

マルチキャスト通信におけるメッセージ到着順序保証*

4 P-2

永石 裕二[†] 吉田 隆一[‡]九州工業大学 情報工学部[§]

1 はじめに

マルチキャスト通信は、いくつかのホストをグループ化したマルチキャストグループと呼ばれるものに対して通信をおこなう。グループ化されたホストには、同じメッセージを同時に送信することができ、個々のホストは複数のグループに属することができるために、複数のグループに対しても同時に送信可能となる。このような通信を担うマルチキャストプロトコルは、メッセージ送受信に関して、信頼性のある通信を保証しなければならない。メッセージの到着順序保証もその一つである。アプリケーションの中には、グループ内のホストに対して届くメッセージについて、データの一貫性や処理の一貫性のために処理順序が重要になってくる場合がある。

そこで本稿では、[Birman 87]のアルゴリズムにもとづいた新しいメッセージ到着順序を保証するアルゴリズムを提案する。本アルゴリズムでは、[Birman 87]のアルゴリズムと比べて順序保証のためのメッセージ数を削減することができ、より高速な処理が可能になると考えられる。

2 メッセージ到着順序と順序問題の分類

一般に分散処理系でのメッセージ通信は、通信遅延の不確定性のために到着順序は保証されない。そのために順序を確実に保証したい場合、何らかの保証手続きが必要になる。

マルチキャスト通信で考えるメッセージの到着順序保証手続きとは、あるマルチキャストグループに（以後、グループと称する）次々と発せられたメッセージがグループ内のホストすべてに同じ順序に到達することを保証することである。この順序付け問題は、次の二つに分けることができる。

1) あるホスト s_1 からの送信メッセージ m_1 と、別のホスト s_2 からの送信メッセージ m_2 があるグループに送られた場合、グループ内のホストに対してすべて同じ順序で受信されなければならない。(total ordering)

2) あるホスト s_3 からメッセージ m_3 と m_4 が送信され、かつ m_3 と m_4 には $m_3 \rightarrow m_4$ という生起順序関係[Lamport 78]があるとき、受信順序も $R(m_3) \rightarrow R(m_4)$ を満たさなければならない。(causal ordering)

本稿では、[Birman 87]の ABCAST アルゴリズムを

もとに 1) の total ordering を保証するアルゴリズムを提案する。

3 Total ordering 保証

様々なホストから送信されたメッセージが、あるグループで受信される時、メッセージごとに各ホストでの受信時刻が異なるので、ホスト間で処理が同じなとは限らない。また、あらかじめ受信時刻印を指定して送信しても、別のホストからのメッセージが不確定的に到着するので、やはり処理が同じになる保証はない。

そこで、グループ内の一部のホストに時刻決定権 token を持たせ、そのホスト (token ホストと呼ぶ) を通じてメッセージ通信をおこなうような形にする。つまり、token ホストから時刻決定のメッセージ (時刻決定メッセージと呼ぶ) を、グループ内の全ホストに送信することで各ホストで受信されたメッセージの時刻を揃える。またその際、グループ内に複数の token ホストが存在する場合には、token ホストだけのグループを構成し (メタグループと呼ぶ)、メタグループ内でまず時刻同期をとる処理を行い、その後、時刻決定メッセージを送信する。

3.1 順序保証アルゴリズム

本アルゴリズムは、token ホストの数によって処理が異なってくるので、token ホストの数が一つの場合と、複数ある場合に分けて述べる。

A. token ホストがグループ内に唯一存在する場合。

1. 送信ホストがメッセージを、目的のグループに対して送信する。
2. メッセージは、目的のグループ内の各ホストでそれぞれ受信され、それぞれのキューに入れられる。その際、受理不可のタグが付加される。
3. token を持つホストは、メッセージを受け取ると同様にキューに入れるが、受理可のタグをつける。そして、その受信時刻を、同じグループ内のすべてのホストに対して時刻決定メッセージとして送信する。
4. token を持たないホストは、token ホストからの時刻決定メッセージを受け取ると、キュー内の同じメッセージ ID を持つメッセージを探し、受信時刻を決定し、タグを受理不可から受理可に変える。そして、受信時刻でソートする。

