

3V-2

分散システムにおける 送信データ予約方式による転送方法

山根 義孝
四国化工機(株)

高橋 義造
徳島大学

1.はじめに

階層構造の分散システムが大規模化、高度化するにつれ、システムの信頼性・耐故障性、柔軟性・拡張性が重要になってくる。耐故障性、拡張性を有する分散システムのデータ転送方法について述べる。

2.送信データ予約方式

耐故障性、拡張性を有する分散システムにおける通信方式として、送信データ予約方式による通信方式を提案する。この方式にデータ通信は次のように行う。

データを必要とするサブシステムはデータを生成するサブシステムを含むグループに対しデータの送信要求を放送する。この時、データの中継経路を確定する。送信要求をうけたサブシステムが送信要求のデータを生成した時、送信要求を行ったサブシステムを含む最小のグループに対し確定された中継経路を介してデータを放送する。データが転送されたサブシステムは、予め送信予約したデータだけを選択して受信する。

この送信データ予約方式によるデータ転送は、次の利点がある。

- 1) 転送範囲内でのサブシステムの増加による通信コストの増加が無い。
- 2) データ転送時のシステム状態に応じた最短時間経路でデータが転送される。
- 3) 各PEは分散システム全体の構造がわからなくても、自分に接続されているPE情報だけをもとにデータ転送が可能である。

3.転送方法

分散システムのデータ転送はFig.1に示す3レベルに分けることができる。送信データ予約方式によるデータ転送では通信路の形態を特定していないため、1.サブシステム間の通信方法、2.PE間の通信方法について述べる。

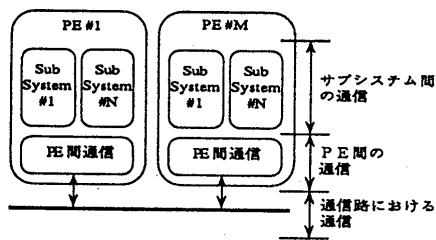


Fig.1 転送方法のレイヤー図

Data Registering Method for Communication
on A Distributed System
Yoshitaka Yamane¹, Yoshizho Takahashi²
¹ SHIKOKU KAKOKI Co.,Ltd.
² Univercity of Tokushima

3.1.サブシステム間の通信

サブシステム間の通信はFig.2に示す手順で行なう。

- 1.送信予約Requestの送信
他のサブシステムの生成データを必要とするサブシステムは最下位グループから最上位グループへ順に転送範囲を広げながら送信予約Requestを放送する。
- 2.送信予約確認Ackの送信
送信予約Requestのデータを生成するサブシステムは、送信予約Requestを送信予約リストに登録し、送信予約確認Ackを送信予約Requestの転送範囲に送信する。
- 3.データの送信
サブシステムが送信予約されたデータを生成した場合、送信予約要求グループにデータを放送する。送信予約要求グループは送信要求Requestの転送した全てのPEを含む最小のグループである。
- 4.データの受信
送信予約確認Ackを受信したサブシステムはデータの受信状態になる。受信状態のサブシステムは送信予約したデータだけを受信し他のデータは無視する。
- 5.送信予約終了Requestの送信
他のサブシステムのデータを利用しているサブシステムで、データを利用しなくなった場合、送信予約終了Requestを送信予約Requestの転送範囲に放送し、送信予約を解除する。
- 6.送信予約の解除
送信予約終了Requestを受信したサブシステムは、送信予約リストをチェックし送信予約がある場合、送信予約リストから送信予約終了Requestの転送範囲を除く。転送PEが無くなった時、送信予約リストから送信予約データの項目を削除する。
- 7.中継リストの更新
送信予約リストで転送範囲が変更になった場合、その旨を変更前の転送範囲に放送する。

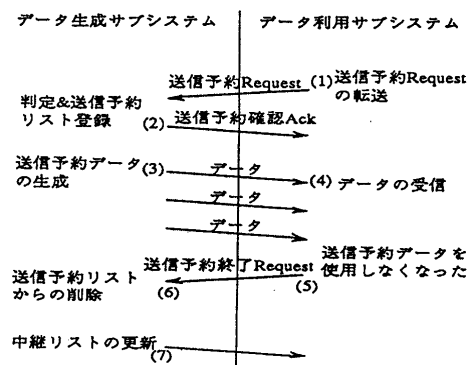


Fig.2 サブシステム間の通信方法

3.2.PE間の通信

PE間の通信は、各PEの通信路情報、中継リスト、パケット情

報をもとに行われる。ここでは、パケットの転送方法で、パケットの中継処理、通信バッファについて考える。

3.2.1.パケット構造

パケット構造をFig.3に示す。

パケット種類	転送範囲	中継情報	データid	データ
--------	------	------	-------	-----

Fig.3 パケット構造

- 1)パケット種類 - パケットには、送信予約Request, 送信予約確認Ack, データ, 送信予約終了Request, 中継リスト更新Requestがある。
- 2)転送範囲 - パケットのどの範囲に放送するかを示す。
- 3)中継情報 - パケットを保存する通信バッファを示す。

