

ストリームシェーパーPureFlow SS10 の開発

Development of Stream Shaper MV0771A PureFlow SS10

茂木正英 Masahide Moteki, 関水宗樹 Muneki Sekimizu

[要 旨]	映像配信サービスにおいて、ストリーミングサーバやリアルタイムエンコーダから配信されるバーストラフィックを出力時間間隔 20μ秒の精度で時間的に一様なトラフィックパターンに整形(シェーピング)するストリームシェーパー PureFlow SS10を開発した。 本装置は、パケット解析を専用のプロセッサで、帯域制御機能にFPGAをそれぞれ使用し、10Gbit/sの処理とマイクロ秒の高精度な帯域制御を実現した。また、音声レベルの8Kbit/sからHDTV信号の伝送を考慮した800Mbit/sまでの幅広い範囲での帯域制御を可能とした。 バーストラフィックをシェーピングにより整形して、高品質で回線効率の良い映像配信サービスを提供できる装置を開発した。本装置は、19インチラックを採用し、高さ2Uで1GbE(ギガビットイーサネット)を5回線収容できるほどに小型化し、省スペース化を図った。
[Summary]	We have developed the PureFlow SS10 stream shaper that shapes burst traffic from a streaming server or a real-time en-coder at a high accuracy of 20 μs. It achieved 10-Gbit/s processing and high-accuracy bandwidth control. Packet analysis is handled by a dedicated processor, and bandwidth is controlled using FPGA, offering a wide range of bandwidth control from voice (8 Kbit/s) to HDTV transmission (800 Mbit/s). We avoid burst traffic by shaping at the transmission side to guarantee QoS, so customers can provide high-quality video distribution at very low cost. The space-efficient design accommodates 5 channels in 19-inch 2U rack.

1 まえがき

FTTHやADSL回線を使った常時接続において、6Mbit/s程度の帯域でDVD並みの高品質映像のストリーミング配信ができるレベルまで達してきた。コンピュータネットワークにおいてもファイル、メール、およびWeb上のデータ転送に加えて、リアルタイム性を有する音声、画像などのストリーミングデータの転送ニーズが高まってきている。特に、TV、ライブ配信、VoD (Video On Demand) サービスが通信事業者、プロバイダ、およびコンテンツ配信事業者の提供サービスとして高まりを見せている。

現状の配信サービスに使用されるストリーミングサーバやリアルタイムエンコーダから配信されるデータ信号は、音声、画像情報の高能率符号化により画像の切り替わりで多量のデータが発生しバースト的に出力されている。配信端末からバースト的に発生されるトラフィックにより、接続先のIPネットワークが輻輳して、ストリーミングを構成するパケットの消失(パケットロス)を引き起こし、映像の乱れや欠落などのサービス品質を低下させる課題を有している。

この課題を解決するために、ストリームシェーパーPureFlow SS10(以下、本装置)は、図2に示すようなストリーミング配信環境において、ストリーミングサーバ群とIPネットワークのルータ間に配置され、送出されたバーストラフィックを、あらかじめ設定した平均レートに基づき、パケット間隔を調整し、時間的に一様なトラ

フィックパターンに整形(シェーピング)することにより、IPネットワークのパケット処理の負荷を軽減して、パケットロスをなくす。また、本装置のシェーピング機能により、IPネットワークに流すトラフィック量を制御できるため、回線のもつ帯域を有効に活用できる利点(帯域制御)が生じる。

図1にPureFlow SS10の概観を示す。

図2にPureFlow SS10を設置した映像配信サービスの例を示す。



図1 PureFlow SS10の概観
External view of MV0771A PureFlow SS10

2 開発方針

本装置の開発方針は以下の3点とした。

また、処理速度に関しては、19インチ、1Uの大きさで現状のハードウェアでの最大処理能力を実現することとし、10Gbit/sを目標とした。

(1) 高精度帯域制御処理

音声信号の 8Kbit/s から HDTV 信号の転送を考慮した 800Mbit/s までの幅広いデータ処理を、複数チャネル同時に行うため、帯域制御としてはビット当りの処理速度を 10Gbit/s とし、μ秒レベルの時間間隔にパケットの配置割付処理を行う。ハードウェアには、高速化処理ができる FPGA を使用する。

(2) 柔軟なパケット解析処理

- ・ 音声、映像のパケットに対応できること。
- ・ IPv4/IPv6 パケットを認識できること。
- ・ フラグメント処理されたパケットにおいてもフラグメント情報を識別し、シェーピング処理が行えること。

