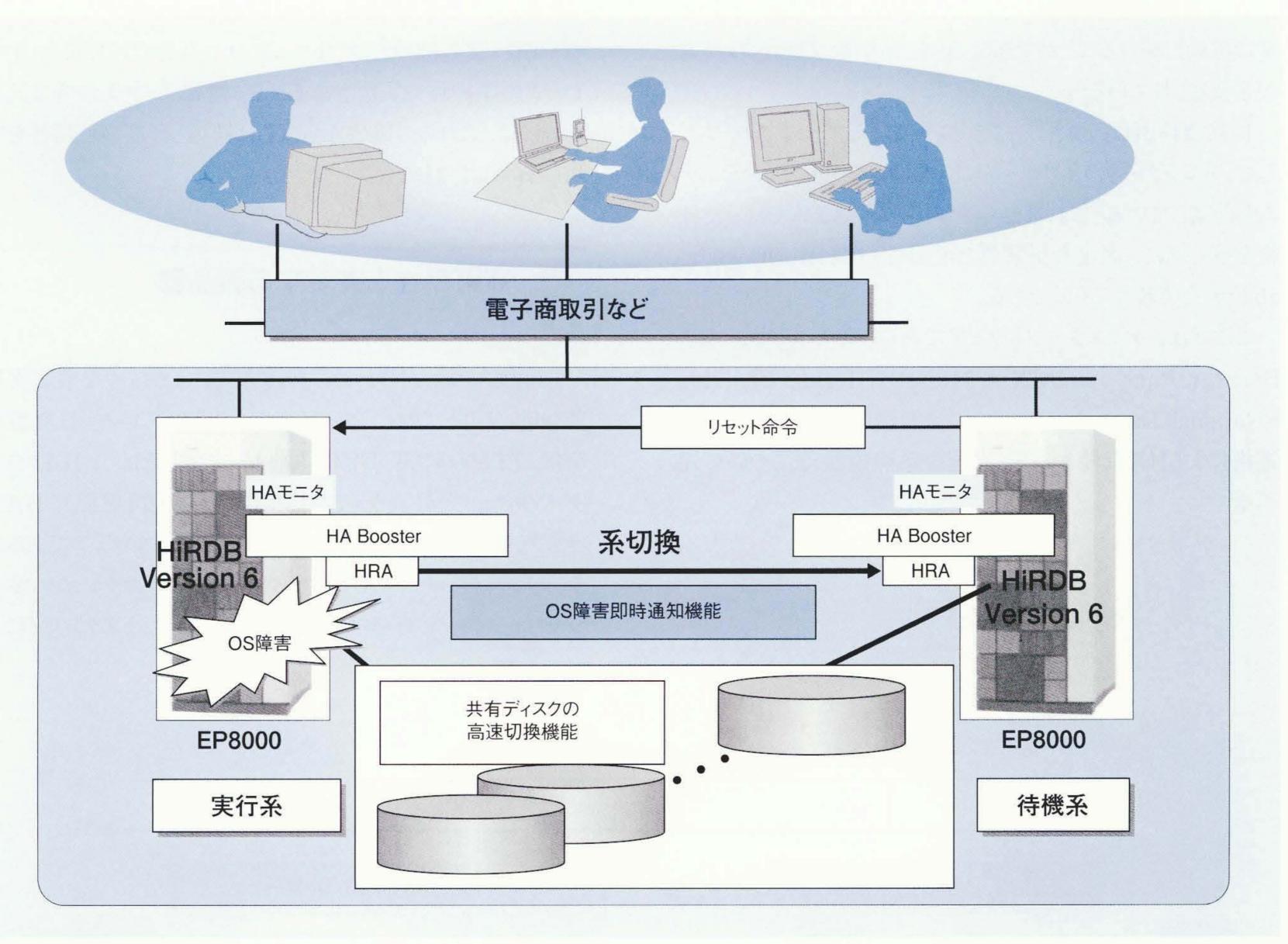
「EP8000シリーズ」による 高可用性システムソリューション

High-Availability System Solutions with "EP8000"

