

ビジネスの多様化と変化に応える “柔軟性”を強化した 最新バージョン7.2をリリース

日立のリレーショナルデータベース「HiRDB（ハイアールディービー）」は、顧客のシステムに対する信頼性への要求に応えるために、無停止運用の可能性を徹底して追求してきた。2005年3月には、その最新バージョンとなる「HiRDB Version7.2」がリリースされ自律運用機能が大幅に強化された。これによりビジネスの多様化と変化に応える“柔軟性”が、これまで以上に高まったのだ。さらに電子文書データの管理に適したセキュリティ機能や、高パフォーマンスなディザスタリカバリなど、さまざまな面からの機能拡張を続けている。本稿ではこれらの新機能を中心に、具体的な機能拡張のポイントを紹介する。

フリーランスライター 岩崎史絵 IWASAKI,Shie

無停止運用を追究するHiRDBが 新バージョンで3つの機能を拡張

日立では、顧客が煩雑な運用管理に頭を悩ませることなく本来のコアビジネスに集中できる環境を、サービスプラットフォームコンセプトであるHarmonious Computing（ハーモニアスコンピューティング）のもとで実現している。これはITに自律性を持たせることで、ビジネス環境の変化への迅速な対応を可能にするという考え方だ。特にユーザーが必要とするタイミングでITサービスを

利用できるようにするには、「ダウンしない」「たとえダウンしても速やかに復旧する」という高い信頼性/可用性が求められる。この実現に向け、日立はハード/ソフトウェアからサービスまでをトータルに提供できる総合力を活かし追求している。そのデータ管理基盤となるのが同社のリレーショナルデータベース製品「HiRDB Version7」である。

「もともとHiRDBは、初期バージョンからParallel Serverをリリースするなど、スケラビリティを重視した設計がなされてきましたが、バージョン7では無停止運用の実現に取り組みました(図1)」と日立製作所 ソフトウェア事業部 DB設計部部長の阿部 淳氏は話す。

日立には、メインフレームで長年培ってきた高可用性技術のノウハウが豊富にある。これらをオープン系RDBMSに投入することで、バッチ処理やバックアップ、データベースの再編成時でも、オンライン性能を十分確保したままサービスを継続できる。また万一、サーバーに障害が発生した場合でも、クラスタ構成にしておくことで数秒～十数秒程度の短時間でフェイルオー

バーすることが可能だという。さらにHiRDBを設計してきた技術者が、国内にいることも重要だ。データベースを熟知した設計者が、必要に応じて直接サポートできる体制を整えている。つまりHiRDBは、機能とサービス双方の側面から、ノンストップを追及してきたと言える。

今回、このHiRDBがバージョン7.2となり、大きく次の3つの点で機能拡張されている。

- ポリシーに従った自律運用機能
- e-文書法を見据えたセキュリティ機能
- 広域災害を想定したディザスタリカバリ機能

以下に、それぞれ具体的に見ていく。

HiRDB Version7.2の 機能拡張ポイント①

ポリシーに従った自律運用

日々のメンテナンスやバックアップ運用など、定型的な運用サイクルはある程度自動化でき、運用管理者を悩ませることは少なくなった。だが、中長期的な運用サイクルはどうだ

うか。データベースサーバーを例に挙げると、データエリアの拡張やデータベース再編成、サーバーマシンの構成変更など、日々の業務負荷に応じてシステムの状態を適切に保つためには、さまざまな運用管理を行なうことが必要になってくる。

HiRDB Version7.2では、このような中長期的運用サイクルまでも視野に入れ、運用管理自体をシステムが自動的に行なうための機能を強化している。

オンラインスケールアウトで ITリソース配分を最適化

自律運用に向けた機能として、はじめに注目したいのが「オンラインスケールアウト」である。これは事前に決めておいたポリシーに従って、DBサーバーの構成を動的かつ自動的に変更できるという機能だ。例えば昼間はオンライン5台+予備機3台で構成されているDBサーバーを、夜間にはオンライン3台+バッチ5台に変更したり、急激な負荷増大時にはオンラインの台数を増やす、といったサーバーの構成変更をサービスを停止せずに自動実行できるのである(図2)。

「従来は個々の業務サーバーごとに、ピーク時を想定してITリソースを割り当てる必要がありました。このオンラインスケールアウトの実現により、システム全体のITリソース配分を最適化できるので、急激な負荷増大に対応できるだけでなく、遊休リソースの削減も可能です」(日立製作所 ソフトウェア事業部 DB設計部技師 石川太一氏)。

