

情報システムの大規模化, 多様化に対応する MパラレルDB/DC

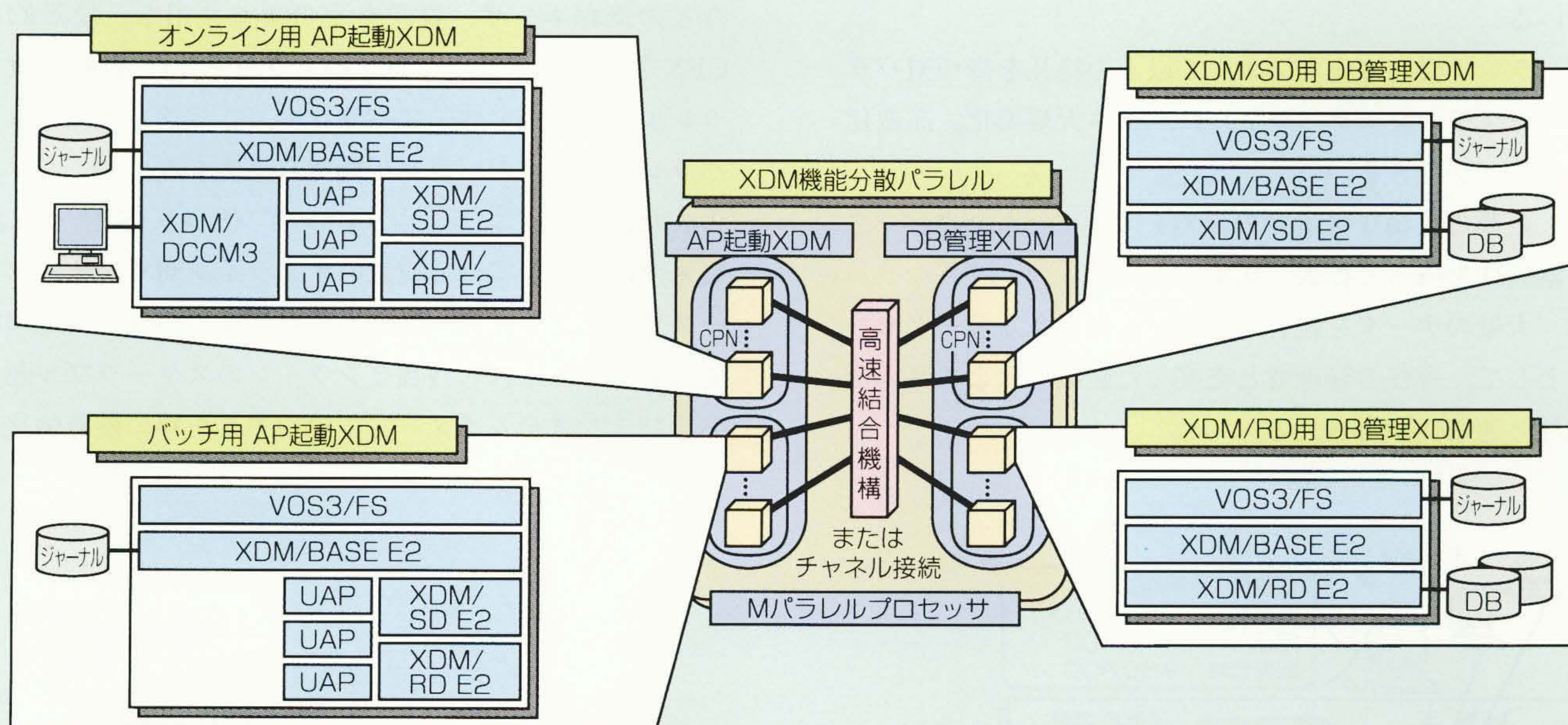
M Parallel DB/DC for Large Scale and Flexible Information Systems

山根利夫* Toshio Yamane

加藤公章* Kimiaki Katō

松下一昌* Kazumasa Matsushita

渡辺澄夫* Sumio Watanabe



注：略語説明

AP (Application Program), VOS3/FS (Virtual Storage Operating System 3/FOREFRONT System Product)
XDM (Extensible Data Manager), XDM/BASE E2 (XDM/Base Extended Version 2), CPN (Central Processing Node)
XDM/SD E2 (XDM/Structured Database E2), XDM/RD E2 (XDM/Relational Database E2)
XDM/DCCM3 (XDM/Data Communication and Control Manager 3), UAP (User Application Program)

XDM E2機能分散パラレルの概要

複数のCPN (Central Processing Node) を結合してデータベース処理などの機能を分担処理することにより、アプリケーションプログラムを変更することなしにシステム全体のスループットを大幅に向上させることができる。

企業の基幹システムとしてのデータセンタには、大量の情報が蓄えられており、これを管理しているのが「DB/DC (Database/Data Communication) システム」である。Mパラレルが提供するDB/DCシステムでは複数のCPN (Central Processing Node) を連携させることにより、増大するデータに対してスケラブル (段階的) にシステムを拡張することができる。