パケット解析処理は上記の処理を考慮し、柔軟性を確保するため、プログラム可能な専用ハードウェアを使用して、10Gbit/s 処理を実現する。

(3) 拡張性

今後、映像配信サービスが拡張されストリーミングサービスが増設される場合を考慮して、1 スロット当りの処理能力

を向上させ、本装置に拡張スロットを設ける。今回は、処理能力 10Gbit/s、1GbE (ギガビットイーサネット) を 5 回線収容可能とした。

3 構成

本装置のブロック図を図 3 に示す。

本装置は、19 インチラック搭載型で、パケット処理ユニット、2 スロット実装可能なラインカード、マネジメントユニット、電源ユニットから構成される。1 スロットにラインカードとパケット処理ユニットを実装し、1 スロットで 10Gbit/s を処理可能とした。装置全体として、高さ 2U の省スペースで 1GbE を 5 回線収容可能とした。

(1) パケット処理ユニット

パケット処理ユニットは、パケット処理エンジンと帯域制御エンジンを有し、パケット処理エンジンでストリーミングデータの解析を行い、ストリーミングデータの 1 連のパケットを 1 つのデータフローとして扱う。帯域制御エンジンは、このデータフローごとにキューを割り当ててシェーピング処理を行う。本ユニットは 10Gbit/s のパケット転送処理を行

