

入出力バッファ型 20 Gbps ATM バックボーンスイッチアーキテクチャ

Input and Output Buffered Switch Architecture for High-Speed ATM Backbone Nodes

釜谷 幸男
Y. Kamatani

加沼 安喜良
A. Kanuma

ATM (非同期転送モード: Asynchronous Transfer Mode) は高速のマルチメディアネットワーク技術として注目されている。近年、当社をはじめ国内外のベンダから LAN 向けの ATM スイッチ、インタフェースなど製品が売り出され、企業、大学などで運用されはじめています。より高速で大規模な企業またはキャンパスネットワークを構築するには、LAN 間の相互接続を実現する大容量のバックボーンスイッチノードが必要になる。

当社では、このようなニーズにこたえるために、622 Mbps を基本リンク速度とし、トータルスループットが 5 Gbps から 20 Gbps まで拡張でき、かつ、最新の QOS (Quality of Service) 機能、マルチキャスト機能などを搭載するスイッチ部のアーキテクチャを開発した。現在、それに基づくチップセットの一部は試作済みであり、その他のチップについても開発中である。

Asynchronous transfer mode (ATM) is a promising bearer transmission technology for high-speed multimedia LANs. Recently, ATM LAN systems have been applied as high-speed multimedia networks. To realize a large-scale corporate or campus ATM network, high-speed multimedia backbone nodes interconnecting the ATM LANs will be required.

Toshiba is focusing attention on a switch architecture for multimedia backbone nodes which supports an expandable high throughput exceeding 10 Gbps, multiple qualities of service (multi-QOS), multicast, and 622 Mbps interface functions. This paper describes this architecture, as well as an ATM switch chip set based on this architecture.

1 まえがき

ATM は、高速マルチメディアネットワーク技術として有望である。その主な理由は、高速であることと、複数のサービスクラスをサポートできることにある。これらの機能により、動画像をはじめとする広帯域の実時間トラフィックから、ファイルなど巨大データのバーストトラフィックまで一元的に転送サービスできる。

近年、当社の AX1500 をはじめとする ATM-LAN 用スイッチあるいはインタフェースなど、ATM-LAN 用の製品群が国内外の多数のベンダから販売されており、企業、研究機関などで導入され始めている。図 1 に示すように、フロア、あるいは組織などを単位とする各ワークグループ LAN は 25 Mbps スイッチや、155 Mbps スイッチで構成される。それらは、さらにバックボーンスイッチで構成されるバックボーンネットワークで相互接続される。

ワークグループスイッチの容量はだいたい 2.5 Gbps 以下であるから、それらを相互接続するためのバックボーンスイッチは 10 Gbps 以上のスループットが要求される。また、ATM フォーラムなどで決定される ATM 標準仕様に準拠することも重要である。われわれは、これらの要求に対応できるバックボーンスイッチアーキテクチャとして、入出力バッファ型スイッチアーキテクチャを開発した。このアーキテクチャに基づき、3 チップからなる高速 ATM スイッチ

