

CS シリーズ (MCP) データベースの有効活用 ——DMS II データベース連携技術と今後の展望

Effective Utilization of CS Series (MCP) Database

——DMS II Database Access Methods and Future Technologies

小 森 芳 司, 竹之内 朋子

要 約 近年企業内の情報系システムは Windows 環境で構築されることが多い。特にオフィス環境においては、情報の利用方法が大きく変わってきた。今までのように一部の限られた人々が情報を利用するのではなく、社外/社内を問わず、エンドユーザが直接データを参照できるようなシステムへと変革している。

ClearPath Plus サーバ CS シリーズでは、メインフレームの既存データベースを Windows 環境から自由にアクセスできる機能を強化し、データベース連携ソフトウェア群を提供している。

本稿では、メインフレームの既存データベースの Windows 環境への複製ソリューションおよびデータベース連携ソフトウェアの技術解説を行う。

Abstract In the recent years, the Windows environment has become a popular one for building information systems within the enterprise. This transformation is especially seen in the offices, where the use of the data and information, formerly in the hands of system administrators and managers are gradually shifting to the end users, who require easy access to data from in and out of the enterprises.

To meet this need, the ClearPath Plus Server CS Series offers much functionality to enable access to databases on the mainframe from the Windows environment. The CS Series also has a vast lineup of software products to support database access.

This paper introduces the solution to replicate mainframe databases on the Windows environment and the software to access mainframe databases from the Windows environment, and explains technologies to realize them.

1. はじめに

CMP ClearPath Plus サーバ CS シリーズ (以下、CS シリーズ) では、MCP 環境と Windows 環境を同一システム内にそれぞれ複数のパーティションで構成することができ、CS シリーズのシステムとしての価値は、その両環境のデータ連携の巧みさに依存するところが多い。Windows 環境とのデータ連携においては、CS シリーズの MCP 環境の主要データベースである DMS II データベースが、おもなデータ供給源となる。

企業にとって重要な情報はメインフレームの基幹系データベースに多く存在し、情報系システム構築では、必要とするデータを基幹系データベースから抽出・加工し、そのデータを基に、メインフレーム内に情報系データベースを構築して、情報系のアプリケーションあるいはツールなどで利用する、というのが一般的であった。

しかし、PC の企業内での普及、Windows を中心とするオープン技術の進展、イン

ターネットあるいはイントラネットなどネットワークの浸透により、エンドユーザ・コンピューティングが普及し、必要なデータは、必要な時に、最新のデータを、という要求が非常に高まっている。このような状況の中で、限られた人に、限定された方法で提供されていたメインフレーム内のデータを、オープン環境の中で自由に利用できるような解決策が必要とされる。

メインフレーム・システムのオープン化には、データ、アプリケーションなど、すべてを一気にオープン・システムに移行するというダウンサイジングが90年代に進んだ。しかし、長年蓄積したメインフレームの財産を、一度にダウンサイジングして新たなシステムに構築し直すことは、投資コストやリスクが高いことが多く、基幹系データベースに存在するデータをオープン環境で有効的に活用して、オープン化に対応する解決策が望まれている。

CSシリーズでの解決方法のひとつは、MCP環境の基幹系データベースDMSⅡをWindows環境のリレーショナル・データベース(RDB)に複製(レプリケーション)してWindows環境で利用する方法である。DMSⅡはいわゆる階層型のデータベースであるが、必要なデータを抽出しRDB型に複製することによって構造の違いを吸収し、Windows環境で提供される豊富なRDB対応のソリューションの利用を可能とする。DMSⅡデータベースをRDBに複製することにより、DMSⅡデータベースのオープン化対応を実現し、さらにデータウェアハウスとしての利用を可能としている。複製RDBのデータの鮮度を保つためには、基幹系DMSⅡデータベースに行われる更新を、RDBに逐次反映する差分更新機能を提供する事で対応している。導入に関しては、従来の基幹系システムの変更がまったく不要であり、大量データ検索処理の負荷分散、投資コストの削減、開発リスクの回避などの利点がある。