近年、データセンタを取り巻くネットワーク上のパソコン、ワークステーションからデータを取り出し、意思決定のための情報として活用するなど、企業データの利用方法も多様化してきている。MパラレルDB/DCシステムでは、基幹システムのデータを

集めてデータセンタ内のXDM/RD E2に統合、管理し、データベースパラレル機能を利用して、非定型な検索を高速に処理できる。このXDM/RD E2を国際標準のSQL (Structured Query Language) でネットワークに接続して自由に利用することにより、蓄積されたデータを意思決定などのための情報として最大限に活用できる。

重要な共通データはセンタでしっかり管理するという“FOREFRONT with Cyberspace”のコンセプトに基づいて、MパラレルDB/DCシステムは基幹システムの情報を戦略的に活用するデータセンタの実現を目指す。

*日立製作所 ソフトウェア開発本部

1 はじめに

従来、業務追加などでCPUの処理能力が限界に達してくと、より高い処理能力のCPUにアップグレードすることで対応していた。しかし、1台のCPU能力には限りがあり、スケーラブルなシステム拡張などを目的としたパラレル処理システム環境構築へのニーズが高まってきた。

パラレル機能の充実によって以下の特長を持つMパラレルシステムをスケーラブルに、より大規模化、高速化することができる。

- (1) 企業の重要な財産であるメインフレーム上のデータを継続性を持って拡張、活用できる。
- (2) 大量のデータを統合的に集中管理できるデータセンタとして、専任の管理者と充実した製品群による高い信頼性と運用性が得られる。

- (3) 過去を含めたすべてのデータを利用可能とするデータウェアハウスとして大量のデータを蓄積し配布するなど、パソコンWSからの統一的情報利用が可能となる。

これらのニーズにこたえるため統合型データマネジメントシステム“XDM E2”^{※1)}では、従来提供していたコンカレントアクセス^{※2)}、マルチサブシステム、分散DB、分散AP等の機能を拡張して、UAP、データベースなどの既存ユーザー資産をそのまま使用し、段階的にCPNを追加して処理能力を向上させる、(1)機能分散パラレル、(2)データベースパラレル、(3)トランザクションパラレルの三つのパラレル機能を提供する。

また、金融オンライン向けのTMS-4V/SP^{※3)}では、従来サポートしている複数のオンライン間でのデータベース共用による、同一勘定系業務処理のパラレル処理機能に加えて、それら複数オンラインのスケーラブル運用および1システムイメージで運用するための機能強化を