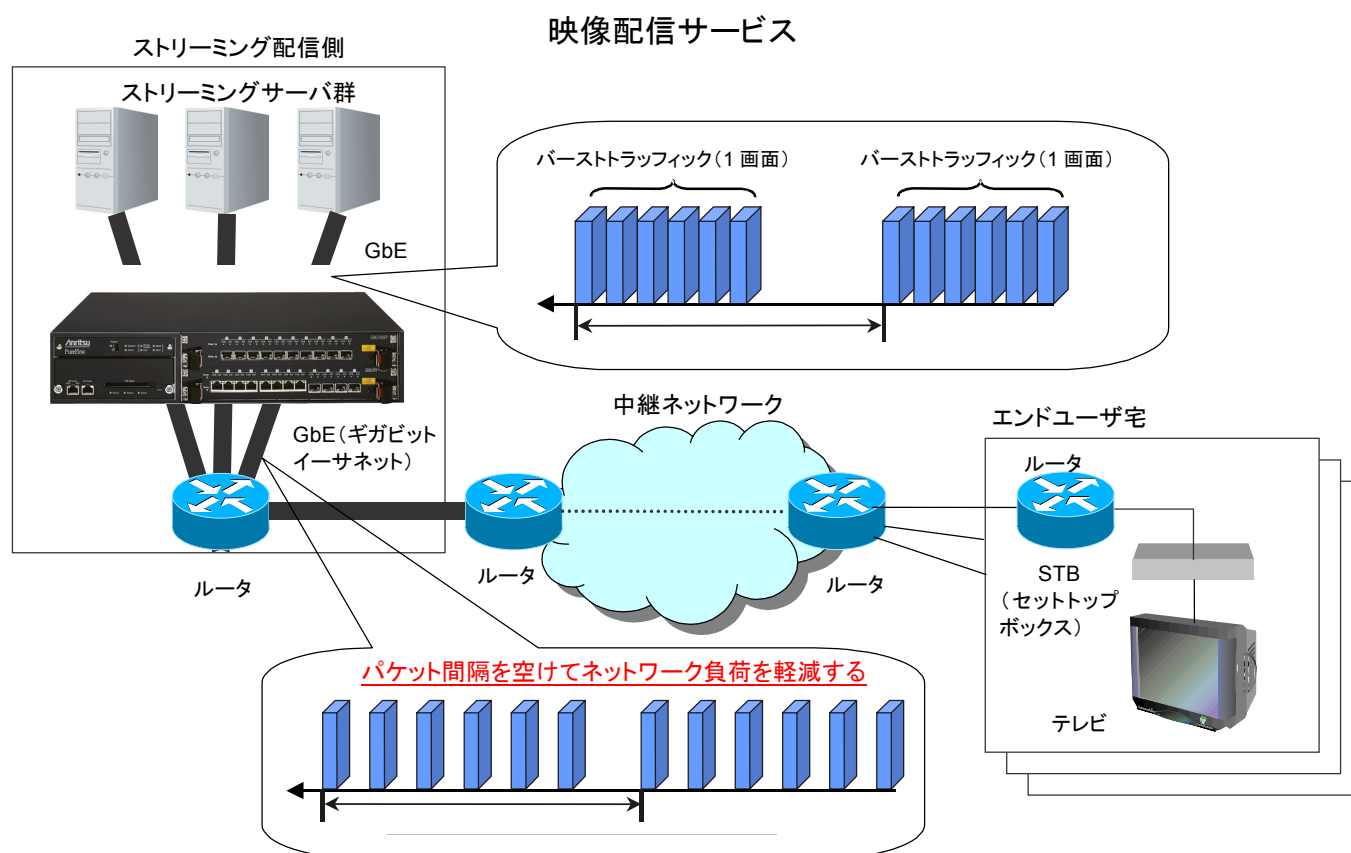


図 2 ネットワーク上の位置づけ
Location in network

うためハードウェアで実現した。(4章で詳説する。)

(2) ラインカード

ラインカードは、1000BASE-X の光インタフェースを 10 ポート有するタイプと 1000BASE-X の光インタフェースを 4 ポート、1000BASE-T の電気インタフェースを 8 ポート有するタイプの 2 タイプ用意した。電気と光のインタフェースを有するタイプは、接続する機器に柔軟に対応するため、メディアコンバータとして使用することも可能とした。

