

(51)Int.Cl.⁶
H04L 12/56

識別記号

F I

H04L 11/20

102

Z

審査請求 有 請求項の数 1 O L (全6頁)

(21)出願番号 特願平10 - 71814

(22)出願日 平成10年(1998)3月20日

特許法第30条第1項適用申請有り 1997年9月24日 社
団法人情報処理学会発行の「第55回(平成9年後期)全
国大会講演論文集(3)」に発表

(71)出願人 391012327

東京大学長

東京都文京区本郷7丁目3番1号

(72)発明者 斉藤 忠夫

神奈川県横浜市神奈川区松が丘55 - 4

(72)発明者 相田 仁

神奈川県川崎市宮前区宮崎2 - 12 - 1 宮
崎台プラザビル203号

(72)発明者 青木 輝勝

埼玉県上福岡市北野1 - 7 - 2

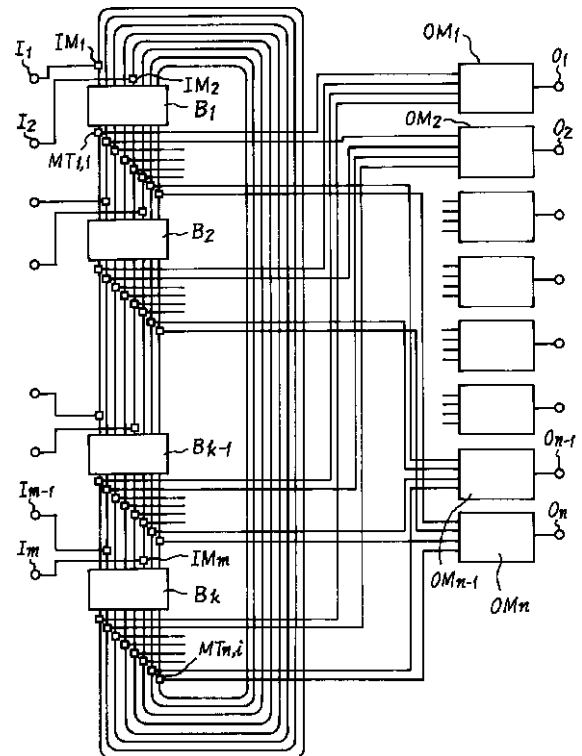
(74)代理人 弁理士 杉村 暁秀 (外8名)

(54)【発明の名称】大容量可変長パケット交換に適したパケットスイッチ方式

(57)【要約】

【課題】 パケットのスイッチング遅延および遅延揺ら
ぎが減少するように適切に構成した大容量可変長パケ
ット交換に適したパケットスイッチ方式を提供する。

【解決手段】 複数k個のバニヤン網を少なくともn本
の並列ラインによってリング型に接続し、入力ポートを
入力モジュールを介して各バニヤン網の入力部に分散し
て接続し、出力ポートを出力モジュールを介して各バニ
ヤン網の対応する出力部に設けたミスルートタグチェ
ック部に接続し、前記バニヤン網を構成する単位スイ
ッチを、パケットが希望する出力部に出せない場合に
はミスルートタグをセットし、このパケットを空いて
いる出力部に出すように構成し、前記ミスルートタグ
チェック部を、入力されるパケットのミスルートタグ
をチェックし、これがセットされている場合には次段
のバニヤン網側へ、セットされていない場合には前
記出力モジュール側へパケットのスイッチングを行
うように構成した。