解決方法のもうひとつは、Windows環境との間で標準インタフェースを介して、Windows環境からDMSⅡデータベースに直接アクセスするデータベース連携である。Windows環境をクライアント、MCP環境をデータベース・サーバとしたクライアント/サーバ型のシステムであり、業界標準インタフェースであるODBC(Open DataBase Connectivity)を使用した製品^{*1}はCSシリーズで提供済である。ODBCはリレーショナル・データベースへのアクセスを目的としたデータ連携技術であるが、マイクロソフト社はさらに、非リレーショナルな様々な企業全体に存在する情報源へのアクセスを目的とした、OLE DBというインタフェースを提唱し、現在、このOLE DBがデータ連携の業界標準のひとつとなっている。OLE DBインタフェースでは非リレーショナル型データベース、電子メール、ファイル、テキストさらにビジネス・オブジェクトなどへの情報源へのアクセスを可能としており、CSシリーズでもWindows環境からDMSⅡへの新しいデータ連携としてOLE DB対応インタフェースを推奨している。OLE DBはODBCに比べてより高速なアクセスを提供し、DMSⅡデータベースを分散データベースとして使用したり、2PC(Two Phase Commit)機能を用いた異機種データベース間でのトランザクション処理も可能である。OLE DBをMCP環境で使用する場合も、既存のシステムの変更は一切必要ないため、投資コスト、開発リスクを大幅に削減し、基幹系DMSⅡデータベースを容易にオープン化対応できる。

本稿では、以上で述べた「データベース複製ソリューション」と「データベース連携」の機能と技術および利点を解説する。

2. MCP 環境から Windows 環境へのデータベース複製ソリューション

2.1 DataExtractor II /NX DB レプリケータの機能と特長

DataExtractor II /NX DB レプリケータ (以下、DB レプリケータ) は、CS シリーズ上の基幹系 DMS II データベースを、Windows 環境の RDB に複製するデータベース複製ソリューションである。複製対象となる RDB は SQL Server (マイクロソフト社) および Oracle (オラクル社) データベースであり、複製データベースは、Windows 環境で情報系データベースとして、豊富な開発ツール、ソリューションを利用して、Windows 環境に情報系システムを構築できる。

さらに、最近、広く使われているデータウェアハウス構築のための ETL (Extraction, Transformation, Load) ツールとしても利用可能であり、データ抽出 (Extract)、データ変換・加工 (Transformation)、データ投入 (Load) という機能を提供する。

複製データベースはリレーショナル・データベースであるため、アクセスは標準の SQL が基本であり、SQL に対応した Windows 環境の豊富な開発ツール、OLAP などのソリューションが利用でき、情報系データベースやデータウェアハウスが短期間で構築できる。また、大量データ検索に関しては、MCP 環境のリソースを必要としないため、負荷分散にも有効である。MCP 環境の DMS II データベースおよびアプリケーションには追加・変更は一切不要であるため、開発投資コスト、開発リスクの大幅な削減も期待できる。

DB レプリケータはさらに次に挙げる主要な機能を提供する。

- ・自動的にデータの鮮度を保持する差分更新機能
- ・DMS II と SQL Server および Oracle データベースとの構造的な違いの吸収 (オープン化)
- ・ユーザの既存 DMS II データベースのスキーマの変更が不要
- ・ユーザのアプリケーションへの変更・追加などの影響を与えない
- ・稼働中のデータベースを停止することなく、業務中およびデータベースのバックアップ中でも影響なく使用できる (24 時間 365 日連続運転対応)

2.2 初期データ構築モジュール

初期データ構築モジュールは、DMS II データベースから複製対象のデータを選択、抽出、変換し、複製先の Windows 環境にデータの送出行を行い、SQL Server あるいは Oracle データベースに初期の複製データベースを構築する。初めて複製データベースを構築するとき、あるいは定期的に DMS II データベースからデータを抽出し、複製データベースをリフレッシュするときに行う。初期データ構築モジュールの実行構造を図 1 に示し、以下にモジュールの働きおよび特徴を説明する。

- 1) MCP 側のレプリカ (複製) 対象選択プログラムは、複製データベースの対象となる DMS II のデータセット (リレーショナル・データベースの表に相当) や項目 (リレーショナル・データベースの列に相当) に関連する情報を、DMS II スキーマから読み出し、レプリカ情報ファイルに出力する。ユーザが定義する情