ではオンラインスケールアウトの自動化方法について、具体的に見ていこう。

HiRDBは、BES(バックエンドサーバー)^{*1}というコンポーネント単位にデータベース

^{*1}注1: DBサーバープロセス、DBバッファキャッシュから構成されるDBアクセスコンポーネント。

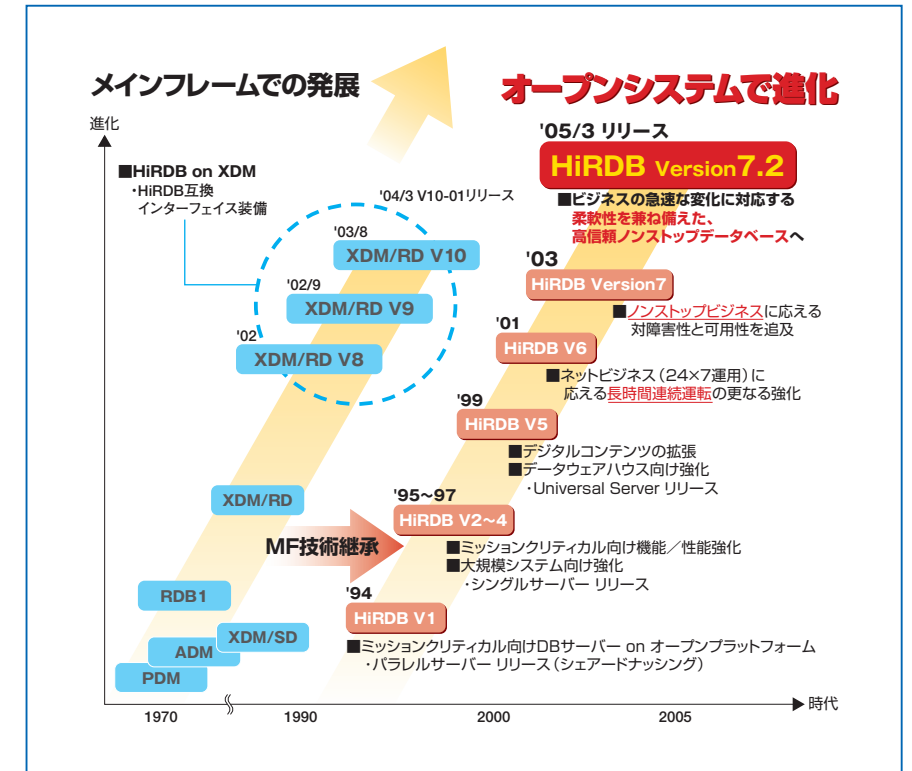
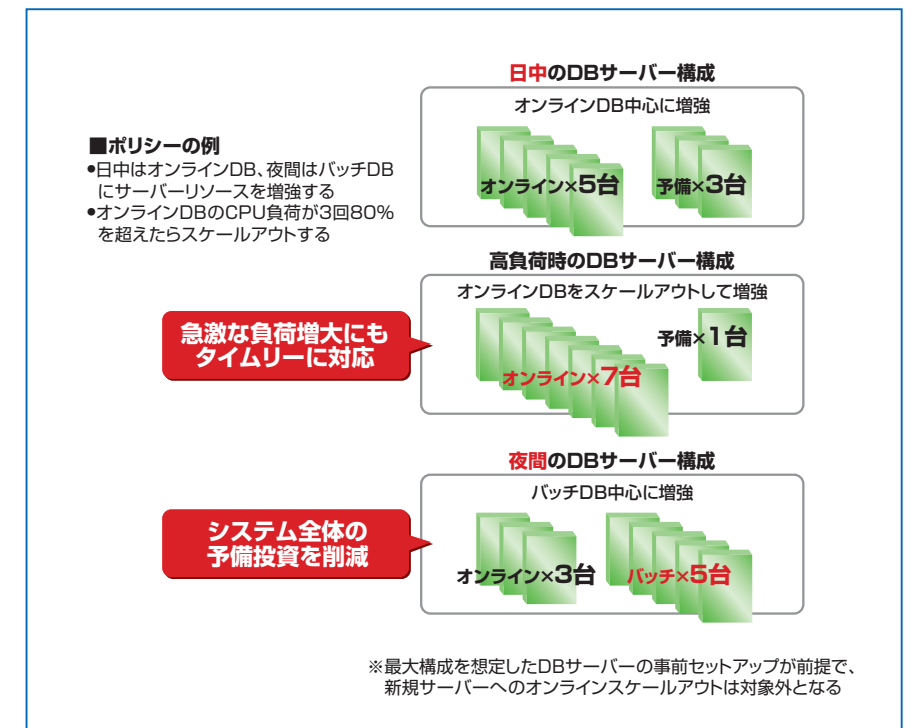


図1: これまでのHiRDBの取り組み



※最大構成を想定したDBサーバーの事前セットアップが前提で、新規サーバーへのオンラインスケールアウトは対象外となる

図2: オンラインスケールアウト運用例



日立製作所 ソフトウェア事業部 DB設計部部長 阿部 淳氏

リソースを細分化して管理している。図2の例では、ブレードサーバー上に8台のサーバーモジュールが組み込まれており、これらのサーバーモジュールに対して、BESを均等に割り当てることで負荷分散を行なっている。

例えば、日中はオンラインサービスで利用するBESグループを5台のサーバーモジュールに割り当てておき、夜間はこれを3台に縮小して再配置し、バッチサービスで利用するBESグループを残り5台のサーバーモジュールに割り当てるといった具合だ。

