

Windows NT における TP モニタ BeTRAN の実装と評価¹

5Q-3

佐藤 将夫

新田 淳

(株)日立製作所 システム開発研究所

1 始めに

TP モニタ BeTRAN は、UNIX マシン上で Online Transaction Processing System を構築するためのミドルレイヤソフトウェアである。TP モニタは、Online Transaction Processing System に必要な機能を実現するため、OS が提供するプロセス管理、プロセス間通信、ファイル管理など数多くの機能を使用する。本稿では、UNIX と Windows NT という全く異なるパラダイムに基づいて開発された2種類の OS について提供機能の差異をまとめた上で、TP モニタ BeTRAN を例に OS の違いを吸収するための実装方式を説明し、Windows NT 上に構築した BeTRAN プロトタイプについて UNIX 版 BeTRAN との比較および評価報告を行なう。²

2 Win32 API

Windows NT は複数のサブシステムを持ち、それぞれのサブシステムが独自の API を提供する。Windows NT の最も主要なサブシステムは Win32 であり、他のサブシステムでは Windows NT の提供する機能を部分的にしか使用できない。表1に各サブシステムの提供機能を比較する。

表1 サブシステム機能比較

項目	Win32	POSIX	Win+Win32s
プロセス	あり	あり	なし
スレッド	あり	なし	なし
ソケット	WinSock	なし	WinSock
セマフォ	あり	なし	なし
Win イベント	あり	なし	あり
共有メモリ	あり	なし	あり

2.1 Win32 API の限界

Win32 API のみでは、Windows NT が提供する全機能を使用できない。OS/2 や POSIX サブシステム特有の機能を Win32 API で実現できないのは当然であるが、その他にも、一部、Win32 API で提供されておらずアプリケーションの作成が困難な機能が存在する。

2.1.1 Windows NT への login

ftp, telnet などのソフトは、指定されたユーザ名称で OS に login しプロセス / スレッドにそのユーザの持つ権限 (uid) を付与する API が必要である。Windows NT にも FTP などのサーバが存在するが、本機能を提供する API は提供されていない。

2.1.2 non-preemptive スレッド

スレッドは、non-preemptive なスレッドと preemptive なスレッドに分類できる。non-preemptive なスレッドは、ユー

ザレベルの context switch が可能なため、ディスパッチタイミングをユーザプログラムで制御可能であり高速な動作ができる。Windows NT は、16bit Windows サブシステムでのみ non-preemptive スレッドを提供する。16bit Windows サブシステムは、Win32 サブシステムの特別な形態として実現されているが、ベースの Win32 サブシステムでは、non-preemptive スレッドを使用できない。

2.1.3 POSIX, OS/2 サブシステムとの通信

Windows NT の Win32, POSIX, OS/2 各サブシステムは、それぞれが独自のプロセス間通信機能を提供する。

Windows NT は、これらのプロセス間通信を下位レイヤで一元的に LRPC として実装しているが、どのサブシステムも、LRPC を行なう API を提供していない。このため、各サブシステム間を跨る通信は、ファイル I/O やプロセスの起動など低速な機能に頼らざるを得ない。

2.1.4 ソケットを含むマルチプルウェイト

一般にサーバを構築する場合、プロセス外からの各種イベント (セマフォ、シグナル、ファイル I/O、パイプ、ソケット、..) を同一の API でマルチプルウェイトでできることが重要である。特に、UNIX 環境との通信を行なうアプリケーションではソケットと他のイベントのマルチプルウェイトが重要になるが、Windows NT のソケットライブラリ WinSock は、ソケット以外のイベントとのマルチプルウェイトを提供しない (表2)。

表2 マルチプルウェイト

関数名	semaphore	signal	pipe	socket
WaitForMul.. ³	○	○	○	×
select	×	×	×	○
select (UNIX)	×	○	○	○

3 BeTRAN の実装と評価

Windows NT 上に TP モニタ BeTRAN のプロトタイプを作成し、TPC-C をベースの簡易モデルを使用した性能測定と評価を行なった。

3.1 NT 版 BeTRAN の構造

図1に Windows NT 版 BeTRAN プロトタイプのレイヤ構造を示す。Windows NT 版 BeTRAN プロトタイプは、Win32 API を用いて作成され、独自の non-preemptive スレッドを持ち、WinSock ライブラリを使用した BeTRAN RPC Service により UNIX 版 BeTRAN と通信を行う。

BeTRAN は、各 UNIX OS 間の仕様差異を吸収し、上位に仮想的な OS を提供する Kernel Service と Thread Service を持つ。Windows NT 版プロトタイプでは、この両 Service の下に、UNIX から Win32 API への関数セマンティクスの変換を目的とした Win32 Connector を設け、上位層に与える影響を最小限にとどめた。

¹ "TP-monitor ProtoType BeTRAN on Windows NT"

Masao SATO and Jun NITTA

Systems Development Laboratory, Hitachi, Ltd.

