引き継がせることで、システムの運転を継続し、可用性を高めている。

RACには、共有ディスク方式をベースとした「キャッシュフュージョン アーキテクチャ」を採用している。共有ディスク方式とは、複数のノードでディスクを共有し、一体のシステムとして稼働させる方式である。従来の共有ディスク方式では、二つ以上のノードで同じデータブロックを更新しようとするときに、「待ち」が発生していた。しかし、ORACLEデータベースでは、RACのキャッシュフュージョンにより、この問題を解決している。キャッシュフュージョンは、インターコネクト(ノード間通信用LAN)を使用し、キャッシュデータを高速に他ノードに転送して、ノード間のデータの同期をとるものである。これにより、さまざまなノードの能力を統合して利用でき、システムのスケラビリティを向上させることができる。また、一つもしくは複数のノードに障害が発生した場合にも、残ったノードで負荷を分散することが可能である(図1参照)。

2.2 データベースとストレージの連携によるディザスタリカバリ

停電や地震などの災害によるミッションクリティカルシステムの停止は、ビジネスに大きな影響を与える。近年、このような事態に備えるための手法として、ディザスタリカバリが重要視されるようになってきた。ディザスタリカバリとは、致命的な災害などによる障害からの復旧を意味する。データベースシステムでは、データのバックアップなどにより、あらかじめ災害に備えることが可能であるが、バックアップを同じ場所(建物、地域)にとると、広域災害が起こったときにバックアップデータも失わ