この、日中/夜間のサーバー構成は「シナリオ」として日立の統合システム運用管理ソフトウェア「JP1」に登録し、構成変更を自動化

できる。つまり、ポリシーに従った自律運用は、あらかじめ作成したシナリオによる自動運転で実現するのだ。

シナリオの作成はJP1のGUIで簡単に行なえる(画面1)。HiRDBでは、BESの割当変更を行なう「シナリオテンプレート」を提供しており、運用管理者はこのテンプレートをカスタマイズしてBESのサーバー配分を変更するシナリオを作るだけで良い。カスタマイズ操作は簡単で、移動するBESに対して、移動先のサーバー情報となるホスト名称など、わずか6つのパラメータを設定するだけで(画面2)。

この操作を移動対象となるBESの分だけ繰り返す。最後に作成したシナリオを、JP1の「ジョブネット」として展開し、運用をスケジューリングする(画面3)。ジョブネットは自動展開されるので、運用管理者はジョブの詳細を意識する必要はない。例えば18:00になったら、夜間のサーバー構成に変更するジョブネットを自動実行するという具合に、登録されたジョブネット単位に運用を組み立てるだけで良い。

また、急激な負荷増大時にはオンラインのサーバー台数を増やすといった設定も、JP1のGUIで簡単に行なえる(画面4)。例えばサーバーのCPU利用率が80%を3回超えたら、アラートするよう「しきい値」を設定してお

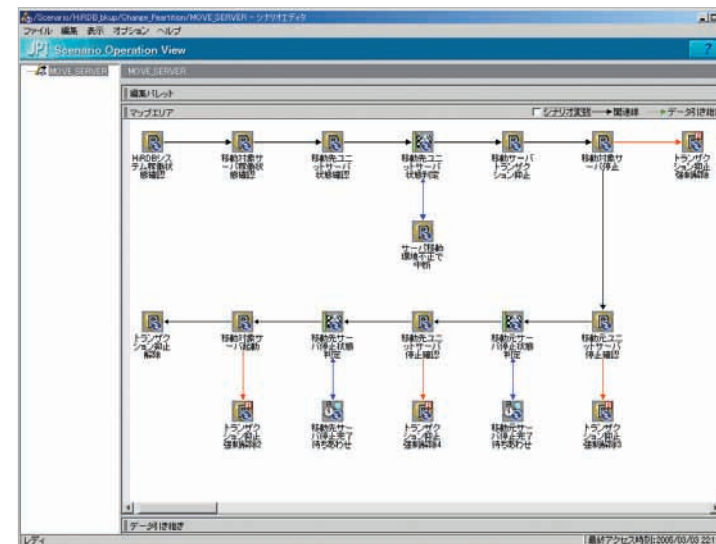
く。さらに該当アラートを検知した場合、サーバー構成を変更するジョブネットを自動実行するという具合だ。

● 新開発のActive-Activeクラスタ機能でサーバーの拡張性と可用性を両立

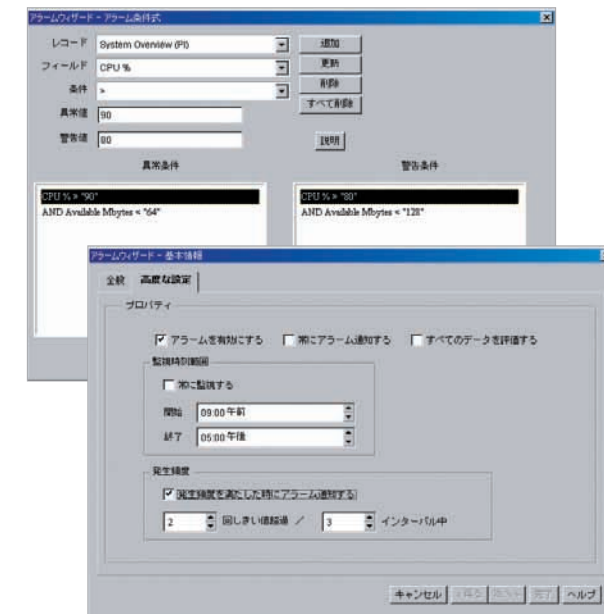
オンラインスケールアウトは、「Active-Activeクラスタ機能」が基盤技術となっている。前述の例では、あらかじめ8台のサーバーをクラスタ化し、論理的に1つのHiRDBサーバーとして並列稼働させている。このクラスタサーバー上に、オンラインで利用するBESグループやバッチで利用するBESグループを動的に割り当てるといったわけだ。ここではこのActive-Activeクラスタ機能に注目したい。

HiRDBはアーキテクチャとしてShared Nothing方式を採用しており、スケラビリティを重視してきた。この方式は各サーバーの独立性を高め、並列処理の効率を高めていくというアーキテクチャだ。このためサーバー台数にリニアに比例した高スループットを発揮できる。