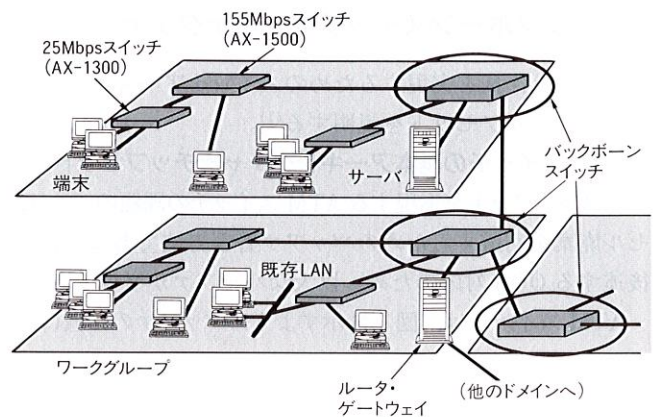


図1. 企業ネットワークのシステムインテグレーション 企業ネットワークの典型的な構成とそこにおけるバックボーンスイッチの設置位置を示す。

System integration of corporate network

チップセットを開発中であり、これらを必要に応じて組み合わせることで 5 Gbps から 20 Gbps までの ATM スイッチが構成できる。

2 バックボーンスイッチの主な仕様

ATM バックボーンスイッチの主な要求仕様を以下に示す。

- (1) 高スループットとその拡張性 現在、製品となっている ATM-LAN 向けスイッチで構成されるワークグループネットワークのスループットは 1.5 Gbps から 2.5 Gbps である。これらのネットワークを相互接続するバックボーンスイッチのスループットとして、5 Gbps 以上 20 Gbps までの拡張性を実現し、高い帯域利用効率を実現する。
- (2) マルチ QOS 対応 マルチメディア応用を収容するスイッチとして必要な品質、具体的には、ATM のデファクト標準決定機関である ATM フォーラムの QOS 仕様に対応した 5 クラスをサポートする。
- (3) マルチキャスト対応 イーサネット^(注1)など従来のメディア共有型 LAN のエミュレーション、あるいは、放送型サービスを提供するネットワークにはブロードキャスト、あるいはマルチキャスト機能が有効である。このスイッチでは高スループットのマルチキャストをサポートする。
- (4) 多様なリンク速度の収容 バックボーンスイッチは帯域の効率的な運用が必要である。そこで、基本リンク速度は 622 Mbps として、すでに使われている 155 Mbps, 51 Mbps, 25 Mbps インタフェースなどを効率よく多重収容する。

3 バックボーンスイッチアーキテクチャ

上述した仕様を実現するためのスイッチアーキテクチャと、そのチップセットを説明する⁽¹⁾。

3.1 スwitchの基本アーキテクチャとチップセット

高スループットを実現する ATM スwitchの課題の一つは、セル廃棄を回避するためのバッファ管理方法である。特に後述する QOS 対応のため、巨大なバッファが必要である。

ATM スwitchは、図 2 に示すようにバッファの位置によ

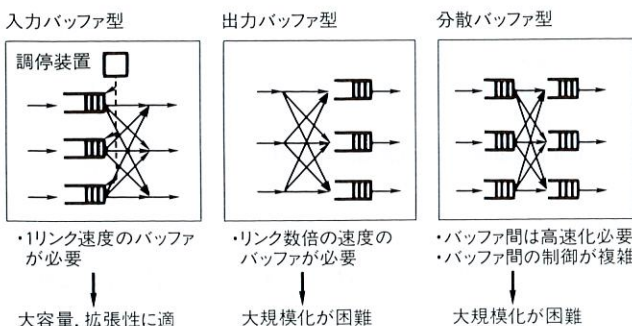


図 2. ATM スwitchのバッファ管理方式 ATM スwitchのバッファ管理方式は、大別して、入力バッファ、出力バッファ、分散バッファの3とおりあるが、入力バッファ型を選択した。

Buffer management for ATM switches

(注1) Ethernet (イーサネット) は、富士ゼロックス株の商標。

って、入力バッファ型、出力バッファ型、分散バッファ型に分類できる。この中でわれわれは 10 Gbps を超えるような大容量のスイッチとして入力バッファ型を選択した。それは、分散型と比べるとバッファを極力集中して管理することにより制御が簡単になり、出力バッファ型と比べると各バッファに必要なスループットが小さくて済むからである。

しかし、入力バッファ型どうしの出力が出力リンクで競合しないようにセル送出を調停する必要がある、2章で述べたスループットの拡張性と高いリンク利用効率の実現には、特にこの調停機能に対して以下のような課題がある。

- (1) 高スループット高リンク帯域利用効率の調停機構 調停装置としては、スループットに応じて拡張でき、帯域を極力むだなく利用する調停機構が必要である。
- (2) 調停装置と協調してセル送出する入力バッファ 調停装置に蓄積セル情報を伝達し、調停装置から指示されたセルを送出する機能が必要である。

これらの課題に対し、われわれは、セル交換出力機能と調停機能を兼ね備えたスイッチ機能 SE (Switch Element) と、出力リンクごとにバッファ管理するインタフェース機能 SAM (Switch Access and Multiplexer/demultiplexer) と、拡張機能 DA (Distributor/Arbiter) とを用いた入出力バッファ型スイッチアーキテクチャを開発した。

