COMDEXへの米国ユニシス出展事例をベースとした大規模トランザクションシステムの構築

Case Study of Building Large-Scale Transaction System Based on Unisys Experience in COMDEX '99

馬 場 功 二

要 約 Windowsシステムを基盤としたシステム開発が急増する中で，ミッションクリティカルな大規模システム開発事例はなかなか急増しない状況にある．本稿では，米国ユニシスがMicrosoft社，EMC社等パートナー企業10社の市販製品を使用してCOMDEX '99で実証したWindows 2000を基盤としたOnLineTransaction Processing（OLTP）システムとData WareHouse（DWH）システムをWindowsシステムの大規模システム構築事例として紹介することで，Windowsシステムがミッションクリティカルな大規模システムの構築に耐えうるものであることを明らかにする．

Abstract Though the Windows-based system development is booming, there exist the conditions that the cases of the system development of the large-scaled mission-critical system have not significantly increased as expected.

This paper introduces the online transaction processing (OLTP) and the data warehouse (DWH) systems, based on Windows 2000, which Unisys demonstrated at COMDEX '99 using the marketable products from ten alliance partners including Microsoft Corporation and EMC Corporation, as the system building case of a large-scaled mission-critical application system, and provides evidences that Windows system satisfies the requirements for the system building of large-scaled mission-critical applications fully.

1. はじめに

Windowsシステムを基盤としたシステム開発はWebシステムを中心に急速に増加しているが，ミッションクリティカルな大規模システムの開発案件はなかなか急増しない状況にある．その主な原因は，Windowsシステムの「優れた柔軟性」や「高いコストパフォーマンス」は理解できても，「Windowsシステムの不安定さ」に対する潜在的な問題意識が強く，大規模Windowsシステムの事例が少ないこととあいまって，「本当にメインフレーム並みの基幹システムがWindowsシステムで構築可能なのか」の疑問が解けないためと推測される．

そのような状況の中で，アメリカ最大の株取引会社NASDAQ社の「株取引監視システム」をWindowsシステムで開発し，またメインフレームで稼働していたNorwest社の銀行業務をオープンシステムとして再構築するなどミッションクリティカルシステムの構築に多くの実績を持つ米国ユニシスが，パートナ企業10社と協力して21世紀の企業システムをイメージした「次世代のデータセンター」をCOMDEX '99に出展し，「市販されているパートナ企業の製品を集めて構築する形態でも，可用性と安定性を兼ね備えた大規模Windowsシステムが構築できること」を実証した．本稿では，Windowsシステムの大規模システム構築事例として，米国ユニシスがCOMDEX

’99 に出展した「次世代のデータセンター」を紹介する．

米国ユニシスの World-Wide Enterprise NT Center Of Excellence（WWENTCOE[*1]）が COMDEX ’99 展示システムの構築を行うにあたり，当社からも WWENTCOE が蓄積した大規模システム構築技術を修得しつつ共同でシステム構築を実施するためにメンバーが参加した．残念ながら当社のメンバーはシステム構築の全工程を経験することはできなかったが，米国ユニシスの構築責任者や構築メンバから確認した内容，自ら構築したミニ OLTP システムの構築経験，システムテスト段階で構築メンバーから都度確認した内容，帰国後に評価・確認した新技術の結果などをまじえて，米国ユニシスが構築した「OLTP システム」と「Data WareHouse（DWH）システム」を紹介すると共に，Windows ベースで大規模システムを構築するのに有効な技術や種々考慮点などを述べる．

2. 出展システム概要

出展された「次世代のデータセンター」のシステム概要（ハードウェア，ソフトウェア，構築方法，運用など）を簡単に紹介する．

データセンター内に設置されたハードウェアは，ユニシスの Unisys e-@ction Enterprise Server ES 5000 56 台（予備サーバ 4 台を含む）と指紋認証装置付きキーボード，EMC 社の SYMMETRIX 3000 ディスク装置 28 TB[*2]（28,000 GB），Cisco 社の Router 2 台，Giganet 社の cLAN 装置 12 台，StorageTek 社の Tape Library 装置 3 台，Imation 社のカートリッジテープ約 1000 本など，パートナ企業が市販している最高の製品を使用した．また，ソフトウェアはマイクロソフト社の Windows 2000 Advanced Server（RC 2）を基盤に IIS[*3]，ISAPI[*4]，ASP[*5]，VC＋＋，COM＋[*6]，SQL Server，Active Directory，Terminal Service 等と EMC 社の Time Finder，EDM（EMCDataManager），Mercury Interactive 社 の LoadRunner と NetIQ 社 の AppManager を使用した（附図 1，附図 2）．

展示システムの内訳は，1 日 3 億件の Web ヒット要求を処理する OLTP システム，7 TB の大規模 DWH を高速検索する情報検索システム，5000 万項目で構成された Active Directory を使った人事情報検索とプリンター保守処理，指紋認証付きキーボードを使ったセキュリティシステムから構成される（附図 1）．システム構築は，世界中で活躍するユニシスの WWENTCOE 技術者約 10 人が 5 ヶ月をかけ，製品知識の習得，機能確認や性能確認の単体評価，Windows 2000 との相性確認，システムインテグレーション，性能と安定性を確認するシステムテスト等の作業を実施した．構築段階で苦労したのは，初めて使用する製品が多く，システムにどの様に組み込むべきかに関して悩んだり，β 版ゆえに期待した機能のリリースが延期され代替機能を作り込まねばならなかったり，或いは補完機能を作り込んでも期待通りの動きをしてくれず時間切れで断念せざるを得なかったことなどであった．

運用に関しては，展示場の 2 階に操作室を設け，数台の PC から 52 台のサーバを集中操作する形態とした．データベースを最新に保持するため，DTS[*7] 相当のユーティリティを開発し夜間に OLTP システムからデータを吸い上げ，DWH システムのデータベースに転送した．また，OLTP 及び DWH システムのデータベース（各々

1.4 TB と 3〜7 TB）は，1 時間に 1 TB の速度でテープ装置に高速バックアップを行った．

3. OLTP システム

OLTP システムは，米国で年間を通じて最もインターネット取引量が多い「クリスマス時期：11 月末からクリスマス迄」を上回る 1 日 3 億件の Web ベース取引が処理出来ることを要求仕様として構築され，3 日目の朝に起こった「突然の停電」にもびくともせず，秒約 100 件の処理能力で 5 日間連続稼働を達成した．