【特許請求の範囲】

【請求項 1】 複数 m 個の入力ポートから複数 n 個の出力ポートへパケットをスイッチングするパケットスイッチ方式において、複数 k 個のバニヤン網を少なくとも n 本の並列ラインによってリング型に接続し、前記入力ポートを入力モジュールを介して各バニヤン網の入力部に分散して接続し、前記出力ポートを出力モジュールを介して各バニヤン網の対応する出力部に設けたミスルートタグチェック部に接続し、前記バニヤン網を構成する単位スイッチを、パケットが希望する出力部に出せない場合にはミスルートタグをセットし、このパケットを空いている出力部に出すように構成し、前記ミスルートタグチェック部を、入力されるパケットのミスルートタグをチェックし、これがセットされている場合には次段のバニヤン網側へ、セットされていない場合には前記出力モジュール側へパケットのスイッチングを行うように構成し、前記入力モジュールを、前記入力ポートからのパケットおよびミスルートタグチェック部からのパケットを格納し、これらのパケットを選択的に次段のバニヤン網に入力させるように構成したことを特徴とするパケットスイッチ方式。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】本発明は、複数の入力ポートから複数の出力ポートへパケットをスイッチングするパケットスイッチ方式に関するものである。

【 0 0 0 2 】

【従来の技術】共通バス方式は I P ルータとして現在最も主流となっている方式であり、ほとんどの商用 I P ルータで用いられている。この共通バス方式のパケットスイッチの簡略図を図 1 に示す。各入力ポートはラウンドロビン方式によって共通バスの使用権を獲得し、この共通バス上にパケットを送り出す（ブロードキャストする）。出力ポートは、このブロードキャストされたパケットのうち自分宛のものだけを受け取る。したがって、ルータ内のパケット損失を 0 にするためには、共通バスに（各ポートの伝送速度）×（ポート数）以上の動作速度が必要になる。

【 0 0 0 3 】共通バス方式は、可変長パケットの交換が容易であり、ルータを構成する回路素子数を少なくすることが可能であるなどの利点があるものの、入出力ポート数に比例して共通バスの伝送速度を大きくしなければならぬ。したがって、この共通バス方式で交換容量が 1 T b p s (T e r a b i t p e r s e c o n d) を越えるような大容量スイッチを実現することは、素子の物理的限界から非常に困難である。

【 0 0 0 4 】バニヤン網は図 2 に示すように単位スイッチを多段接続することにより大容量スイッチを実現する方式であり、各入力ポートからすべての出力ポートに対して 1 本ずつの経路が存在している。

【 0 0 0 5 】バニヤン網は網内の各単位スイッチが並列的に動作するため、高速大容量化を実現する方法として適している。しかしその反面、図 3 に示すように、パケットの同一入線によって内部ブロッキングが発生してしまう問題点がある。なお、図 2、3 では単位スイッチとして 2 × 2 単位スイッチを用いた場合を取り上げたが、4 × 4、8 × 8、1 6 × 1 6 などの単位スイッチを用いても同様の接続が可能である。

【 0 0 0 6 】タンデムバニヤン網はバニヤン網における内部ブロッキングを低減するための一手法であり、図 4 に示すように複数のバニヤン網を一行に配置したスイッチである。タンデムバニヤン網ではパケットをすべて 1 段目のバニヤン網（図 4 の一番左のバニヤン網）に入力させるが、もしバニヤン網で内部ブロッキングが発生した場合には、そのパケットのミスルートタグをセットし、空いている出力ポートに出力させる。そして 1 段目のバニヤン網を通過後に続けて 2 段目のバニヤン網に入力させる。一方、1 段目のバニヤン網で無事に希望する出力ポートに出られたパケットは、2 段目以降のバニヤン網を通らずにそのまま出力バッファに入れられる。

【 0 0 0 7 】以上のように、内部ブロッキングが発生したパケットは希望する出力ポートに出られるまで次段のバニヤン網に入力することができるため、バニヤン網で発生する大部分の内部ブロッキングを回避することが可能である（タンデムバニヤン網において内部ブロッキングが発生するのは 1 段目から最終段まですべてのバニヤン網でミスルートタグがセットされた場合のみである）。

【 0 0 0 8 】

【発明が解決しようとする課題】タンデムバニヤン網方式は、バニヤン網における内部ブロッキングを低減するための手段として有効である。しかし、タンデムバニヤン網方式では、はじめすべてのセルを 1 段目のバニヤン網に入力させ、1 段目で内部ブロッキングが発生したパケットは引き続き 2 段目以降のバニヤン網に入力させてゆく。したがって入力トラヒックが大きい場合には、1 段目のバニヤン網での内部ブロッキングの発生率が高いため、2 段目以降のバニヤン網に入力するパケットの確率がかかなり高くなってしまい、パケットのスイッチング遅延および遅延ゆらぎの増加を引き起こす。さらに、遅延ゆらぎの増加に伴い、パケットの順序狂い率も増加してしまうという問題がある。

