

# 通信網ワイド分散処理プラットフォーム PLATINA における アウトラインメッセージ通信機構

2T-6

田中 聡 丸山 勝巳

NTT ネットワークサービスシステム研究所

## 1 はじめに

パーソナル移動通信などの高度通信サービスを効率良くかつ柔軟に導入するための通信網ワイドでの協調分散処理をサポートし、厳しい実時間性、高い多重処理性能、高信頼性を満足する、実行環境（プラットフォーム）が必要である。これらの要求条件を満たす分散処理プラットフォーム PLATINA [1] の検討を行ってきた。

PLATINA におけるメッセージ通信機構は、メッセージを格納する領域（メッセージバッファ）をすべてのオブジェクトからアクセス可能な空間に配置し、メッセージバッファへのポインタのみを送信する。通信サービス制御プログラムでは、メッセージは小規模のものが大部分なので、メッセージバッファは短い固定長のものを想定している。しかし、小さいメッセージバッファを用いてファイルなどの大きなメッセージを通信するためには、メッセージバッファへのコピーと、通信回数が増大によるオーバーヘッドが大きい。そこで、メッセージバッファ外に置かれたメッセージ（アウトラインメッセージ）をも通信できるメッセージ通信機構を検討している。

本稿では、PLATINA におけるアウトラインメッセージ通信機構について述べる。

## 2 PLATINA メッセージ通信機構

通信網を構成するノード上におかれる PLATINA の構成を図 1 に示す。PLATINA は、並列オブジェクトモデルの実行環境を提供するカーネル、交換リソース制御、ファイル管理などの共通的な機能を提供するサーバ群、各種通信サービスを実現する応用プログラム群から構成される。

カーネルは、能動オブジェクトの実行スケジューリング、メッセージ通信、メモリ管理、論理メモリ空間制御機能を提供する。PLATINA では、ユーザ空間、共有空間、カーネル空間の論理空間がサポートされる。ユーザ空間は複数個生成され、応用プログラム、サーバが動作する。共有空間はノードで 1 面だけあり、どのユーザ空間からもアクセス可能である。カーネル空間もノードで 1 面だけで、カーネル、サーバが動作する。

並行オブジェクト指向による通信サービス制御プログラムを実行するため、高スループット、低遅延のオブジェクト間メッセージ通信を実現する必要がある。そのため、ノード内のオブジェクト間では、メッセージバッファを応用プログラムから共通にアクセスできる共有空間に置き、メッセージバッファへのポインタを送信する方式をとる。メッセージバッファの大きさを比較的短い（数百バイト）固定長とすることにより、管理のオーバーヘッドを小さくしている。

## 3 アウトラインメッセージ通信への拡張

メッセージバッファは短い固定長のため、ファイルの読み書きに必要な大きなメッセージを通信するにはオーバーヘッドが大きい。そこで、通常メッセージバッファとは別に割り付けられたアウトラインメッセージも通信できるようにメッセージバッファを図 2 のように用いる。

メッセージバッファはヘッダ部とデータ部から構成される。ヘッダ部は、メッセージを制御するために用いるリンクフィールド、送受信オブジェクト識別子フィールド、

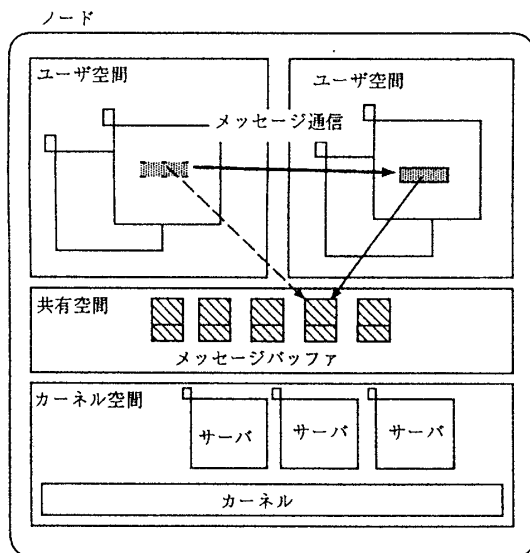


図 1: PLATINA メッセージ通信機構

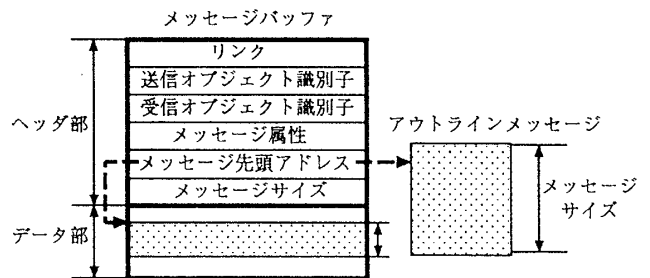


図 2: メッセージの指示

ルド、属性を示すメッセージ属性フィールド、メッセージの先頭の論理アドレスを指示するメッセージ先頭アドレスフィールド、メッセージの大きさを示すメッセージサイズフィールドなどから構成される。