*Message ordering on Multicast Communication

[†]Yuji NAGAISHI

[‡]Takaichi YOSHIDA

[§]Kyushu Institute of Technology, Iizuka

5. キュー内のメッセージはキューの先頭にあり、かつタグが受理可の時に限り受理される。

B. token ホストがグループ内に複数存在する場合。
(上記アルゴリズムの3を以下に変更する。)

- 3'-1. メタグループの各ホストは、メタグループ内の時刻決定権を持つホスト (super token ホスト) に対して受信時刻を知らせる。(図1, Step 1 参考)
- 3'-2. 各ホストから受信時刻のメッセージを受けとった super token を持つホストは、受け取った時刻の中で最も大きい時刻を時刻決定メッセージとしてそのグループ内のすべてのホストに対して送信する。(図1, Step 2 参考)

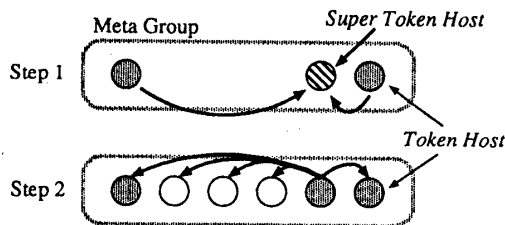


図1: メタグループの処理

3.2 token ホスト決定アルゴリズム

次に、前節で述べた token ホストとメタグループをどのように決定するのかを述べる。あるホストがあるグループに加入しようとする時、以下のアルゴリズムによって token ホストやメタグループを決定する。

例えば、図2においてホスト a_1 が、グループに G に加入する場合を考える。

1. まず、加入しようとするホストに対して token を与える。(この場合 a_1 にあたる。図2, Step 1 参考)
2. 次に、加入した新しいグループ (この場合 GROUP G) にある token ホストに対して token 移動処理を行なう。(図2, Step 2 参考)
3. 加入する前のグループ (この場合 GROUP A) にある token ホストに対して token 移動処理を行なう。(図2, Step 2 参考)

token 移動処理とは、以下の手続きを指す。

1. 加入したホストの所属グループが token ホストの所属グループを含んでいる場合、token ホストの token を削除する。
2. token ホストの所属グループが加入したホストの所属グループを含んでいる場合、加入したホストの token を削除する。

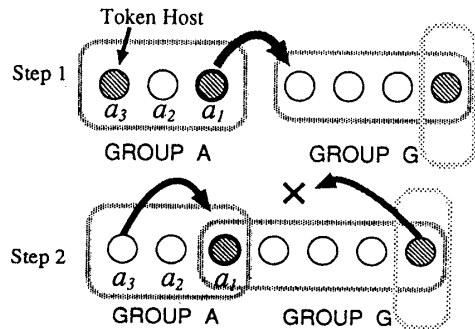


図2: token ホストの決定

3. グループ内に二つ以上の token ホストができた場合、メタグループが形成されてなければ、token ホストだけのグループ (メタグループ) を形成させ、加入したホストにメタグループでの時刻決定権として super token を与える。(但し、super token はメタグループ内でただ一つのみ存在する。)
4. 既にメタグループが形成されている時は、メタグループに加入させる。

4 評価

本アルゴリズムによってどのくらいメッセージが削減できるかを示す。グループ内の全ホスト数が N で token を持つサイト数が n の時、あるホストからある別のマルチキャストグループに対して送信した場合、[Birman 87] の ABCAST アルゴリズムと比べて、順序保証に必要なメッセージが以下のように削減できる。

プロトコル	通信形態	メッセージ数
ABCAST	point to point	$3N$
NEW	point to point	$2N + n - 2$

5 おわりに

本稿では、total order を保証するアルゴリズムを述べた。今後は、causal order も考慮してさらに拡張していく予定である。

参考文献

- [Lamport 78] L. Lamport, *Time, Clocks, and the Ordering of Events in a Distributed System*, Commun. ACM, Vol. 21, No. 7, July 1978
- [Birman 87] K. Birman, T. A. Joseph, *Reliable Communication in the Presence of Failures*, ACM Trans. Comput. Syst., Vol. 5, No. 1, Feb. 1987