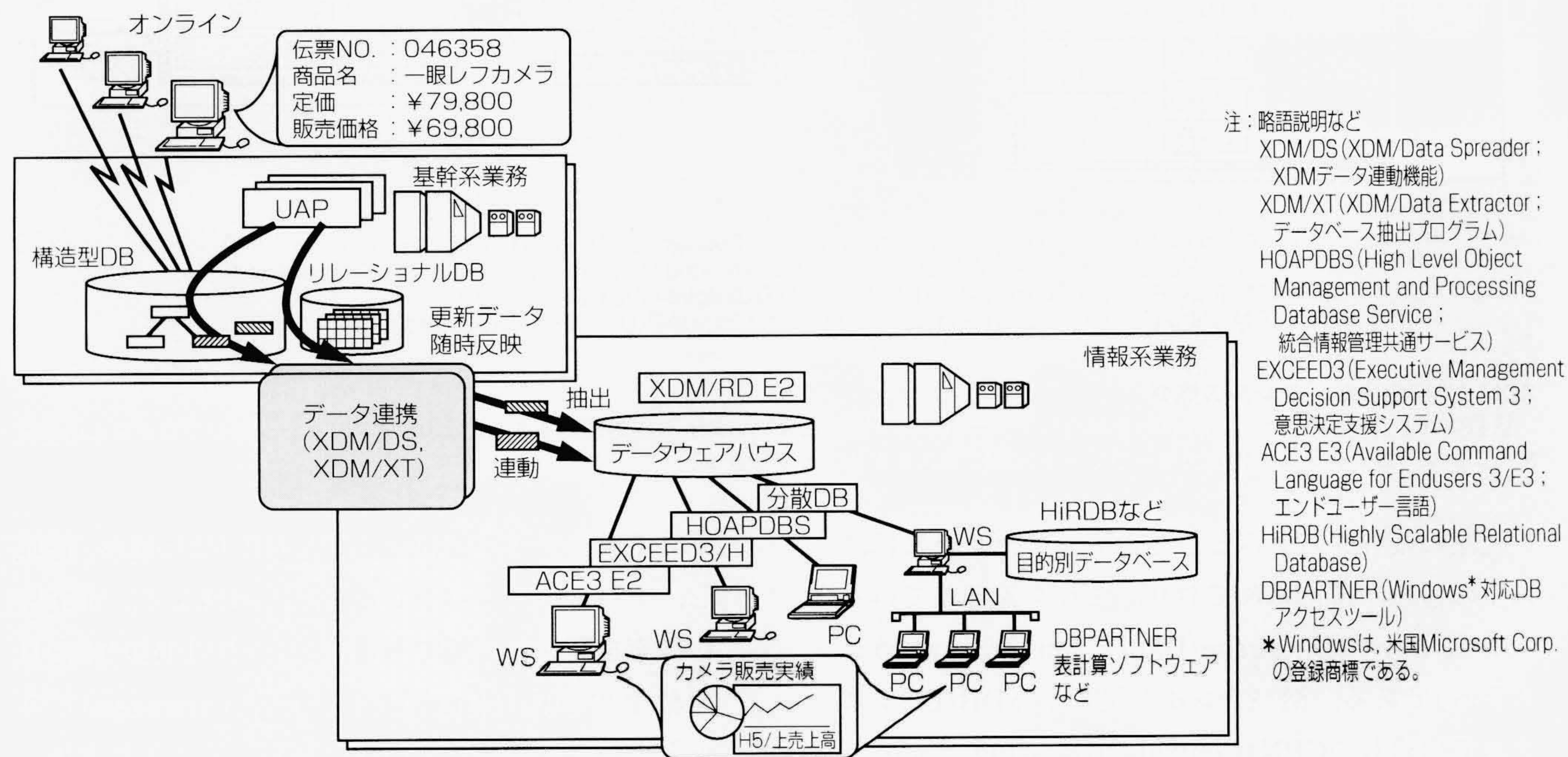


図1 XDM/RD E2によるデータウェアハウスの構築

基幹系システムのデータをリレーショナルデータベースと連携させてデータウェアハウスを構築する。データセンタ内で情報を統合、集中管理することによって信頼性の高いシステムが構築できる。

※1) XDM E2は、構造型データベースのXDM/SD E2、リレーショナルデータベースのXDM/RD E2、DC機能のXDM/DCCM3など各種サブシステムを配したマルチサブシステム統合型DB/DCシステムである。

※2) コンカレントアクセスは、システム共通領域にXDMの一部を常駐させ、UAP空間からのデータベース操作要

求時にXDM空間との空間連絡を不要とする方式である。各空間が独立して稼動するために負荷が少なく、特にマルチプロセッサ環境に最適である。

※3) TMS-4V/SPは、金融機関の勘定系オンラインシステム向けDB/DCシステムであり、階層型データベース機能を持っている。

行う。

ここでは、運用が容易なデータセンタで、ますます増大するデータを集中して管理する「MパラレルDB/DCシステム」について述べる。

2 XDM E2パラレル機能

2.1 機能分散パラレル

既存の大規模オンライン業務、データベースアクセスを大量に行うバッチ業務を、複数のCPUで処理するために分割することは、一般的に困難である。そのため、XDM E2は次の特徴を持つ機能分散パラレルを提供する。

2.1.1 機能の分割・分散によってスループットを向上 —従来のユーザー資産は継続して使用可能

機能分散パラレルではシステム環境に応じて機能を分割し、複数のCPNに分散させて配置する(37ページの図参照)。トランザクション処理のうちデータベース処理の比率が高い場合には、データベース処理を実行するXDM(DB管理XDM)をアプリケーションプログラムを実行するXDM(AP起動XDM)とは別のCPNに配置できる。また、オンライン処理にバッチ処理の影響を与えたくない場合には、バッチ処理実行のXDMをオンライン処理実行のXDMとは別のCPNに配置できる。

アプリケーションプログラムが発行するデータベースアクセス要求は、AP起動XDMがデータベース定義に従って目的のDB管理XDMに転送する。DB管理XDMはデータベース処理を実行し、検索結果などの処理結果をAP起動XDMに返す。CPU負荷の高いデータベース処理をアプリケーションプログラムとは別のCPNで実行することにより、アプリケーションプログラムの変更なしにシステム全体のスループットを向上させることができる(例：各XDMのCPU負荷が均等になるようにデータベースを配置した場合、4XDMで約3倍のスループット)。

2.1.2 DB管理XDM間の排他処理なしにアプリケーション間でデータベースを共用

CPN間でデータベースを共用して排他処理を行う方式(Shared Everything)は、排他のオーバーヘッドが大きい。そのため機能分散パラレルでは、DB管理XDMごとにデータベースを配置し、DB管理XDM間での排他処理を行わない方式(Shared Nothing)を採用している。

また、複数のAP起動XDMで実行されるアプリケーションプログラムからは、アクセス要求の転送によって必要なデータベースを利用することができるので、データ

ベース共用も実現できる。

2.1.3 性能要件に応じてXDM間通信装置を選択

各ノード間の通信を高速に行うことが要求されるシステムには高速結合機構を使用する。なお、コストパフォーマンスの観点からチャンネル接続を選択することもできる。

2.1.4 1システムイメージの運用

オペレータが個々のXDMを意識しなくてもよいように、運用コマンドは特定のXDMが受け付けて各XDMへ振り分ける。システムの開始、終了や状態表示などは1システムのイメージで運用できる。データベース回復などを各XDMで個別に実行できるように、回復処理に使用するジャーナルは各XDMで独立して取得する。なお、各XDMで取得したジャーナルをマージすることも可能である。

