

3G-8 OZ+ : オブジェクト指向開放型分散システム¹ — OZ+ 通信系の拡張 —

塚本 亨治 (電総研)² 水谷 功 (住友電工)³ 梶浦 広行 (シャープ)⁴ 篠原 弘樹 (松下電器)⁵

1 はじめに

現在、OZを拡張・改良して、複数のサイトにまたがって複数のユーザがオブジェクト交換しながら処理を進めることのできるOZ+の開発を進めている[1][2]。OZ+においては、キャッシュデータの更新、オブジェクトの所在検索、仮想計算機(VM)の所在検索、VMの管理、情報収集など、複数オブジェクトに対する処理が必要である。これを効率よく行うためにはマルチキャスト機能が必要となる。本稿では、これまでに拡張・実装を行ったOZ+通信系のシステム構成、及び、信頼性のあるマルチキャスト通信とそのアドレス管理について述べる。

2 通信系のシステム構成

2.1 ハードウェア構成

OZのネットワークはブロードバンドLAN、キャリアバンドLAN(IEEE802.4)をMACブリッジで接続したものであった。OZ+では、これにさらにFDDI-LAN及び、CSMA/CD(IEEE802.3)-LANをブリッジやルータにより接続しネットワークを拡張した。図1にOZ+通信系のネットワーク構成を示す。

2.2 CLNP、ES-ISの実装

OZのネットワーク層にはCLNP(ISO8473)のインアクティブサブセットを使用していたが、OZ+ではCLNPのフルセットを実装した。また、ネットワークアドレスの管理としてES-IS(ISO9542)を実装した。

2.3 CSMA/CD、FDDIへの対応

既存のネットワーク(イーサネットメディア)を活用した利用形態としてCSMA/CD(IEEE802.3)の実装を試みた。CSMA/CDを実装することにより、市販のFDDI-Ethernetブリッジを用いてFDDI上のOZ+エンドシステムとの通信をすることにした。図2にOZ+通信系のエンドシステム構成を示す。

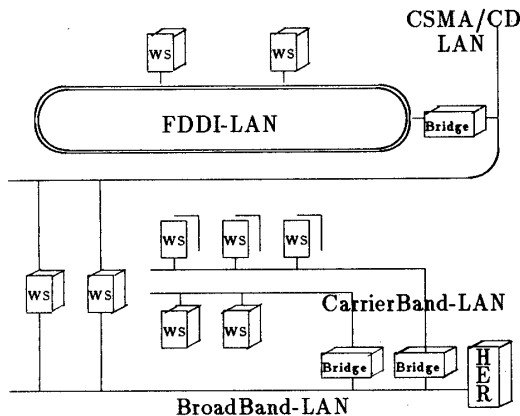


図1 OZ+通信系のネットワーク構成

2.4 RBTの拡張

OZでは、通信し合うオブジェクトの数が非常に多く、しかも参照関係が逐次変化する。それに対処するためOZではACSE以下全層にコネクションレス型を採用し、その信頼性を向上させるためにACSEの上位層としてバルクデータ転送プロトコル(RBT)を実装した[3]。OZ+ではRBTをさらに拡張してマルチキャスト機能を設けた。

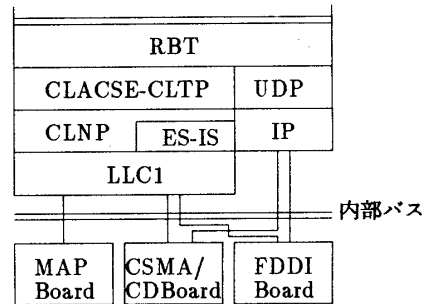


図2 OZ+通信系のエンドシステム構成

3 RBTマルチキャスト

RBTマルチキャストに要求される機能要件を以下に挙げる。

- 信頼性のあるマルチキャスト
- セグメントを意識しないマルチキャスト
- ルータの拡張は行わない

3.1 RBTユーザインタフェース

3.1.1 アドレス情報(VMID)