3.1 全体構成図

COMDEX 会場で次世代データセンターの説明用に掲示されたハードウェア構成図（図 1）を使って OLTP システムのハードウェア概要を説明する．

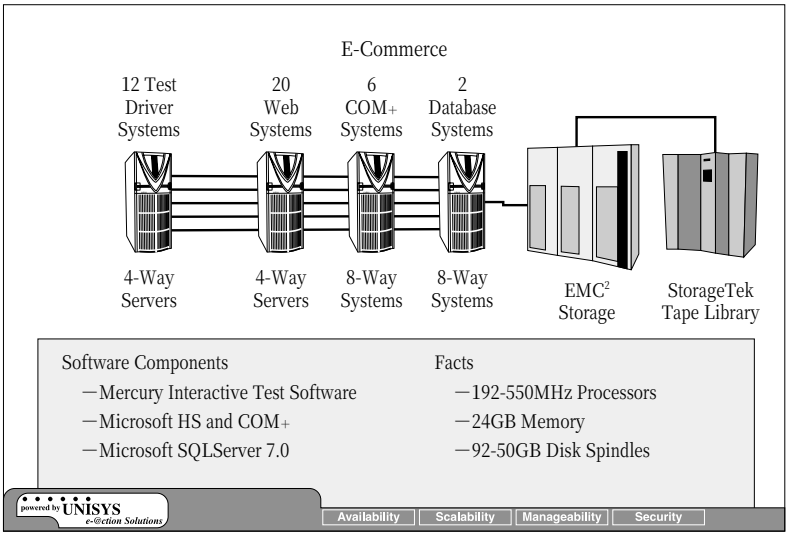


図 1　COMDEX 会場の説明用 50 インチ画面に表示されたハードウェア構成図

OLTP システムは 28 台のサーバを 3 階層（NW 層，AP 層，DB 層）に構成し，12 台のテストドライバシステム（Mercury　Interactive 社の LoadRunner）がクライアント・シミュレータの役割をする．

ディスク装置は 4.6 TB（50 GB×92 スピンドル）の物理容量を備えた EMC 社のエンタープライズストレージが 6 本のファイバー・チャネルで DB サーバに接続される．単一 SQL　Server で構築したデータベースは RAID 1 ミラーリングと Business Continuity　Volumes（BCV[*8]）を採用して 3 重化され，BCV を切り離してオンラインを停止させることなく StorageTek 社のテープライブラリ装置（25 GB×320 カートリッジテープ：Imation 社製）にバックアップできるように構成している．

ネットワークは Cisco 社の Catalyst 6500（二重化電源）で 100 MB/秒の LAN 環境を構成し，テストドライバシステムから Web サーバへのデータ投入とサーバ全体の管理に使用する．また，3 階層のサーバ間のデータ転送（Web 層―AP 層―DB 層）

には Giganet 社の cLAN（1.25 Gbits/秒，Full Duplex）を採用して，3 階層構成に起因して発生するかもしれないネットワーク遅延を抑えている．

サーバはユニシスの Unisys e-@ction Enterprise Server ES 5000（4 Way 及び 8 Way）を採用し，負荷が一番掛る AP サーバと DB サーバに 8 way サーバを配置する．サーバのメモリー容量は，Windows 2000 Advanced Server の最大メモリーサイズである 8 GB に統一され，大規模 OLTP 処理に耐え得るリソースを具備している．

3.2 トランザクション処理

図 2 はトランザクション処理を実現するためのシステム構成と必要なソフトウェアの配置について述べている．

当初の計画では Windows 2000 で提供される 3 種類のクラスタリングサービス（NLB[*9]，CLB[*10]，MSCS＋SQL Server）を採用して「拡張性」「可用性」「保守性」を備えた 3 階層の OLTP システムを構築する予定であった．しかし，NLB はシステムテスト段階で 20 台の Web サーバの負荷バランスが均等に取れないため，代わりに 1 台の LoadRunner から 2 台の Web サーバにデータを投入する静的な負荷分散方法に変更せざるを得なかった．また，CLB 機能は RC 2 版でドロップされ，後日「Application Center 2000」としてリリースされることになったため，Web サーバ上に顧客番号をベースに 6 台の AP サーバへ処理を振り分ける補完処理を開発してトランザクションの分散を実施した．データベース層は MSCS（MicroSoft Clustering Server）によるフェールオーバクラスタを採用している．

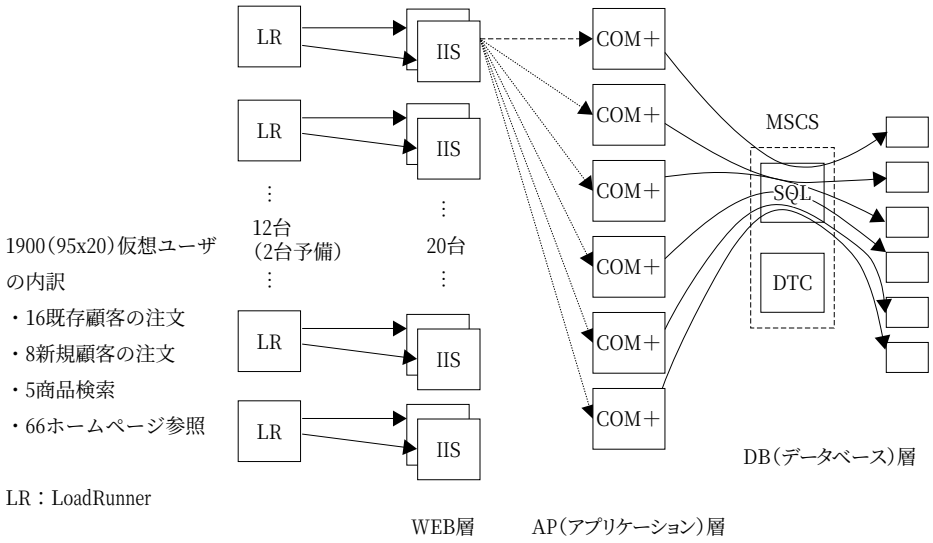


図 2　トランザクション処理でのソフトウェア配置

なお，NLB 機能，CLB 機能に関しては，後日，当社で独自に実施した 3 階層 OLTP システム（4×NLB，4×CLB，MSCS＋SQLServer）の評価テストでは，「NLB 機能の負荷分散と対象となる WEB サーバの動的拡大・縮小」及び「CLB の負荷分散と対象となる AP サーバの動的拡大と縮小」が正しく働くことを確認している．