「十数台のサーバーで並列稼働させている事例も多数あります。ビジネスの拡大に伴って、管理するデータやアクセス数が増加して



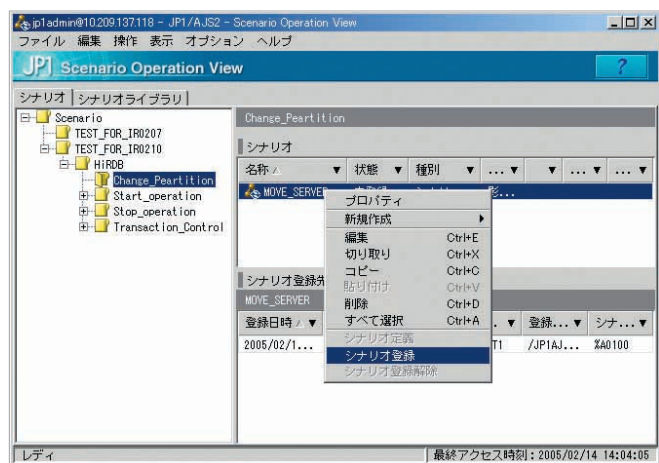
画面3: 自動展開されたジョブネットの例



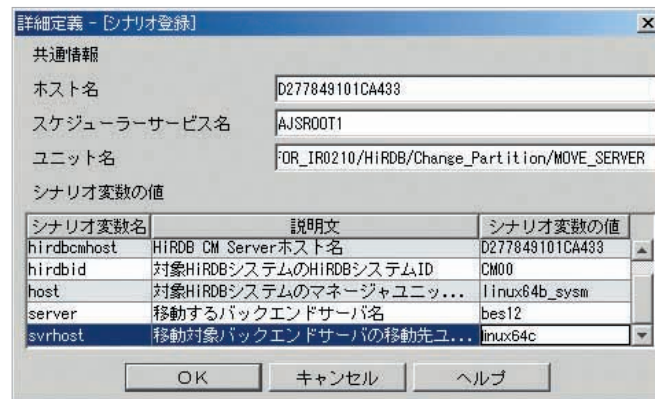
画面4: サーバーのCPU利用率が80%を3回超えたらアラートするよう設定



日立製作所 ソフトウェア事業部 DB設計部 技師 石川太一氏



画面1: シナリオの作成画面



画面2: シナリオテンプレートのカスタマイズ画面

も、サーバー台数を増やしてシステムを拡張できる。まさにShared Nothingアーキテクチャは、ブレードサーバーの考え方に適していると言えます(阿部氏)。

一方でサーバーの独立性が高いShared Nothingアーキテクチャは、可用性の面での問題点が指摘されていた。クラスタを構成する一部のサーバーがダウンした場合、負荷分散のためのサーバー構成変更が難しいという問題だ。これがHiRDBでは、Active-Activeクラスタ機能により解決されたという。つまり、一部のサーバーがダウンした場合でも、そのサーバーが担当していたコンポーネント(BES)を、動的にほかのActiveなサーバーに移行できるようになったのだ。

加えて、この移行に必要な時間も非常に迅速だ。「ミドルウェアがダウンしたケースなら数秒程度、もしハードウェアがダウンした場合でも数秒〜十数秒で完了します(阿部氏)。日立には高信頼化システム監視機能「HA

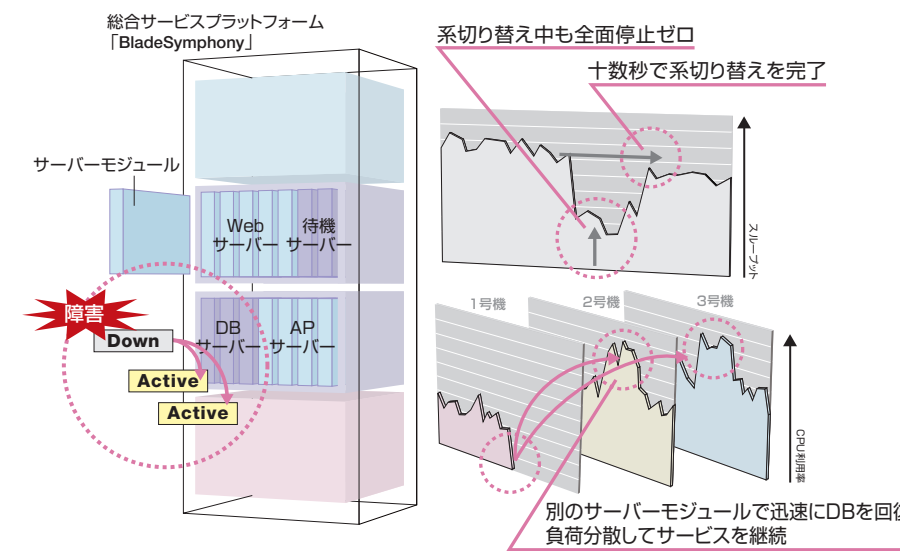


図3: Active-Activeクラスタの例

この例では3台のクラスタサーバーでHiRDBを稼働させている。1号機がダウンすると、その直後に2号機、3号機で回復処理が始まる様子を、それぞれのCPU利用率で確認できる。また、ダウンした1号機のCPU負荷も均等に分散されており、DBサーバー全体のスループットを維持している。ポイントは、1号機ダウンからわずか数秒で負荷分散が完了した点、そしてサーバー構成変更時、全面ダウンにならない点だ。

モニタ」というクラスタソフトウェア製品があり、これが常にサーバーの稼働状態を監視しているのだ。このため監視対象の動作に異常が発見された場合には、直ちに系切り替え処理が実行される。これまでもHiRDBは、Active-Standbyクラスタでは数秒～十数秒の系切り替えを実現していたが、今回これをActive-Activeクラスタに応用することで、サーバー障害への対応力をさらに高めることが可能になったのである。