報はこれだけであり、この情報は、データの抽出、複製先データベースのスキーマ生成に使用する。

- 2) 初期データ抽出プログラムは、レプリカ情報ファイルに書かれた情報をもとに、複製対象となる DMS II データの選択を行い、抽出を実行する。抽出されたデータは抽出ファイルに一括出力される。
- 3) 同時に初期データ抽出プログラムは、Windows 環境に複製データベースを作成するために、DMS II スキーマ情報を取りこみ、レプリカ情報ファイル内の情報を元にして、複製先の SQL Server あるいは Oracle データベースのスキーマ (DDL) を SQL/DDDL ファイルに出力する。

キーとなるスキーマ変換技術に関しては、次の「2.3 節の DMS II スキーマから RDB スキーマへの変換」で、さらに詳しく説明する。

- 4) Windows 側のデータ転送ユーティリティは、MCP 環境に作成されたファイル群を取り込み、SQL/DDDL ファイルを使って、複製データベースを初期化する。次に、抽出ファイルに出力されている複製データを読み込み、SQL Server ユティリティあるいは Oracle SQL * Loader を使用して初期データの投入を行う。DMS II と SQL Server あるいは Oracle データベース間のコード変換、データ様式の変換も、このユーティリティによって行われる。この処理が完了すると、Windows 環境の複製データベースが利用可能となる。

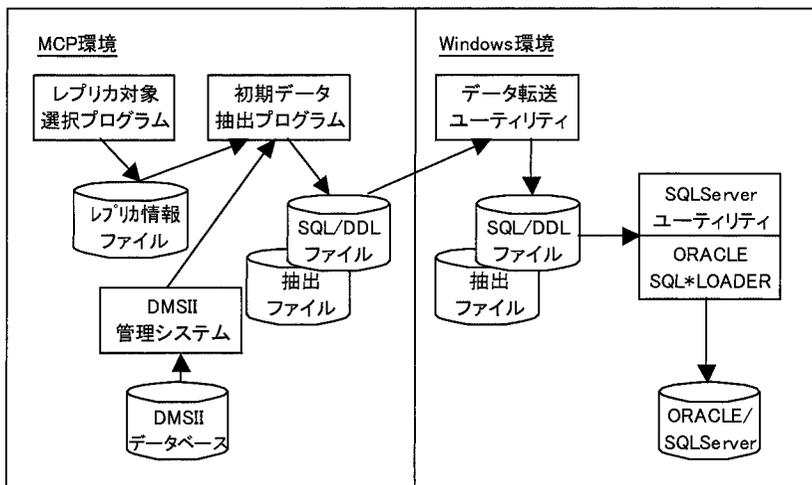


図 1 初期データ構築モジュール

2.3 DMS II スキーマから RDB スキーマへの変換

DMS II データベースはいわゆる階層型データベースであり、複製対象のリレーショナル型データベースとはスキーマ構造が大きく異なり、ユーザにスキーマの変換 (マッピング) 作業を委ねることは大きな負担となり、さらに誤りの原因となる。DB レプリケータは、DMS II データベースからのスキーマの変換、複製対象データの抽出を汎用的に処理できる機能を提供しており、ユーザはレプリカ対象選択プログラムを

使用して、複製対象とする DMS II のデータセットおよび項目を指定するだけで、スキーマの変換、データの抽出を自動的に実行するように設計している。

スキーマ変換を実行する初期データ抽出プログラムは、ユーザが記述したスキーマそのものではなく、生成済の抽象化されたスキーマ情報を直接に読み込む方法を使う。DMS II データベース管理システムのソフトウェアを生成する特殊用途のコンパイラ (DMALGOL コンパイラ^{*2}) を使い、初期データ抽出プログラムをコンパイルする時点で、複製元の DMS II スキーマ情報を組み込ませる手法をとっている。この方法により、スキーマ情報はスキーマ名を指定するだけで、コンパイラが自動的に取り込むので、複製元の DMS II データベースに対応した、効率の良い初期データ抽出プログラムが簡単に生成できる。

初期データ抽出プログラムは、スキーマ変換と同時に、複製対象となるデータセットおよび項目データを、複製元のユーザ・データベースから全件読み込み、抽出ファイルに一括して出力する。大量データを一括出力するため、高効率な処理が必要であること、通常のアプリケーション・レベルのプログラムではアクセスできない制御データを読み取る必要があり、DMS II データベース管理システムの提供する特殊な機能を使って、DMS II のデータを読み取り、抽出ファイルに出力させる技術を採用した。