1） テストドライバシステムの処理

テストドライバシステムは，Mercury Interactive 社の LoadRunner（Ver 6.0 β）で構成され，顧客がインターネットを使って物品を注文している環境をシュミレーションしている．10 台で構成したテストドライバシステムには 1900 仮想ユーザ（1 Web サーバ当たり 95 仮想ユーザ）を登録し，画面操作手順をスクリプトで記述して実行する．メニュー画面には，「商品注文」，「商品情報検索」，「メニュー画面参照」，「新規ユーザ登録」などがあり，テストドライバシステムからは，480（24×20）仮想ユーザが「注文処理」を，100（5×20）仮想ユーザが「商品情報検索」を，そして 1320（66×20）仮想ユーザが「メニュー画面参照」を実行する．また，注文処理で選択する商品数は乱数で 2 から 10 に変化させている．なお，大量なデータを発生させるため，1 件のトランザクション処理が終了した後に「応答待ち時間無し」で次のトランザクションを発生させている．

（1 Web サーバ当たり 95 仮想ユーザの内訳）

・16 仮想ユーザが「既存ユーザからの注文」
・8 仮想ユーザが「新規ユーザからの注文」
・5 仮想ユーザが「商品情報検索」
・66 仮想ユーザが「メニュー画面参照」

図 3 に既存ユーザからの注文を受けた場合の出力表示画面のシーケンスを示す．

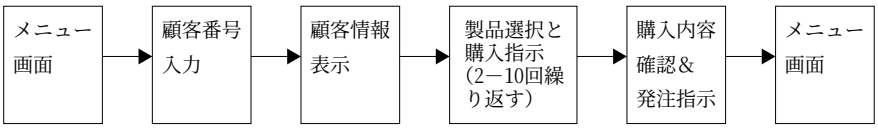


図 3 既存ユーザの注文処理の画面表示順番

2） ネットワーク層の処理

ネットワーク層の処理方法は，以下のように行う．

・「メニュー画面参照」依頼は，IIS/ISAPI で処理する．
・「商品情報検索」依頼は，ISAPI が，AP サーバ上の製品検索用データベースでデータを取り出して結果を表示する．この処理は，ISAPI のキャッシュでヒットする可能性が高く，ヒットすると秒 800 件，ヒットしないと秒 120 件の処理を行う（詳細は，5）参照）．
・「既存ユーザ」の注文処理は，ISAPI が「顧客情報」「商品選択」「商品購入」「注文指示」を処理する COM＋コンポーネントを呼び出し，データベースに対する検索，挿入処理を行う．
・「新規ユーザ」の注文処理は，「ユーザ登録」後「既存ユーザ」と同じ処理を行う．

Web サーバでの処理プログラムは，ASP より高速に処理できる ISAPI で開発され，IIS と ISAPI は処理効率を上げるためインプロセスモード（同じプロセス空間）で処理している．また，テストドライバと Web サーバ間は LAN を経由

してデータを転送するが，Web サーバと AP サーバ間はデータ転送専用ネットワーク Giganet を使って転送する．Giganet はユニシスの単体評価テストで「伝送遅延が 5 ms 程度の Low-Latency」，「TCP/IP 経由より CPU 使用率が低い」，「Gigabit Ethernet の約 2 倍の速さ」等の効果があることが確認され，システムテストでも LAN を経由するより応答時間が 10 倍速くなっている．この結果から判断すると Giganet は大規模システム構築時の近距離サーバ間のデータ転送に有効な製品と言える（図 4）．

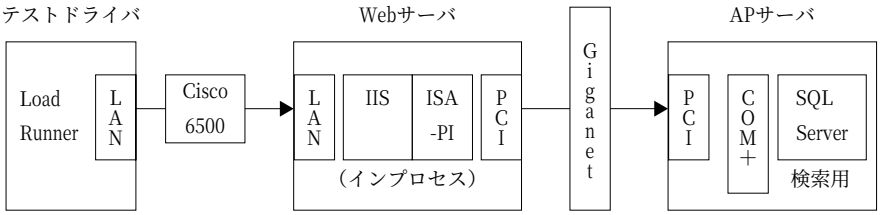


図 4　ネットワーク層と AP（アプリケーション）層の処理

3）　AP サーバ層の処理

AP サーバ上には VC＋＋で開発した COM コンポーネントが登録され，それぞれの COM コンポーネントを独立したトランザクションとして処理している．複数の COM コンポーネントに対する処理を一つのトランザクション（一連の処置）として処理する方法もある．この場合，SQL Server での処理時間に相違は無いがテーブルのデッドロックが発生し易くなる．すなわち，注文処理を例に取ると，「顧客情報」，「商品選択」，「商品購入」，「注文処理」を並行して処理することでデータベースのロックを避けている．また，商品情報を検索する処理は，AP サーバ上で IMDB（In Memory DB）機能の代わりに SQL Server を使用した（図 5）．

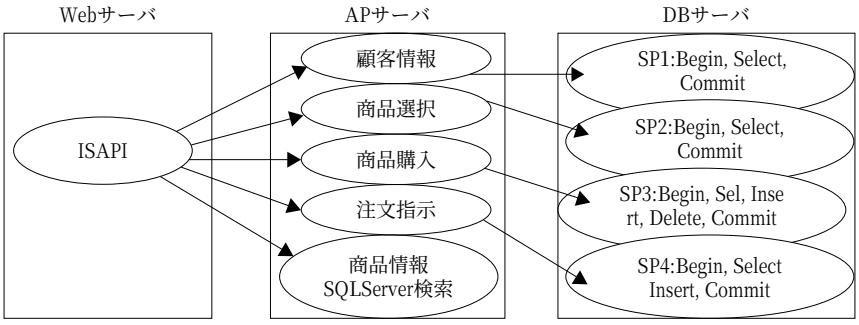


図 5　注文処理の AP（アプリケーション）層と DB（データベース）層

4）　DB サーバ層の処理

EMC 社 SYMMETRIX 3000 ディスク装置は複数の物理デイスク（50 GB 単位）

を一つの論理的な塊（Meta Volume）として作成するので，その塊をロジカルユニット[*11]（論理ドライブ）として扱った．当初 23 個のロジカルユニットが必要であったが，ドライブ文字を使い切った（E—Z の 22 文字）ため，22 個のロジカルユニットとし，テーブルに E—T を，インデックスに U—X を，トランザクション・ログに Y，Z を割り当てた．必要な 17 個のテーブルは同時に参照しないテーブル同士を組み合わせて 16 個のロジカルユニットにまとめ直した．各テーブルのインデックスは，参照系のテーブルにはクラスタ化インデックス（テーブルとインデックスを同じテーブル内に置く），更新系には非クラスタ化インデックス（テーブルとインデックスを別のテーブルに置く）を作成した．