【 0 0 0 9 】本発明の目的は、上述した従来のタンデムバニヤン網方式における問題を解決し、内部ブロッキングの発生率を抑えることにより、パケットのスイッチング遅延および遅延揺らぎが減少するように適切に構成した、大容量可変長パケット交換に適したパケットスイッチ方式を提供することにある。

【 0 0 1 0 】

【課題を解決するための手段】上記の目的を達成するた

め、本発明によるパケットスイッチ方式は、複数 k 個のバニヤン網を少なくとも n 本の並列ラインによってリング型に接続し、前記入力ポートを入力モジュールを介して各バニヤン網の入力部に分散して接続し、前記出力ポートを出力モジュールを介して各バニヤン網の対応する出力部に設けたミスルートタグチェック部に接続し、前記バニヤン網を構成する単位スイッチを、パケットが希望する出力部に出せない場合にはミスルートタグをセットし、このパケットを空いている出力部に出すように構成し、前記ミスルートタグチェック部を、入力されるパケットのミスルートタグをチェックし、これがセットされている場合には次段のバニヤン網側へ、セットされていない場合には前記出力モジュール側へパケットのスイッチングを行うように構成し、前記入力モジュールを、前記入力ポートからのパケットおよびミスルートタグチェック部からのパケットを格納し、これらのパケットを選択的に次段のバニヤン網に入力させるように構成したことを特徴とする。

【 0 0 1 1 】本発明によるパケットスイッチ方式においては、図 5 に示すように複数のバニヤン網をリング状に接続し、入力ポートを各バニヤン網に分散している。これにより、入力トラヒックが分散され、各バニヤン網への見かけ上の入力トラヒック量を低減することができる。例えばバニヤン網の平面数が N であるとする、タンデムバニヤン網方式では総トラヒック量が 1 段目のバニヤン網に入力するのに対し、本方式では各バニヤン網にその $1/N$ のトラヒックを入力させればよいことになる。この分散効果により、各バニヤン網での入力トラヒック量が見かけ上減少するため、1 段目のバニヤン網での内部ブロッキング発生率がタンデムバニヤン網と比較して減少し、パケットのスイッチング遅延および遅延ゆらぎを小さくすることができる

【 0 0 1 2 】

【発明の実施の形態】図 6 は、本発明のパケットスイッチ方式によるパケットスイッチの一実施形態を示すブロック図である。このパケットスイッチ 1 は、入力ポート I_h ($h = 1, 2, \dots, m$) に結合した入力モジュール $I M_h$ と、バニヤン網 B_i ($i = 1, 2, \dots, k$) 4 と、ミスルートタグチェック部 $M T_{h,i}$ と、出力ポート O_j ($j = 1, 2, \dots, n$) に結合した出力モジュール $O M_j$ とを具える。この図に示す実施形態においては、 $m = n = 8$ 、 $k = 4$ としているが、他の数としてもよいことはもちろんである。

【 0 0 1 3 】バニヤン網 B_i は、すでに図 2 に示したように、単位スイッチを多段接続することにより大容量スイッチを実現する。各単位スイッチは、通常のバニヤン網とは異なり、パケットが希望する出力ポート O_j に通じる出力部に出せない場合にはミスルートタグをセットし、空いている出力部にこのパケットを出す。図 2 では、 2×2 単位スイッチを用いた場合のバニヤン網を示

しているが、 4×4 、 8×8 、 16×16 、または 32×32 単位スイッチを用いても同様の多段接続が可能である。バニヤン網 B_i の入力部および出力部の数は、少なくとも出力ポートの数 n だけなければならない。これらのバニヤン網を少なくとも n 本のラインでリング状に接続する。この図においてパケットは、リング状に接続されたバニヤン網を反時計方向に進む。

