

# 特殊なシステムにおける高効率な パケットスイッチスケジューラ

大森一智<sup>1</sup> 兪明連<sup>1</sup> 横山孝典<sup>1</sup>

本研究では、限られた資源であり高価である通信インフラをより高効率に利用するために、従来手法である 2DRR (The two-Dimensional Round Robin) スケジューラを基にした Re-2DRR スケジューラの閾値変動による特性を明らかにし、Re-2DRR (The Repetive two-Dimensional Round Robin) スケジューラにて追加され負担と成る動作をなるべく減らす。また、特殊なシステムを定義し、それにより高効率にて通信を行えるスケジューリングアルゴリズムを 2 種提案する。

## 1. はじめに

近年、コンピュータ間ネットワーク (LAN : Local Area Network) や衛星通信等は多様化したシステムのため大規模化複雑化している [1][2]。しかしながら、通信路のインフラ設備は高価で貴重な資源である。通常、複数の端末間にて 1 つの伝送路を共有して通信を行おうとすると混信し、正しい通信を行うことが出来なくなる。そのため、通信の多重化を行い限られた通信路を共有し、より多くの端末等が混信を起こさずに通信出来る様にする必要がある。限られた貴重な伝送路を有効活用するための技術の 1 つとして、通信の多重化の研究・開発がされている。

その多重化技術として 2DRR スケジューラ [3] を元にし、動作の一部を改良した Re-2DRR スケジューラ [4] がある。これは、基と成る 2DRR スケジューラより、高スループットにて通信を可能にするパケットスイッチスケジューラであるが、改良した動作である繰り返し動作が多く動作されるため、実装の負担に成ってしまう事が考えられる。そのため、Re-2DRR スケジューラで追加された動作を制御するための 2 つの閾値の変動による特性を明らかにする。

また、本研究では特殊なシステムでの通信を定義し、その定義した特殊なシステムでより効率よくパケットスイッチを行い、パケットを送ることが可能なアルゴリズムを Re-2DRR スケジューラの閾値変動の特性を踏まえた上で 2 種新しく提案する。

## 2. $N \times N$ パケットスイッチ

$N \times N$  パケットスイッチとは計算機ネットワーク、通信網間接続、パソコン通信ネットワークなど多様に使われている多数対多数の通信に使われるものである [5]。入力と出力が  $N$  個あり、入力されたパケットを一時的に蓄積させるメモリ領域 (バッファ) を持ち、各入力が入出力に論理的に接続されている。スイッチにより各入力からどの出力へ通信を行うかで、入力から任意の出力へ通信を行う事が出来る。

よって、物理的に個別にネットワークを作らずとも、パケットスイッチを使うことで論理的な多数対多数のネットワークを構築することが出来る。

ただし、パケットスイッチの特性上、同時刻に 1 つの入力から複数の出力へのパケット送信、また同時刻に複数の入力から同一の出力へパケットを送信することは出来ない。

## 3. The two Dimensional Round-Robin Scheduler

The two-Dimensional Round-Robin (2DRR) スケジューラとは、出力別 FIFO バッファを持つ  $N$  入力  $N$  出力パケットスイッチを用いて通信を多重化するためのスケジューラである。2DRR スケジューラは 4 つのマトリクスを 2 次元ラウンドロビンにて動作させることで、全入出力の公平性を保ちながら高いスループットを実現しているスケジューラである。ただし、2DRR スケジューラが正しくスケジューリングを行うためには入出力数を  $N$  としたとき、 $N+1$  が素数でなければ成らない。この条件を満たしていれば、 $N$  がどのような値でも公平性を保ちながらスケジューリングを行う事が出来る。