RBTユーザはVMIDを宛先アドレスとしてRBTに送信を依頼する。マルチキャストを実現する場合、論理的なグループ識別子を使用し通信するのが一般的である。OZ+では、下記の理由によりグループ識別子をVMIDのリスト型で実現した。

1. OZ+アプリケーションはオブジェクトをアドレス情報として通信を分散カーネルに要求する。オブジェクトはVM間を自由に(RBTの管理外で)移動できるため、アプリケーションがマルチキャストを要求する範囲(グループ)のアドレス情報(VMID-P-Address)を、RBTがオブジェクトの移動、生成、消滅ごとに管理、変更することは困難である。
2. OZ+ではVMが動的に生成、消滅、または、移動を起すため、VMをグループ化し、管理することが難しい。

3.1.2 タイプ、及び結果通知

以下の2つの型を用意した。

- ALLRECEIVED型:
RBTは、マルチキャストした全パケットの送達確認を受けると、その結果(送達確認を受信した数)をRBTユーザ(以降ユーザ)に通知する。何らかの異常で送達確認が揃わない場合、ユーザまたはシステムが設定した有効時間が経過した時、ユーザに受信した送達確認の数を通知する。

¹OZ+ : Object Oriented Open Distributed System — Extension of OZ+ Communication Architecture—

²Michiharu TUKAMOTO (Electrotechnical Laboratory)

³Isao MIZUTANI (SUMITOMO Electric Industries, Ltd.)

⁴Hiroyuki KAZIURA (SHARP Corporation)

⁵Hiroki SHINOHARA (MATSUSHITA Electric Industrial Co., Ltd.)

- ONERECEIVED 型:

RBT、はマルチキャストに対する最初の送達確認をユーザに通知する。その後受信した送達確認はRBT内で廃棄する。また、再送中の仮想コネクションがあれば中断する。

3.2 シーケンス

ブロードキャストメカニズムだけでは上記の機能要件は満足させられない。そこでRBTで高信頼、しかも効率のよいマルチキャストを実現することにし(図3)、宛先VMIDを宛先エンドシステム(P_Address)ごとにまとめ、ホスト単位で同報する方式をとった。パケットを受信した各エンドシステムのRBTは各VMIDをセクタとして各ユーザに通知する。また、必要に応じ、送信元に対し送達確認、または、エラーを応答する。送達確認はセクタで識別されるユーザごとにあるが、RBTはそれらの送達確認を1つの送達確認パケットにまとめとして送信元へ戻す。ユーザデータが大容量の場合には、RBTで分割して送信する。RBTでは分割されたパケットのシーケンス番号を管理しており、送達確認時には受信したパケットのシーケンス番号を通知する。もし、マルチキャストを行ったパケットの中で到達しなかったパケットが検出された場合、到達しなかった宛先RBTに対して選択的再送を行う。

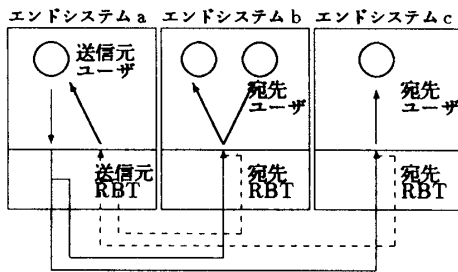


図3 RBT マルチキャストの動き

4 アドレス管理

4.1 RBTによるアドレステーブルの更新

RBTはVMIDとプレゼンテーションアドレスとのアドレステーブルを保持している。RBTは、アドレステーブルを参照し、P_Addressを獲得したのち、パケットを送信する。P_Addressが見つからない場合、自局管理のアドレス検索機能(局管理[4]はサイト内の他のすべての局管理に対し、そのVMIDに対応するP_Address獲得要求をマルチキャストする)を用いてP_Addressを獲得し、テーブルを更新した後、パケットを送信する。一方、RBTはパケットを受信した際には、送信元VMがアドレステーブルに登録されていなければ、送信元のアドレス情報(VMID、P_Address)をアドレステーブルに登録する。