三瓶 英智 Hideaki Sampei

真矢 Yuzuru Maya

中村 市川 正也 Masaya Ichikawa Yutaka Nakamura



注:略語説明ほか

HiRDB (Highly Scalable Relational Database), HA (High Availability), HRA (Health Check and Reset Adapter), HA Booster (HA Booster Pack for AIX*), OS (Operating System) *AIXは、米国における米国International Business Machines Corp.の登録商標である。

"EP8000"による高可用性システムの概略構成

情報ライフラインを支えるEP8000システムでは、HAモニタ、"HA Booster Pack for AIX"、および"HiRDB"の連携により、ミッション クリティカル システムに必須の高可用性を実現する。

いつでも、どこでもコンピュータが利用でき、インター ネット上のさまざまなサービスが人々の生活を支えてい る社会では、情報は欠かすことのできないライフライン となる。サービスの中断や停止は、ビジネスチャンスの 喪失や社会的信用の失墜に直結することから、ミッショ ンクリティカル業務を任せることができる高信頼・高可 用なプラットフォームが切望されている。

日立製作所は、この情報ライフラインを担うために、

"EP8000"を中心とするシステムソリューションを提案 している。メインフレームを利用したシステム構築の経 験やノウハウを生かし、HA(High Availability)クラス タソフトウェアである「HAモニタ」を中心に、"HA Booster Pack for AIX", "HiRDB"などのミドルウェ アを連携させることにより、メインフレームに匹敵する 信頼性と可用性を実現し、ミッション クリティカル シス テムへの適用を可能としている。

はじめに

インターネットの普及に伴い、ビジネスの時間的・地理的制 限が取り払われ、いつでも情報システムにアクセスできること が、ますます大きな意味を持つようになっている。電子商取引 などのミッション クリティカル システムで処理される業務は企業 活動そのものであり、わずかな時間の停止であっても、社会 的信用やビジネスチャンスの喪失など、大きな損失につながる。 情報は産業・経済活動にとってのライフライン(命綱)であり、 常に稼動し続けることができる、堅牢(ろう)なプラットフォーム が必要とされている。

日立製作所は、情報ライフラインを支えるプラットフォームと して、「エンタープライズサーバ EP8000」を中心としたシステ ムを提案している。これは、メインフレームシステムで培った技 術をベースに、高度な信頼性と可用性(HA:High Availability)を実現するものである。

