

# OZ: 対象指向開放型分散システムアーキテクチャ

6H-3

--- コネクションレス型通信アーキテクチャの実装と評価 ---

塚本享治 (電総研)

吉江信夫 (住友電工)

近藤貴士 (シャープ)

中込昌吾 (ABC)

田中伸明 (松下電器)

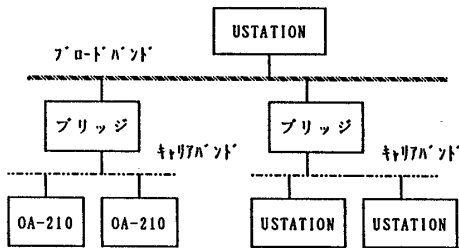
## 1. はじめに

近年、分散システムの研究開発が盛んに行われているが、筆者らはオブジェクト指向の分散システムOZを提案し、現在、そのプロトタイプの実証評価を行っている。分散システム下の通信は、単純な手順による高速で基本的なデータ転送サービスを土台に、分散OSと密に融合することで、高度な分散環境の構築を可能にする。OZでは、OSI標準の枠の中かで、分散システムに適していると思われるコネクションレス型通信サービスを全層に採用し、通信をモデル化した。[1], [2]

本稿では、OZのコネクションレス型通信アーキテクチャの実装とその評価について報告する。

## 2. 実装

実装には、共同研究各社より提供されたUNIX系ワークステーションをベースにプロトタイプとしての機能の実現を重視した。本評価に用いたネットワークは、図1に示す様にブロードバンドとキャリアバンドをブリッジで接続した構成をとっている。



2. 1 エンドシステム

図1. ネットワーク構成

実装上の問題点と対応を以下に述べる。

### ① エンドシステムの構成

通常、ワークステーションでは、フロント・エンド・デバイス(通信ボード)の形で下位層の通信機能を提供している。ここでは、通信ボードにどの層までを実装するかということが一つの問題となる。既存のOSIボードは、MAC副層からセッション層までを様々な形態で実装している。この実装を決定づける要因としては、使用目的、効率、ボードとメインCPUボード(ホスト)間のインタフェース、各層の機能等が上げられる。OZで使用するLLCタイプ3は、MACの即時応答機能を使用する関係上、メディア(MAC)への依存度が高く、ボードへの実装が適する。また、CLTP、CLSPは、実質的には各セクタでなにを区別するかにより、実装位置を決定できる。

今回、OZでは、図2に示す様に通信ボードには、最小構成のLLCまでを実装した。将来は、トランスポートセクタをマルチCPUボードへの対応や、コネクション型との識別に使用する予定でありCLTPまでのボード内実装を検討している。また、3層以上のプロトコルは、すべてホスト側のドライバ内

に実装しており、セッションセクタで、プロセス(パッケージ)を選択している。

### ② サービス・パラメタ

コネクションレス通信では、アドレスがサービスの主パラメタとなる。アドレス長は、標準の上では200バイト以上になる可能性があり、その扱いは実装の1つのポイントになる。OZでは、アドレスをホスト内のDI B (Dictionary Information Base) に一括保存しており、送信時には、そのエントリ

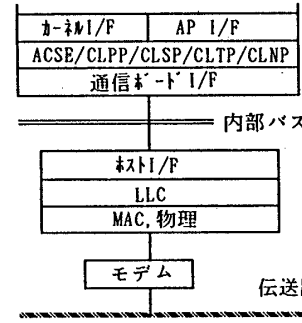


図2. エンドシステム構成

へのポイントだけが下位層へと渡される。受信時には、アドレスエリアが確保され、そこへのポイントが上位層へと渡される。

### ③ ホスト・ボード間インタフェース

ホストと通信ボード間では、ボード上のデュアル・ポート・メモリを介してパラメタやバッファが渡される。この間のパラメタの受渡しはシェイクハンドで行っており、1要求に2回の割り込みが発生する。また、バッファを送信用、受信用で区別することにより、ホストとボード間でのバッファ管理(排他制御)を容易にするとともに、バッファの独占を防止している。