データ抽出完了後に、スキーマ情報を使って、Windows 環境の SQL Server あるいは Oracle データベースに合わせたスキーマを SQL/DDDL (Data Definition Language: スキーマ定義言語) に生成してファイルを自動的に抽出するため、ユーザは複製先のデータベースのスキーマを作成する必要がない。さらに Windows 側のデータベースの機能を使って、抽出ファイルを一括コピーするバッチ・スクリプトを生成し、ユーザは SQL/DDDL ファイルとバッチ・スクリプトを実行するだけで、複製データベースへのデータ投入を完了できるように設計している。

2.4 差分データ転送モジュール

差分データ転送モジュールは、DMS II データベースに対する更新 (作成/変更/削除) ログを記録する追跡記録ファイルから、複製対象となる差分データを抽出し、Windows 環境に転送する。Windows 環境では受信した差分データを複製先の SQL Server あるいは Oracle データベースに適用し、データの新鮮度を、複製元と同期して保持する。図 2 に差分データ転送モジュールを示し、以下に各プログラムの働きを説明する。

- 1) 追跡記録読み込みプログラムは、複製対象となる DMS II データベースへの更新ログ情報が書かれている追跡記録ファイルから、対象のデータセットおよび項目への更新データを抽出し、Windows 側のデータ受信プログラムに更新データを送信する。
- 2) データ受信プログラムは MCP 側から受信した更新データを、差分更新プログラムにマイクロソフト社の MSMQ インタフェースを介して送信する。
- 3) 差分更新プログラムは複製対象のデータをフィルタリングし、コード変換、データ様式の変換を行い、更新のための SQL DML 文 (SQL データ操作文) を生成し、SQL Server あるいは Oracle データベースを更新し、DMS II データベースと同一の更新状態に複製データベースを維持する。差分更新の詳しい操作に関

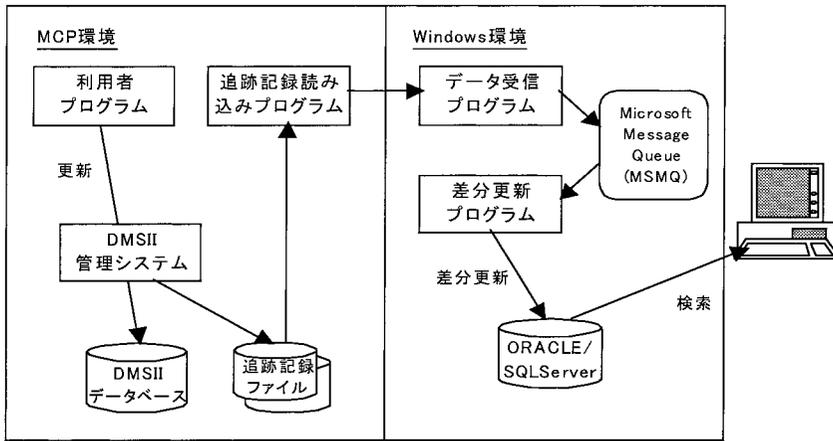


図 2 差分データ転送モジュール

しては、次の「2.5 節の差分更新技術の実現」において詳しく説明する。

2.5 差分更新技術の実現

DMSII データベース管理システムは、データベースへのすべての更新ログ情報(差分データ)を追跡記録ファイルと呼ぶファイルに書き出している。追跡記録ファイルには、レコードの更新情報、インデックスの更新情報、データベース管理のための制御情報などが存在するが、DB レプリケータは、複製対象となるデータセットのレコードの更新ログ情報のみを選択する。更新ログ情報は、図3に示すように作成、変更、削除の3タイプの情報が存在し、それぞれ、データセット内のレコードの一意な物理的なアドレスを持つ AA ワード (Absolute Address Word^{*3}) とレコードの更新イメージを含んでいる。DB レプリケータは、この情報をブロックにまとめて、シリアルな番号をつけて Windows 環境に転送し、差分更新が完了したシリアル番号を複製データベースに格納する。

作成レコード	制御語	AA語	作成後イメージ	制御語	
変更レコード	制御語	AA語	変更前イメージ	変更後イメージ	制御語
削除レコード	制御語	AA語	削除前イメージ	制御語	