さらに、HiRDBでは「わずか数秒というサーバー構成変更の瞬間も、受け付けた新規トランザクションを保持するトランザクションキューイング機能を備えています。また、障害部位を局所化しており、正常なサーバーは処理を継続していますので、サーバー構成変更時の全面ダウンもありません」(石川氏)という。これによ

り、オンラインサービスを利用するエンドユーザーにはサーバー障害さえも気づかせないとしている。もちろんこの技術を基盤とするオンラインスケールアウトでも、サーバー構成変更は高速である。この具体例をデモンストレーションで見ることができた(図3)。HiRDBのホームページでも公開)。

オンラインとバッチでスムーズなデータ連携を実現

前述のオンラインスケールアウトで挙げた、日中と夜間でのサーバー共有によるメリットは、ITリソースの配分を最適化できることだけに留まらない。例えば、取引履歴のようにオンラインDBに逐次格納されるデータに対してバッチ集計処理する場合、一般的にはオンライン

サービスを停止してDB更新を抑制する必要がある。これがHiRDBでは、オンラインサービスを止める必要がなく、バッチ処理も並行して実行できる。

「HiRDBでは、ある時点のオンラインDBを高速に複製し、その複製DBをバッチ集計のためにアクセスできるインナレプリカ機能を提供しています。オンラインとバッチでスムーズなデータ連携を実現し、バッチアプリケーションからは複製DBをアクセスして集計処理などを行なえるため、I/O競合が発生せず、オンライン性能へ影響を与えません(図4)」(石川氏)。