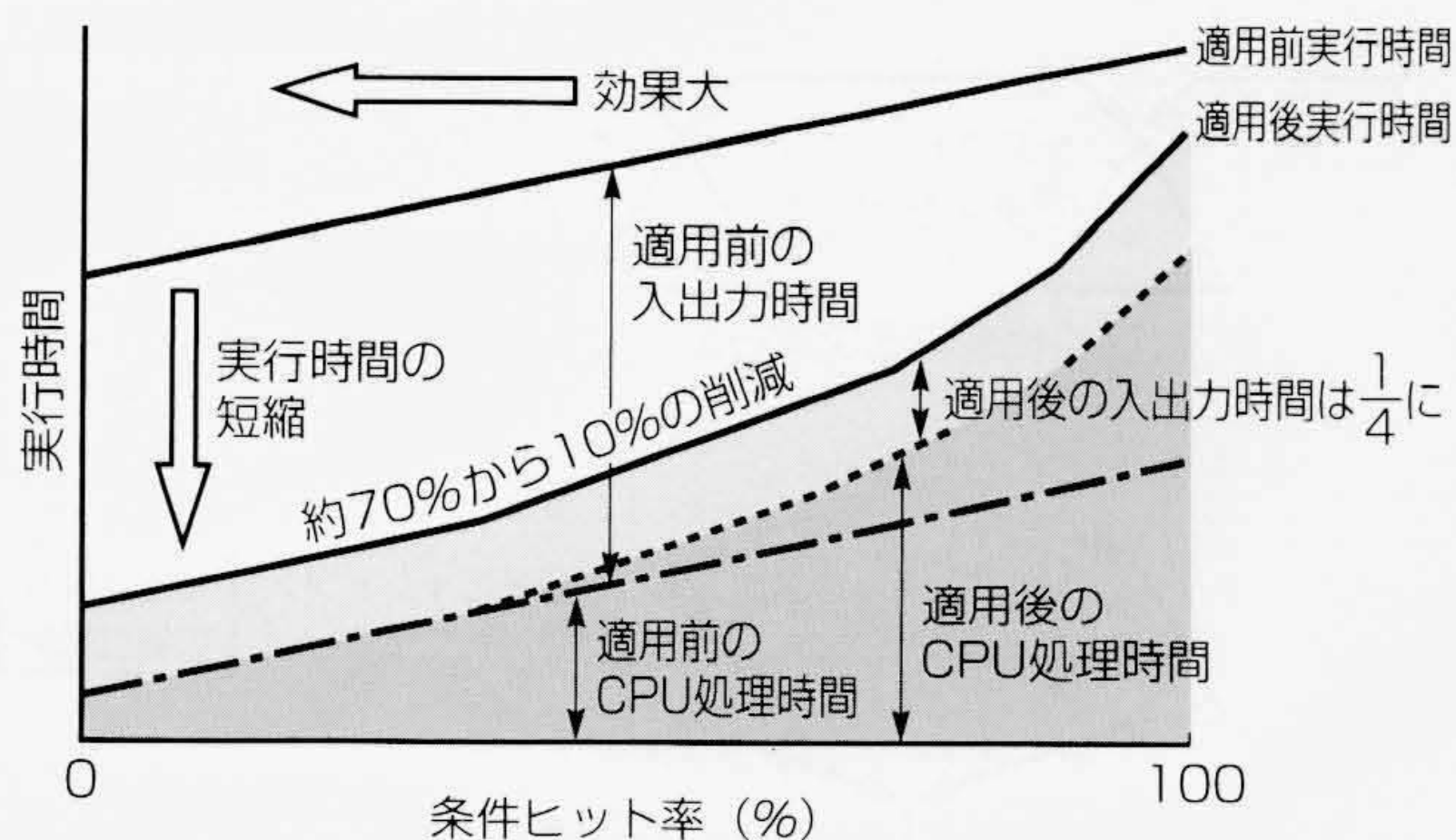
2.1.5 システム環境の変化に柔軟に対応

システムの大規模化、利用範囲の拡大などによってシステム環境が変化した場合には、アプリケーションの変更なしに、柔軟にXDMの分割・追加・統合ができ、スケラブル(段階的)にシステム構築ができる。

2.2 データベースパラレル

2.2.1 概要

Mパラレルでは、スループット向上を目的とする機能分散パラレルに加え、大幅なレスポンスタイム向上を可能とするデータベースパラレル機能をサポートしている。データベースパラレルにはXDM/RD E2がサポートする入出力パラレル、SQLパラレル、XDM/RD E2、XDM/SD E2の両方で使用可能なバッチパラレルが



注：前提条件 (1) 表を4ボリュームに分割
(2) 複数の入出力チャンネル
(3) マルチプロセッサ(CPN数：2)

