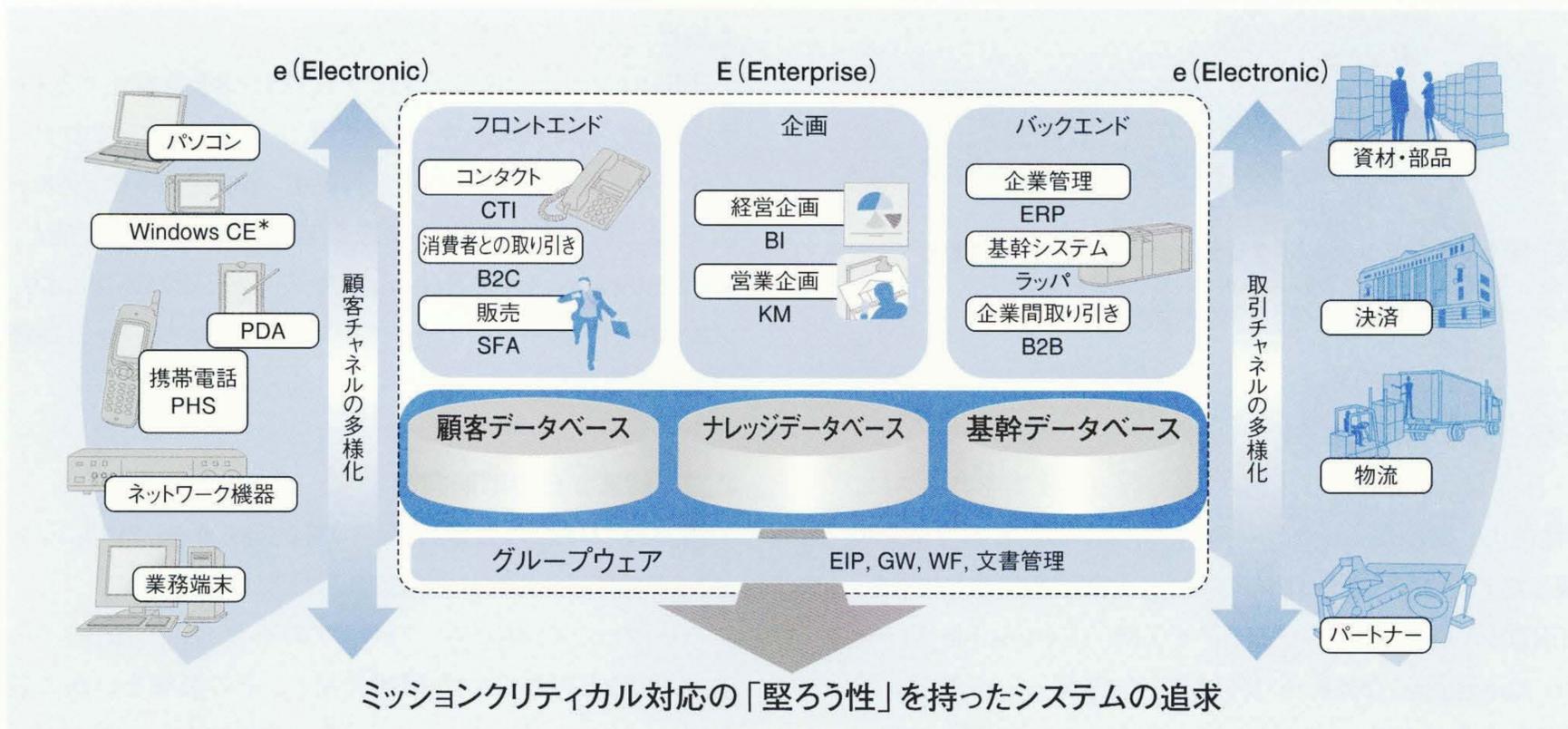


ネットビジネスを支えるミッション クリティカル システムに対応したデータベース

HiRDB Version 6

Hitachi's Database Management System for Network Businesses

原 憲宏 Norihiro Hara 河村 信男 Nobuo Kawamura 阿部 淳 Jun Abe



注：略語説明ほか

PDA(Personal Digital Assistant), CTI(Computer-Telephony Integration), B2C(Business to Consumer), SFA(Sales Force Automation)
 BI(Business Intelligence), KM(Knowledge Management), ERP(Enterprise Resource Planning), B2B(Business to Business)
 EIP(Enterprise Information Portal), GW(Groupware), WF(Workflow)
 *Windowsは、米国およびその他の国における米国Microsoft Corp.の登録商標である。

HiRDB Version 6のねらい

24時間365日稼働環境にある現在のネットビジネスを担う業務システムには、基幹(Enterprise)システムと同等の堅ろう性が求められる。このようなネット ビジネス システムを支えるデータ基盤“HiRDB Version 6”では、ミッション クリティカル システムへの対応を目的に、信頼性をさらに強化した。

サービスの停止が重大なビジネスチャンスの喪失に直結するネットビジネスの世界では、24時間365日稼働し続ける堅ろう性を備えたシステムが求められる。その中でも特に、障害発生時の影響の局所化と迅速なシステム回復が重要である。

日立製作所は、これらのニーズにこたえるデータ基盤として、“HiRDB Version 6”を提供している。HiRDB Version 6では、ネット ビジネス アプリケーショ

