

2DRR スケジューラに基づいた スループットを高めるパケットスイッチスケジューラ

大森一智[†] 兪明連[†] 横山孝典[†]
東京都市大学[†]

1. 研究背景

近年のコンピュータ間のネットワーク (LAN: Local Area Network) や衛星通信網などは、多様化や大規模化している。そのため、限られた貴重な資源である伝送路を使い、より多くの端末間の通信を行うために通信の多重化が研究・開発されてきた。一般的に複数個の端末が限られた伝送路を共有し通信を行おうとすると、混信し正常な通信は行うことが出来ない。そのため、適切にパケットを送信する順序をスケジューリングし通信することが重要である。

2. 従来研究とその問題点

2.1 2DRR スケジューラ¹⁾

2DRR (Two-Dimensional Round Robin) スケジューラ¹⁾とは、 N 入力 N 出力システムの完全スケジュール (静的アルゴリズムに沿ったスケジュール) を実現するスケジューラである。高いスループットと公平性を持っており、単純なアルゴリズムで構成されているためハードウェアへの実装が容易であり、なおかつ高速なスイッチングが可能である。

2DRR 方式は完全スケジュールを可能にするため図1の様に、各入力ポートに出力ポート別に FIFO 方式のバッファをもつシステムを利用する。各入力ポートに届いたパケットは図1の左側の入力ポート側にある出力ポート別のバッファに入れる。同一タイムスロットにて入力ポートより出力できるパケットは1つのみである。また、各入力ポートは図1の右側の様に全出力ポートに繋がっており、出力先が競合しない限り複数の入力ポートから複数の出力ポートへ同時にパケットを送ることができる。

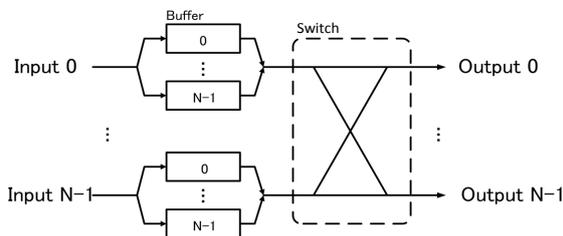


図1 N 入力 N 出力のスイッチ

2DRR 方式では、入力側のバッファからの送信要求を示したリクエストマトリクス (RM), スケジューリング時の順序を示したパターンマトリクス (PM), RM を

参照する場所を示しているスケジューリングマトリクス (SM), 送信許可を記録するアロケーションマトリクス (AM) の 4 種類のマトリクスを利用しスケジューリングを行う。この時の PM は図2の右のマトリクスの様に、マトリクス内で行と列において同じ値が入らないように作成する必要がある。

問題点としては、図3の真ん中と右側の AM の塗りつぶされた部分の様に、リクエストがあるのに送信許可が出ていない部分が生じてしまう。これは、PM により決められたパターン順序で RM を参照していく上で、先に許可を得てしまった部分があるため競合してしまうので許可が出ることができない。

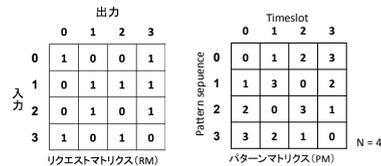


図2 N=4 の RM と PM

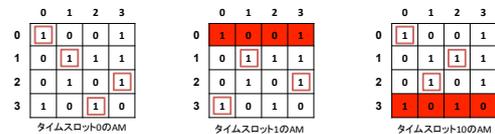


図3 2DRR スケジューラの問題点

2.2 ユニット型低遅延スケジューラ²⁾

ユニット型低遅延スケジューラとは 2DRR スケジューラの利点であるポート間の公平性を損なわずに、AM の無駄を減らしてスループットを高めたアルゴリズムである。

手法としてはマトリクスを図4の様にユニットという単位にて分割し、そのユニット毎にスケジューリングしていくというものである。マトリクスを分割することで、パターンによる参照範囲を小さくし参照密度を上げている。ただし、このアルゴリズムではマトリクスを分割するために、2DRR スケジューラにて使われていた PM と同等の役割を果たす MPM という物が必要になる。MPM は PM とは違い、行と列に加え分割したユニットの中でも同じ値が存在しないように作成する必要がある。

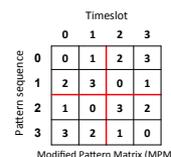


図4 分割したマトリクス例

¹A Packet Switch Scheduler based on 2DRR for High-Throughput

[†]Kazunori Omori, Myungryun Yoo, Takanori Yokoyama

[†]Tokyo City University

問題点としては、ユニット化してしまったため、図5のグラフの様にポート数であるNが大きくなると、必要なMPMの作成にPMより大きい計算が必要になり実装が困難である。また、MPMの特性である、ユニット内で同じ値が存在しない様にという制限があるために、分割できるNの値には制限がある。そのため、Nの値の自由度が2DRRスケジューラより低くなってしまっている。