まず、4 つのマトリクスについて説明する。

### (1) リクエストマトリクス (RM)

リクエストマトリクス (以下 RM) は、出力別 FIFO バッファを持つ  $N$  入力  $N$  出力パケットスイッチにある、各入力側にある出力別の FIFO バッファに送信要求を出しているパケットがあるかを示しているものである。RM [i, j] と示すとすると、i は入力であり、j は出力である。また、RM に入っている値が 0 だと送信要求のあるパケットはバッファに無い事を示し、1 だと送信要求のあるパケットがバッファに存在するという事を示す。RM は各バッファにある送信要求のあるパケットの存在の有無を示しているマトリクスなため、タイムスロットが進むと新しいパケットが入力されるか、バッファにあったパケットが送信されるかの動作が起きる。よって、各タイムスロットで RM は書き換えられることになる。

<sup>1</sup> 東京都市大学

Tokyo City University, 1-28-1 Tamazutsumi, Setagaya-ku, Tokyo, Japan

(2) パターンマトリクス (PM)

パターンマトリクス(以下 PM) は、次に説明するスケジューリングマトリクス (SM) の作成や、スケジューリング動作時に各タイムスロットでど SM を使うか、との順で RM を参照するか等を示すものである。そのため、様々な順番を生成するために PM では同じ行と列に同じ値が存在しないように作成する必要がある。

よって、PM 作成には下記の式を用いて、各値が同じ行と列に同じ値が存在しないように作成する。

$$PM[I,J] = (I + I \times J + J) \bmod (N + 1)$$

例として、図 1 に N = 4 の時の PM を作成した。各値が同じ行と列に同じ値が存在しないように作成されている。これにより、動作時に様々な順番が使用できる事になる。図 1 では、数値別に色分けしてある。観て解るように行と列にて同じ値(色)が存在しないことが解る。

また、提供するテンプレートファイルは、**エラー! 参照元が見つかりません。**に示す通り、3 つのセクションから構成している。

Pattern sequence	Timeslot			
	0	1	2	3
0	0	1	2	3
1	1	3	0	2
2	2	0	3	1
3	3	2	1	0

図 1 パターンマトリクス

(3) スケジューリングマトリクス (SM)

スケジューリングマトリクス (以下 SM) は、スケジューリング動作時に送信要求を示している RM を参照するパターンを示しているものである。SM は PM を元に N 個作成され、各タイムスロットで PM が示す順序で扱うことで、様々な順番で RM を参照出来る。図 2 は N = 4 の時の SM 例である。この場合、N = 4 であるため、4 つの SM が作成される。

	0	1	2	3
0	0	1	2	3
1	3	0	1	2
2	2	3	0	1
3	1	2	3	0

Scheduling Matrix (SM)  
Step 0 : Time Slot 0 ~ 3

	0	1	2	3
0	3	0	1	2
1	1	2	3	0
2	0	1	2	3
3	2	3	0	1

Scheduling Matrix (SM)  
Step 1 : Time Slot 4 ~ 7

	0	1	2	3
0	2	3	0	1
1	0	1	2	3
2	1	2	3	0
3	3	0	1	2

Scheduling Matrix (SM)  
Step 2 : Time Slot 8 ~ 11

	0	1	2	3
0	1	2	3	0
1	2	3	0	1
2	3	0	1	2
3	0	1	2	3

Scheduling Matrix (SM)  
Step 3 : Time Slot 12 ~ 15

図 2 スケジューリングマトリクス

(4) アロケーションマトリクス (AM)

アロケーションマトリクス (以下 AM) は、スケジューリング動作時にタイムスロット毎に出された送信許可を記録したものである。AM の行と列の意味合いは RM と同じであり、送信許可が出た部分を 1 とし、許可が出ていない部分は 0 となる。ただし、パケットスイッチの特性上、同タイムスロットに 1 入力から 1 出力への送信となるため、AM の行と列には値 1 が重複することは無い。

2DRR スケジューラの動作は、まずタイムスロット毎にバッファに蓄積されている送信待ちのパケットの状況を RM として作られる。経過タイムスロットを基に PM を参照し、このタイムスロットでどの N 個ある SM を用いるか、どの順序で SM のパターンを使い RM を参照するかを決め、AM に送信許可を書き込む。最後に作成された AM を用いてパケットを送信する。動作概要は図 3 に示す。