ン構築のための先進技術を採用するとともに、ノンストップサービスを実現するための信頼性と運用性を強化し、ミッション クリティカル システムへのいっそうの対応を図った。「HAモニタ」との連携やホットスタンバイ、トランザクションキューイングによる高速系切り替えなどの優れた機能により、24時間365日ノンストップのネットビジネスを実現する。

1 はじめに

今日の24時間365日のビジネス環境では、システム障害などのサービス停止によって重大なビジネスチャンスを逃しかね

ない。システムの停止が、業務の中断だけでなく、多大な金銭的損失や社会的信用の失墜につながる危険性があり、止まることが許されないシステムを「ミッション クリティカル システム」と言う。ビジネスチャンスを逃さず、ネットビジネスの世界で成功するためには、ミッション クリティカル システムとしての高



図1 安心できるミッションクリティカルシステムの仕組み

高い信頼性と運用性を実現する技術と、自社開発ならではのサポートサービスにより、24時間365日の安定稼働を実現する。

い堅ろう性が求められる。

日立製作所は、メインフレームで培った技術をベースに、「性能」、「信頼性」、「スケーラビリティ」といった基本性能を重視したデータベース管理システム“HiRDB”を提供してきた。HiRDB Version 6では、Java[®]やXML (Extensible Markup Language) などネット ビジネス アプリケーション基盤に求められる先進技術への対応はもちろんのこと、24時間365日ノンストップのビジネスを支えるため、ミッションクリティカルシステムへの対応としていっそうの高信頼化を図った(図1参照)。

ここでは、ビジネスのノンストップ化を実現するためのアプローチとHiRDB Version 6の対応機能、特に高速系切り替え機能について述べる。

2 ノンストップ化実現へのアプローチ

ノンストップ化実現へのアプローチには、ハードウェアの二重化などによる障害発生防止以外に、(1) 計画停止時間の短縮、および(2) 障害からの早期回復がある。

2.1 計画停止時間の短縮

第1のアプローチは、定期的なデータベース再編成やシステム拡張などを実施するためにデータベース マネージメントシステム(DBMS)を停止させる、いわゆる「計画停止」の時間を極力短縮することである。

※) JavaおよびすべてのJava関連の商標およびロゴは、米国およびその他の国における米国Sun Microsystems, Inc.の商標または登録商標である。