では次に、このオンラインと並行してバッチ処理を実行する運用手順について具体的に見ていこう。

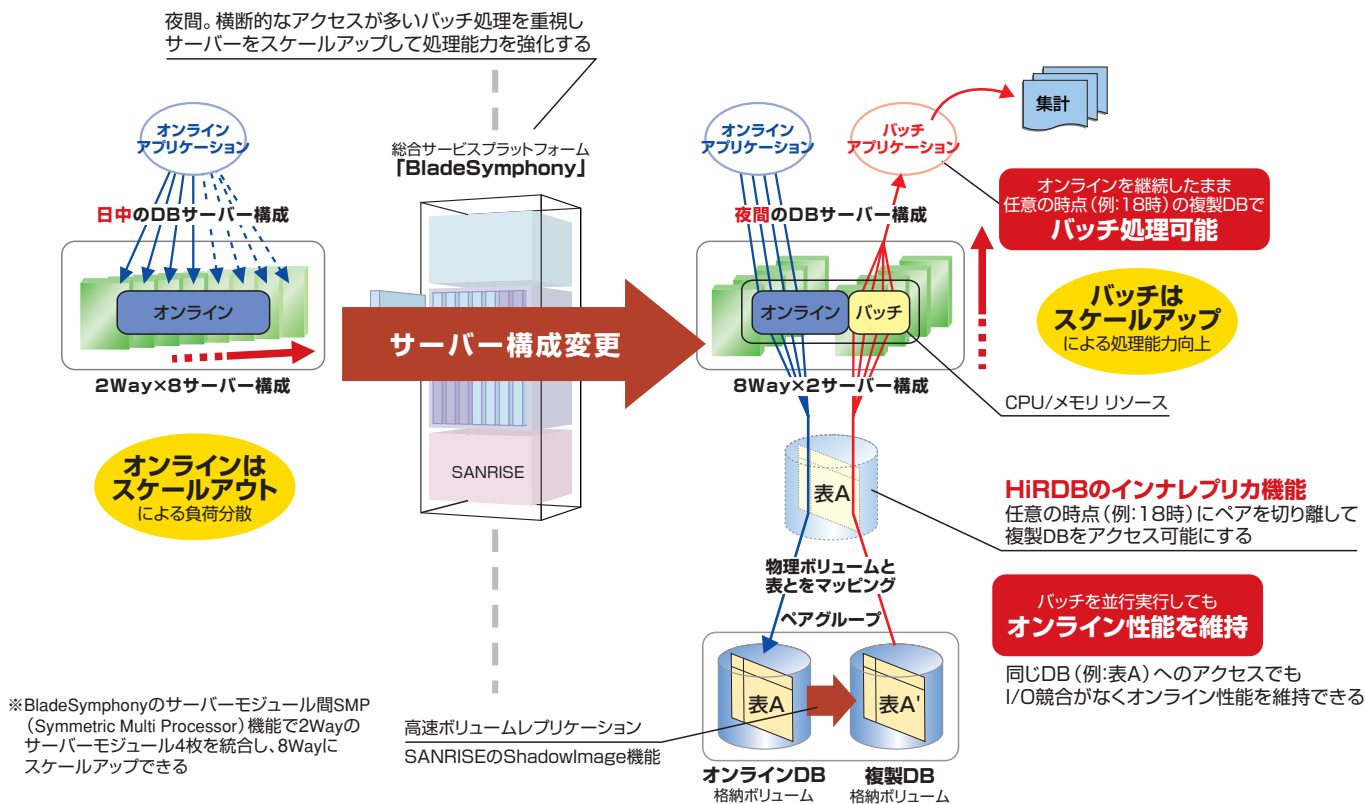


図4: オンラインとバッチでスムーズなデータ連携を実現

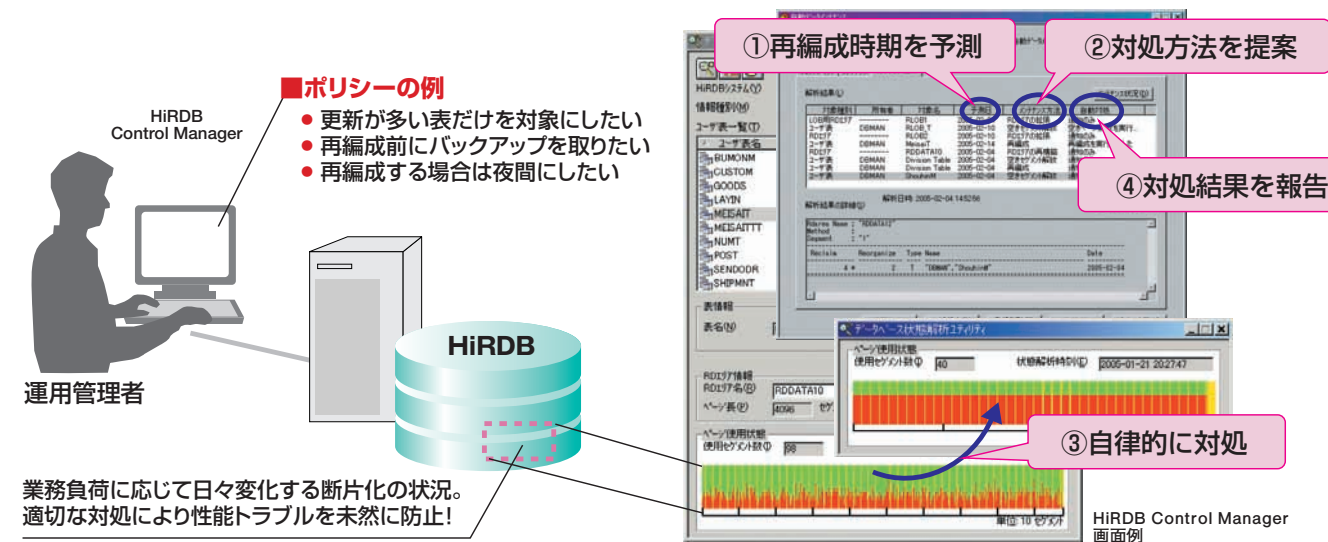


図5: データベースの再編成時期を予測、自動対応も可能

Image」(高速ボリュームレプリケーション)機能を利用する。事前にオンラインDB格納ボリュームとバッチDB格納ボリュームをペアグループに設定しておけば、サーバーのCPUを使わずにストレージの機能で、双方の更新状態を一致させることができる。これにより、テラバイト級の大容量データベースでも高速に複製DBの作成が可能だ。HiRDBではこの複製DBを利用してバッチ処理を実行する。

あらかじめHiRDBのインナレプリカ機能で、オンラインDB格納ボリュームとバッチDB(複製DB)格納ボリュームをデータベースの表とマッピングし、同じ表へのアクセスでも、バッチアプリケーションからはバッチDB格納ボリュームへアクセスするよう定義しておく。例えば18時までのオンライン取引履歴情報をバッチ集計したい場合は、18時にボリュームのペア状態を切り離せばよい。この切り離しはごくわずかな時間で完了するので、オンラインサービスを停止する必要はない。

また、ランダムアクセスの多いオンライン処

理では、サーバーをスケールアウトして負荷を分散したほうが性能向上効果が高い。一方、横断的なアクセスが多いバッチ処理では、サーバーをスケールアップして処理能力を強化したほうが効率的な場合が多い。日立的統合サービスプラットフォーム「BladeSymphony」では、「サーバーモジュール間SMP(Symmetric Multi Processor)機能」によって、2Wayのサーバーモジュールを最大4枚統合することによるスケールアップが可能になっている。この機能を利用すれば、日中はオンラインを重視してスケールアウトし、夜間はバッチを重視してスケールアップするなど、マシンパワーを最大限に利用できる。

「サーバーやストレージ統合によって、異なる業務サービス間でITリソース配分を共有する動きが加速していますが、その際難しいのはデータ連携方式の問題です。HiRDBは、この問題をDBのレイヤで解決し、オンラインスケールアウトやスケールアップによるITリソース共有のメリットを最大限に引き出し

す」(阿部氏)。つまりHiRDBは、Blade Symphonyによるシステム統合のメリットを最大限に引き出すミドルウェアと言えるだろう。

データベースの再編成も自動化 管理者の負担を大幅に軽減

HiRDBの自律運用への取り組みはオンラインスケールアウトだけに留まらない。データベースの中長期的な運用サイクルを自動化するためには、データエリアのメンテナンス運用の自動化が不可欠だ。常に適切なパフォーマンスを保つためには、日々の業務負荷に応じて変化するデータベース格納状態(断片化の状況)を考慮して、適切なリソースを適切な時期に、適切な方法で再編成する(断片化を解消する)ことが必要となる。