このようにして作成したデータベースであったが，システムテスト段階でロック状態が発生し処理能力低下が判明したため，ロックを発生させている二つのテーブルをそれぞれ六つに分割し，各 AP サーバ毎のテーブルを用意してロックの発生を押さえた．

そのほか，Distributed Transaction Coordinator（DTC[*12]）は，トランザクション・ログ処理などで CPU 負荷が大きいため，クラスタリング（MSCS）の待機サーバ側に配置した．また，AP サーバ（COM＋）上の DTC の処理負荷を軽減させるため，DB サーバ上の DTC を参照するように設定を変更した．ちなみに，DB サーバ上の DTC のパフォーマンスモニターにより，秒 4000 件の COM＋トランザクション処理が実行されていることが判っている．また，DTC のトランザクション・ログファイルのサイズを標準値（4 MB）から 500 MB に拡張した結果，DTC サーバの CPU 使用率は 75% から 25% に下がった．

5） 検索処理システム効率測定

システムテスト中の OLTP システムを使って「検索処理」の効率測定した時の，効率結果と分析結果について述べる（図 6）．

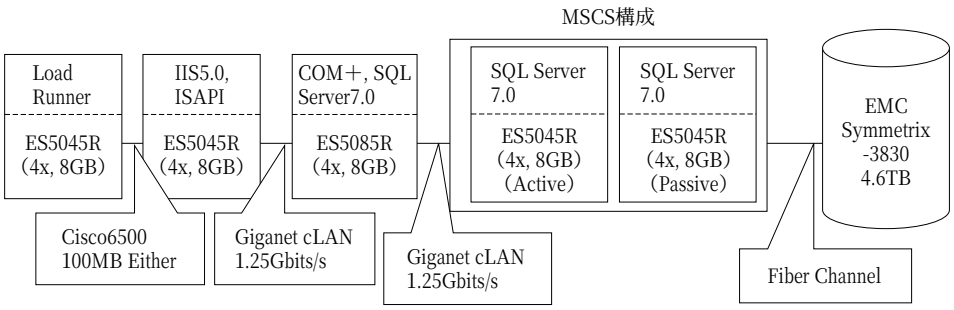


図 6　検索処理測定システムの構成図

（構成）

・LoadRunner,Web サーバ，AP サーバ，DB サーバ（MSCS 構成）を各 1 台で構成した．

（測定方法）

・LoadRunner から製品の情報を検索する「商品情報検索」処理を連続して実行

・各サーバの CPU 使用率をパフォーマンスモニターで監視

・IIS ログから「1 秒間の TRX 処理件数」と「応答時間」を算出

（結果）

・表 1 に効率測定結果を示す.

表 1　検索処理の効率測定結果

仮想ユーザ数（LR）	20 ユーザ	40 ユーザ	80 ユーザ	120 ユーザ
TRX 処理件数（s） キャッシュヒット	820 TRX/S	800 TRX/S	790 TRX/S	740 TRX/S
応答時間（ms）	数 ms	数 ms	数 ms	数 ms
CPU 使用率（LR）	70%	80%	100%	100%
（IIS）	20%	20%	40%	40%
（COM＋）	0%	0%	0%	0%
TRX 処理件数（s） キャッシュミス	120 TRX/S	120 TRX/S	120 TRX/S	120 TRX/S
応答時間（ms）	40–170 ms	40–350 ms	650–700 ms	1000 ms
CPU 使用率（LR）	40%	40%	40%	40%
（IIS）	20%	20%	20%	20%
（COM＋）	40%	80%	80%	80%

（考察）

・IMDB（In Memory DB）の代わりに商品検索用 SQL Server 7.0 を AP サーバに置いたが，Web サーバ上の ISAPI キャッシュでヒットするとデータベース（SQLServer）を全く参照しない．この結果，AP サーバの CPU 使用率は 0% となり，TRX 処理件数は秒 800 件に達している．また，応答時間は数ミリ秒である．

・一方，ミスヒットの場合は，AP サーバ上の COM＋の CPU 使用率が 80% に達し，TRX 処理件数は秒 120 件である．また，仮想ユーザ数を増加させるとテストドライバ（LoadRunner）の CPU はまだ余裕がある（40%）が，処理依頼件数が増加するため処理件数が変わらないものの 1 件の応答時間が悪化している．

6）　運用・管理

①　サーバオペレーション管理

データセンター内で稼働している 52 台のサーバは，展示場 2 階の操作室から数人の技術者によって，Windows 2000 Advanced Server の Terminal Service 機能を使って集中操作された．このサービスを使用すると同じセグメント内に接続したクライアントからサーバのスクリーンと全く同じ操作が実施できる．もちろん，リブートやシャットダウンも可能である．

②　サーバのリソース管理

52 台のサーバの稼働状況やリソースの使用状況は，NetIQ 社の AppManager を使用して監視された．Windows 2000 のパフォーマンスモニター機能を使う

ことで複数のサーバの稼働状況を監視する事は可能であるが，3 次元のグラフィック画面で一箇所から集中監視できることは操作員にとって大変有効な手段である．

③　セキュリテイ管理

ミッションクリティカルシステムでは「セキュリティ」も必須機能である．今回の展示では，Identicator 社の製品を使って開発した認証アプリケーションと指紋認証装置付きのキーボードを利用して，管理者からログイン許可を与えられた人だけが説明用 PC を操作できるようにセキュリテイ設定を行った．このため操作員は，事前に左右の人差し指の指紋を登録した．