※2) HACMPは、米国における米国International Business Machines Corp.の商標である。

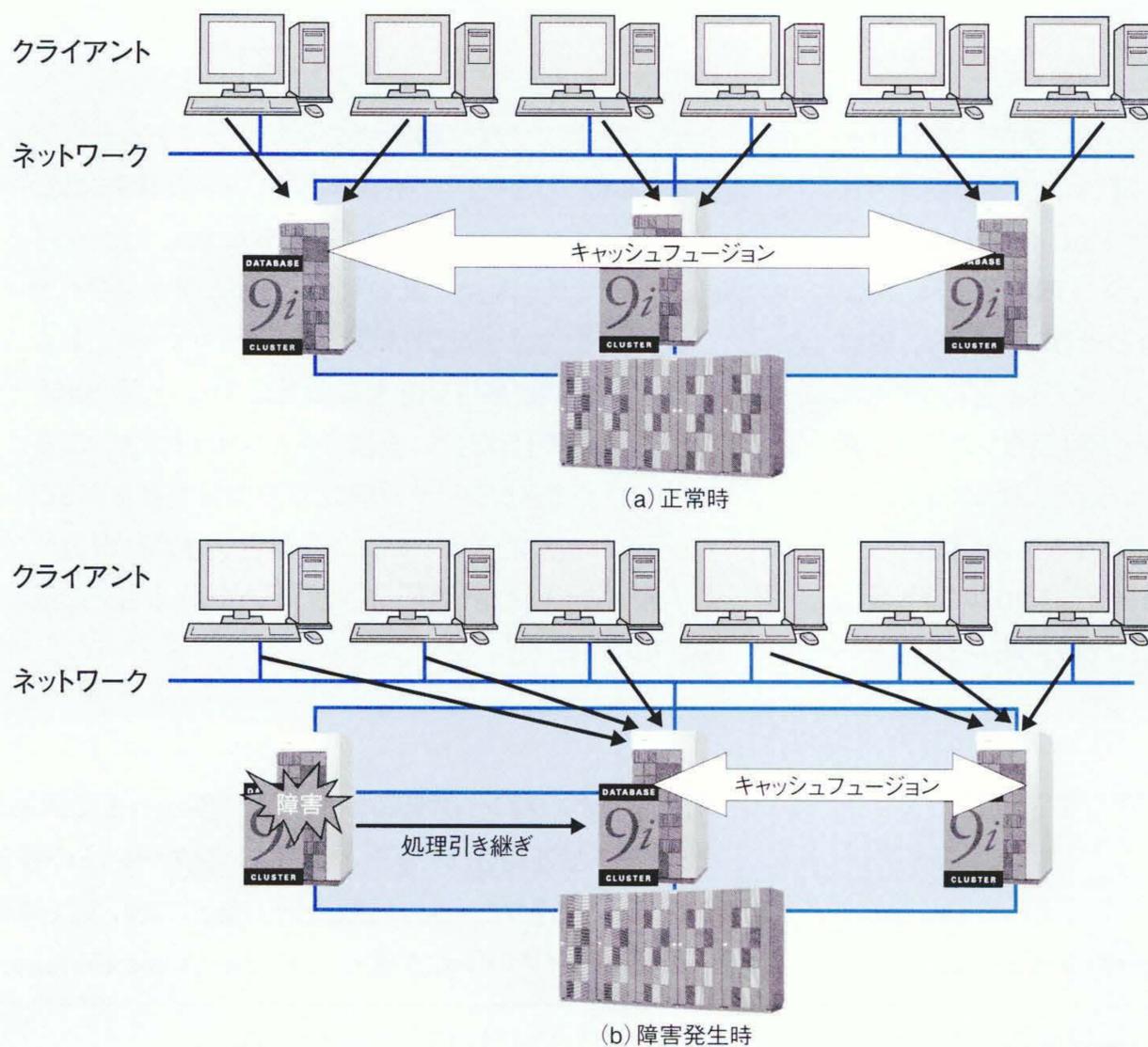


図1 HACMPとRACの組み合わせによるクラスタ構成例

HACMPとRACを組み合わせることにより、正常時には負荷分散ができ、スループットを向上させることができる。障害時には、障害が起きていないノードに処理を引き継がせる。

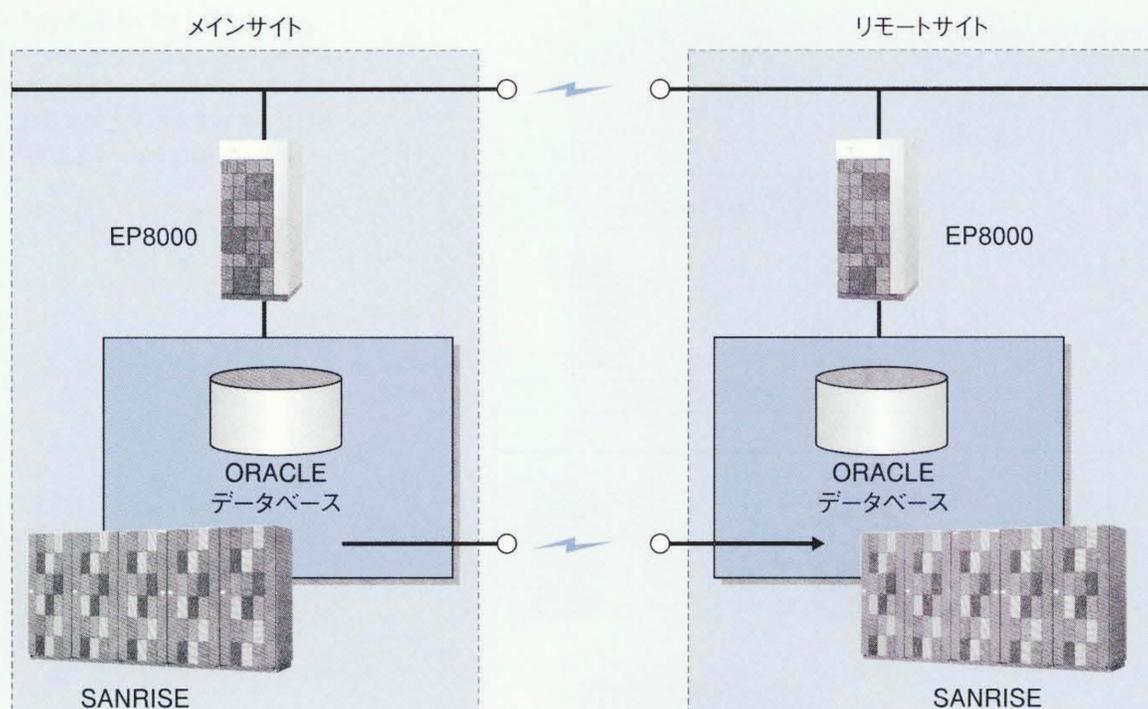


図2 ディザスタリカバリの構成例

遠隔地でデータの同期をとることにより、広域に及ぶ自然災害などに備える。

れる可能性がある。そのため、ネットワークを経由して、遠隔地にバックアップデータを保管することが有効である(図2参照)。

“Oracle 9i Data Guard”は、メインサイトのデータベースのリアルタイムな複製をリモートサイトに保持することにより、障害時に備えるORACLEデータベースの拡張機能である。

日立製作所の“TrueCopy”は、サーバフリー(非経由)でSANRISEストレージ間のリモートコピーを行う機能である。両者は単体でもディザスタリカバリでのバックアップができるが、これらを組み合わせることにより、システムに合ったさまざまなバックアップ方法が選択できるようになる。また、フォールバック(通信品質の調整)を適切かつ高速に行うこともできる。