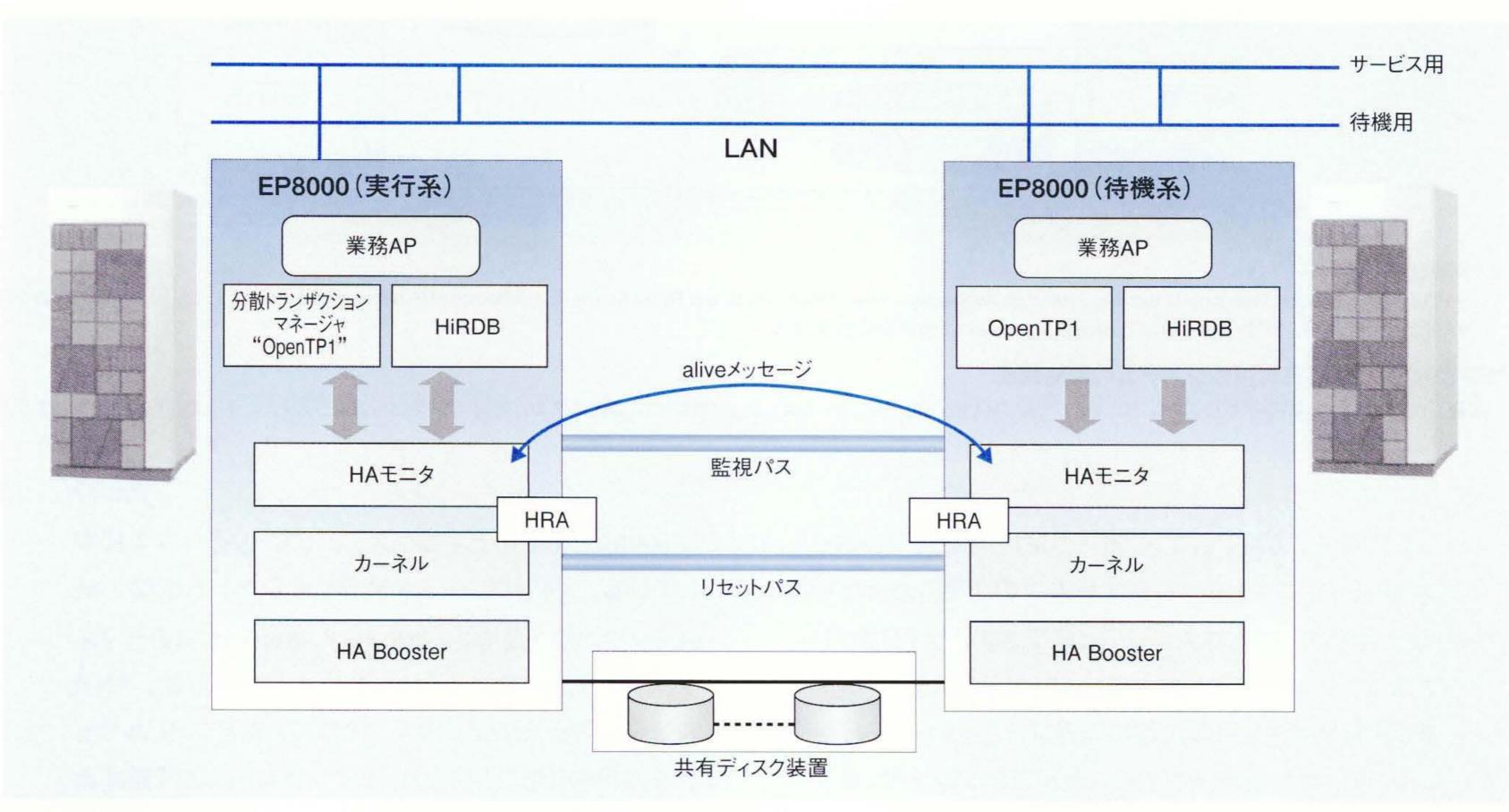
ここでは、クラスタソフトウェアであるHAモニタと、"HA Booster Pack for AIX"や"HiRDB(Highly Scalable Relational Database)"などの日立製作所のミドルウェアとの 連携によるHA機能として, 高速な系切換機能について述 べる2)~4)。

情報ライフラインに求められるもの

情報ライフラインを支えるシステムには、(1) サービスを停止 させないことのほか、(2) サービスが停止しても直ちに回復す ることが要求される。(1)はハードウェアの冗長化によって障 害発生を防止することや、ハードウェアの拡張、交換などに必 要な「計画停止」時間を短縮することであり、(2)は、発生し てしまった障害に対し、その影響を局所化して即座にサービ スを再開させることである。サービス停止につながる障害には、 ハードウェアの故障やソフトウェアの不具合だけでなく、オペ レータの誤操作や災害も含まれる。情報ライフラインを支える ためには、これらの障害から直ちに回復し、高い可用性を実 現しなければならない。

高可用性を実現する製品群

EP8000では、ホットスタンバイ方式による高速な系切換機 能を用いることにより、高い可用性を実現している。日立製作 所は、EP8000の高可用化ソフトウェア製品として、HAクラス タソフトウェアとミドルウェアを提供している(図1参照)。HAク ラスタソフトウェアである「HAモニタ」と、ミドルウェアであるス ケーラブルデータベース"HiRDB"や分散トランザクションマネー ジャ"OpenTP1"を連携させることにより、ホットスタンバイによ



注:略語説明

AP (Application Program), LAN (Local Area Network)

図 1 ホットスタンバイ構成によるミッション クリティカル システムの高信頼・高可用化

実行系と待機系、および共有ディスクで構成する。HAクラスタソフトウェアとして、HAモニタと"HA Booster"を搭載し、ミッション クリティカル システムを実現する。

る高速な系切換が可能となる。

HAクラスタソフトウェアであるHAモニタでは、EP8000の独自機能である"HA Booster Pack for AIX"と組み合わせることで、系切換処理をさらに高速化することができる。これら各製品の機能について以下に述べる。