3.3　OLTP システム構築時の考慮点

1）　AP のメモリーリーク追求

開発中の ISAPI を使って，構築中のミニ OLTP システムでシステム稼働テストを実施している時，パフォーマンスモニターのデータより，「コンテキストスイッチの値が大き過ぎる」ことが判明した．この値は，通常秒 2000 回位が適切だが，秒 65000 回を示していた．システム効率を低下させる要因となるため AP 開発担当者が追求を始めたが，なかなか原因が判明せず苦労した．原因は，「ISAPI と COM＋で共通に使っていたクラスでハンドルのメモリーリークが発生していた」ためであった．ちなみに，追求方法は「どぶ板作戦」で行い，ハンドル数を関数の呼び出し前後でダンプして増加している箇所を特定することができた．アプリケーションレベルでのメモリーリーク除去は必須アイテムである．

2）　ISAPI と COM＋のセキュリテイ

ミニ OLTP システムの構築中，単体テストからシステムテストを開始した時期に「ISAPI と COM＋が接続出来ない現象」に悩まされた．原因はセキュリテイで，「ISAPI を Out-of-Process に設定し，そのアカウントを COM＋のサーバがアクセスできるユーザに設定する」ことで解決できた．セキュリテイの考え方は基本的に WindowsNT　Server 4.0 と同じだが，Windows 2000 では NT 4.0 で別々に設定していたことが「COM＋Manager」に統一されロール機能が追加されたことと，設定を変更するだけで有効になるパラメータとリブートしないと有効にならないパラメータが混在したことに惑わされたためであった．

3）　ディスクファイル破壊

Windows 2000 Advanced Server RC 1 を使って MSCS 機能確認を行っていた時には，「ファイル破壊」にも悩まされた．ロジカルユニットを OS に反映する際「ベーシックディスク」と「ダイナミックディスク」が混在していることに気がつき，HELP を参照しながら「ダイナミックディスク」を指定した．ところが，ファイルを作成しただけで「ファイル内のデータが破壊された」との表示がされ「片方のノードからしかアクセスしていないのにどうしてファイルが破壊されるのか」理由が分からなかったが，インターネットの MS サポート情報より「現時点では，MSCS 環境ではダイナミックディスクは使用できない」ことが判明した．当社でのその後の確認で，MCSC 環境でのダイナミックディスクは「依然として制限が解除されていない」ことが判明している．

4. DWH システム

4.1 全体構成図

COMDEX 会場に掲示された HW 構成図（図 7）を使って DWH システムのハードウェアの概要を説明する．

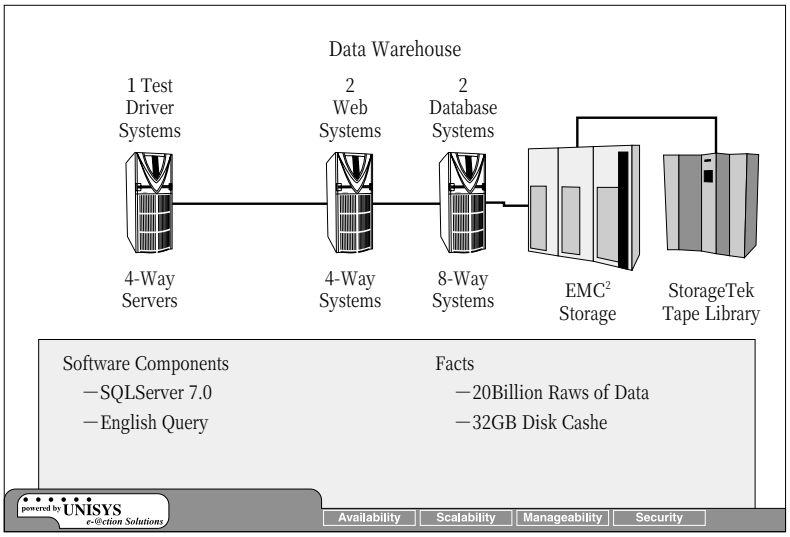


図 7　COMDEX 会場の説明用 50 インチ画面に表示された DWH システムのハードウェア構成図

DWH システムのハードウェアは 4 台のサーバを 2 階層（NW 層，DB 層）で構成している．デイスク装置は EMC 社のエンタープライズストレージ（SYMMTRIX 3930-50）が 2 台，容量的には 19.8 TB（50 GB×(384＋12 ホットスペア)）が 8 本のファイバーチャネルで DB サーバに接続される．データベース（単一 SQL Server）は RAID 1 と BCV を採用して 3 重化され， SQL Server の処理を停止することなく，BCV を切り離して 1 時間に 1 TB の速度でバックアップできる様に構成されている．バックアップソリューションは，EMC 社の EDM（EMC Data Manager）システムと Time Finder 及び StorageTek 社のテープライブラリ装置（9740 Module）で構成され，EDM システム（UNIX と EDM-SW）に接続された 10 本の Ultra-SCSI チャネルを経由して Tape Library 装置内の 10 台のテープ制御装置を並行稼働させて，1 時間に 1 TB の高速オンラインバックアップを実現している．Imation 社のカートリッジテープには約 25 GB のデータが記録でき，ロボットによるテープ交換時間は 5 秒である．サーバは，ユニシスの Unisys e-@ction Enterprise Server ES 5000(4 way と 8 way）を使用し，データ検索の応答時間維持の要となる DB サーバに 8 Way を配置し，メモリーサイズは Windows 2000 Advanced Server の最大メモリーサイズである 8 GB に統一している．

4.2 検索処理

システム構造は，IIS/ASP を使って SQL Server の 7 TB のデータベースを検索して結果を表示するごく一般的な DWH システムである．特徴は，7 TB（7000 GB）と

いう大容量でありながら，数億のデータを約 30 秒で高速検索可能なことである．実際の比較データはないが，米国ユニシスの開発担当者も一様に処理速度は確かに高速だ，と評価している．

当初，SQL Server を MSCS 機能を使ってクラスタ構成にする予定であったが，障害テストで Windows 2000 の新機能である「Mount 機能」を使って配置した 62 個のロジカルユニットの Mount 情報がクラスタリソースとして引継がれないため，フェールオーバできないことが判明し，断念した（図 8）．