図2 入出力パラレルの実行時間短縮例

入出力時間が並列数分短縮されるため、条件ヒット率が低い検索で特に効果大きい。

ある。

近年、データベースに蓄えられるデータ量は増大の一途であり、特に意思決定のためには、大量データに対する非定型検索を高速に処理する必要がある。基幹システムをXDM/RD E2で構築して、そのままパラレル検索することができるうえに、分散したデータをデータ連携機能を使用してXDM/RD E2に集め、データウェアハウスとしても利用することができる(図1参照)。

Mパラレルシステムでデータを集中管理することにより、基幹システムとの統一的運用、利用が可能であり、保守性に優れ、信頼性の高い情報システムを構築することができる。

蓄積されたデータは、パソコン上の表計算ソフトウェア、ODBC(Open Database Connectivity)対応ソフトウェアから自由にアクセス可能となる。また、長期間蓄積された時系列データを扱うようなシステムでは、XDM/RD E2をHiRDBと連携させてワークステーションやパソコンでデータを分析することもできる。

2.2.2 入出力パラレル

表を複数のボリュームに分割格納すれば、アプリケーションプログラムを変更することなく、入出力処理を並列実行できる。入出力パラレルは、大量のデータから一部のデータを取り出すような(検索条件へのヒット率が小さく、CPU負荷が少ない)非定型的な業務に特に大きな効果がある(図2参照)。

例えば、1,000万件のデータから、1,000件を取り出す

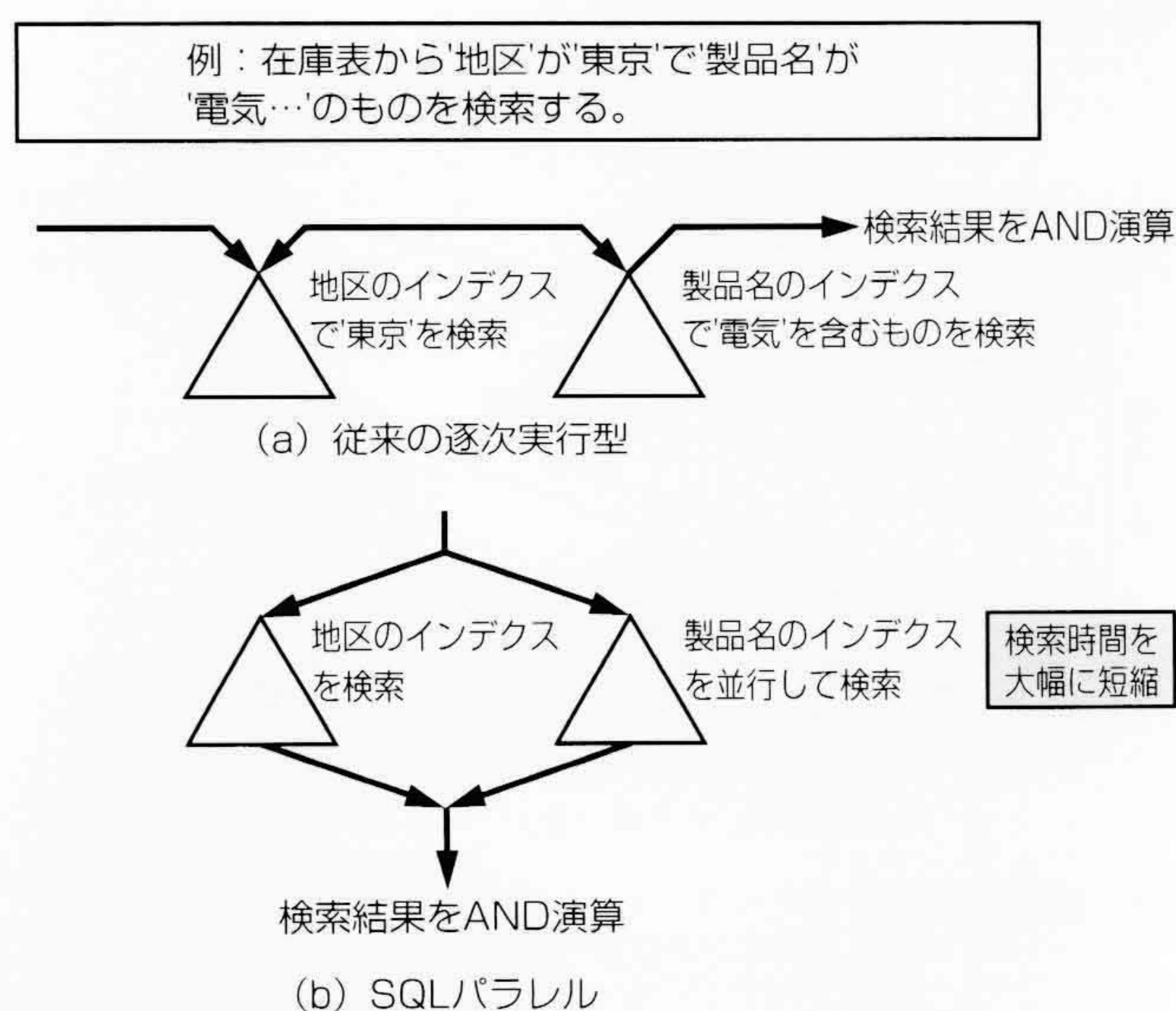


図3 SQLパラレル(複数インデクス利用検索)