(3) マネジメントユニット

本ユニットは装置の運用設定情報(コンフィグデータ)の設定制御、登録記憶等の装置管理を行う。また、常時、装置状態を監視して、異常時の検出とそのログ記録を行う。

装置監視、設定制御インタフェースとして、外部との通信を行う Ethernet インタフェースおよび RS-232C シリアルインタフェースを有する。

(4) 電源ユニット

24 時間運用と信頼性確保のため、2 重化冗長構成を採用し、障害発生時の活線挿拔が可能である。

4 設計の要点

図 3 に本装置のブロック図を示す。主要なブロックについて設計の要点を述べる。

(1) パケット処理エンジン

パケット処理エンジンは、複数の専用のプロセッサコアを搭載したプログラム可能な LSI で実現した。

パケット処理エンジンは、パケットごとに処理を行い 10Gbit/s の処理能力を実現するため 1 つのパケットの処理を約 67ns で行っている。

パケット処理エンジンは、下記の 3 つのブロックを持ち、主に IPv4/IPv6 パケット解析、フラグメント処理を行う。

- ・ パケット解析部
- ・ 検索部
- ・ 結果分析部

(a) パケット解析部

パケット解析部では、IP アドレス、プロトコル番号、TCP/UDP ポート番号(5tuple)を抽出しテーブル検索

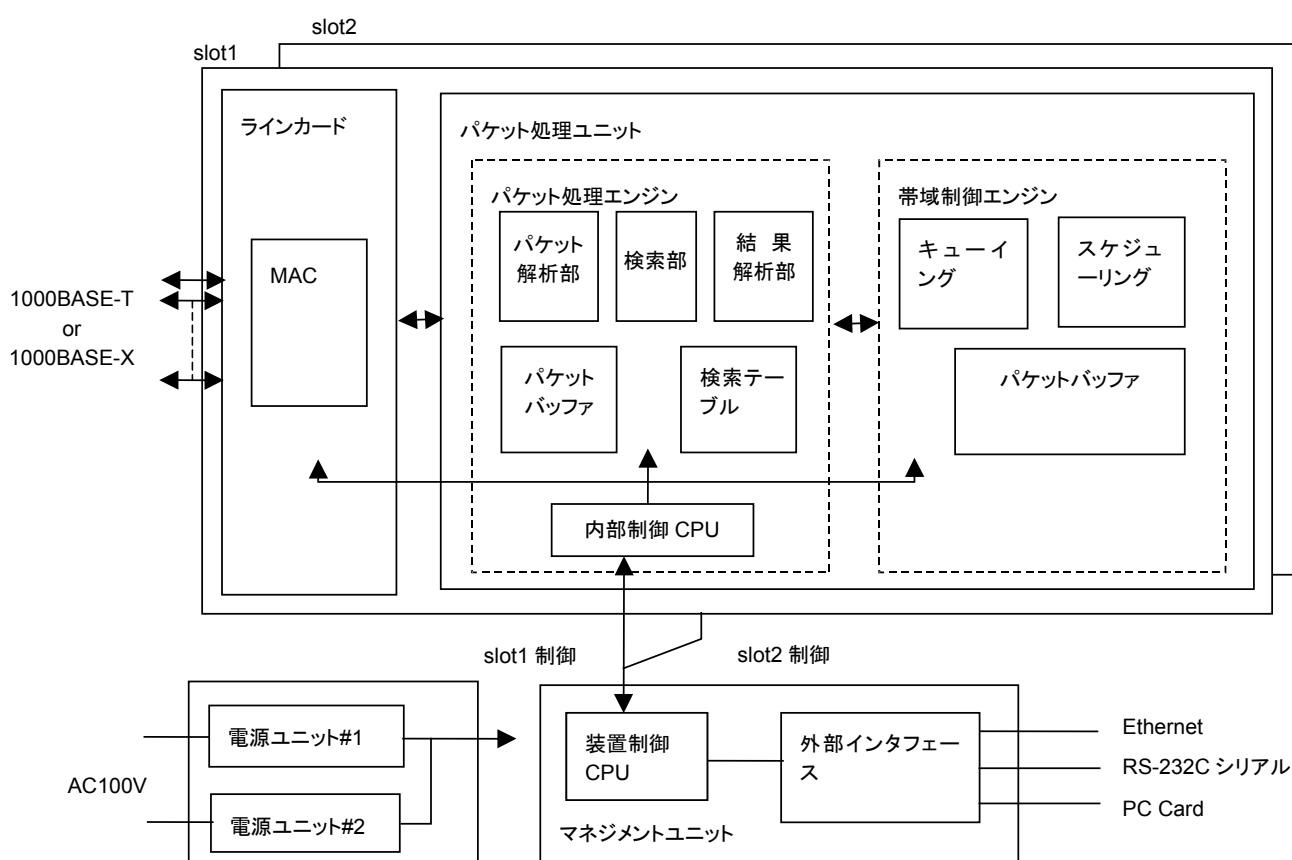


図 3 ブロック図
Block diagram

用のキー情報を作成する。

IPv4 だけでなく、次世代映像配信ネットワークとしての IPv6 パケットの識別も可能とした。

(b) 検索部

パケット解析部で抽出したキー情報を使用してテーブルの検索を行う。

テーブルの登録、管理は内部制御 CPU で行う。

帯域制御エンジンのキュー情報との関連付けをテーブル情報に登録する。

(c) 結果分析部

検索結果情報を取得し、検索結果に従いパケット廃棄、帯域制御エンジンへのキュー情報の転送を行う。

フラグメント処理用のテーブルの追加、削除も結果分析部で行う。

(d) フラグメント処理

ストリーミングデータ等の UDP を使用したアプリケーションでは、大きな UDP データグラムをネットワークで転送可能な長さに分割するフラグメント処理が行われる。

フラグメントデータは、ポート番号等の UDP ヘッダ情報が先頭パケットのみにしかないため、分割されたパケットを 1 連のデータとして扱うためには、後続のパケットを識別する方法が必要となる。図 4 にフラグメント処理のブロックを示す。フラグメント処理では、先頭フラグメントパケット、中間フラグメントパケット、最終フラグメントパケットでそれぞれ異なった処理を行う。

先頭フラグメントパケットでは、IP アドレス、プロトコル番号、および TCP/UDP ポート番号の 5tuple 情報を抽出し、テーブルを検索する(このテーブルをフローテーブルと呼ぶ)。それと同時に IP ヘッダ内のフラグメント ID を抽出し、フローテーブルの結果情報を使用して、新たにテーブルを生成する(このテーブルをフラグメントテーブルと呼ぶ)。後続のフラグメントパケットは、IP アドレス、プロトコル番号、フラグメント ID を抽出し、先頭フラグメントパケットで生成したフラグメントテーブルを検索することにより、先頭のフラグメントパケットと同一のキュー情報を使用することができる。最終フラグメントパケットでは、中間フラグメントパケットと同様に、フラグメントテーブルを検索する。また、最終フラグメントパケットでは、ハードウェアのリソース量を削減するために該当するフラグメントテーブ