図 3 追跡記録レコードのイメージ

リレーショナル・データベースの場合、必ずユニークなキー項目が必要であるが、DMSII では、必ずしもユニークなキーを必要としない。したがって、複製データベースにマッピングする場合には、ユニークなキーを指定するか、ユニークなキーがない場合は、新たに付加する必要がある。これを解決する方法として DB レプリケータは、<データセット名> AA という名前の項目を追加し、各レコードに AA ワードの値を含ませている。

Windows 側の差分更新プログラムは、各レコードから必要な項目を抽出し、コード変換を行い更新用の SQL DML 文を作成する。レコード作成の場合は、AA ワードによる重複存在確認を行い、INSERT 文を実行する。レコード変更の場合は、追跡記録の中の変更前イメージと変更後イメージを比較し、変更された項目のみをデータ適用する UPDATE 文を実行する。レコード削除の場合は、AA ワードをキーとした DELETE 文を実行する。すべての更新は、DMS II の提供するリカバリ処理と同様に AA ワードをキーとしているため、DB レプリケータ稼働中にどちらかの環境に障害が発生しても、データの不整合は発生しない。再起動時には、複製データベースに格納されている前述のシリアル番号を MCP 環境に報告し、そのシリアル番号からの更新情報の再送で同期をとる仕組みである。

3. MCP 環境と Windows 環境のデータベース連携

3.1 OLE DB Data Provider の機能

OLE DB インタフェースを使用することによって、ODBC に加えて、ユーザは一般ファイル、表計算ソフトのシート、電子メール、ビットマップや動画など、非リレーショナル型データに対して標準のインタフェースを用いてアプリケーションの開発が可能となった。マイクロソフト社の SQL Server データベースにおいては、OLE DB 準拠の異種データベースと連携するリンクサーバ機能が導入され、OLE DB 準拠の異機種データベースへのアクセスが、SQL Server の提供するツールおよび機能で行える。

DMS II データベースにおいては、OLE DB 規格に準拠したソフトウェアを OLE DB Data Provider と称して提供している。このソフトウェアによって、Windows 環境のアプリケーションから DMS II に対するアクセスが SQL Server と同様の容易さで行なえる。また、MCP 環境の一般ファイルも OLE DB インタフェースの対象となり、従来の COBOL および ALGOL などの言語を使用した MCP 環境での開発から、Windows 環境の VB や VC++ などを使って、MCP 環境のデータをアクセスするビジュアルなソフトウェアの開発が可能になった。

3.2 OLE DB のコンポーネント

マイクロソフト社の提唱する OLE DB は、様々なデータ管理サービスをカプセル化するコンポーネント・オブジェクト・モデル (Component Object Model: COM) ベースのインタフェースである。統一的な方法によってアプリケーションから異機種間にまたがるデータのアクセスおよび共有が可能である。OLE DB コンポーネントは、図 4 にあるように、コンシューマおよびプロバイダと大きく二つに分類できる。コンシューマとは、OLE DB インタフェースを使用するシステムまたはアプリケーションを指し、開発ツール、プログラミング言語、高度なアプリケーションなどがコンシューマである。

OLE DB プロバイダは更に二つのコンポーネントに分類することができる。一つはデータプロバイダであり、データを外部に行セットとして表形式にて公表する。データプロバイダには、リレーショナル・データベース、記憶管理、スプレッドシート、一般ファイル、電子メールなどがある。ユニシスが提供する OLE DB Data Provider

製品は OLE DB の仕様でいうデータプロバイダの一つである。

プロバイダの二つ目の分類は、サービスプロバイダである。このプロバイダはデータを保持せず、OLE DB インタフェースを介してデータを生成・利用することによって、データベース管理システムの機能をサービスとしてカプセル化する論理オブジェクトである。このため、サービスプロバイダは、プロバイダと同時にコンシューマにもなりうる。