パラレル実行では、複数のインデクスを同時に検索することによって大幅なレスポンスタイム向上が可能となる。

業務で表を12ボリュームに分割した場合、経過時間は約 $\frac{1}{10}$ となる。

2.2.3 SQLパラレル

XDM/RD E2では、表結合など負荷の高いSQLを、それぞれの表を操作するサブSQLに分割して実行するSQLパラレル機能を提供する。個々の表を分割して並行に処理する入出力パラレルを併用することにより、さらに高速化できる。

また、複数のインデクスを利用した検索では、それぞれのインデクスに対する処理を並行して実行する。複数インデクスの処理例を図3に示す。

さらに、機能分散パラレル環境では、分割されたサブSQLを複数のCPNに振り分けて負荷を分散することができる。

2.2.4 バッチ ユーティリティ パラレル

XDM/SD E2のスキーマ、XDM/RD E2の表が複数のエリアに分割格納されているとき、DBの再編成などのジョブをエリア単位に独立して並列実行することができる。また、PREST(Parallel Reference and Synchronous Transfer：並列同期転送)機能を併用することもできるので、大量データのローディングなどのバッチ業務処理時間が大幅に短縮できる。

2.3 トランザクションパラレル

オンラインシステムの大規模化、多様化に伴って従来のXDMシステムでは、複数のXDMシステムに分割することで対応してきた。しかしこの対応では、各XDMシステムごとに接続端末を固定したり、端末オペレータが接続先のXDMシステムを選択する必要があった。そのため、XDM/DCCM3でのトランザクションパラレル機能の提供によってこれらの問題を解決した。

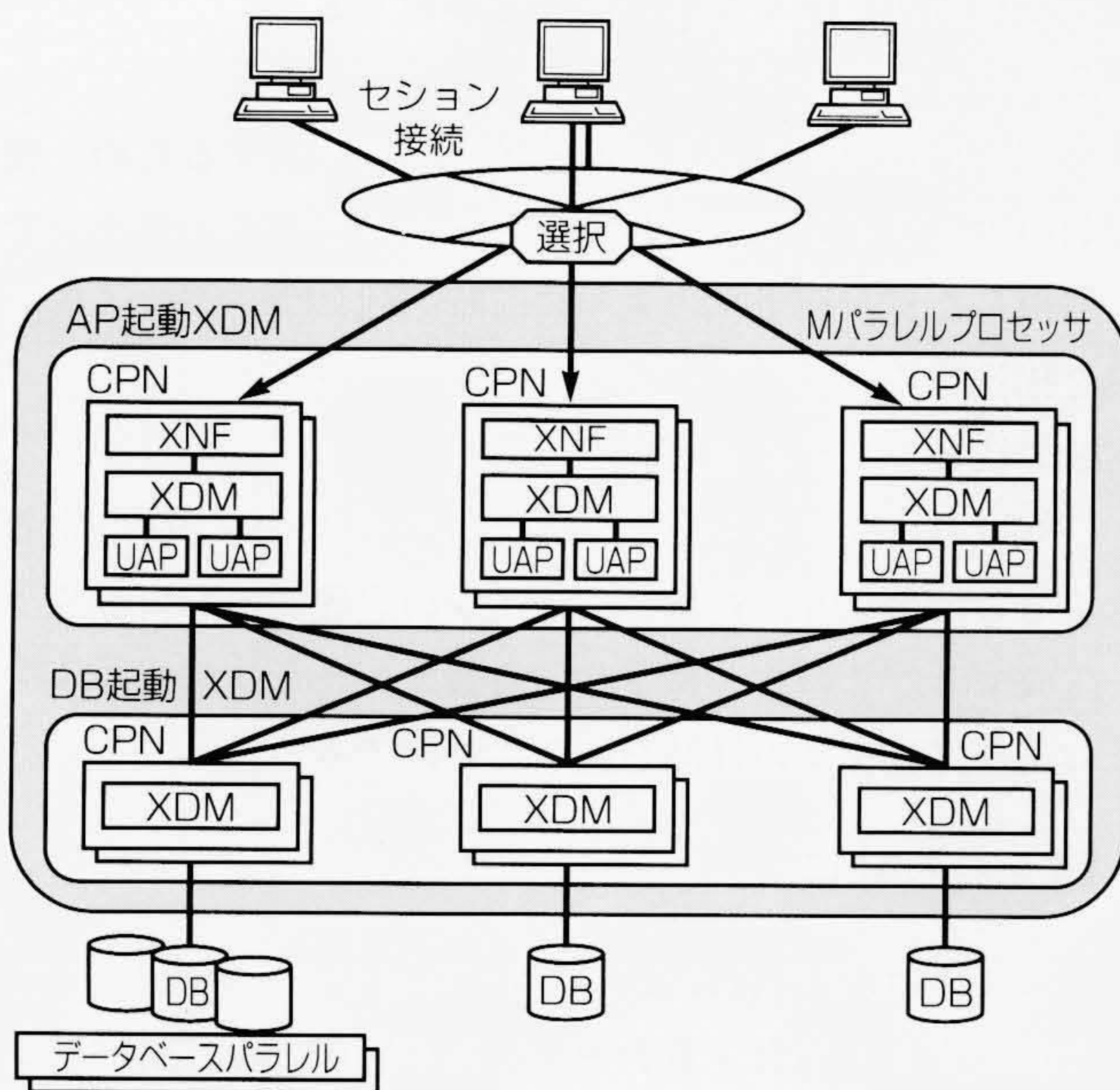
トランザクションパラレルとは、トランザクション入力元である端末とのセッションを、AP起動XDMを構成する各XDMシステムに分散させる機能である。