ルのデータを削除する。

これらの一連の処理を、ソフトウェアを介さずパケット処理エンジン内で行っているため、フラグメント処理も 10Gbit/s の処理で行うことが可能である。

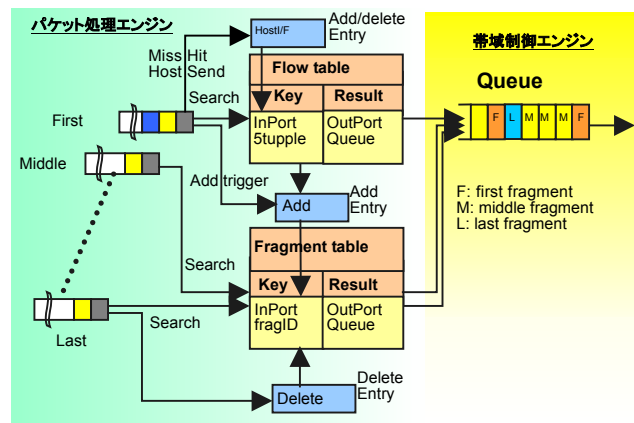


図 4 フラグメント処理ブロック図

Block diagram of fragment process

(2) 帯域制御エンジン

帯域制御エンジンは、主に 2 つの機能を持つ。

- ・ キューイング
- ・ スケジューリング

帯域制御エンジンは、受信パケットをパケットバッファに格納し、パケット処理エンジンから転送されたキュー情報をキューイングブロックに転送する。スケジューリングブロックで、キューイングブロックからキュー情報を読み出し、出力可否を判断し、パケットバッファからパケットを読み出して出力する。

帯域制御エンジンの論理的なブロックを図 5 に示す。キューイングブロックは、16 ブロックが並列に構成され、1 ブロック内に 256 キューが配置されている。1 つのキュー当たり 2000 パケットまでキュー情報を管理する。

1 つのキューに対してパケットの出力可否を制御するための固定長の値(トークンと呼ぶ)を周期的に分配する。このトークンが入力パケット長分溜まった場合に、パケットが出力可能となる。1 つのキューに対するトークン分配処理を 60ns 以内で処理することにより 1 つのブロックの 256 キューに対して 16μs で 1 巡の処理を行う。これにより、μ秒レベルでの高精度な帯域制御が可能となる。

また、ストリーミング配信では、1 スロットあたり 4095 本のストリーミングデータが同時配信可能である。

スケジューリングブロックは、キューイングブロックからキュー情報を読み出し、パケット出力の判断を行う。

スケジューリングブロック内に、高優先の遅延・帯域保証クラス、中優先の帯域保証クラス、低優先のベストエフォートクラスの 3 クラスを有する。キュー情報は、高中優先の帯域保証クラスがラウンドロビンで順番に読み出される公平優先制御で処理を行い、出力可否の判断後、後段のブロックに転送される。後段のブロックでは、高優先のクラスが優先的にパケットを出力し、高優先のパケットがなくなったときに低優先のパケットが出力される固定優先制御で処理を行う。

図 6 にパケット出力制御の方式を示す。トークン分配周

期を $20\mu\text{s}$ とした場合、出力レートを 40Mbit/s に設定すると、トークンは $20\mu\text{s}$ ごとに 100 バイトずつ与えられる。1000 バイトのパケットが入力された場合、 $200\mu\text{s}$ で 1000 バイトのトークンが溜まり、その結果パケットが出力される。パケットを出力したら、出力したパケット長分の溜まったトークンを減算する。1000 バイトトークンが溜まった状態で、1000 バイトのパケットが出力された場合、トークンは 0 バイトとなる。そのため、パケットがバーストで入力されてもトークンが再度パケット長になるまでキュー内で待たされるため、出力パケット間隔がレートに従って一様に整形された出力パターンとなる。