ビジネス規模の拡大やネットビジネス参入企業の増大などに伴い、DBMSが扱うトランザクション量とデータ量は膨大なものとなっている。例えば、大量かつ頻繁なデータの追加と削除により、データベース内のデータの断片化、いわゆる「フラグメンテーション」が増大する。定期的にデータを再格納(データベースの再編成)することにより、フラグメンテーションを解消し、安定性能を確保する必要がある。しかし、このようなデータベースのメンテナンスのためにサービスを止めたりDBMSを停止したりすることは、ユーザーにとって重要なビジネスチャンスの喪失につながる危険性がある。また、サービスを止め、バッチ処理を実行する場合も同様である。

HiRDB Version 6では、(1) データベースを多重化するインナレプリカ、(2) データベース静止化、(3) オンラインでのデータベース再編成、データベース再配置、データ・インデックス一括登録、(4) データ格納領域の自動拡張などの機能と、SAN(Storage Area Network)やストレージとの連携により、システムを停止せずにデータベースメンテナンスとバッチ処理の実行を実現した。

2.2 障害からの早期回復

第2のアプローチは、障害発生時にできるかぎり迅速にシステムを回復し、サービスを再開することである。

ハードウェアの故障やソフトウェアの不具合、誤操作、さらに、災害などのさまざまな障害に対し、その影響をいかに最小限にとどめ、一刻も早くシステムを回復させるか、または、いかにサービスを止めずに提供し続けるかが重要である。

HiRDB Version 6では、高速系切り替え機能により、このような障害からの早期回復が行える。

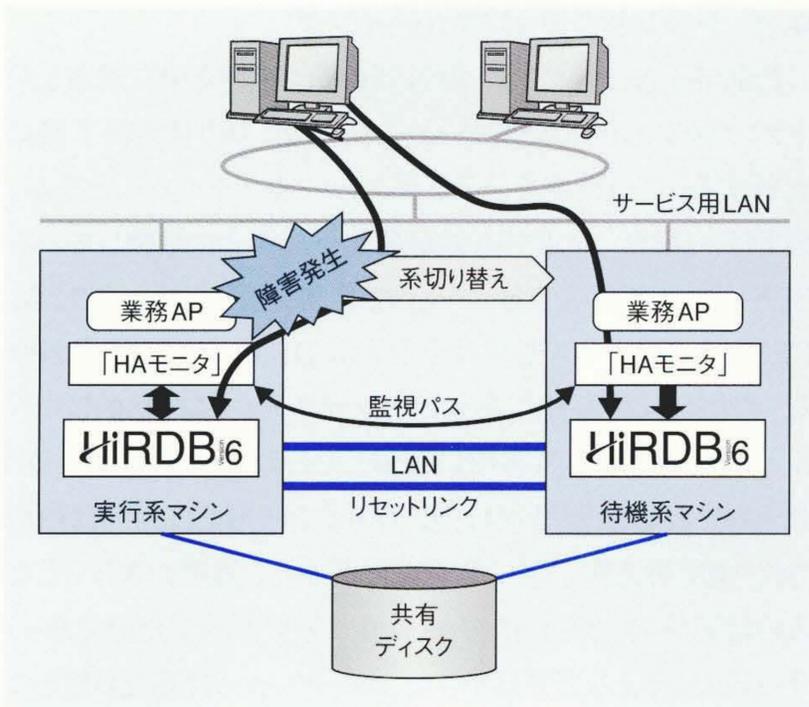
3 高速系切り替え機能

系切り替えでは、サービス実行用の実行系マシンと、待機系マシンを用意しておく。サービス実行中に障害が発生すると、クラスタソフトウェアと連携して待機系マシンに自動的にサービス実行を切り替え、データベースを回復した後にサービスを続行する(図2参照)。障害検知と系の切り替え制御をクラスタソフトウェアが行い、HiRDBが切り替え後のデータベース回復処理を行う。

HiRDB Version 6では、(1) 後述する「HAモニタ」との連携による障害検知と、(2) ホットスタンバイによるHiRDBの高速再開により、系切り替えに要する時間を大幅に短縮した。