3.2.2.パケットの中継処理

各PEにおけるパケットの中継処理は、通信路の結合リスト、データの中継リストをもとに行なう。

通信路の結合リストは、自分と結合しているPEのリストで、PE番号/グループ番号を示す。データの中継リストは、データid毎に転送する通信路,転送範囲及びデータ転送の有無を示す。各PEは、パケットの中継処理を次のように行う。

- 1)中継リストへの中継登録
送信予約Requestを受信した場合、中継リストにデータid, 転送範囲, 転送通信路を仮登録する。送信予約確認Ackを受信した場合、仮登録データを正式登録する。
- 2)中継リストの中継削除
送信予約終了Requestを受信した場合、中継リストからデータidに関する項目を削除する。
- 3)受信パケットの転送
中継リストの転送通信路で、未転送の通信路に対し受信パケットを転送する。転送後、中継リストを更新する。

3.2.3.仮想通信路

送信データ予約方式による通信方式は、放送機能を用いてデータを転送する方式である。この通信方式は、階層構造の各レベル毎にそのレベルに存在するグループ/PE専用の仮想通信路[3]を設け、パケットの中継情報により示される仮想通信路を通じて各PEに転送(放送)する方法をとる。仮想通信路は、Fig.4に示すように各レベル毎にそのレベルに存在するグループ/PE数だけ存在する。

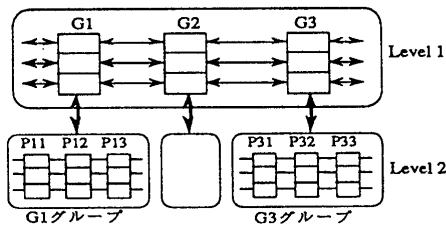


Fig.4 仮想通信路

各PEは通信バッファとして、同一レベル間用通信バッファと異なるレベル間用通信バッファを持つ。それぞれのバッファ数は次のようになる。

- 1)同一レベル間用通信バッファのバッファ数はそのレベルに存在するグループ数/PE数である。
- 2)異なるレベル間用通信バッファのバッファ数は、下位レベルに存在するグループ数/PE数-1である。

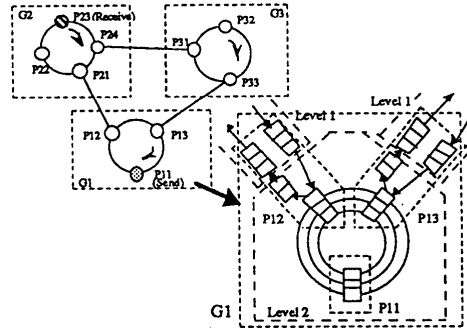


Fig.5 バッファ構成

Fig.5に仮想通信路に対応したバッファを示す。

3.2.4.中継情報

中継情報は、各レベルで使用する仮想通信路を示す。つまり、中継情報の各レベルの値は、パケットが各レベルのグループで最初に受信したPE, または、パケットを生成したPEのPE番号である。中継情報は、最初はパケットの生成PE番号で、他のPEに転送される毎に以下の手順に従い更新される。

- Step1 受信パケットを送信したPEのPE番号と自分のPE番号を各レベル毎に上位レベルから比較し、最初に異なるレベルを決定する。このレベルはPE間の通信路が決定された時点で決定され、通信路が変更がない限り不変である。
- Step2 Step1で決定したレベル以下の中継情報の値を自分のPE番号の値で置き換える。

中継情報の更新例を、Fig.6に示す。Fig.6の中継経路の下線部分はPE番号が最初に異なるまでのレベルを示す。

中継経路: P2111 → P2130 → P2132 → P1231
 中継情報: 2111 → 2110 → 2110 → 2231

Fig.6 中継情報の更新例

3.3.通信路の追加

通信路が追加された場合、通信路が追加されたPEは、次の処理を行なう。

- 1)追加された通信路に結合されているPEのPE番号/グループ番号を調べ、通信路の結合リストに登録する。
- 2)中継リストをもとに追加された通信路に対する中継データを決定する。
- 3)通常のパケットの中継処理を行なう。

4.まとめ

・階層構造の分散システムにおける送信データ予約方式による転送方法について述べた。この方法には、1)転送途中のPE/通信路が故障してもその時のシステム状態に応じた最短時間経路でデータが転送される。2)パケットの中継処理がPE間の結合情報だけで行える利点がある。

[参考文献]

[1] 伊藤, "自律分散システムはいかにして構成されるか", 計測と制御, vol.29, no.10, 1990
 [2] 森, "自律分散システムと制御分野での実用例", 計測と制御, vol.29, no.10, 1990
 [3] William.J.Dally, "Virtual-channel flow control", IEEE Trans. Parallel and Distributed Syst., vol.3, no.2, 1992