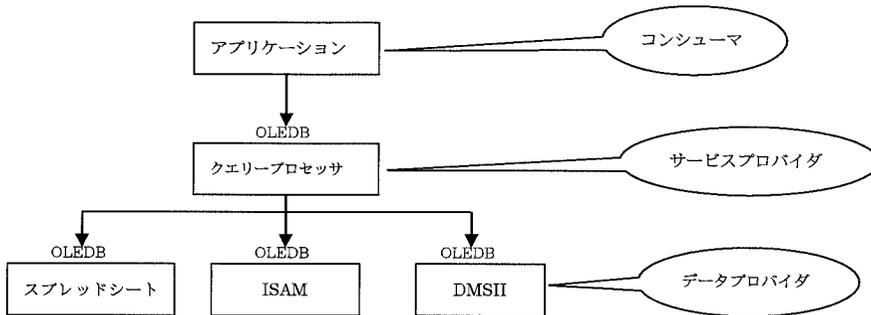


図 4 OLE DB コンポーネントの分類

3.3 OLE DB Data Provider の実装と利点

CS シリーズにおける OLE DB インタフェースの実装は、図 5 のように Windows 環境の OLE DB Provider および MCP 環境の OLEDB/SERVER および OLEDB/WORKER の各モジュールより構成される。Windows 環境の OLE DB Provider は、OLE DB 構文の解析から DMS II 構文への変換および文字コード変換などを行なう。MCP 側の OLEDB/SERVER は、Windows 環境より要求を受け、セッションの管理を行い、またセッション毎に OLEDB/WORKER を立ち上げる。DMS II データベースにアクセスするのは OLEDB/WORKER となる。

ODBC は標準 SQL 文を用いてデータアクセスを行うが、OLE DB のデータアクセスは、データプロバイダおよびサービスプロバイダ側に大きく依存する。例えば ICommand (OLE DB の拡張インタフェースのひとつ) インタフェースを支援する SQL Server などのデータプロバイダに対しては、SQL 文を用いデータアクセスを行なうことも可能である。

DMS II に対する効率面では、OLE DB のインタフェース言語を DMS II の操作言語

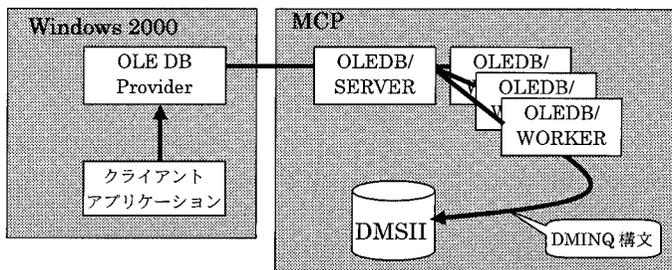


図 5 MCP 環境の実装

に置き換える方が SQL 文の解析などが発生しない分、ODBC 処理より早い。ODBC はリレーショナル・データベースとしての制約があることに對し、OLE DB では、レコード単位での操作、Index の指定など、従来の DMS II に対するプログラミングとも類似しているため、DMS II の利用者は理解しやすい利点がある。

3.4 OLE DB インタフェースの利用

SQL Server のツール、データ変換サービス (Data Transformation Service : DTS) を使用し DMS II データセットのデータを SQL Server へ流し込むインポート、または、その逆で DMS II 内にデータを落とし込むエクスポート機能を使用することも可能である。例えば、毎朝、DMS II データベースのマスタ系のデータセットを SQL Server のテーブルとしてインポートし、アプリケーションから SQL Server を参照することができる。データ変換サービスにはスケジューリング機能もあり、オペレータ介入なしに自動運用することもできる。

SQL でのインタフェースが必要となる場合、マイクロソフト社の SQL Server のリンクサーバ機能によって、DMS II のデータベースを SQL Server 内のデータベースとして見せることにより、SQL Server をクエリープロセッサ (SQL 構文の解析を行うサービスプロバイダ) として扱うことができる。このような定義を行っておくことにより、SQL Server 独自のツール、例えばクエリーアナライザなどのツールが DMS II データベースに対して直接使用でき、エンドユーザのインタフェースはすべての SQL Server のツールから可能となる。

OLE DB Data Provider を用いることによって、開発者は Windows 環境の GUI ツールなどを使用したアプリケーションから DMS II のデータをアクセスすることが容易に行える。例えば Windows 環境の Web アプリケーションや VB アプリケーションより、Active Data Object (ADO) インタフェースを使用する場合、DMS II 固有の構文などをまったく意識することなく、ADO の標準インタフェースにてデータを検索および更新することが可能となる。