【 0 0 1 4 】ミスルートタグチェック部 $M T_{h,i}$ は、 1×2 スwitchの機能を持つ。入力パケットのミスルートタグをチェックし、これがセットされている場合には次段のバニヤン網 B_i 側へ、セットされていない場合には出力モジュール $O M_j$ 側へパケットのスイッチングを行う。

【 0 0 1 5 】入力モジュール $I M_h$ は、図 8 に示すように 2 つの $F I F O$ ($F i r s t I n F i r s t O u t$) バッファ 1 1 および 1 2 とセクタ 1 3 とを具える。2 つの $F I F O$ バッファ 1 1 および 1 2 はそれぞれ、入力ポート I_h またはミスルートタグチェック部 $M T_{h,i}$ から転送されてきたパケットを格納する。セクタ 1 3 は、2 つの $F I F O$ バッファ 1 1 および 1 2 の一方からパケットを選択し、次段のバニヤン網 B_i に入力させる。本実施形態では、入力ポート I_h からのパケットを低優先、ミスルートタグチェック部 $M T_{h,i}$ からのパケットを高優先と定義しているため、セクタ 1 3 は、ミスルートタグチェック部 $M T_{h,i}$ からのパケットが $F I F O$ バッファ 1 1 に存在する場合にはそのパケットを優先的に次段のバニヤン網 B_i に入力させる。

【 0 0 1 6 】また、入力ポート I_h 、ミスルートタグチェック部 $M T_{h,i}$ の双方から 100% の入力トラヒックが存在する場合には、 $F I F O$ バッファ 1 1 および 1 2 がオーバフローを起こす可能性がある。したがって、入力モジュール $I M_h$ をはじめとするバニヤン網のリング内部のすべての装置を、入力ポート I_h の 2 倍の伝送速度で動作させる。

【 0 0 1 7 】出力モジュール $O M_j$ は図 9 に示すように、タイミング調整バッファ 2 1、共通バス 2 2、制御部 2 3 および出力バッファ 2 4 から構成されている。タイミング調整バッファ 2 1 は、各バニヤン網 B_i から出力したパケットを格納する。制御部 2 3 は共通バス 2 2 を管理し、各タイミング調整バッファ 2 1 にラウンドロビン方式によって共通バス 2 2 の使用权を与える。タイミング調整バッファ 2 1 に格納されたパケットは、この使用权を獲得した時に出力バッファ 2 4 に移される。

【 0 0 1 8 】パケットを入力ポート I_h から入力し、スイッチングして、希望する出力ポート O_j に出力するまでの流れは次の通りである。

【 0 0 1 9 】まずパケットは入力ポート I_h から入力され、入力モジュール $I M_h$ 内の $F I F O$ バッファ 1 2 に格納される。入力モジュール $I M_h$ では、入力ポート I_h からのパケットよりもミスルートタグチェック部 $M T$

h_i からのパケットを優先するため、このパケットはミスルートタグチェック部 $M T_{h_i}$ からのパケットが F I F O バッファ 1 1 内に存在しない場合にのみセクタ 1 3 を通ってバニヤン網 B_i に入力される。

【 0 0 2 0 】バニヤン網 B_i に入力したパケットは、単位スイッチ間をスイッチングしてゆくが、もし希望する出力ポート O_j に通じる出力部が使用中の場合には、パケットヘッダ内のミスルートタグをセットされ、使用されていない出力部に送られる。バニヤン網 B_i を通過したパケットは続いてミスルートタグチェック部 $M T_{h_i}$ に渡される。

【 0 0 2 1 】ミスルートタグチェック部 $M T_{h_i}$ では、パケットにミスルートタグがセットされているかどうかをチェックし、セットされている場合には引き続き次段のバニヤン網 B_i に入力させる。一方、ミスルートタグがセットされていない場合には、出力モジュール $O M_j$ に転送される。

