

# 7T-02 適応的全順序マルチキャストの改良

細谷 篤 † 水野 忠則 ‡ 佐藤 文明 ‡

† 静岡大学大学院情報学研究科 ‡ 静岡大学情報学部

## 1 はじめに

共有仮想環境などの分散アプリケーションにおいて全順序グループ通信サービスは必要不可欠なインフラであり、近年ではマルチキャスト通信によるグループ通信サービス[1]について多数の研究がなされている。本研究ではスケーラブルなシステムを目的とした全順序プロトコルとして Adaptive Multicast Protocol(AMP)[2]を提案している。このプロトコルでは空のメッセージであるダミーメッセージが必要となる。本稿ではこのダミーメッセージ数を削減するために行った改良とその評価について報告する。

## 2 提案方式

### 2.1 AMP のシステム構成

AMP では、マルチキャストグループ内のすべてのクライアントをいくつかのドメインに分ける（図1）。また、すべてのクライアントは同一の乱数生成器を持っていることを前提としている。

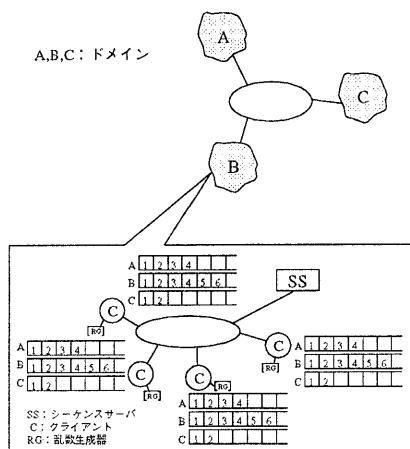


図 1: AMP のシステム構成

各ドメインにはシーケンスサーバを 1 つずつ設置し、ドメイン内の順序管理はそれぞれのシーケンスサーバが行なう。シーケンスサーバはクライアントの送信メッセージそれぞれにドメイン内でユニークなシーケンスナンバーを与えることで順序管理を行う。

またドメイン間の順序付けは乱数生成器が発生する乱数シーケンスによって決定される。各クライアント

はメッセージを受信すると、それを送信元ドメインごとに分けられたバッファに格納する。クライアントは乱数シーケンスによって送信元ドメインを選択し、そのバッファのメッセージをシーケンスナンバーの順に配信する。クライアントの持つ乱数生成器はグループ全体で同一のものであり、乱数シーケンスも同一のものを発生する。つまり選択されるドメインの順序もすべてのクライアントで同じになるのでドメイン間の順序付けは保証される。

以上のようにメッセージの順序はドメイン内、ドメイン間でそれぞれ保証されているのでマルチキャストグループ全体の順序が保証されていることになる。

AMP では送信頻度に応じて各ドメインに重みを割り当てる、ドメインの識別子とそれに対応する重みをリストにしてすべてのクライアントに配布する。このリストのことを分布と呼ぶ。送信頻度が大きいドメインには大きな重みが割り当てられる。

分布はクライアントが配信の際にドメインを選択するのに使われ、重みが大きいドメインほど選択されやすくなる。

分布はグループ内のサーバひとつが定期的に作成、配布するためクライアントはその時々の送信頻度状況に応じた配信が可能である。つまりグループ全体のトラフィックの変動に適応的である。

### 2.2 ダミーメッセージ

あるドメインが急激にその送信頻度を変化させると配信がブロックされやすくなり配信遅延を引き起こす。変化した後すぐに新たな分布が作成されればこの状況はすぐに回避されるが、それほど頻繁に分布を作成、配布することは難しい。そこで各ドメインのシーケンスサーバはトラフィックを監視し、自分の管理するドメインの送信頻度の全体から見た割合が最新の分布の重みに達していない場合、ダミーメッセージという空のメッセージを送信する。空になった配信用のバッファをダミーメッセージが埋めることでブロックを回避することができる。

## 3 これまでの課題と改良点

ダミーメッセージはアプリケーションの立場からすると無意味なメッセージである。帯域を効率良く使うためになるべくその数は少ないほうがよい。ダミーメッセージ数の削減はこれまでの大きな課題のひとつであった。これまでの方式ではシーケンスサーバはトラフィックを監視しつつ定期的にダミーメッセージの送信が必要かどうかを調べ、必要とされている個

Improvement of adaptive totally ordered multicast

Atsushi Hosoya, Fumiaki Sato, Tadanori Mizuno

† Graduate School of Information, Shizuoka University

‡ Faculty of Information, Shizuoka University

数分だけグループ全体に送信していた。

しかしダミーメッセージ 1 個が複数個分の意味を持つことが可能ならば何個必要とされていても実際に送信されるダミーメッセージは 1 個で済む。そこで今回ダミーメッセージに限り複数のシーケンスナンバーを持つことができるよう改良した。