以下、図 3 を用いて説明する。SE 内に小型の出力バッファを出力リンクごとに設け、各出力バッファの輻輳(ふくそう)状態を SAM 内の入力バッファにリンクごととフロー制御信号としてフィードバックする。入力バッファはそのフロー制御信号を受けて、非輻輳リンク行きのセルをみずからのバッファの中から選択して送出する。

図 3 のように SE をマトリクス状に並べ、各 SE からの

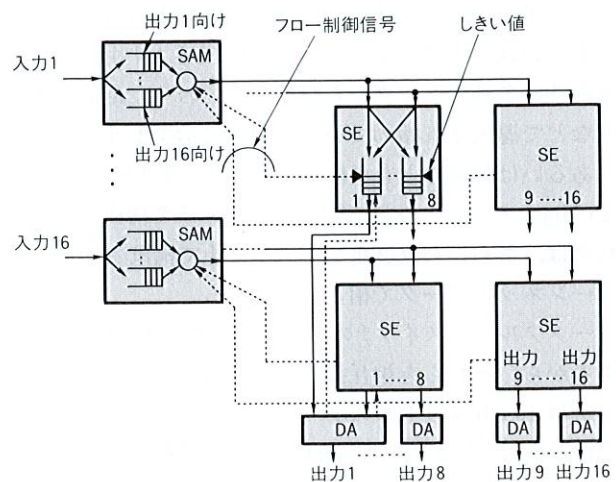


図 3. ATM バックボーンスイッチアーキテクチャ バックボーンスイッチは、巨大入力バッファとしての SAM, その調停装置としての SE と DA で構成される。

ATM backbone switch architecture

フロー制御信号を入力バッファに接続すれば、出力リンク数と調停機能の拡張が同時に果たされる。つまり、リンク数に応じた調停装置の拡張が容易になる。なお、図3のカラム方向に並ぶSEどうしもDA内の出力バッファからのフロー制御信号によって調停される。また、各入力バッファは負荷の小さい出力リンクに対してセルを転送することになるからリンク帯域のむだな空き状態を避け、高いリンク帯域利用率を実現する。

われわれは、スイッチアーキテクチャに用いる上述の三つの要素、SE、SAM、DAをおのおの一つのLSIとして3チップで構成されるチップセットを開発している。

複数のSAMと一つのSEとで5 Gbpsまでのスイッチを構成する。さらに、5 Gbps以上に拡張するにはSEをマトリックス状に配置し、SE間を調停するDAで接続することにより、20 Gbpsまでの拡張性が確保される。図4には拡張時のチップセットの接続を示す。チップセットのうち、5 Gbps (622 Mbps 8リンク分)のCMOS SEチップをすでに試作した⁽²⁾。

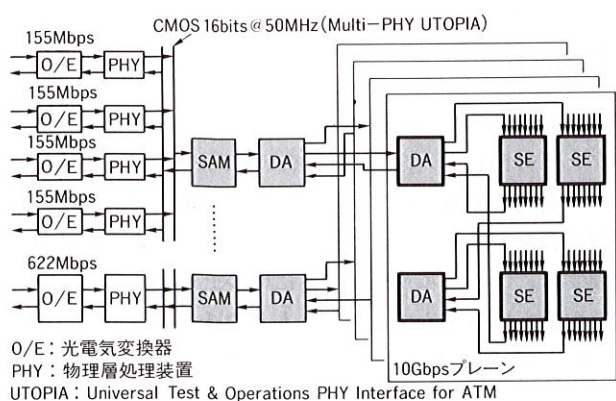


図4. 20Gbps 32×32 ATMスイッチ構成 20 Gbpsの構成にしたときのおのおのチップの接続状態を示す。
20 Gbps 32×32 ATM switch configuration