3.1 ミドルウェアの高可用性機能

HiRDBとOpenTP1では、HAモニタとの連携で以下の機能が有効となり、正確な障害検知と系切換時間の短縮が可能となる。

(1) プロセスの事前起動機能

待機系側のHiRDBでは、事前にプロセス開始処理を完了させて、待機させておく。これにより、系切換処理でのプロセス起動時間を短縮することができる。

(2) プロセス障害検出機能

HiRDB, OpenTP1, およびHAモニタでは,プロセス間で定期的に稼動情報を伝達している。HAモニタでは稼動情報を監視し,規定時間内に情報の伝達がない場合には障害と判定して,待機系への切換を行う。通常のプロセス監視とは異なり,プロセスの挙動を監視することができるので,プロセスの異常終了のほか,ハングアップ状態を検出することもできる。

3.2 HATTS

(1) 両系の稼動状況の監視機能

実行系と待機系では、互いの障害を検出するために、「aliveメッセージ」を交換する。規定時間が経過しても相手系からaliveメッセージを受信しない場合、相手系で障害が発生したと認識する。この障害検出時間はaliveメッセージの間隔とリトライ回数に依存するが、少なくとも数秒を要する。

(2) 実行系の障害検出時でのリセット機能

待機系のHAモニタでは、実行系の障害を検出すると、障害の発生した実行系を確実に閉そくさせるために、実行系にリセット信号を送信し、リセットする。

3.3 HA Booster Pack for AIX

HA Booster Pack for AIX(以下, HA Boosterと略す。) では, 系切換時間をさらに短縮させるために, 共有ディスクの高速切換機能とOS(Operating System)障害即時通知機能を提供する。

3.3.1 共有ディスクの高速切換機能

共有ディスクの高速切換は、制御グループ方式と疑似オフライン方式によって提供される。

(1) 制御グループ方式

共有ディスクへのアクセス制御を行う単位として、「制御グループ」という概念を導入する。この制御グループは、一つまたは複数のボリュームグループで構成する。共有ディスクへの

アクセスが許可されるか禁止されるかを,制御グループごとに 設定することができる。これにより,ボリュームグループの一括 切換が可能となり,切換時間はボリュームグループ数に比例 せず,一定時間となる。

(2) 疑似オフライン方式

実行系と待機系では、システム起動時に、共に共有ディスクをオンライン状態に設定する。共有ディスクへのアクセスは HA Boosterによって制御され、実行系のHA Boosterでは共有ディスクへのアクセスを許可するが、待機系では、デバイスドライバを呼び出さないことによって疑似的なオフライン状態とし、アクセスを禁止する。

このように、共有ディスクを両系からオンラインのまま排他することにより、切換時間を0.1秒以下に短縮する。

3.3.2 OS障害即時通知機能

実行系のHA Boosterでは、OS障害を検出すると、系切換機構(HRA:Health Check and Reset Adapter)を経由して障害発生を直ちに待機系に通知する。待機系ではaliveメッセージの途絶を待たずに切換処理を開始し、障害の検出時間を1秒以下に短縮する。HRAとは、実行系と待機系のHAクラスタソフトウェア間で、aliveメッセージ、リセット信号、およびOS障害情報を高速・確実に通信するためのハードウェアのHA製品である。

4

ホットスタンバイシステムの 処理の流れと効果

ホットスタンバイシステムとしてHAモニタとHA Booster, HiRDBを連携した場合の切換処理と、その効果について次に述べる。

待機系のHiRDBは、系切換前にプロセスを起動し、プロセス開始処理を完了させた状態で待機させる。また、実行系と待機系では、共に共有ディスクをオンライン状態に設定するが、サービス中は、疑似オフライン方式により、実行系のHABoosterは共有ディスクへのアクセスを許可し、待機系はアクセス禁止とする。その後、実行系で障害が発生すると、HABoosterは、OS障害即時通知機能により、瞬時に障害を通知し、サービスを引き継ぐ待機系のHABoosterが共有ディスクをアクセス許可に設定する。待機系のHiRDBプロセスは待機状態から活動状態となり、処理を再開する(図2参照)。