【 0 0 2 2 】出力モジュール $O M_j$ に入ったパケットはタイミング調整バッファ 2 1 に格納され、共通バス 2 2 の使用権を獲得するまで待機する。そして共通バス使用権が回ってきたら、パケットは共通バス 2 2 を通じて出力バッファ 2 4 に入れられ、出力ポート O_j を経て本パケットスイッチ 1 から出力される。

【 0 0 2 3 】従来のタンデムバニヤン網方式のパケットスイッチと本方式のパケットスイッチとをシミュレーションにより評価した結果を示す。ここで、使用するバニヤン網を、図 1 0 に示すような 4×4 単位スイッチを用いた $6 4 \times 6 4$ バニヤン網とし、ポート伝送速度を $1 5 5 . 5 2 \text{ M b p s}$ とし、入力負荷を $9 0 \%$ とし、I P データグラムを $2 1 \sim 1 5 0 0 \text{ o c t e t}$ の可変長とした。

【 0 0 2 4 】図 1 1 にバニヤン網の平面数 - パケット損失率特性を示す。図 1 1 では、バニヤン網の段数の増加に伴ってタンデムバニヤン網方式、本方式ともにパケット損失率が減少してゆく様子が示されているが、特に本方式でははじめからトラヒックの分散が行われているため、タンデムバニヤン網方式と比較してパケット損失率が小さくなっている。また、本方式はタンデムバニヤン網方式と比較してパケット損失率の収束が早いため、少ないバニヤン網の段数で高性能化を図ることができる。これは設計の立場に立った場合、大きな長所となるであろう。

【 0 0 2 5 】一方、図 1 2 では、6 段のバニヤン網を用いたタンデムバニヤン網方式と本方式に対して、入力パケットが何段目のバニヤン網で出力できたかを示している。

【 0 0 2 6 】図 1 2 では、本方式ではほとんどのパケットが 3 段目以内のバニヤン網でスイッチングを完了しているのに対し、タンデムバニヤン網方式では 5 段目のバニヤン網に入力するパケットが $1 0 \%$ 近くもあり、本方

式と比較して、遅延および遅延ゆらぎが大きくなっていることがわかる。

【 0 0 2 7 】

【発明の効果】本発明によれば、入力ポートをあらかじめ分散しておくことによって、入力トラヒックの分散を図り、各バニヤン網への見かけ上の入力トラヒック量を低減することができる。これにより、例えばバニヤン網の平面数が N であるとする、タンデムバニヤン網方式では総トラヒック量が 1 段目のバニヤン網に入力するのに対し、本方式では各バニヤン網にその $1 / N$ のトラヒックを入力させればよいことになる。この分散効果により、各バニヤン網での入力トラヒック量が見かけ上減少するため、1 段目のバニヤン網での内部ブロッキング発生率がタンデムバニヤン網と比較して減少し、パケットのスイッチング遅延および遅延ゆらぎを小さくすることができる。したがって本発明によれば、大容量可変長パケット交換に適したパケットスイッチ方式が提供される。