2.3 サーバとストレージの連携による統合ディスク管理

EP8000とSANRISEの連携では、「EP8000高信頼化体系」に基づき、高信頼性、高可用性の強化を図っている。具体的には、「日立ファイバチャネルアダプタ」により、トレース機能・障害検出機能を強化して信頼性を高めている。さらに、「SANRISE専用デバイスドライバ」を追加することにより、ディスクドライバのレベルでも入出力監視強化やトラブルシュート強化が可能となり、信頼性と保守性をいっそう向上させることができる。

3 事前検証による高信頼・高可用システムの実現

2000年5月、日立製作所と日本オラクル株式会社は、共同で日立-オラクルSANソリューションテクノロジーセンター(以下、SSTCと言う。)を開設した¹⁾。SSTCは、SANRISEとORACLEデータベースを組み合わせることにより、24時間365日動き続けるビジネスに柔軟に対応しながら、日々増加し続けるビジネスデータを、効率的に管理、運用する企業情報シ

ステムの構築を支援する目的で設立された。

SSTCでは、EP8000をはじめとする「各種プラットフォーム+ORACLEデータベース+SANRISE」のクラスタ構成で、さまざまな検証を行っている。

例えば、障害発生時でのシステムの挙動検証として、想定される障害を模擬し、サーバ・データベース・ストレージがそれぞれどのような動きをするのかについて確認を行っている。想定される障害には、ノードダウン、インターコネクト障害、ファイバチャネル障害などがある。また、実システムでは、運用上、障害からの復旧時間を短縮することが非常に重要となる。これらの要求に対応するため、障害時でのフェイルオーバー(リモートサイトへの切替)時間の測定や、障害復旧時の動きなども詳細に検証している。

SSTCでの検証では、サイジングデータの取得を目的として、TPC-C業務モデルをベースに日立製作所が開発したベンチマークツールを使用している。TPC-Cとは、TPC(Transaction Processing Performance Council)が提供するベンチマークテストの規格であり、在庫管理のオンライントランザクション業務をモデルとしたものである^{2),3)}。したがって、このツールでは、実システム上で動作する業務システムに近いトランザクションの種類や負荷をシミュレーションすることができる。これにより、実際の運用を想定したチューニング方法の検討や、RACのスケラビリティの検証、高負荷での障害発生時での挙動の確認などが可能となる。これらの検証で得られた結果を基に、実際のシステム構成支援を行っている。