トランザクションパラレルの使用によって各XDMシステムの負荷が平均化されるため、システム全体のスループットを向上させることができる(図4参照)。

トランザクションパラレルでの利点は以下の2点である。

(1) 端末ユーザーからは複数を意識させないシステム構成

端末ユーザーからは一つのXDMシステムとして見える。そのため端末ユーザーは、複数のXDMシステムで構成していることを意識することなく、従来と同様に単一



注：略語説明

XNF [Extended HNA (Hitachi Network Architecture)
Based Communication Networking Facility]

図4 XDM/DCCM3のトランザクションパラレル

端末とAP起動XDMの間のセッションを動的に分散させてシステム全体のスループットを向上させる。

のシステムとして扱うことができる。

(2) システムでの最適な動的負荷バランスによるユーザー負担の軽減

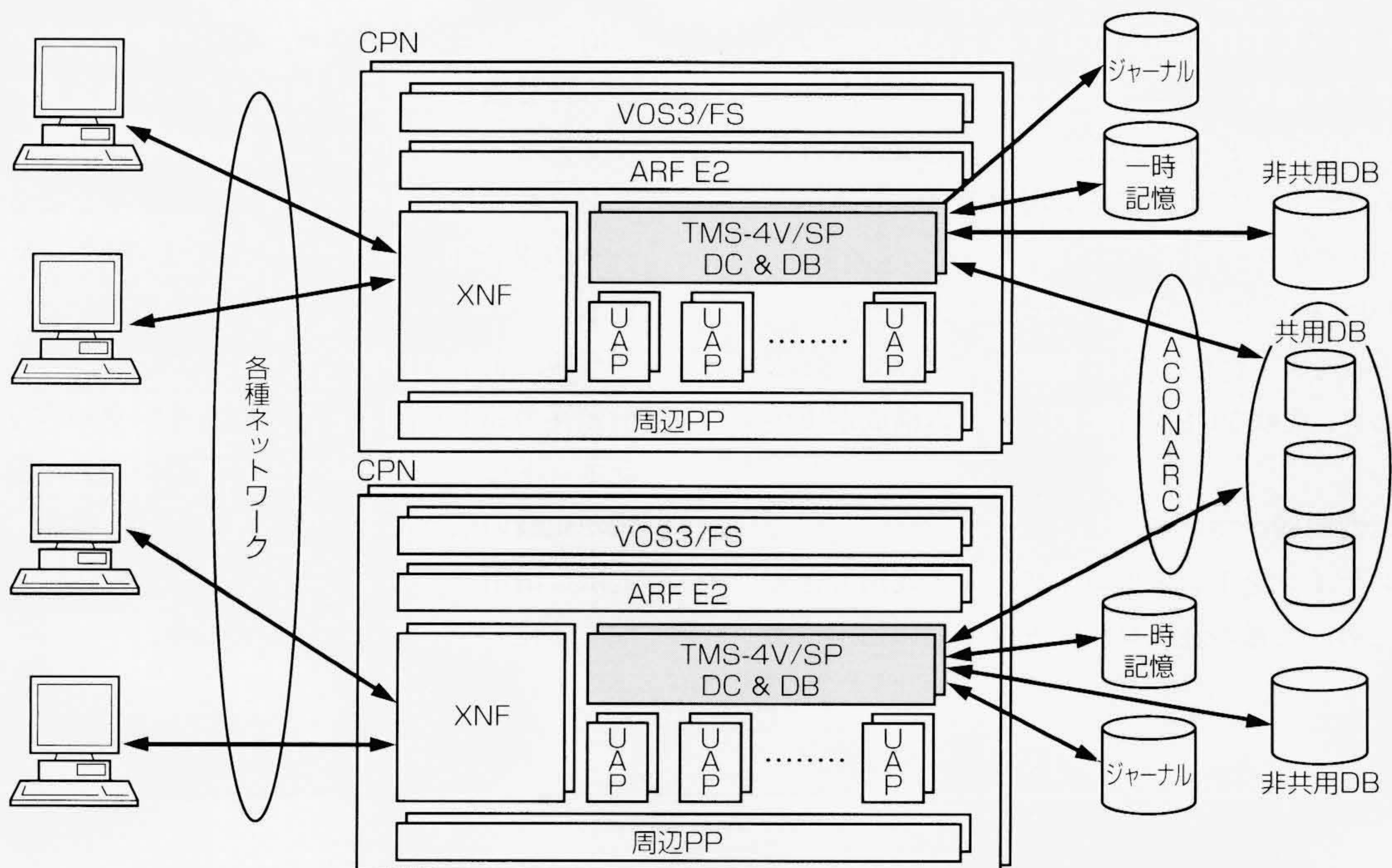
端末とのセッション接続時に各XDMシステムの負荷が平均化するように、XDM/DCCM3が自動的に接続先XDMシステムを選択する。負荷の平均化は高速結合機構を使用し、CPU使用状況およびセッション接続数によってXDMとXNF、およびVOS3/FSのシステム間負荷分散機能が連携して行う。したがって、ユーザーは各XDMシステムのシステムチューニングを行う必要がなく、ユーザーの負担を軽減することができる。