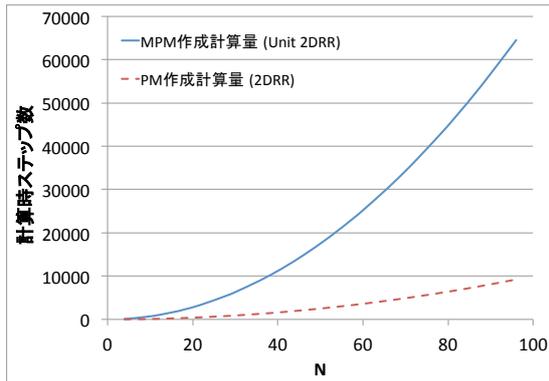


図5 MPMとPMの作成時計算量グラフ

3. 研究目的

2DRRスケジューラの動作改良のみを行い、ユニット型低遅延スケジューラより計算量の少ない手法にて、2DRRスケジューラより高いスループットを実現する。

4. 提案手法

各タイムスロットにてAMの作成が完了時に、あらかじめ指定した閾値より多い無駄が生じていた場合に参照に用いたSMを置き換え、次のタイムスロットにて使うSMを参照しサブのAMを作成する。もし、そのサブで作成したAMが初回のAMより無駄が少なかった場合、新しく作ったサブのAMを使いパケットを送信する。また、この繰り返し動作は、あらかじめ決めた値を上限とし、その上限まで繰り返しても初回のAMより無駄の少ないAMが作成されなかった場合は、初回のAMにて送信を行う。

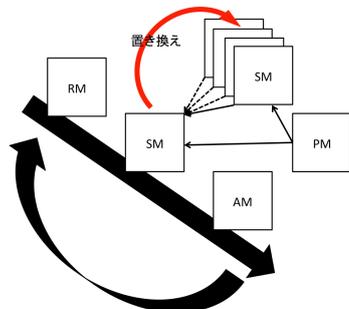


図6 Re2DRR動作概要図

5. シミュレーション

シミュレーションは、Nが4,6,16において行った。パケット数はNが4と6の場合は10000タイムスロット分(10000 × N × N)生成し、Nが16においては3000タイムスロット分(3000 × N × N)を生成した。表の2DRRが2DRRスケジューラ、Unit 2DRRがユニット型低遅延スケジューラであり、Re2DRRが提案手法のスケジューラである。aveが各パケットが各入力ポートにあるバッファに入ってから送信が完了するまでの時間であり(平均バッファ待機時間)、endtimeは準備されたパケットが全部送り終えた時の時間である。単位はどちらもタイムスロットである。Re2DRRの括弧の値は、Cが送信されていないポートの閾値であり、Rは繰り返す上限回数である。

表1 N=4のシミュレーション結果

スケジューラ名	ave	endtime
2DRR	5.883	10006
Unit 2DRR	3.624	10001
Re 2DRR	2.552	10000

表2 N=6のシミュレーション結果

スケジューラ名	ave	endtime
2DRR	11.098	10029
Unit 2DRR	134.265	10496
Re 2DRR (C:1, R:5)	6.673	10010
Re 2DRR (C:1, R:2)	8.371	10023

表3 N=16のシミュレーション結果

スケジューラ名	ave	endtime
2DRR	44.171	3168
Unit 2DRR	32.925	3103
Re 2DRR (C:1, R:6)	16.439	3049
Re 2DRR (C:1, R:1)	26.295	3070

6. 考察

現段階のシミュレーション結果では、提案手法がスループットに関してはどのNの値でも2DRRスケジューラより優れている値が出ている。また、Nが6の場合において、ユニット型の低遅延スケジューラは適切なスケジューリングが行われておらず、2DRRスケジューラの利点が損なわれている。

謝辞

本研究はJSPS科研費24500046の助成を受けたものです。

参考文献

- Richard O. LaMaire.: Two-Dimensional Round-Robin Schedulers for Packet Switches with Multiple Input Queues, IEEE/ACM Transactions on Networking, Vol.2, No.5, pp.471-482 (1994).
- 川上 浩只, 浅谷 耕一: ポート間公平性を考慮した低遅延スケジューリングアルゴリズムの提案, 電子情報通信学会論文誌 B, Vol. J88-B, No.10, pp.2052-2061 (2005).