これにより、HiRDBの再開始処理を含めた系切換時間³⁰は、 従来数十秒かかっていたが、HAモニタとHA Boosterの適 用により、十数秒程度に短縮することができ、情報ライフライン のプラットフォームとして十分な高可用性を提供することが可 能になった(図3参照)。

このソリューションは、すでに、金融系振替システムや情報系システム、証券系システムなどに可用性の高いプラットフォームを提供するために採用されている。これらのミッション クリ

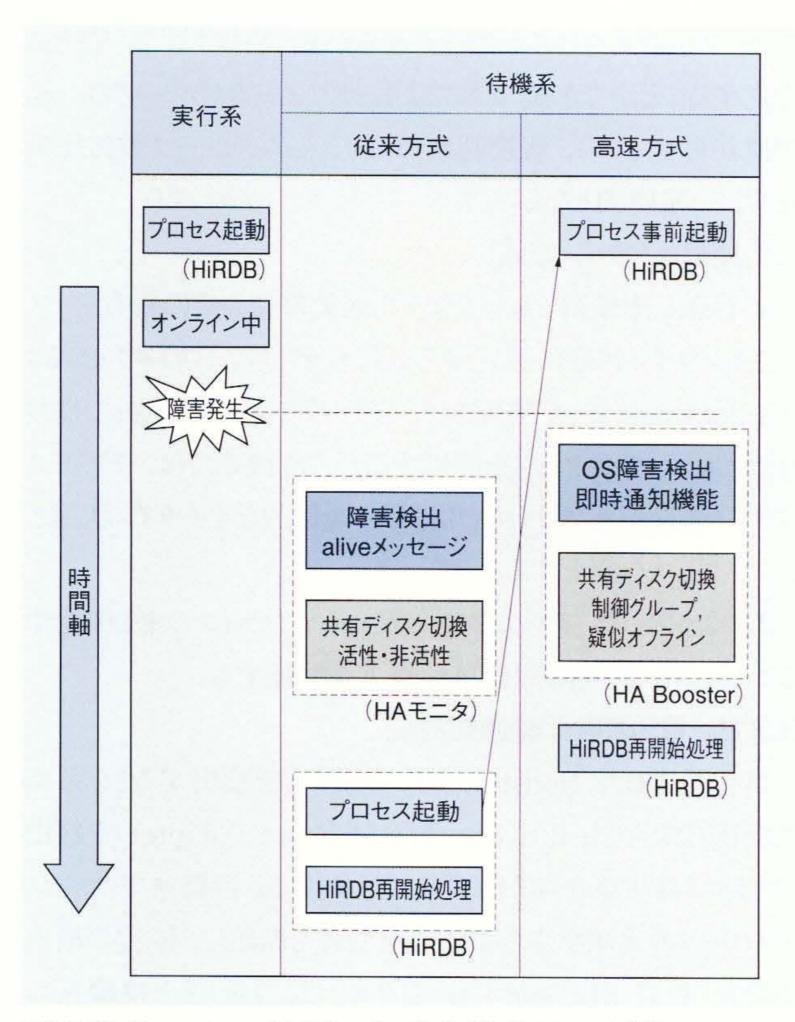


図2 HA Booster, HAモニタ, およびHiRDBの連携

HAモニタにより、HiRDBプロセスを事前に起動させる。HA Boosterで障害を瞬時 に検出し、また、共有ディスクの即時切換により、系切換を短時間で行う。

ティカル システムにおけるサービス中断を極小化することで, 情報ライフラインの高信頼化に貢献する。

おわりに

ここでは、情報ライフラインに求められる高信頼・高可用性 を実現する日立製作所の「エンタープライズサーバ EP8000 のHA製品と「日立ミドルウェア」の連携、および効果について 述べた。