### 2. 3 ブリッジ

ブリッジは、MACフレームの中継という基本機能のほかに、付加的な機能として、中継フレームのフィルタリングとそのデータベースの学習、管理情報の設定と獲得、フレーム滞留時間の監視等を実装している。今回の実装では、汎用WSを用い、モデルを忠実に実現し、特に管理機能を重視した。

## 3. 性能評価

各々の転送要求が独立に扱われ、単一のPDUでデータが転送されるコネクションレス型通信では、転送に要する時間は、(1)式の様に表示する。

$$\text{Time}(x, \alpha) = ax + b + \alpha \quad \text{---(1)}$$

a: 単位データ(1バイト)の実質的な転送に必要な時間

x: データ長

b: データ長に依存しない転送処理時間(定数)

α: データ長に依存しないモジュール間の遅延時間

ここでいうモジュールとは、入出力機構をもち、各々独立で動作するオブジェクトをさす。本システムは、図3に示す4つの

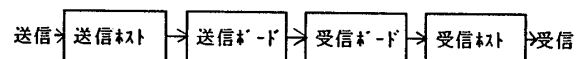


図3. 転送モジュール構成

OZ: Object Oriented Open Distributed System Architecture

-- Implementation and Evaluation of Connectionless-mode Communication Architecture --

Michiharu TUKAMOTO<sup>(1)</sup>, Nobuo YOSHIE<sup>(2)</sup>, Takashi KONDO<sup>(3)</sup>, Shogo NAKAGOME<sup>(4)</sup>, Nobuaki TANAKA<sup>(5)</sup>

(1)Electrotechnical Laboratory (2)SUMITOMO Electric Industries, Ltd. (3)SHARP Corporation

(4)ABC Co., Ltd. (5)MATSUSHITA Electric Industrial Co., Ltd.

モジュールから成る。LLCタイプ3を使用した場合のモジュール間の転送は、すべてシェークハンドで行われる。

評価テストでは、データ長を変数として、エンドシステム (USTATION) 間の一方向転送テスト、対向エコーバックテストと送信モジュールでの折り返しテストを、他のトラフィックがない理想的な状態で行った。この評価テストの目的は、各部の処理時間を割り出し、現状を把握し、効率向上への改善を行うことである。本条件下では、 $\alpha$  は0になると考えらる。また、OZではLLCにタイプ3を使用しているが、一般的なタイプ1を使用した場合も測定し、比較検討を行った。

3. 1 評価

①セグメント内転送

図3における送信から受信までの所用時間の測定値を図4に示す。

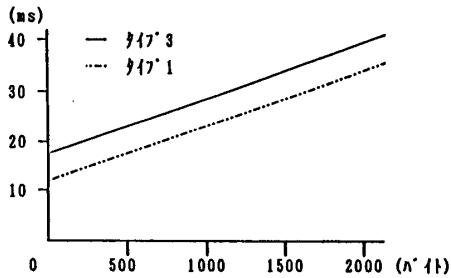


図4. セグメント内転送時間

この結果から与えられるタイプ1とタイプ3の各々の評価式は以下の通りである。

$$Time_1(x) = 0.012x + 13 \quad \text{---(2)}$$

$$Time_3(x) = 0.012x + 18 \quad \text{---(3)}$$

(単位--x : バイト、時間 : ms)

タイプ3は、タイプ1より5msのオーバーヘッドをもつが、これはLLCタイプ3のプロトコル自体のオーバーヘッド以外にMACの即時応答をソフトウェアで行っていることや、同一相手への連続送信を可能にする為に自LSAPの選択を行っていることによる。