3.1 HAモニタとの連携による障害検知

HiRDBでは、各種プラットフォーム上のクラスタソフトウェアと連携し、系切り替えを行うことができる。特に、日立製作所のクラスタソフトウェアである高信頼化システム監視機能「HA



注：略語説明 AP(Application)

図2 系切り替えの仕組み

クラスタソフトウェアで稼動状況を監視し、障害発生時に待機系へ切り替える。切り替え先のHiRDBでデータベースが回復し、サービスを再開する。

モニタ]との連携では、共用メモリを介したHAモニタとの生死監視により、HiRDBのダウンやスロウダウンを検知する。また、実行系と待機系との間で、HAモニタが通信経路を用いて互いの正常動作を確認するハートビート信号により、系の障害を検知する。

HAモニタ連携による系切り替えの流れは、(1) HAモニタによる障害の瞬時検知、(2) 待機系へのリソース(ネットワーク、ディスク)の切り替え、(3) データベースの回復、(4) トランザクションの決着(ロールバック)、(5) サービスの再開のようになる(図3参照)。

HiRDBでは、HAモニタに対し、共用メモリを介して定期的に生死情報を伝達している。HAモニタでは、メモリの更新状況を参照し、規定時間に更新されていない場合には障害と判断し、待機系に切り替える。障害の検知は瞬時であり、他のクラスタソフトウェアを用いた場合に比べ、系切り替えに要する時間を短縮することができる。

データベースの回復では、障害発生時にメモリ上のデータベースバッファからデータベースに書き込まれていなかった更新を回復する。これは、データベース更新ログを使い、更新処理を再実行することによって行う。HiRDBでは、データベース更新ログを、データベースを構成するディスクごとに分け、複数のプロセスで再実行を並列に処理することにより、データベースの回復を短時間で行う。

トランザクションの決着では、障害発生時に処理中であったトランザクションをロールバックする。データベースの回復後、このトランザクションの決着が完了するのを待たずに新トランザクションの受け付けを開始することにより、早期のサービス再開を実現している(デイレードラン)。