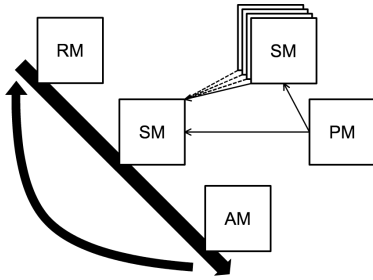


図 3 2DRR 動作概要図

#### 4. The Repetitive two Dimensional Round-Robin Scheduler

2DRR スケジューラに動作を加えて改良を行ったアルゴリズムが Re-2DRR スケジューラである。加えた動作には、2種の閾値が必要であり、Sub-AM というマトリクスが追加された。

まず、Re-2DRR スケジューラでは閾値 C の空パケット容認数と、閾値 R の Sub-AM を再度作成する上限回数である、2種類の閾値を追加し改良した動作に使っている。

図 4 は Re-2DRR スケジューラの動作概要図であり、赤く示している部分が追加された動作である。Re-2DRR スケジューラはまず、2DRR スケジューラと同様の手順で AM を作成する。その作成した AM に閾値 C 以上の空パケットがある場合、次のタイムスロットで使用する PM の参照順で複数個ある SM の選定と SM のマスキングパターンの参照順を利用し Sub-AM を作成する。この Sub-AM が AM より少ない空パケット数であった場合、スケジューラは Sub-AM にてパケットを送信する。もし、Sub-AM が AM より空パケット数が多いか、同じ個数である場合は再度別の PM の参照順にて Sub-AM を作り直し AM と比較する。ただし、この再度 Sub-AM を作成し直す回数には上限があり、その上限数が閾値 R である。もし、閾値 R 回 Sub-AM を作成しても AM より良く成らなかった場合は、スケジューラは AM を用いてパケットを送信する。

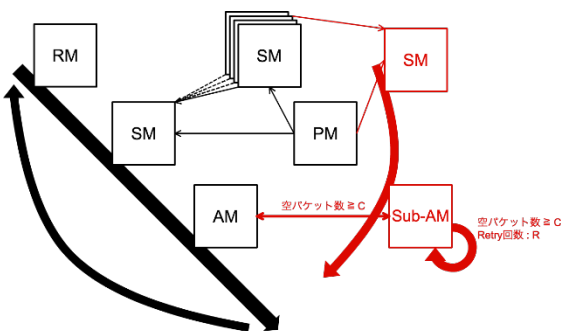


図 4 Re-2DRR 動作概要図

#### 5. Re-2DRR の閾値分析

入出力が 4 の場合の表 1 より、Sub-AM を再度作成する回数の上限である各閾値 R の設定で、Making Sub-AM と Sending by Sub-AM の回数を抑制しても、End Time と Ave Packet Wait Time に大きな変動は観られない。これより、一様でパケットを生成した場合、閾値設定にて Sub-AM が作成される回数を抑制しながら通信を行うことができる。

表 1 N=4 の時の Re-2DRR スケジューリング結果

C (容認する 無駄の数)	R (Sub-AM を再度 作成す る上限 回数)	Making Sub-AM	Sending by Sub-AM	End Time	Ave Packet Wait Time
1	1	18833	4315	277192	138556
	2	33922	5494	277004	138433
	3	49911	6227	277325	138635
	4	61212	6655	276782	138430
2	4	29519	1593	277045	138485
3	4	16608	386	277130	138545

#### 6. 特殊なシステム

本研究では、以下の種類の特特殊なシステムを定義する。

- (1) N より少ないポート間、または決められたポート間に通信が多い

まず、N より少ないポート間に通信が多いというのは、図 5 の左側(a)のマトリクスを例に挙げると、1 が示している箇所の通信が多く、そこにトラフィックが集中しているシステムである。これは、1 対 1 通信などを想定し、常時、1 対 1 が高負荷にて通信をすることを仮定する。

また、決められたポート間に通信が多いというのは、図 5 の中央(b)のマトリクスのように、N 個のポート間の通信が決められて多く、トラフィックが集中しているというシステムである。これは、N 対 N 間の通信を想定し、常時、N 個の端末から N 個の端末へ個別に固定の通信を高負荷にて行っていることを仮定する。