「主にデータ参照を行なう表と、データが頻繁に更新される表では再編成の時期が異なりますし、業務負荷によって再編成が必要になる時期も変わってきます。また、再編成

だけでは問題が解消されないようなケースでは、データベース格納領域自体を拡張する必要もあるでしょう」(石川氏)。

このようにパフォーマンスを保つためのメンテナンス運用は、定型的な運用として自動化するのが困難だった。HiRDB Version 7.2の自律化機能では、この問題も解決している。具体的に見てみよう。

HiRDBでは、運用管理GUI「HiRDB Control Manager」で、日々の運用操作を簡素化するためのさまざまな機能を提供してきた。このうちV7.2で追加した自動データメンテナンス機能では、以下の機能を提供している。

- データベース格納状態を監視する機能
- 断片化の状況に応じて、メンテナンスの時期を予測する機能
- 適切なメンテナンス方法を提案する機能
- メンテナンスを自動対応する機能

また、データエリアのメンテナンス方法としては、以下の3つを断片化の状況に応じて運用管理者に提案してくれる(図5)。

- 使われなくなった空きエリアを集め、断片化状態を軽減するガベージコレクション機能
- データベースの詰め替えを行ない、断片化状態を解消するデータベース再編成機能
- データエリアの拡張機能

図5の例では、MEISAITという表を2005年2月14日までに再編成する必要があることを示しており、さらにこれを自動的に再編成して対処したことを示している。このように、どのデータベースリソースが、いつまでに、どんな方法でメンテナンスする必要があるのかを運用管理者に提案し、さらにはその運用自体を自動化できるのだ。

HiRDB Version 7.2の機能拡張ポイント②

e-文書法を見据えたセキュリティ対策

強固なWORM機能でデータの改ざんを防止

セキュリティに関する機能拡張のポイントとしては、「WORM(Write Once Read

Many) 機能」の実装が挙げられる。

「これは、まるでCD-Rにデータベースを焼き込むように、一度書き込んだデータの変更や削除を一切行なえなくする機能です。この機能を使えば、一般ユーザーはもとより、たとえばデータベース管理者であっても一度書き込んだデータの書き換えはできなくなるので、データを不正改ざんから守れます」(阿部氏)。

民間事業者に対して保存などが義務づけられている文書に関しては、2005年4月に施行される「e-文書法」によって電磁的記録での保存が認められることになり、今後飛躍的にビジネス書類の電子保存が進むことになるだろう。しかし同時に、電子データの原本保証の仕組みがこれまで以上に要求される。つまりデータ改ざんなどの不正な行為がなされていないということを保証する手段を伴うことで、はじめて書類の電子化の道が開けるのだ。

ではWORM機能の具体的な使用例を見てみよう(図6)。例えば取引データを格納する表をWORM化する場合、表定義(create table)時のオプションに「INSERT ONLY」という指定を追加するだけで良い。保存期間を10年にしたい場合は「INSERT ONLY WHILE 10 YEARS BY JINSERTDATE」のように設定できる。以降、この表へはINSERTのみが許可される。たとえばデータベース管理者のパスワードが盗まれても、データ登録後10年間は、データをUPDATEしたり、DELETEすることができないのだ。

またHiRDBでは、データベースへのアクセス履歴を記録する監査証跡機能も提供している。接続したユーザー名称や接続時刻、発行したSQL情報などのすべてをありのままに記録でき、これらの記録からデータベースへの不正行為が検出できるようになっている。この監査証跡自身もWORM化できるのは言うまでもない。

こうした強固なセキュリティ機能を適用すれば、企業の取引データや機密情報、病院のカルテといった分野まで幅広く電子保存化の道が開けるだろう。

データ転送方式	メインサイト	バックアップサイト	メインサイトの性能	データ損失
同期	DB ログ	DB ログ	距離に依存して著しく低下	◎ データ損失ゼロ
非同期	DB ログ	DB ログ	低下しない	× データ損失が発生
ハイブリッド	DB ログ	DB ログ	低下が少ない	◎ データ損失ゼロ

図7：性能維持とデータ損失ゼロを両立した、ハイブリッドデータ転送方式

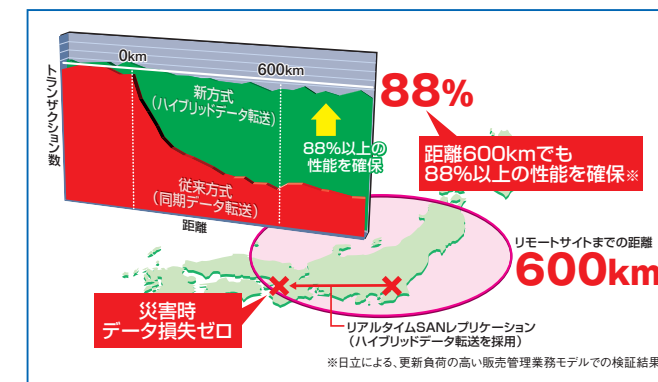


図8：600km離れたサイト間でデータ損失ゼロ、オンライン性能88%を実現

HiRDB Version 7.2の機能拡張ポイント③

万が一の広域災害からデータを守るディザスタリカバリ

HiRDB独自のハイブリッドデータ転送によりオンライン性能の維持とデータ損失ゼロを両立

3点目の機能拡張ポイントであるディザスタリカバリでは、ストレージ機能との連携が強化されている。

「サイトそのものが倒壊するような大規模災害に対しては、数100km離れた遠隔地にバックアップサイトを設置する必要があります。さらに災害復旧を迅速化するためには、データ損失ゼロでのバックアップが要件となりますが、バックアップサイトまでの距離が離れると、オンライン性能の低下が大きな問題となります」(阿部氏)。