4 トータルソリューションの提供

日立製作所と日立OPSSは、「EP8000+ORACLEデータベース+SANRISE」の連携をサポートすることにより、さまざま

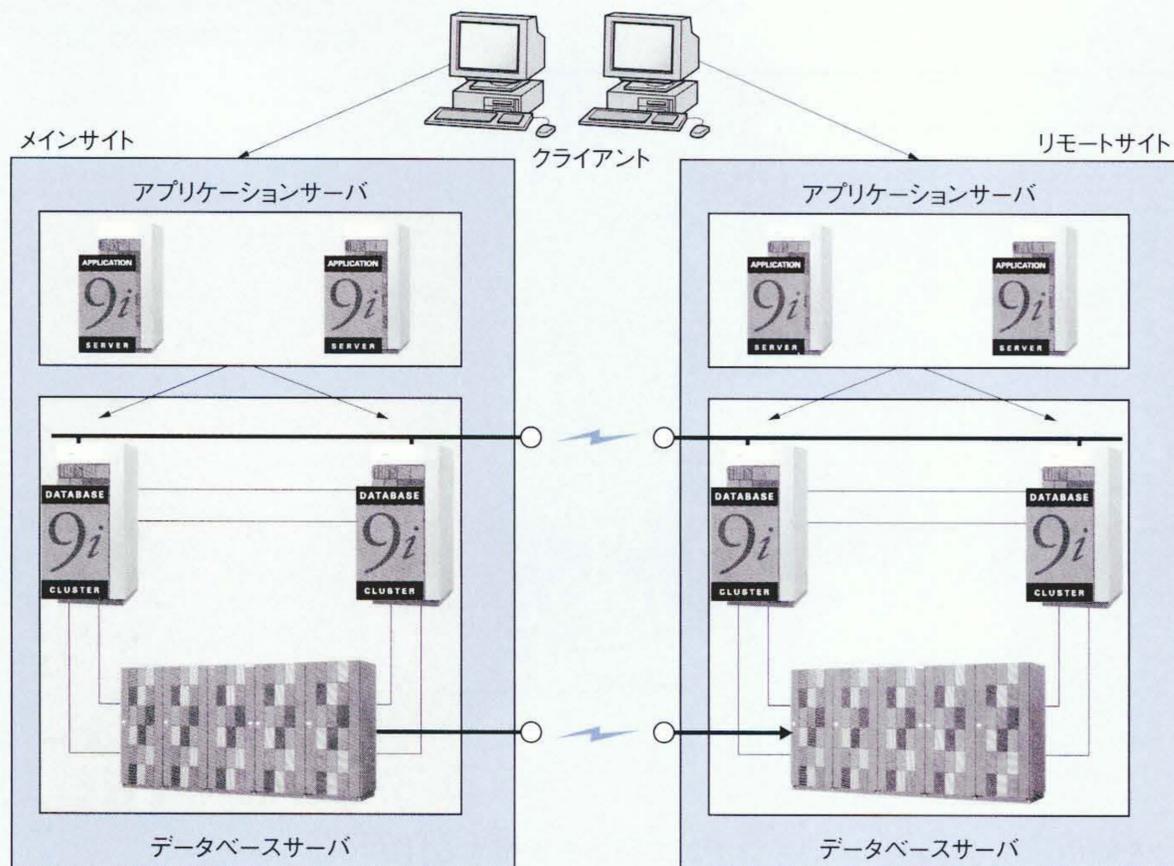


図3 高可用性システム構成例

メインとリモートサイト両方のデータベースサーバ、アプリケーションサーバをクラスタ構成にし、サーバとストレージ間のファイバチャネルを2本以上にするなどにより、いっそう高可用性システムを実現する。

な組み合わせのサービスを提案している。例えば、RACによるクラスタ構成とディザスタリカバリを組み合わせ、メインサイトのデータベースとリモートサイトのデータベースをそれぞれクラスタ構成にすることも可能である。これにより、災害だけでなく、システム障害に対しても可用性を高めることができる(図3参照)。

また、日立製作所と日立OPSSは、日本オラクル株式会社と連携し、SSTCで、EP8000をクラスタ構成にした環境で事前検証を行っている。この事前検証によって得られたストレージの構成法、クラスタ構成のデータベース構築、パフォーマンスチューニング、障害発生時のリカバリ手順などの実践的なノウハウを、高信頼・高可用性のシステムソリューションに活用している。

これら一連の活動により、導入時(データベースの構成、データベース・ストレージのチューニングなど)から導入後(障害時のサポートなど)までをトータルにサポートする、高付加価値なソリューションを提供している。

5 おわりに

ここでは、高性能サーバ・データベース・ストレージを連携させたシステムソリューションとクラスタ構成による高信頼・高可用性の検証について述べた。

今後ますます増大する高信頼・高可用性システムへのニーズを満たすためには、従来以上にさまざまな障害・災害への入念な対策を図った堅牢(ろう)なシステムが必要となってくる。このような状況を念頭に置きながら、日立製作所と日立

OPSSは、サーバ・データベース・ストレージの連携をさらに深め、さまざまな検証を行いつつ、顧客の要求にこたえるトータルサポートの提案といっそうのサービスの拡充を図っていく考えである。

参考文献など

- 1) SSTC (<http://www.sstc.gr.jp/>)
- 2) TPC (<http://tpc.org/>)
- 3) TPC BENCHMARK C, Standard Specification Revision 5.1 (2002.12)

執筆者紹介

福田 弘 忠



1988年日立製作所入社、情報・通信グループ ソフトウェア事業部 オラクルビジネス統括センタ 所属
現在、株式会社日立オープンプラットフォームソリューションズでオラクル関連製品の評価・販売支援業務に従事
E-mail: fukuta_h@hitachi-opss.com

矢田 龍 太郎



2001年日立製作所入社、情報・通信グループ ソフトウェア事業部 オラクルビジネス統括センタ 所属
現在、株式会社日立オープンプラットフォームソリューションズでオラクル関連製品の評価・販売支援業務に従事
E-mail: r_yada@hitachi-opss.com

新田 英 之



2002年日立製作所入社、情報・通信グループ ソフトウェア事業部 オラクルビジネス統括センタ 所属
現在、株式会社日立オープンプラットフォームソリューションズでオラクル関連製品の評価・販売支援業務に従事
電子情報通信学会会員
E-mail: h_nitta@hitachi-opss.com