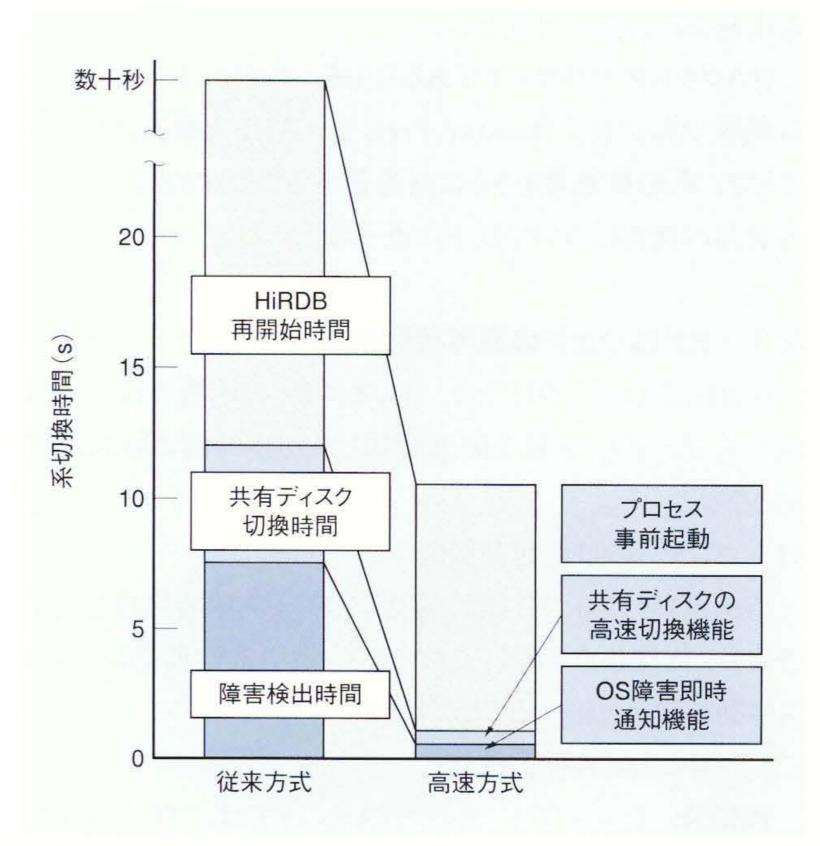


図3 系切換時間

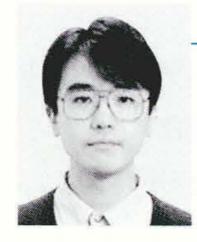
従来の系切換時間は数十秒であったが、HAモニタ、HA Booster、およびHiRDB の連携により、十数秒程度に短縮できる。

日立製作所は、今後も、高信頼・高可用プラットフォームの 実現を目指し、EP8000、HAモニタ、HA BoosterなどのHA 製品のいっそうの強化と、高度なシステムソリューションの提 案により、顧客を支援していく考えである。

参考文献など

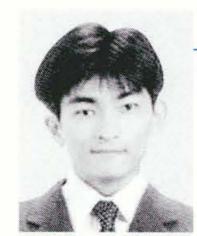
- 1) 吉岡,外:進化するEビジネスと企業情報システム,日立評論,84,9, $559 \sim 562 (2002.9)$
- 2) 市川,外:高可用化ミドルソフトHA Boosterの開発,第65回情報処 理学会全国大会, 3F-6(2003.3)
- 3) 原, 外:ネットビジネスを支えるミッション クリティカル システムに対応した データベース—HiRDB Version 6—, 日立評論, 84, 9, 571~574 (2002.9)
- 4) http://www.hitachi.co.jp/Prod/comp/EP8000/index.html

執筆者紹介



三瓶英智

1993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 基本ソフトウェア本部 第10S設計部 所属 現在, HAクラスタソフトウェアの開発に従事 E-mail: sanpei_h@itg. hitachi. co. jp



市川正也

1999年日立製作所入社,システム開発研究所 第3部 所属 現在, HAクラスタソフトウェアの研究開発に従事 情報処理学会会員

E-mail: m-ichika@ sdl. hitachi. co. jp



真矢 讓

1980年日立製作所入社,システム開発研究所 第3部 所属 現在, HAクラスタソフトウェアの研究開発に従事 工学博士

情報処理学会会員, 電子情報通信学会会員 E-mail: maya @ sdl. hitachi. co. jp



中村 豊

1993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 第1ネットワークソ フト設計部 所属

現在, HAクラスタソフトウェアの開発に従事 E-mail: nakamuyu@itg. hitachi. co. jp