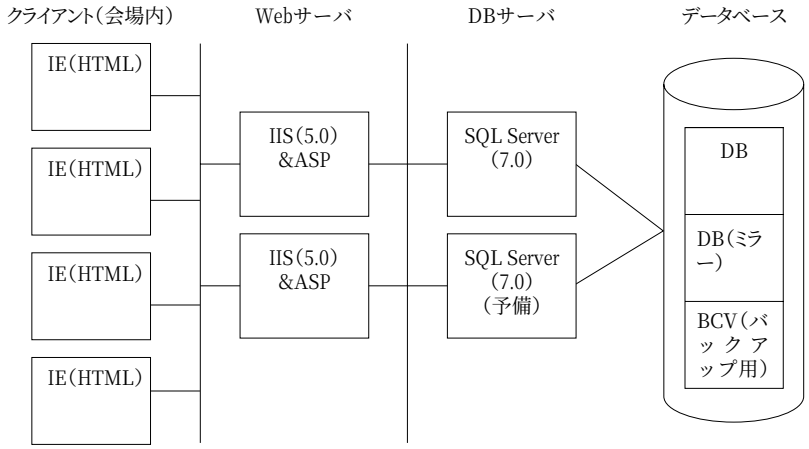


図 8　DWH システムのソフトウェア概要図

1）　ネットワーク層の処理

データベースを検索する仕組みは図 9 のようになっている．ASP を介して DB サーバ上に準備した SP（Stored Procedure）を呼び出す一般的なものである．検索処理は 2 種類用意されており，事前に用意した七つの検索処理をパラメータを指定して実行するものと，検索したい内容をキーワードで指定し，SQL Server の English-Query 機能を使って SQL 文を自動生成して検索するものである．

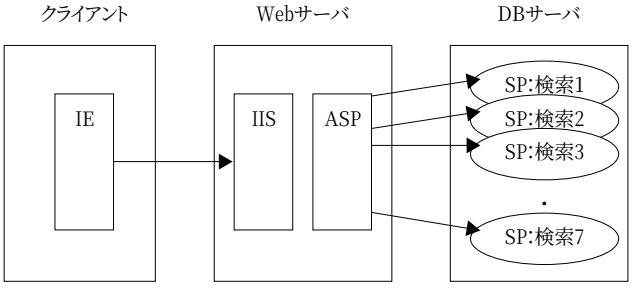


図 9　DWH の検索処理

2）　データベース層の処理

データベースは 62 個のロジカルユニット上に 152 個のテーブルが作成される．

ドライブレターが 22 個（A–D は使用済）しか使用できないため，Windows 2000 の新機能である「Mount 機能」を使って一つのドライブ文字に複数のロジカルユニットを所属させ，各ロジカルユニットは独立して参照できるように構成している．初期のデータベースのデータ量は 3 TB で，開催中毎日 OLTP システムから 200〜300 GB の注文データを DTS と同様なユーティリティを使って拡張した．なお，6 TB のデータベース構造の作成に 13 時間，約 3 TB のデータの生成に 78 時間かかった．

4.3 運 用 管 理

1） バックアップ

大規模 DWH の運用で一番の難題はバックアップ方法の確立である．この DWH システムのデータベースは 7 TB（7000 GB）の大きさがあるので，仮に毎時 70 GB の速度でテープ装置にフルバックアップしたとしても 100 時間かかる計算になり，オンライン処理への影響が甚大となる．今回は 1 時間に 1 TB のバックアップ能力がある EMC 社のバックアップソリューション（図 10）を使用したことで，オンライン処理を停止することなくバックアップを取ることが可能になった．

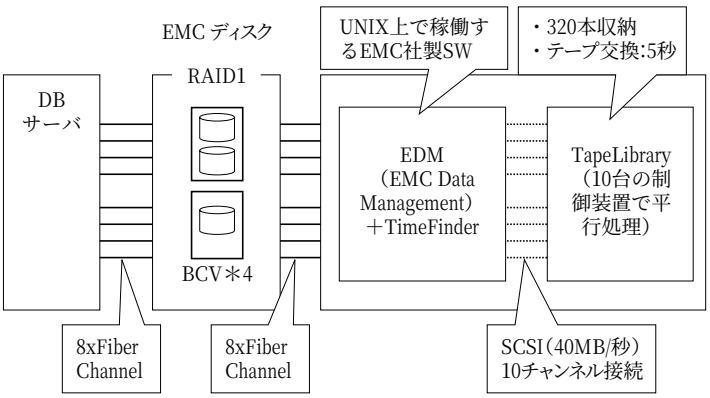


図 10　DWH のバックアップ処理構成

4.4 DWH システム構築時の考慮点

1） 大量ディスク構成での MSCS 環境構築

Windows 2000 の「Mount 機能」を使って 62 個のロジカルユニット上に 152 個のテーブルを作成したが，DB サーバの障害テストでフェールオーバ時に「Mount 情報」が引き継がれないことが判明したため，Cluster-Resource-Kits を使って情報を引き継げるように対応した．しかし，小規模ディスク構成では成功したが，実際に 397 個のディスクを使って実施すると，原因不明のディスクエラーが発生し，許された時間内で解決する事ができなかった．この DWH システムでは特に I/O 効率が重要視されため，データベース層のクラスタリング構成（MSCS）構築を断念した．

大規模ディスク構成のクラスタリング機能は，日本ユニシスで評価・確認中である．

5. お わ り に

今回ユニシスはCOMDEX '99に「次世代のデータセンター」を出展し，Windows 2000を基盤とするミッションクリティカルな大規模システムが充分に構築可能であることを実証した．但し，時期的な関係もありWindows 2000はベータ版を利用せざるを得ず，予定した機能を総て使用することは出来なかった．今後は正式リリース版のWindows 2000新機能を有効に活用することでWindows 2000を基盤とした大規模システム構築はもっと実現しやすくなるであろう．

また，日本ユニシスが2000年3月8日に発表したUnisys e-@ction Enterprise Server ES7000は，ミッションクリティカルな基幹システムを構築するために必要とされる信頼性，可用性，拡張性等のメインフレーム属性を兼ね備えている．間もなく正式リリースされるWindows 2000 Datacenter Serverと組み合わせることで，メインフレーム並みの安定したミッションクリティカルな大規模システムが容易に構築できると確信している．