- (2) 1 つのインポートから多数のアウトポート、または多数のインポートから 1 つのアウトポートに通信が多い

これは、図 5 の右側(c)のマトリクスのように、1 つのインポートから全てのアウトポート、または全てのインポートから 1 つのアウトポートに通信が多く、トラフ

ックが集中しているシステムである。これは、ブロードキャストなどを仮定し、1 個の端末から不特定多数の端末へ高負荷にて通信を行うことを仮定する。

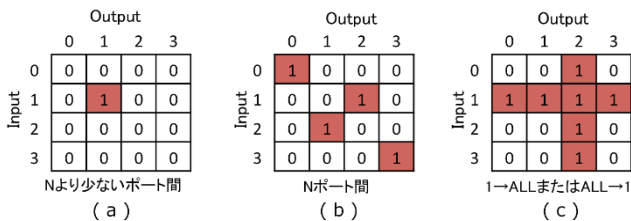


図 5 特殊システムにて負荷のかかるポート

## 7. 提案手法

本研究では、提案手法として下記の 2 種を提案する。

### (1) Sub-AM の候補群の競争的な選択

この手法では、Re-2DRR スケジューラで追加した動作の Sub-AM を作る部分を改良する。Re-2DRR スケジューラでは、AM で空パケット数が閾値 C より多かった場合 Sub-AM を 1 つ作成するが、この手法では、Sub-AM を新しく設定する閾値 S 個分 Sub-AM を作成し、その複数個 Sub-AM から空パケット数がより少ないものを選び利用する。また、この手法では複数個の Sub-AM は 1 度作成するのみであり、元の Re-2DRR スケジューラのように再度 Sub-AM を作成することは無い。そのため、閾値 R は必要無くなる。

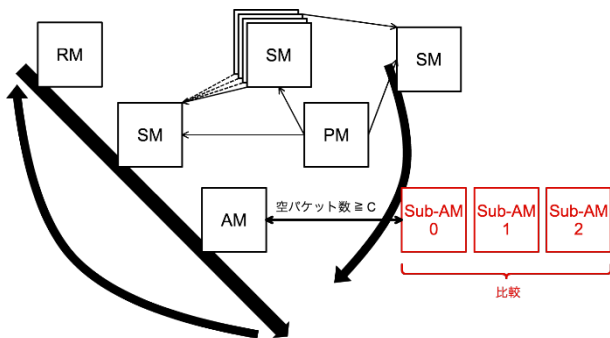


図 6 Sub-AM の候補群の競争的な選択手法概要図

### (2) 優先ポート設定

この手法は、優先ポートを設定し、そのポートのリクエストが優先的に送信許可を得られる様に、Re-2DRR スケジューラの AM または Sub-AM を選択する条件を変更したものである。今までの条件は、より空パケットが少ない方を選択するというものであるが、この手法では任意に指定した優先ポートに送信許可が得られるならば、空パケット数が多くてもその AM または Sub-AM を選択するという条件に変更する。

## 8. シミュレーション結果及び考察

上記にて定義した特殊なシステムにて、Re-2DRR と提案した新しい 2 種類の手法にてシミュレーションを行った。

表 2, 表 3, 表 4 は N=4 の時のそれぞれの特特殊なシステムで特定のポートを負荷をかけた状態でのシミュレーション結果である。結果としては、N ポートより少ないポート間の通信では、優先ポート設定の手法を用いると、負荷がかかっているポートのスループットを向上することが出来る。また、Sub-AM の候補群の競争的な選択手法を用いると、Re-2DRR にて負担と成っていた繰り返し動作を無くしながら、Re-2DRR と同等の性能を出していることが解った。

また、N ポート間での特殊なシステムの通信では、優先すべきポートが増えてしまったため、優先ポート設定手法は性能を向上させることが出来ないが、Sub-AM の候補群の競争的な選択手法を用いると、負担と成っていた動作をなるべく押さえながら、Re-2DRR の性能に近かったことが確認出来た。