## 4 シミュレーション

### 4.1 シミュレーションモデル

改良前と改良後のシステムについて同じ条件でシミュレーションを行い、ダミーメッセージ数の比較を行った。また改良後のシステムについてダミーメッセージに与えるシーケンスナンバーの数を要求されている数の 100%, 75%, 50% としてシミュレーションを行い、メッセージの受信から配送までの最長時間を取りた。

共通のパラメータを以下のように設定した。

- ・ メッセージ平均発生間隔 0.5 s/クライアント
- ・ 分布更新間隔 5 s
- ・ ドメイン内通信の平均通信遅延 100ms
- ・ ドメイン外通信の平均通信遅延 300ms
- ・ クライアント総数 100
- ・ ドメイン数 2,4,5,10,20,25,50
- ・ 送信頻度急変イベントなし, 1 回 /33.3s, 1 回 /20s
- ・ シーケンスサーバの平均処理時間 1ms

クライアントは各ドメインに均等に分かれているとした。クライアント総数は 100 と一定なのでドメイン数が変化するとドメインあたりのクライアント数も変化する。

送信頻度急変イベントとは人為的に送信頻度の変動を起こすイベントで、今回は総数の 10% のクライアントをランダムに抽出し、そのクライアントの送信頻度を 10 秒間に 3 倍にする。

### 4.2 結果

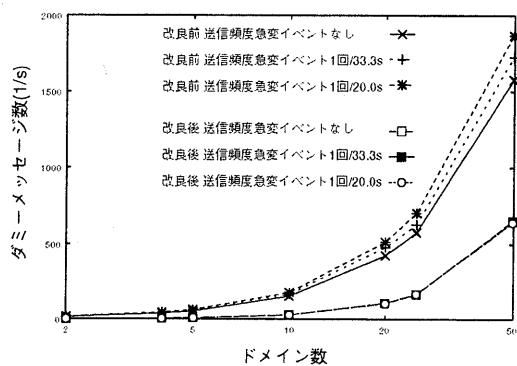


図 2: ダミーメッセージ数

図 2 のグラフの通り改良後はダミーメッセージの削減に成功した。改良後のダミーメッセージ数はシ

ケンスサーバがダミーメッセージが必要かどうかを調べる回数のみに依存するため、改良前と違い送信頻度急変などの障害が起きても増加することはない。

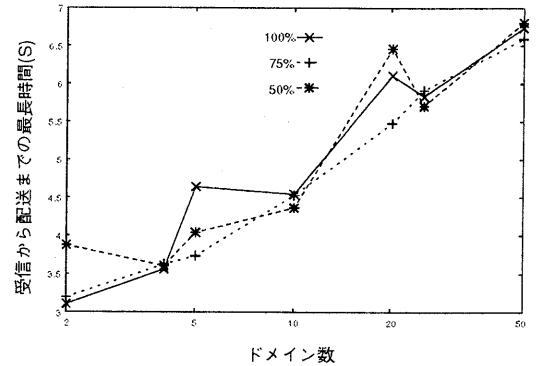


図 3: 受信から配送までの最長時間

空メッセージであるダミーメッセージに多くのシーケンスナンバーを与えることはその後に続く本物のメッセージを長い時間待たせてしまう危険性も孕んでいる。図 3 のグラフは送信頻度急変イベントを 1 回 /20s の割合で起こしているときのものである。このグラフからは 100%, 50% に比べ 75% のものの方が安定しているように見える。つまり要求された個数そのままを返すのではなく多少抑え気味にしたほうが効率が上がるといえる。ただし抑え過ぎるとダミーメッセージの不足でブロックが起こりやすくなり、それが 50% のグラフの不安定さを産んでいると考えられる。

## 5まとめ

今回は AMPにおいてダミーメッセージの数を削減するという改良を行った。1 個のダミーメッセージに複数のシーケンスナンバーを持たせることでダミーメッセージの数を削減することに成功した。

また要求されたダミーメッセージ数をそのまま送信せず工夫をすることで配送の効率化ができることがわかった。

今後はさらにシミュレーションを行いヒストグラムなどを取ることでより正確な情報を得て、どのようにダミーメッセージを送信するのが効率的なのか調べていくのが課題である。

## 参考文献

- [1] G.V.Chockler,N.Huleihwi and D.Dolev, "An Adaptive Totally Ordered Multicast Protocol that Tolerates Partitions", The 17th ACM Annual Symposium on Principles of Distributed Computing (1998).
- [2] 細谷篤, 佐藤文明, 水野忠則, "適応的全順序マルチキャストの拡張", 情報処理学会論文誌 掲載予定 (2001).