＊1 WWENTCOE（World Wide Enterprise NT Center of Excellence）ユニシス社のエンタープライズサーバ向けシステム構築を専門に手がける世界規模の組織
＊2 TB（Terabytes）ディスク容量を表現する単位．1 TB＝10^3 GB＝10^6 MB＝10^9 KB＝10^{12} Bytes
＊3 IIS（Internet Information Service）インターネットサーバとして機能させるためのサービス
＊4 ISAPI（Internet Server Application Programming Interface）インターネットサーバ用の関数セット．アプリケーションは，VC++でISAPIを使って開発する
＊5 ASP（Active Server Page）IISに搭載されるサーバ側のスクリプトエンジン．VBScriptやJavaScriptを実行可能
＊6 COM+（Component Object Model）Windows 2000からサポートされた仕様で，COMを効率的に作成でき，また簡単に利用できる様にした技術
＊7 DTS（Data Transformation Service）SQLServer 7.0のデータ変換サービスツール
＊8 BCV（Business Continuity Volumes）ミラーリングされた本番用ボリュームに加え，バックアップ用に第3のミラーボリューム（BCV）を装備できる．BCVボリュームは専用ソフトウェア（TimeFinder）を使って本番用ボリュームに影響なく切り離し，再接続が可能
＊9 NLB（Network Load Balance）WindowsNT 4.0からの機能で，最大32台のネットワークサーバでクラスタを構成し，クラスタへの全入力を構成サーバで負荷分散する．また，クラスタを構成するサーバが停止しても残りのサーバでクラスタを再構成出来るので，可用性に優れている
＊10 CLB（Component+load Balance）Windows 2000 OSの機能で，最大8台のアプリケーションサーバでクラスタを構成し，クラスタへの全入力を構成サーバで負荷分散する．また，クラスタを構成するサーバが停止しても残りのサーバでクラスタを再構成出来るので，可用性に優れている
＊11 ロジカルユニット（Logical Unit）ディスク装置内の複数の物理ディスクドライブを一つにまとめた論理的な装置．WindowsOSはこの論理装置を1台のディスクドライブとして扱う
＊12 DTC（Distributed Transaction Coordinator）COM+，MSMQ，SQLServerを使って分散トランザクションを実行する時のトランザクションマネージャー

附表 1　COMDEX '99 に出展した「次世代のデータセンター」に使用した製品一覧

ハードウェア製品

装置	出展システム	
サーバ（Unisys 製品）	UnisysES 5000 シリーズ 8 Way	16 台（LAN-Card は 2 枚 or 3 枚），DB サーバは EMC ディスク装置（Fiber Channel）と最高 8-Channel Module（Qlogic 社製）で接続されている
	UnisysES 5000 シリーズ 4 Way	40 台（Web サーバ，Load-Runner 用，System-Manager に使用）
	ES 5000 サーバを 4 台単位でラックに設置．メモリーは 8 GB.OS は Windows 2000 Advanced Server RC 2（Release Candidate 2）を使用.	
ディスク装置（EMC 社製品）	OLTP	SYMMETRIX 3830×1.　50 GB×96＝4.8 TB（1.6 TB×3）
	DWH	SYMMETRIX 3830×1.　50 GB×396＝19.8 TB（6.4 TB×3）
	AD（Acitve Directory）	SYMMETRIX 3830×1.　36 GB×96＝3.456 TB
ネットワーク制御（Cisco 社製品）	Catalyst 6500（24 ポート）×2．全体の NW,VLAN（Virtual-LAN）に使用.	
データ転送（Giganet 社製品）	GigaNet（8 Ports）×12（予備 2 台を含む）．サーバ間のデータ転送，DB サーバの MSCS クラスタリングのホスト間専用通信装置として使用.	
バックアップ/リストア（EMC 社＋Storage Tek 社製品）	EMC 社の EDM（EMC Data Manager）ソリューションの構成装置として使用．StorageTek 社の Tape Library 装置（9740 LibraryStorageModule）は 320 本のカートリッジテープ（Imation 社製）を収納.	

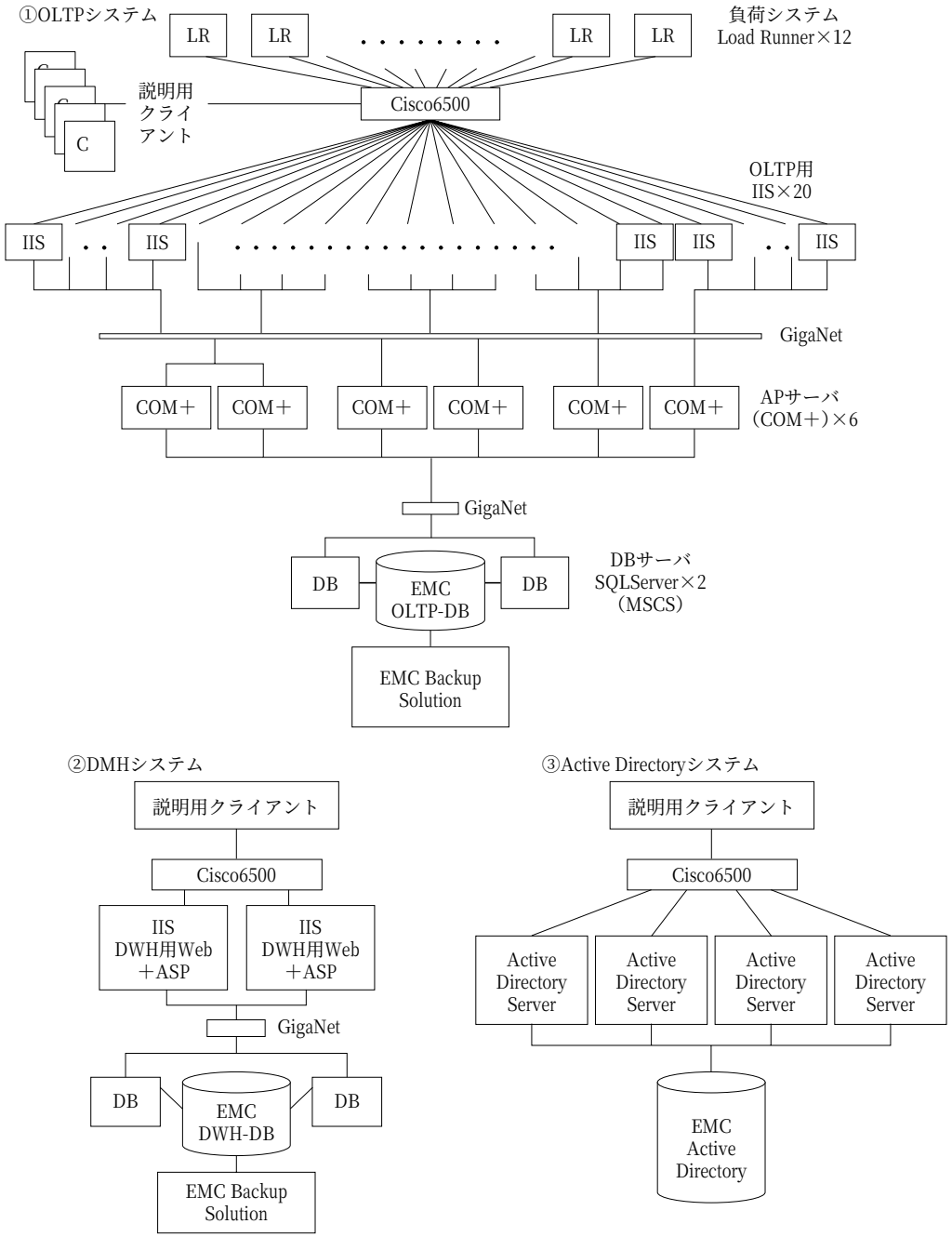
HW スペック　CPU : Pentium 550 MH,　メモリー：8 GB　HDD : 18 GB×2（OS 用ミラー構成）SCSI : UltraSCSI