また、送信元が送達確認を受けるまでの評価式は、以下のよう求められる。

$$Time_{ack}(x) = 0.007x + 19.5 \quad \text{---(4)}$$

評価テスト時には、伝送路上のエラーやLLCタイプ3の再送は発生しておらず、これらの評価式は $\alpha = 0$ の場合の(1)式に合致する。

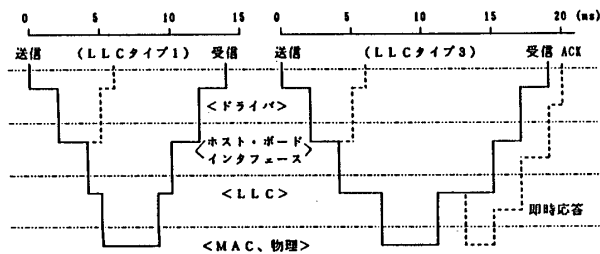


図5. セグメント内転送タイムチャート(データ長:1000バイト)

図5にデータ長1000バイト(PCI:80バイト)の場合のタイムチャートを示す。送信ホストモジュールの評価式は(5)式の様になり、このモジュールが全体のボトルネックになっている。

$$Time_{sn}(x) = 0.005x + 4.0 \quad \text{---(5)}$$

これより、最大スループット値は約1.5Mbpsとなる。

また、LLCタイプ3の送信側は、(4)式と(5)式の差の間、送達確認を待つことになり、大容量データ転送の場合、この間の連続送信は効率上、有効である。連続送信を行うとフロー制御が問題となるが、OZでは、ACSEの上位プロトコルRBTHで実装している。

②セグメント間転送(ブリッジ経由の転送)

LLCにタイプ3を用い、データ長1000バイトの転送を行った時のタイムチャートを図6に示す。ブリッジの滞留時間は約20msになる。ここでのボトルネックは、中継部のオーバーヘッド(10ms)にある。今回の実装では、UNIXのアプリケーションプロセスでフレームの中継やフィルタリングを行っている為、オーバーヘッドが大きく、これらの中継処理をドライバに実装することで、かなり効率向上が図れると思われる。ブリッジの評価式は、

$$Time_b(x) = 0.018x + 12 \quad \text{---(6)}$$

最大スループットは、約430kbpsになる。

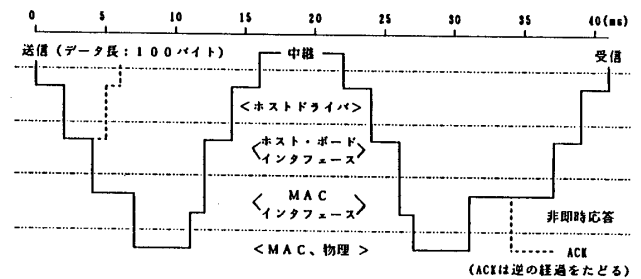


図6. セグメント間転送タイムチャート(LLCタイプ3)

4. 改善方法

オーバーヘッドは、プロトコル自体よりも、むしろ、割り込み処理等のモジュール間インタフェースの処理やバッファ間のデータ転送にある。このことから、次のような改善が考えられる。

- ・同期処理をなくし、割り込み回数を削減することで、ホストとボード間のインタフェースの効率化を図る。
- ・バッファ間コピーの比重が大きい為、コピーを高速化し、回数を減らす。現在、通信ボード内では、メディアとの送受信の為に、内部バッファを使用している。
- ・MACの即時応答は受信側のソフトウェアで起動されているが、これをLSI(トークンバス・コントローラ)で行う。

5. むすび

今回の評価は、理想的な条件下で各部の処理時間に重点をおいて行った。今後は、この評価を踏まえ、効率の改善を図るとともに、 $\alpha > 0$ の通常のN対N通信や、高負荷時の通信の評価を行い、より実環境に即した対応を行う予定である。以上、OZのコネクションレス型通信アーキテクチャの実装と評価について報告した。

参考文献

[1]塚本、吉江、舟渡、四反田、「OZ:対象指向型分散システムアーキテクチャ--LLC3ベースのLANアーキテクチャ」、62年度後期情報処理全国大会  
 [2]塚本、吉江、舟渡、四反田、「OZ:対象指向型分散システムアーキテクチャ--通信系」、63年度前期情報処理全国大会