² 各名称は各社の登録商標または商標です

³ WaitForMultipleObject 関数

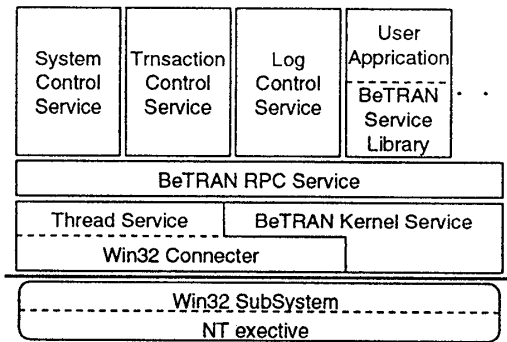


図1 BeTRAN レイヤ構造

3.2 Win32 Connector

3.2.1 スレッド

上述したように Win32 API が non-preemptive スレッドを提供しないため、独自の non-preemptive スレッドライブラリを作成した。本ライブラリは、`alloca()`、`setjmp()`、`longjmp()` 関数で context switch を行ない、次の特徴を持つ。

- 100% C 言語、UNIX 環境とのソース互換⁴
- POSIX pthread Subset+^a

3.2.2 マルチプルウェイト

non-preemptive スレッドのコントローラは、context switch の契機となるイベント全てをマルチプルウェイトする。UNIX 版 BeTRAN は、`select()` 関数を使用して socket, signal, パイプなどをマルチプルウェイトする。一方、WinSock ライブラリの `select()` 関数は、socket 以外のイベントとのマルチプルウェイトを提供しない。このため、各イベントを専用に待つ preemptive スレッドを生成し、これらのスレッドから `select()` スレッドへの連絡を行ない、UNIX と同等のマルチプルウェイトを実現した(図2)。

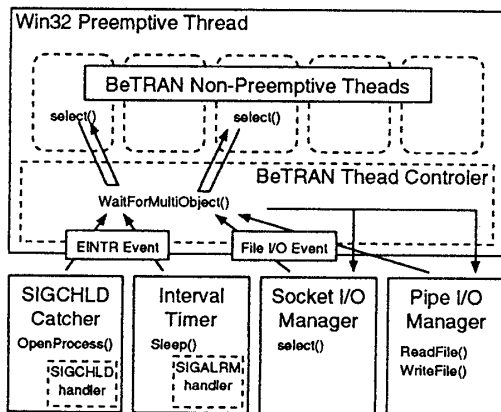


図2 select の拡張

3.2.3 メッセージキュー

Windows NT は、UNIX のメッセージキューに相当するプロセス間通信手段(複数のプロセスで共有可能でリアルな通信手段)を持たない。このため、メッセージキュー相当の機能を共有メモリおよびセマフォを使用して作成した。セマ

⁴OS, コンパイラの仕様により一部アセンブラが必要となる。

フォは、メッセージ待ち状態にある複数のプロセスに対し、メッセージの到着を知らせるために利用される。

3.3 性能評価

Windows/NT 版 BeTRAN プロトタイプについて簡単な性能評価を行なった。測定プログラムは、TPC-C ベンチマーク以下の変更を加え簡略化したものである。表3に測定に用いたプラットフォームを示す。

- transaction を NEW ORDER のみに限定
- DATA 規模 1/10
- transactional file system⁵上に DATA BASE 構築
- スケーリング規則適用せず

表3 測定プラットフォーム

OS	CPU	Memory	Disk
Windows NT	486DX2(66Mz)	96M	SCSI
HI-UX/WE2	PA-RISC(80Mz)	80M	SCSI

表4に測定結果を示す。NT 版プロトタイプの性能は UNIX 版の 2/5 程度であるが、これはハードウェアの性能差が原因と考えられる。ハードウェアの価格比が性能比以上あるため実用上十分な性能であると評価できる。今回の測定は TPC-C に完全準拠していず、他 TP モニタとの正確な比較を行なえないが、TP モニタ・測定プログラム共に性能チューンナップ前であることを考えると、NT 版 BeTRAN プロトタイプはオフコン程度の実力を持つと推測できる。今後は、TPC-C 準拠のベンチマーク測定、同一ハードウェア、異 OS 状態での性能測定を予定している。

表4 測定結果

	TPM
NT 版 BeTRAN プロトタイプ	80
UNIX 版 BeTRAN	200

4 おわりに

Windows NT 上に TP モニタ BeTRAN のプロトタイプを作成し、実務レベルの transaction 性能を持つことを確認した。Windows NT は、統一された内部アーキテクチャを持つが、サブシステム間の通信やマルチプルウェイトなどについて、サーバシステムの構築に十分といえる API を提供していない。今後、これらの機能が拡張サポートされることを望む。

参考文献

- [1] Helen Custer, "INSIDE WINDOWS NT", Microsoft PRESS, 1993
- [2] "Win32 Programmer's Reference", Microsoft PRESS, 1993
- [3] W. Richard Stevens, "UNIX NETWORK PROGRAMMING", PRENTICE HALL, 1990
- [4] "TPC BENCHMARK C", Transaction Processing Performance Council, 1992

⁵BeTRAN/FS/DAM