つまりバックアップサイトにおいてデータ損失ゼロを保障するためには、メインサイトのデータベース書き込みと同期して、バックアップサイトのストレージへの書き込みを保障する必要があります。これを「データ同期転送方式でのバックアップ」というが、距離に応じて大きくなるI/Oレスポンスの遅延が、オンライン性能を大きく低下させてしまうのだ。実際、東京～大阪間で同期転送方式でのバックアップを行なうと、オンラ

イン性能は40%程度にまで低下するという。

こうした課題を解決するのが、HiRDBが提供するリアルタイムSANレプリケーション機能だ。日立ディスクアレイサブシステム「SAN RISE」シリーズの遠隔ボリュームレプリケーション機能(Hitachi TrueCopy)とHiRDBを連携させることで、DBボリュームの転送は非同期、更新ログボリュームの転送は同期で行なう「ハイブリッドデータ転送方式」を実現している(図7)。このハイブリッドデータ転送方式では、サーバーのCPUを使わずにSAN(Storage Area Network)の機能だけでデータを高速にバックアップできる。このため600km離れたサイト間でも、データ損失ゼロを実現できるだけでなく、オンライン性能を十分に確保(更新負荷の高い販売管理業務モデルで88%以上確保)できることを実証できた(図8)。

このハイブリッドデータ転送方式には、文部科学省が実施するリーディングプロジェクト「e-Society基盤ソフトウェアの総合開発」の中で、東大と日立が共同開発した「ストレージ・データベース融合技術」の成果が反映されている。つまり、ストレージとデータベース双方に自社製品を持つ日立ならではの強みが、十分に発揮された成果と言えるのだ。

さらに日立では、HiRDBおよびSANRISEをコアとしたソフトウェアおよびハードウェアに

加え、コンサルテーションから、設計/構築、運用、アウトソーシングまでのトータルソリューションメニュー「ビジネスコンティニュイティ・スイツ」を提供している。データの可用性と信頼性を確保し、ビジネスの継続性を高める上で、間違いなく大きな貢献を果たすはずだ。

「HiRDBは時代のニーズに合わせ、お客様のビジネスを支えるデータ管理基盤を提供してきました。今後も産学連携の研究成果などの最新技術に加え、ユーザー企業からのさまざまな要望に応える製品を提供していきます」と、阿部氏は締め括る。

DBM

取引履歴テーブルをWORM化する表定義例

```
CREATE TABLE JUTYU
(DNO CHAR(6), TCODE CHAR(5), SCODE CHAR(4), JSURYU INTEGER, JDATE DATE, JTIME TIME,
JINSERTDATE DATE NOT NULL WITH DEFAULT SYSTEM GENERATED,
JOPTION CHAR(30) UPDATE ONLY FROM NULL, JNOTE CHAR(60) UPDATE)
INSERT ONLY WHILE 10 YEARS BY JINSERTDATE
```

e-文書法を見据えたセキュリティ対策機能

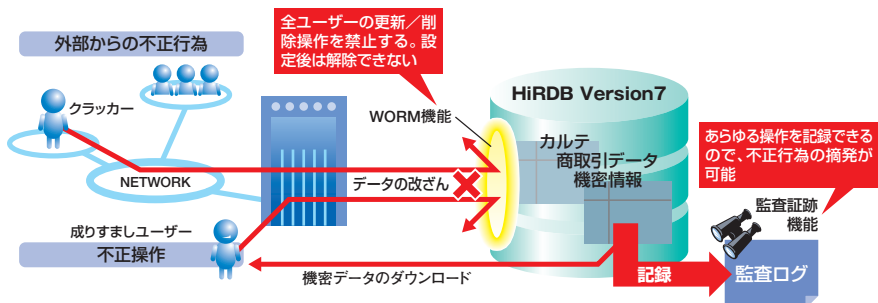


図6：取引履歴テーブルをWORM化する表定義例

- 稼働環境
 HP-UX (HP、日立)、Solaris (Sun)、AIX (IBM、日立)、HI-UX/WE2 (日立)、Red Hat Linux (レッドハット株式会社)、Windows XP、Windows NT、Windows 2000、Windows Server 2003
 ※64ビット版および、Itanium Processor Familyにも対応
- 価格
 1同時接続当たり：12万6000円(税込価格)
 1CPU当たり：189万円(税込価格)
 *最大同時接続数の下限は、UNIX版の場合4、WindowsおよびLinux版の場合2。
- 問い合わせ先
 HMCC
 (日立オープンミドルウェア問い合わせセンター)
 TEL：0120-55-0504
- 製品情報
 URL：http://www.hitachi.co.jp/hirdb/