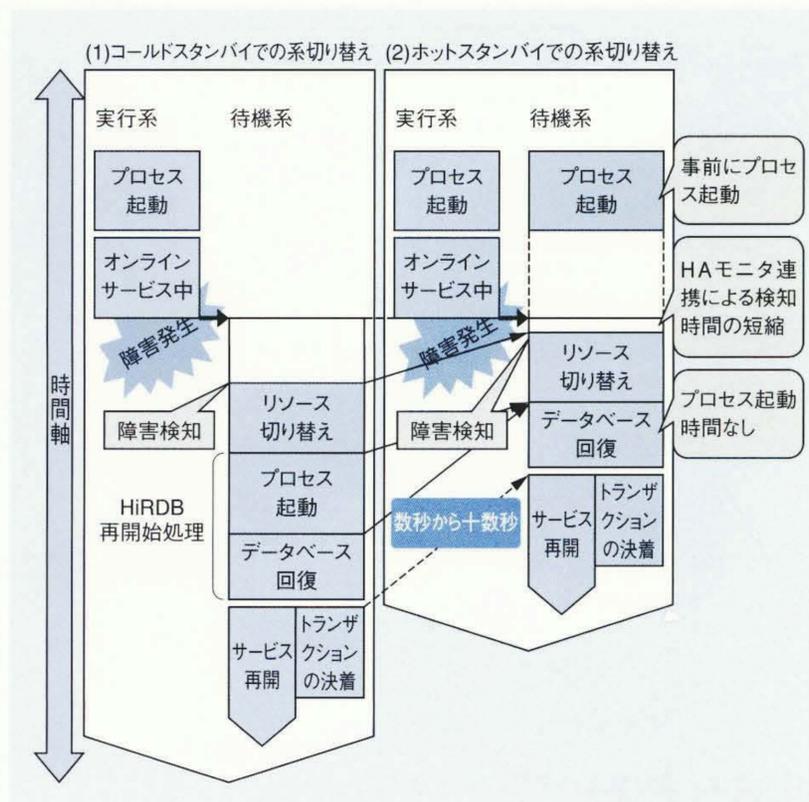


図3 「HAモニタ」との連携とホットスタンバイによる系切り替え時間の短縮の仕組み

日立製作所のクラスタソフトウェアであるシステム監視機能「HAモニタ」との連携によって障害を瞬時に検知し、事前に待機系でプロセスを起動しておく「ホットスタンバイ」を用いることにより、切り替え時のHiRDB再開始を短時間で行う。

3.2 ホットスタンバイによるHiRDBの高速再開始

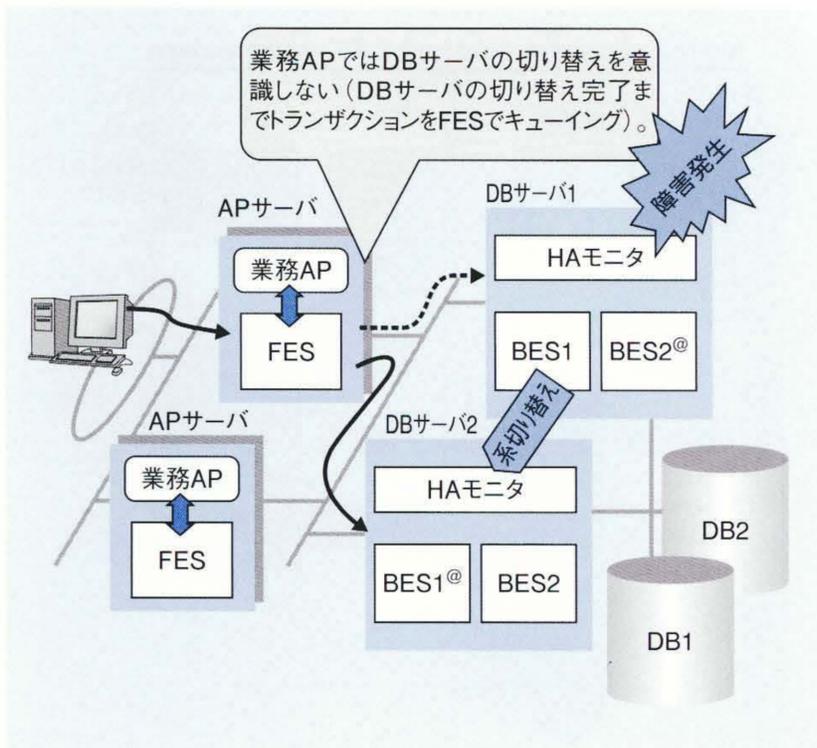
ホットスタンバイ機能により、ユーザーのサーバプロセスとHiRDBシステムプロセスを立ち上げた状態で待機系を準備しておくことができる。このため、障害発生時の系切り替えでは、すでにプロセスが立ち上がっているため、データベースの回復処理からHiRDBを即座に再開することができる。系切り替え後にシステムの起動を行う「コールドスタンバイ」と比較すると、リソース切り替え後に進んでいたプロセス起動に相当する時間の分だけ、系切り替えに要する時間を短縮することができる。

HiRDB Version 6の最新リリース版では、上述した方式により、障害発生後、数秒から十数秒でシステムを回復することができる。

4 切り替え中のトランザクションキューイング

4.1 HiRDB/Parallel Serverの構成

HiRDBで提供する大規模システム用モデルのHiRDB/Parallel Server(以下、Parallel Serverと言う。)では、ノード間でデータベース共用を行わない「シェアードナッシング方式」を採用し、ノードごとの並列処理により、高性能とスケーラビリティを実現している。Parallel Serverは、フロントエンドサーバ(FES)やバックエンドサーバ(BES)などの複数のサーバで構成しており、全体で一つのデータベース マネージメントシステムとして機能する。FESで業務アプリケーションからのサービス要求を受け付け、複数ノードに配置されたBESで並