3 TMS-4V/SPのパラレルサポート

3.1 概要

金融オンライン向けのDB/DC製品であるTMS-4V/SP (Transaction Management System 4V/System Product)でのパラレルサポートについての顧客ニーズは下記の点にある。特に、システムの高信頼性に対する要請は大きい(図5参照)。

(1) システム規模、発生データ量に応じた負荷分散



注：略語説明

ARF E2 (Advanced Reliability Feature Extended 2)

PP (Program Product), ACONARC (Advanced Connection Architecture)

図5 TMS-4V/SPパラレル

各CPNのオンラインの独立性を保持することにより、システム全体の信頼性についてのリスクを低減している。

(2) エンドユーザー(営業店)から見たノードダウンシステム

(3) スケーラブルなシステム運用

(4) 1システムイメージの運用

上記の顧客ニーズにこたえるために下記のパラレル機能を実現する。

3.2 TMS-4V/SPのパラレルサポート機能

(1) データベース共用によるトランザクションのパラレル処理機能

オンラインシステムの規模拡大に対応するために、複数のオンライン間で互いにデータベースを共用使用し、同一の勘定系の業務処理を並列に実行する。このとき、相互のデータベースアクセスでの整合性の確保についてはARF E2(Advanced Reliability Feature Extended 2)による資産の排他制御を行って実現している。

(2) 負荷分散

各種ネットワークで発生するオンライントランザクションを、複数のオンラインによって静的に負荷分散してパラレル処理を実現する。

(3) パラレルセッション機能

複数のオンラインと端末間を複数のセッションで接続することにより、一つのオンラインに障害があっても別のオンラインによって業務処理を続行することができる。

(4) スケーラブル運用

休日の稼動CPN台数の縮退運転など、稼動CPN台数の変更によるスケーラブル運用を実現する。

(5) $n:1$ ホットスタンバイ

複数のCPNで動作しているオンラインシステムのスタンバイジョブを、1台をCPNに複数立ち上げておくことができるので、柔軟なCPN台数のシステム構成が実現できる。

(6) システム間同期コマンド

複数のオンラインを1システムイメージで運用できるようにするため、共通の運用コマンドは1回の指示によ

って複数オンラインで同期して処理する。

(7) ジャーナルマージ

バッチ処理を1システムイメージで実行するため、複数のオンラインで取得したジャーナルを1オンラインで取得したジャーナルのように、時系列にマージする機能を提供する。

(8) 統合システム定義

複数のオンラインを1システムのイメージで統合定義することにより、複数オンラインの定義環境を自動的に生成する。このことにより、複数オンライン環境の生成および保守作業の軽減を図ることができる。

4 おわりに

ここでは、管理者による運用が容易なデータセンタで、今後ますます増大するデータを集中管理する「MパラレルDB/DCシステム」について述べた。

企業での情報の重要性は年々高まり、情報システムの障害、データの喪失は企業の存続にもかかわるものとなりつつある。

データを統合管理するデータセンタと、分散管理するパソコン、ワークステーションを有機的に連携させることにより、最も信頼性が高く、効率の良い情報システムを構築することができる。個別の情報はその所有者のところに分散し、共通の情報はデータセンタの管理者がバックアップも含めて正しく集中管理する。また、データセンタをワークステーション、パソコンに分散されたデータの保管庫、バックアップとして利用することもできる。

今後、企業の心臓であるデータセンタとして、性能、信頼性の向上を図るとともに、インターネット、イントラネットなどの利用環境との連携を深め、蓄積された情報が企業の隅々にまで行き渡って企業の栄養分として最大限に活用されるように、MパラレルDB/DCシステムの機能強化を図っていく考えである。

参考文献

- 1) 松岡, 外: 市場即応型ハイスピードビジネスの実現を目指す新FOREFRONT, 日立評論, 77, 5, 334~338 (平7-5)
- 2) 安部, 外: 新世代大型コンピュータMパラレルシリーズ

の開発, 日立評論, 77, 5, 371~374(平7-5)

- 3) 尾山, 外: 新世代システムによるデータセンタを実現するパラレルソフトウェア, 日立評論, 77, 5, 375~380 (平7-5)