3.2 マルチQOSへの対応

現在、標準となっているQOSクラスは二つに大別できる。すなわち、遅延時間、セル廃棄確率について品質保証する品質保証クラスと、それら保証のないベストエフォートクラスである。この品質要求の違いにより、ベストエフォートトラヒックには品質保証トラヒックの通過を待って転送されるという待合せが生ずる。そのために数万セル分以上のバッファがスイッチ内に必要とされている。

このアーキテクチャでは、入力リンク部に大きなバッファを実装するのでスループット一定のまま容量を拡張できる。また、各入力バッファを出力リンクごと、クラスごとに、出力バッファをクラスごとに管理し、各出力リンクか

らのフロー制御信号をクラスごとに構成し、各入力バッファは輻輳がなくもっとも優先度の高いクラスのセル流を送出する。よって、リンク帯域のむだがなく、かつ優先度順にセルを交換できる。

スイッチの中心部では品質保証とベストエフォートを合計5クラス実装しており、これで、最新のATMフォーラムの五つのQOS仕様、すなわち、CBR (Constant Bit Rate), rt-VBR (real-time Variable Bit Rate), nrt-VBR (non-real-time VBR), ABR (Available Bit Rate), UBR (Unspecified Bit Rate)に対応できる。

3.3 マルチキャスト

このスイッチアーキテクチャでは、高スループットのマルチキャストをサポートする。その機能は図5に示すように、SAMとSEで階層的にセルコピーを生成することで実現される。すなわち、SAMはSEごとにセルコピーし、SEはリンクごとにセルコピーする機能をもっている。SAMとSEのセルコピー機能を組み合わせることで任意の出力リンクの組合せに対応するマルチキャストを実現する。

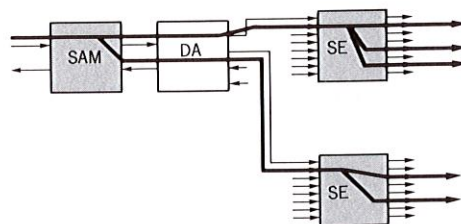


図5. セルマルチキャストサポート このチップセットを用いたマルチキャストのサポートのアーキテクチャ概要を示す。図の例では、SAMで二つにセルコピーし、さらにSEで三つと二つにセルコピーして、合計五つの同じセルが同時に出力される。

Cell multicast support

4 性能評価

ここでは、シミュレーションで求めたスイッチのスループット特性を示す。図6に示すスイッチのモデルと、表1のバックボーンスイッチ設計のための標準トラヒックソースモデルを用い、品質保証クラスとベストエフォートクラスのスループットを評価した。なお、スイッチモデルには低速インタフェース多重用のMUX (Multiplexer), DEMUX (Demultiplexer) 機能も含まれている。パラメータは、スイッチLSIのバッファサイズ B (セル数), フロー制御信号生成のしきい値 Th (セル数), フロー制御信号の伝搬遅延時間 D (セルクロック) である。

4.1 品質保証クラスのシミュレーション

図7にしきい値に対する最大スループットを品質保証クラスの一つであるrt-VBRについて評価した結果を示す。VBR-II型のトラヒックを仮定し、セル廃棄確率は 1×10^{-6}

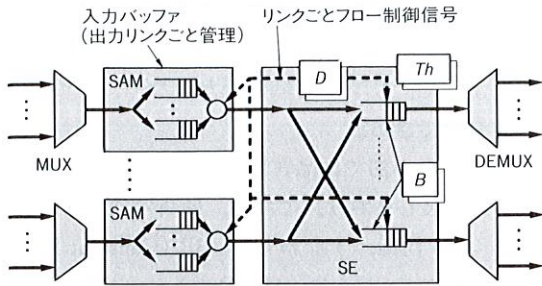


図6. スイッチのシミュレーションモデル バッファサイズ(B), しきい値(Th), フロー制御信号遅延時間(D) がスイッチの性能に及ぼす影響を評価する。

Simulation model for ATM switch

表1. トラフィックソースのシミュレーションモデル

Simulation model for traffic sources

	ピークレート	平均レート	平均バースト長(セル)
VBR-I	1/6	1/24	240
VBR-II	1/25	1/150	500

(レートはリンク帯域で規格化)