注1：@は、切り替え時の待機側を示す。

注2：略語説明 DB(Database)、FES(Frontend Server)
BES(Backend Server)

図4 HiRDB/Parallel Serverでの系切り替え中のトランザクションキューイングの仕組み

バックエンドサーバ(BES)の系切り替え中に到着したトランザクションは、切り替え終了後に正常に実行される。アプリケーションでは、BESの切り替えを意識することなく、業務を継続することができる。

列にデータベース処理を行う(図4参照)。

このシステム構成例では、複数のアプリケーションサーバに業務アプリケーションとともにFESがそれぞれ配置されている。このため、アプリケーションサーバを用いたロードバランシングが可能である。アプリケーションサーバの負荷が増した場合、アプリケーションサーバを追加することにより、負荷を分散する。また、サービス要求に対し、FESで該当するBESを判断するので、業務アプリケーションではBESを意識する必要がない。系切り替えの構成に関しては、DBサーバ1とDBサーバ2との間で相互に系切り替えを行う方式としている。現用では、DBサーバ1上のBES1がデータベース“DB1”に、DBサーバ2上のBES2がデータベース“DB2”にそれぞれアクセスし、業務を行う。

4.2 トランザクションキューイング

Parallel Serverでは、障害時の系切り替え中に到着したトランザクションをFESでいったん待たせ、切り替え終了後に実行する。この機能を「トランザクションキューイング」と呼ぶ。例えば、DBサーバ2で障害が発生した場合、待機しているBES1に切り替わり、DBサーバ2で業務を続行する。このとき、FESからのアクセス要求は自動的にDBサーバ2に切り替わる。障害時に実行中であったトランザクションは自動的にロールバックされるが、障害中に到着したトランザクションは、DBサーバ2でBES1が処理可能になるまでFESで自動的に待たされ、切り替え終了後に処理を再開する。処理がエラーとなるのは障害時に実行中であったトランザクションだけであり、サービスに与える影響は小さく、ユーザーからは系切り替えの時間が見えない。

5 おわりに

ここでは、ビジネスのノンストップ化を実現するためのアプローチ、HiRDB Version 6の対応機能、および障害からの早期回復を実現する高速系切り替え機能について述べた。

HiRDB Version 6では、Javaストアドプロシジャやデジタルコンテンツプラグインなどウェブサービスやブロードバンドに関わる機能もサポートしている。

日立製作所は、今後も、可用性をはじめとするミッションクリティカルシステムへの対応機能をさらに強化し、自社開発ならではのサポートサービスと合わせ、24時間365日稼働し続ける、堅ろう性を備えたシステムを提供することにより、顧客のネットビジネスを支援していく考えである。

参考文献など

- 1) 河村, 外: SAN環境におけるHiRDB, Oracleデータベースシステムの構築, 日立評論, 83, 5, 359~362(2001.5)
- 2) データベース特別取材班: まるごとわかるデータベース読本, 技術評論社(2000.7)
- 3) <http://www.hitachi.co.jp/Prod/comp/soft1/hirdb/>

執筆者紹介



原 憲宏

1992年日立製作所入社, 情報・通信グループ ビジネスソリューション事業部 先端ミドルウェア開発部 所属
現在, データベース管理システムの研究・開発に従事
E-mail: nhara@itg.hitachi.co.jp



阿部 淳

1984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 ネットワークソフトウェア本部 データベース設計部 所属
現在, データベース関連の事業推進に従事
情報処理学会会員
E-mail: abe_j@itg.hitachi.co.jp



河村信男

1981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 ネットワークソフトウェア本部 データベース設計部 所属
現在, HiRDBの開発に従事
情報処理学会会員
E-mail: nokawamu@itg.hitachi.co.jp