ソフトウェア製品一覧

製品	出展システム
MS 社	Windows 2000 AdvancedServer（RC 2），SQLServer 7.0, DB 移行ツール（DTS 高速版）
NetIQ 社	AppManager Suite
Mercury Interactive 社	LoadRunnar 6.0（β 版）
EMC 社	Time Finder, EDM（EMC Data Manager）
Identcator 社	Identcator

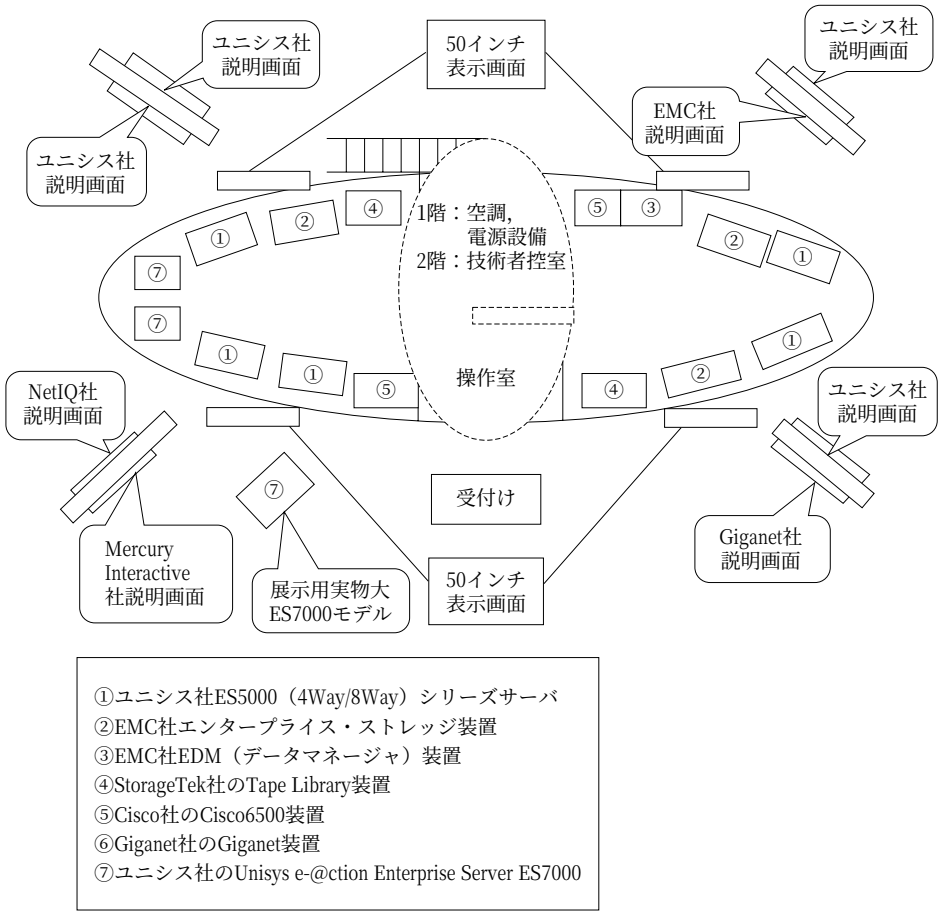
サーバ台数

サーバの種類	台数	CPU	Mem	Cluster	OS	コメント
DB サーバ	4 台	8×	8 GB	MSCS	AS	OLTP のみ使用
Load Runner 用サーバ	12 台	4×	8 GB		AS	負荷システム
Web サーバ（IIS）	20 台	4×	8 GB		AS	
AP サーバ（COM＋）	6 台	8×	8 GB		AS	
Terminal サーバ	2 台	8×	8 GB		AS	
Active Directory 用	4 台	8×	8 GB		AS	
System Manager 用	2 台	4×	8 GB		AS	
Web サーバ （Internal）	2 台	4×	8 GB		AS	COMDEX 会場内のクライアント用
予備サーバ	4 台	4×	8 GB		AS	
合計	56 台					



附図 1　COMDEX '99 に出展した「次世代のデータセンター」のハードウェア構成図

建物の周りはガラス張り（グラスハウス）となっており，見学者が製品を見学できる．



附図 2　COMDEX '99 に出展したユニシス社のブース全体図

執筆者紹介　馬　場　功　二（Koji Baba）

1969 年日本ユニシスに入社．メインフレーム客先でのシステムサービスを経験後に SW 主管部にてオペレーティングシステムのサポートサービスに従事．その後オープン系のシステムサービス業務に転じる．現在，W 2 K テクノロジーセンターに所属．