Windows 環境のトランザクション・マネージャを使用することで、DMS II データベースを分散トランザクションに参加させ、複数のデータベースを同時に更新することも可能である。例えば、図 6 にあるように、Windows 環境で 2 PC プロトコルを実装するマイクロソフト社の分散トランザクション・コーディネータ (Microsoft Distributed Transaction Coordinator : MSDTC) を用いることにより、複数の DMS II データベースに対する更新を一つのグローバル・トランザクションによって管理したり、DMS II を異種分散データベース間で、一つのトランザクションで同時更新を行ったりすることが可能となる。

Windows 環境より 2 PC の更新を行なう場合、MCP 環境でのプログラミングは不要である。DMS II データベースに、トランザクションの独立性および更新データが反映されていることを保証するオプションとグローバル・トランザクションへの参加を表明するオプションの設定により 2 PC が実現できる。

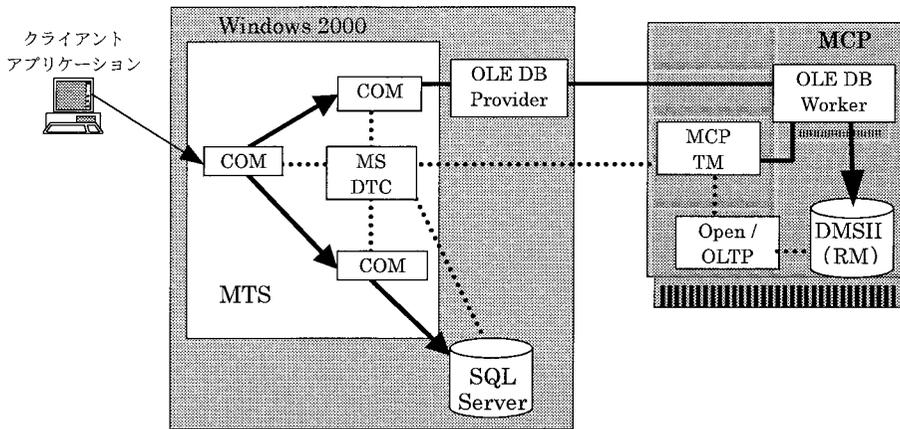


図 6 分散トランザクション・コーディネータとの連携

4. おわりに

本稿では DataExtractor II /NX DB レプリケータおよび OLE DB Data Provider の提供する機能、技術および利点を論じたが、DB レプリケータはデータウェアハウス、データマートへの利用/展開、OLE DB Data Provider はさらなる機能追加を目指したい。

Java 環境とのデータ連携を実現する CS シリーズ上の JDBC (Java Database Connectivity)、最近、注目を浴びている XML (eXtensible Markup Language) 技術を用いた DMS II データベースのデータ交換技術も既に製品として提供しているが、これらの製品に関しては次の機会にその技術を論じたい。

- * 1 ODBC: ODBC とは、仕様に準拠した Windows 環境のツールや言語プログラムより、ODBC 準拠の異機種データベースへの SQL CLI (Call Level Interface) を用いた操作を可能とした規格である。多くのデータベース提供者およびソフトウェアベンダは、ODBC インタフェースに対応した多くの製品を提供している。CS シリーズでは TransIT ODBC およびその後継として、INFOAccess ODBC が製品として提供されている。
- * 2 DMALGOL: DMALGOL は、DMS II システム・ソフトウェア用に機能拡張された ALGOL コンパイラであり、アプリケーション・プログラムでは取得できない DMS II データベースのあらゆる情報を操作できるプログラムのコンパイルが行える。
- * 3 AA ワード: AA ワードは、DMS II のファイル上のデータを管理するためのアドレスであり、各データレコードの物理的な位置を示すものであり、ファイル上の相対的なブロック・アドレスとワード・アドレスによって構成されている。

執筆者紹介 小森 芳司 (Yoshiji Komori)

1982 年近畿大学理工学部経営工学科卒業。同年日本ユニシス(株)入社。A/NX/CS シリーズのデータベース関連ソフトウェアの開発、保守業務に従事する。現在、第二ソフトウェアサービスセンターマイグレーションサービス一室に所属。

竹之内 朋子 (Tomoko Takenouchi)

1990年上智大学外国語学部比較文化学科卒業。同年日本ユニシス^株入社。A/NX/CSシリーズの利用技術サービスに従事した後、同シリーズのデータベース関連ソフトウェアの開発保守業務に従事する。現在、プロダクトサービス部NXソフトウェア室に所属。