4.2 局管理によるアドレステーブルの更新

局管理がVMを削除する時、サイト内のすべての局管理に対しアドレステーブルからそのVMに関する情報削除を要求する。また、局管理は一定時間利用されないアドレス情報を削除する。

5 サイト間通信

サイトとは1つのシステム管理が管理している論理的な領域である。アドレス情報の管理は各サイト内で行われる。したがって、サイトをまたがって通信を行う場合、何らかの手段で相手先のアドレス情報を獲得しなければならない。

5.1 中継RBT

送信元RBTが他サイトにある宛先VMのアドレス情報を持たないことを検出すると、宛先サイトの中継RBT(Network ManagerのあるエンドシステムのRBT)に転送する。中継RBTは宛先VMIDが、自ホスト内のVM宛ならばRBTユーザに通知するが、フォワードすべきものならば最終宛先エンドシステムに転送する。

5.2 送達確認

送達確認は要求元RBTに対して直接送信する。送達確認を受信したRBTはアドレステーブルを更新する。その後の通信は新しいアドレスを用い直接通行される。

5.3 送信エラー

宛先VMIDのアドレス情報を持つ場合には、送信元RBTは直接宛先エンドシステムのRBTに対し通信を行う。パケットを受信したRBTはその上位に宛先VMがない場合(移動、消滅など)通信エラーを戻す。送信元RBTは再度、宛先サイトの中継RBTに対して通信を試みる。その通信にも失敗すると、RBTは宛先VMとの通信を不能だと判断し通信不能エラーを自ユーザに通知する。

5.4 サイト間マルチキャスト

サイトをまたがってマルチキャストを行うときの動きを図4に示す。送信要求中に他サイトにあるVMが検出された場合には、RBTはサイトごとに宛先VMをまとめ宛先サイトの中継RBTに対して転送を行う。パケットを受信した中継RBTはそれを各エンドシステムごとに分配してそれぞれの宛先RBTにフォワードする。送達確認は3.2と同様、エンドシステムごとにまとめて送信する。

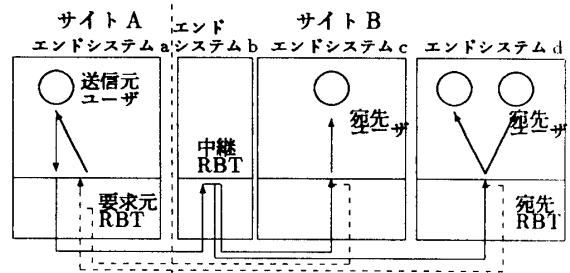


図4 サイト間マルチキャスト通信

6 おわりに

OZ+通信系の実装及び、拡張したマルチキャストプロトコルについて述べた。今後は、5章で述べたサイトの考え方を利用し、WANに対応するシステムに拡張していく予定である。なお、本研究は通産省型プロジェクト「電子計算機相互運用データベースシステム」の一環として行われている。

参考文献

- [1] M. TSUKAMOTO et al., The Architecture of Object-Oriented Open Distributed System: OZ, Interoperable Information Systems ISHS '88, Ohmsha, PP153-166 (1988,11)
- [2] 塚本他, OZ+: オブジェクト指向開放型分散システム - オブジェクトの分散管理 -, 情報処理学会 1990年代の分散処理シンポジウム論文集, PP57-64 (1990,11)
- [3] 塚本他, OZ: オブジェクト指向開放型分散システムアーキテクチャ - LLC 3タイプ3を活用する通信アーキテクチャ、実装、およびその評価 -, 情報処理学会 マルチメディア通信と分散処理研究会, 43-4 (1989,9)
- [4] 塚本他, OZ: オブジェクト指向開放型分散システムアーキテクチャ - OZ+システム管理の基本設計 -, 情報処理学会 第41回全国大会, 6Q-3 (1990,9)