最後に、1 ポートから全ポート、全ポートから 1 ポートにて負荷のかかる特殊なシステムでの通信では、より優先すべきポートが増えてしまったため、優先ポート設定手法は動作が正常に行われていなかった。そのため、優先ポートを減らし負荷のかかっているポートの一部だけを優先ポートとしてシミュレーションした結果、優先ポートとして設定されたポートのスループットをあげることは出来た。だが、全体的な性能が低下してしまっている。

表 2 N ポートより少ないポート間でのシミュレーション結果表 (N=4)

	Making Sub-AM	Sending by Sub-AM	End Time	Ave Packet Wait Time	Make Sub-AM group	(1, 1) Ave Wait	(1, 1) MAX Wait Time
Re-2DRR (C:1 R:4)	543976	6908	398402	1561.236		2667	4456
Re-2DRR (C:1 R:3)	414737	6042	399563	1564.182		2677	4516
Sub-AM group (C:1 G:4)	546784	6077	394756	1553.451	136696	2643	4735
Sub-AM group (C:1 G:3)	418731	5916	397935	1560.104	139577	2662	4541
Priority (R:4)	1134177	136500	396138	1670.850		1106	2876
Priority (R:3)	583186	115229	397203	1659.550		1560	4093

表 3 N ポート間でのシミュレーション結果表 (N=4)

	Making Sub-AM	Sending by Sub-AM	End Time	Ave Packet Wait Time	Make Sub-AM group
Re-2DRR (C:1 R:4)	332910	67320	423527	1993.703	
Re-2DRR (C:1 R:3)	257648	55580	424641	1995.891	
Sub-AM group (C:1 G:4)	386220	39035	423077	1989.144	96555
Sub-AM group (C:1 G:3)	285531	36749	425153	1998.776	139577
Priority (R:4)	90322	73489	440118	2047.618	
Priority (R:3)	90400	73511	445007	2065.108	

表 4 1 ポートから全ポート,全ポートから 1 ポート負荷時の結果表 (N=4)

	Making Sub-AM	Sending by Sub-AM	End Time	Ave Packet Wait Time	Make Sub- AM group
Re-2DRR (C:1 R:4)	1857306	99209	796028	2991.344	
Re-2DRR (C:1 R:3)	1430890	95902	799003	2998.931	
Sub-AM group (C:1 G:4)	2162800	96774	802024	3007.809	540700
Sub-AM group (C:1 G:3)	1640709	95880	807698	3024.896	546903
Priority (R:4)	0	0	801634	3133.948	
Priority (R:3)	0	0	806794	3133.948	
Priority (R:4) limit	73350	3689	799421	3103.385	

## 9. まとめ

本研究では Re-2DRR スケジューラの閾値変動による特性を明らかにし,3 種の特殊なシステムを定義した上で,従来手法である Re-2DRR スケジューラと,その Re-2DRR スケジューラの一部改良したアルゴリズム 2 種を提案し,それぞれの特殊なシステムにおいてシミュレーションを行い,各特殊なシステムの通信において,より効率の良いスケジューラを確認した.

## 謝辞

研究の一部は JSPS 科研費 24500046 の助成を受けたものである.

## 参考文献

- [1] 高城衛, 木村成伴, 海老原義彦: マルチアクセス通信のためのグループ化 2 次元ラウンド ロビンスケジューラの提案, 電子情報通信学会論文誌 B, Vol. J82-B, No. 4, pp. 560-568, (1999).
- [2] Craig Partridge et al., : A 50-Gb/s IP Router, IEEE/ACM Transactions of Networking, Vol. 6, No. 3, pp. 237-248, (1998).
- [3] 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).
- [4] 大森一智, 兪明連, 横山孝典, : 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, 情報処理学会 第 75 回全国大会, 1L-4, pp. 205-206, (2013).
- [5] M. J. Karol, M. G. Hluchyj and S. P. Morgan, : Input versus queueing on a space-division packet switch, IEEE Transaction Communications, Vol.COM-35, No.12, pp. 1347-1356, (1987)