【図面の簡単な説明】

【図 1】従来の共通バス方式のパケットスイッチを示す線図である。

【図 2】バニヤン網を示す線図である。

【図 3】バニヤン網における内部ブロッキングを説明する線図である。

【図 4】タンデムバニヤン網を示す線図である。

【図 5】本発明による、バニヤン網をリング状に接続したパケットスイッチを図式的に示す線図である。

【図 6】本発明のパケットスイッチ方式によるパケットスイッチの一実施形態の構成を示すブロック図である。

【図 7】図 6 のパケットスイッチにおけるミスルートタグチェック部の構成を示すブロック図である。

【図 8】図 6 のパケットスイッチにおける入力モジュールの構成を示すブロック図である。

【図 9】図 6 のパケットスイッチにおける出力モジュールの構成を示すブロック図である。

【図 1 0】 4×4 単位スイッチを用いた $6 4 \times 6 4$ バニヤン網を示す線図である。

【図 1 1】バニヤン網の平面数 - パケット損失率特性を示すグラフである。

【図 1 2】入力パケットが何段目のバニヤン網で出力できたかを示すグラフである。

【符号の説明】

1 パケットスイッチ

1 1、1 2 F I F O バッファ

1 3 セクタ

2 1 タイミング調整バッファ

2 2 共通バス

2 3 制御部

2 4 出力バッファ