メッセージ先頭アドレスフィールドには、メッセージバッファのデータ部にメッセージが配置された場合には、データ部のメッセージの先頭アドレスが設定され、メッセージバッファ外にメッセージが配置された場合には、そのアウトラインメッセージの先頭の論理アドレスが設定される。

オブジェクトはメッセージバッファのメッセージ先頭アドレスフィールドの値を間接参照することでメッセージにアクセスする。これにより、アウトラインメッセージかどうかの区別なく、メッセージにアクセスすることができる。また、問い合わせの結果をアウトラインメッセージとして返す場合、問い合わせメッセージのメッセージバッファを再利用することができる。

アウトラインメッセージ通信の制御では、メッセージの存在する物理ページを送信オブジェクトのユーザ空間から受信オブジェクトのユーザ空間に再配置することによって、メッセージのコピーを削減する(図3(2))。

しかし、アウトラインメッセージの場合には、送信オブジェクトが送信後もアウトラインメッセージにアクセスする必要がある。そこで、送信オブジェクトは送信後もアウトラインメッセージにアクセスするかどうかのフラグもパラメータとしてメッセージ送信時に指定する。このフラグにより、送信後もアクセスすることが指示されていた場合(図3(1))、アウトラインメッセージの存在する物理ページは共有され、コピーオンライト制御[2]が行なわれる。

受信オブジェクトがメッセージを受信する時に、受信オブジェクトのユーザ空間に再配置されたアウトラインメッセージの先頭の新しい論理アドレスがメッセージバッファのメッセージ先頭アドレスフィールドに設定される。

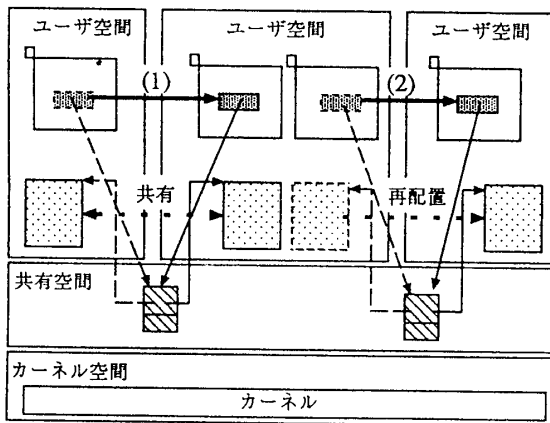


図3: アウトラインメッセージ通信制御

#### 4 ノード間アウトラインメッセージ通信

ノード間メッセージ通信では、ノード間通信制御を行なうオブジェクト(INCオブジェクト)が通信リンク

上でノード間通信プロトコルを使ってメッセージを転送する。アウトラインメッセージも同様に転送される。

送信ノードのINCオブジェクトは、転送の完了まで再送のためにメッセージを保持する必要がある。そのため、送信オブジェクトが送信後もアウトラインメッセージにアクセスする場合には、送信オブジェクトとINCオブジェクトの間でアウトラインメッセージの物理ページが共有され、コピーオンライト制御によりコピーが遅延される。そして、受信ノードのINCオブジェクトから受信オブジェクトのユーザ空間へアウトラインメッセージの物理ページが再配置される。これにより、送受信オブジェクトとINCオブジェクトの間でのメッセージのコピーを削減する。

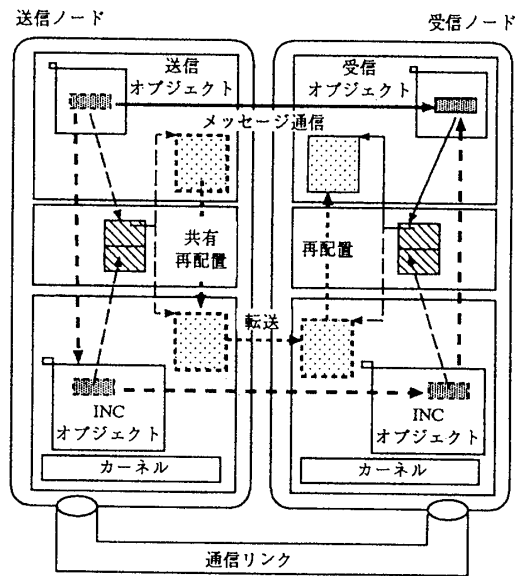


図4: ノード間アウトラインメッセージ通信

#### 5 まとめ

PLATINAにおけるアウトラインメッセージ通信機構について述べた。本機構は、メッセージバッファ上にメッセージがある場合と同様にアウトラインメッセージの通信を可能とし、アウトラインメッセージの物理ページを共有し、通信処理のオーバーヘッドを削減する。

今後は、実機上に実装し、オーバーヘッドの定量評価を行なうとともに、ファイルサーバ、通信網制御ノードのシステム管理、マルチメディア通信制御等への応用について検討する。

#### 参考文献

- [1] M. Kubota, et al. : "Distributed Processing Platform for Switching Systems : PLATINA", Proc. ISS'92.
- [2] R. Rashid, et al. : "Machine-Independent Virtual Memory Management for Paged Uniprocessor and Multiprocessor Architectures", Proc. ASPLOS II, 1987.