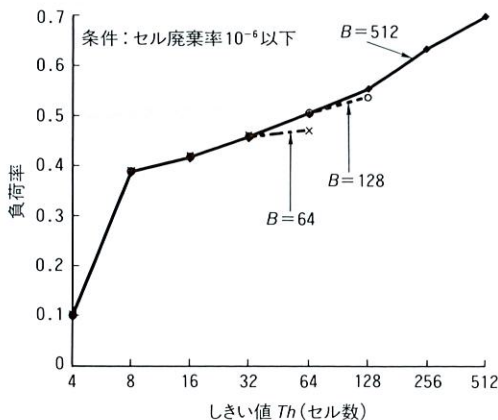


図7. VBRの最大トラフィック量 VBR-IIを用いたVBRトラフィックの最大負荷量を示す。しきい値(Th)を8から256に設定して、負荷率は0.4から0.6程度となる。

Maximum offered load of VBR traffic

以下という条件で注入できる最大トラフィック量をシミュレーション評価した。しきい値を8から256に設定してスループットは0.4から0.6程度となる。VBR-Iでも同様の傾向であった。

4.2 ベストエフォートクラスのシミュレーション

図8には、ベストエフォートクラスについて、その負荷率に対する遅延時間が示されている。負荷は、VBR-II, スイッチLSIのバッファ容量としては各クラスごとに64セル分に設定した場合である。VBR-Iも同様であった。この結果からセル遅延特性は、しきい値, フロー制御信号の遅延

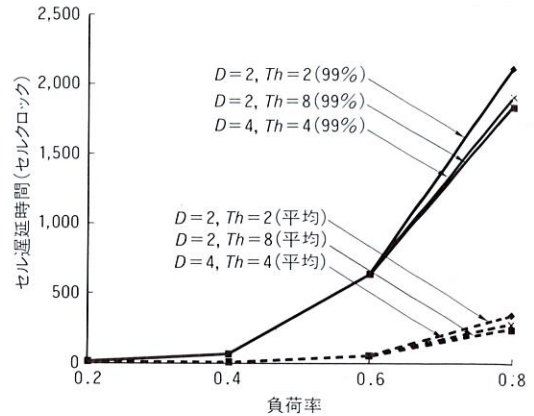


図8. ベストエフォートトラフィックの遅延特性 VBR-IIを用いたベストエフォートクラスの平均と99%遅延時間特性を示す。セル遅延特性は D , Th にほとんど依存しない。

Delay characteristics of best-effort traffic

時間, トラフィックパラメータの違いにはほとんど依存しないことがわかる。

5 あとがき

5 Gbps から 20 Gbps のスループットを実現する入出力バッファ型 ATM バックボーンスイッチアーキテクチャを開発した。高速の出力バッファ型スイッチ機能が各入力バッファ間競合を調停することで、高いスループットと拡張性を実現する。このスイッチはさらにマルチ QOS, マルチキャスト機能も搭載する。このアーキテクチャを, SE, SAM, DA の三つの機能 LSI からなるチップセットとして実現する。このうち SE チップはすでに試作済みであり⁽²⁾, その他のチップについても開発中である。

文献

- (1) Y. Kamatani et al: VLSI-Oriented Input and Output Buffered Switch Architecture for High-Speed ATM backbone Nodes, IEICE TRANS. COMMUN. E79-B, 5, May 1996, pp.647-657
- (2) Y. UneKawa et al: A 5Gbps 8×8 ATM Switch Element CMOS LSI Supporting Five Quality of Service Classes with 200MHz LVDS Interface, ISSCC DIGEST OF TECHNICAL PAPERS, Feb. 1996, FA7.3, pp.118-119



釜谷 幸男 Yukio Kamatani

研究開発センター 情報・通信システム研究所主任研究員。ATM スイッチアーキテクチャの研究開発に従事。電子情報通信学会会員。Communication & Information Systems Research Labs.



加沼 安喜良 Akira Kanuma

マイクロエレクトロニクス技術研究所 システム LSI 技術研究所主幹。マルチメディア用 LSI, 特に ATM 用 LSI の開発に従事。IEEE, 電子情報通信学会会員。System LSI Engineering Lab.