5 0 I_h 入力ポート

IM_n 入力モジュール

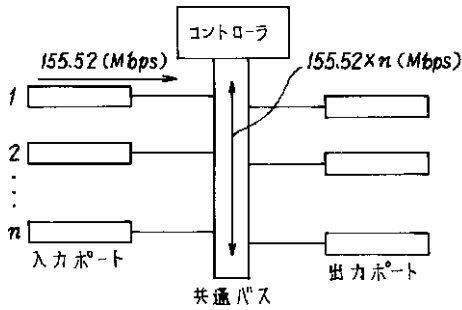
B_i バニヤン網

MT_{n,i} ミスルートタグチェック部

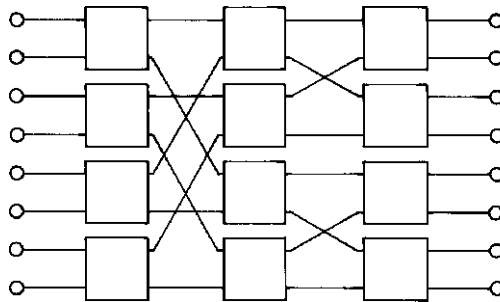
OM_j 出力モジュール

O_j 出力ポート

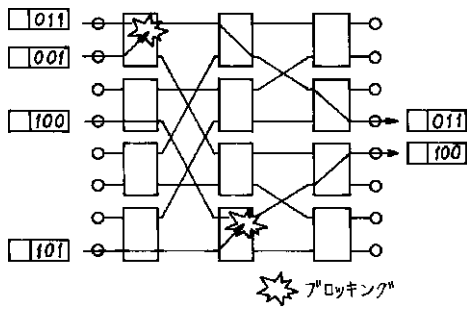
【図 1】



【図 2】



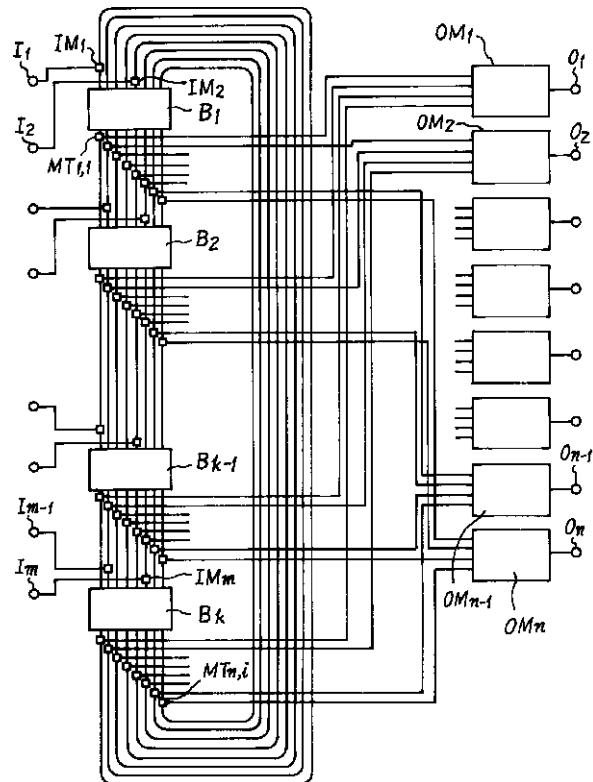
【図 3】



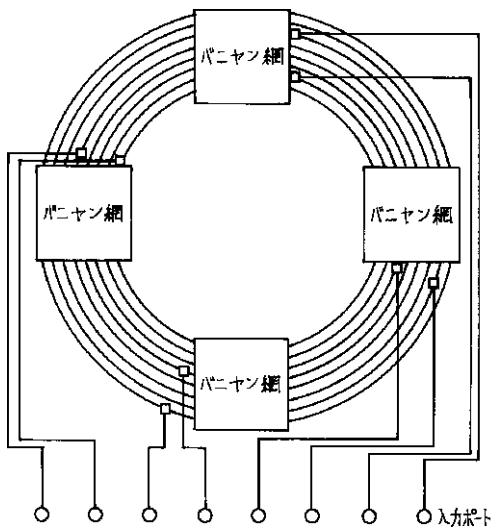
【図 4】



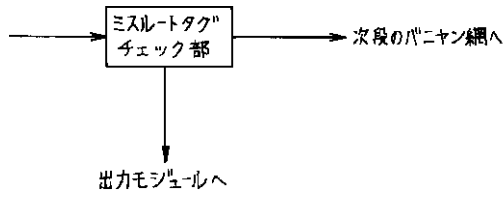
【図 6】



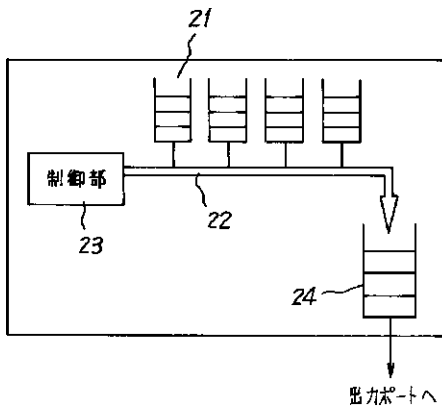
【図 5】



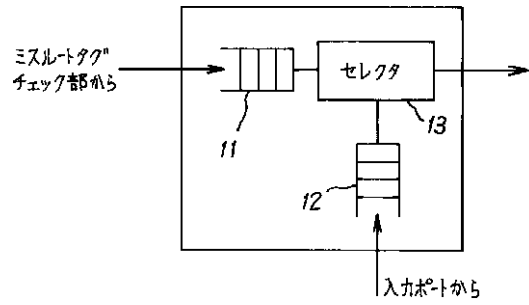
【 図 7 】



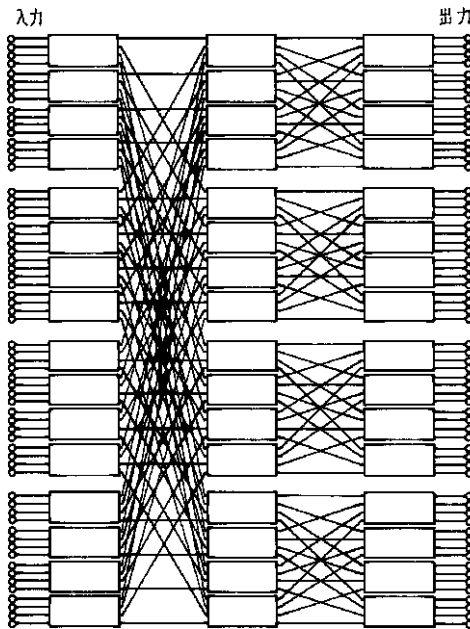
【 図 9 】



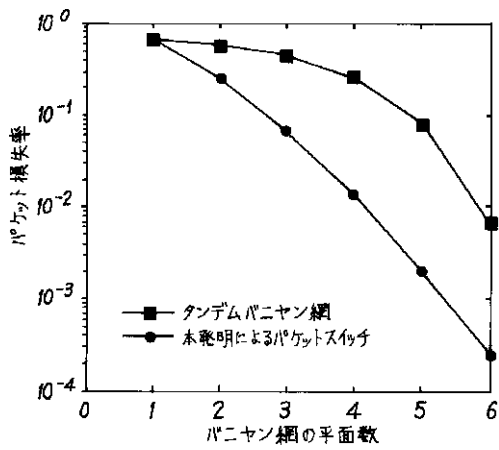
【 図 8 】



【 図 1 0 】



【 図 1 1 】



【 図 1 2 】