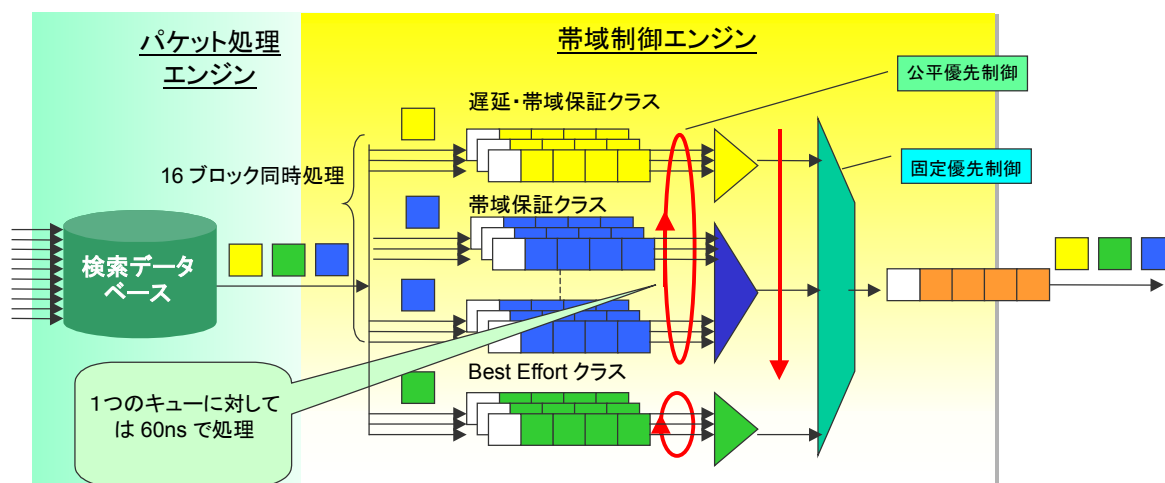


図 5 帯域制御エンジンブロック
Block diagram of bandwidth management engine

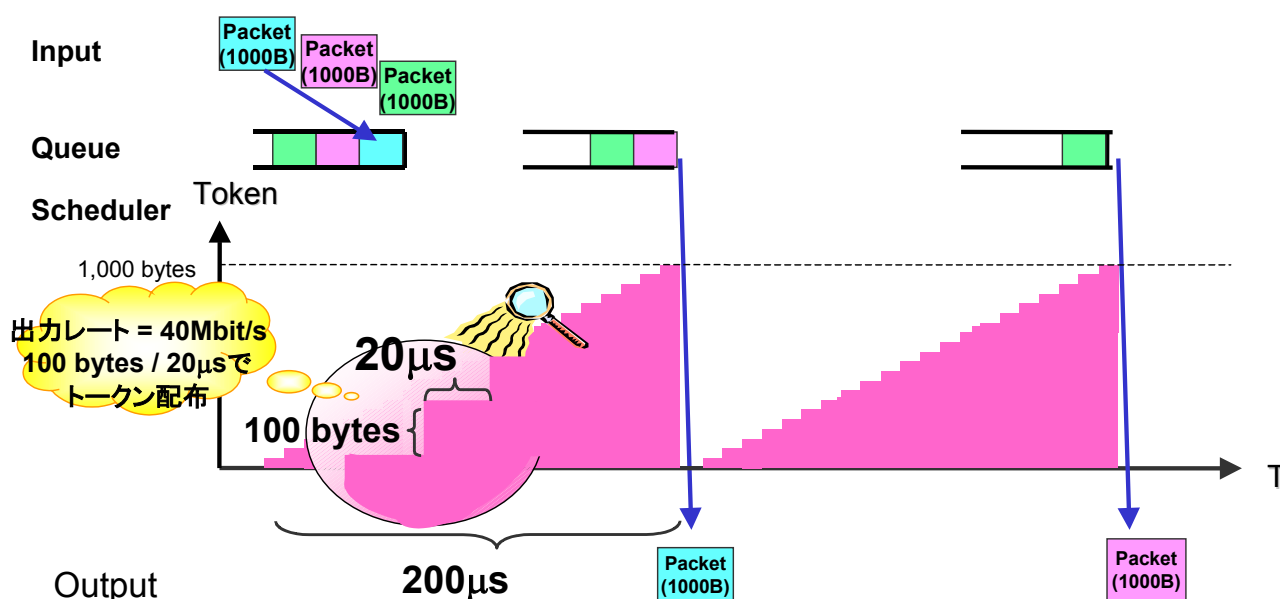


図 6 パケット出力方式
Mechanism of packet output

また、トークン分配周期を 8ms, 800μs, 80μs, および 20μs に設定可能とし、8Kbit/s から 800Mbit/s の幅広いレート設定を可能とした。

5 シェーピングの実施例

図 7 にストリーミングサーバからのバーストラフィックと本装置でシェーピング処理後のトラフィックモニタのグラフを示す。

ストリーミングサーバからの出力 packets を、1ms サンプリングのトラフィックモニタ機能(測定器: アンリツ製 IP ネットワークアナライザ MD1231A1)により観測すると、平均 2Mbit/s のストリーミングデータにおいても、ピークレートで 468Mbit/s の packets がバーストで出力されていた。これは、1400 バイト長の packets を 42 packets 分連続で出力し、約 200ms 後に次の packets を出力していることを示す。このバースト packets を本装置により、シェーピングを行うと設定レート(2Mbit/s)での等間隔な出力パターンとなり、パケットロスの原因となるトラフィックの負荷を軽減することができた。

6 機能

表 1 に本装置の主要機能を示す。

7 むすび

映像配信サービスにおいて、パケットロスの原因および回線使用率を圧迫する原因となるバースト packets に対して、本装置を用いて実際の映像配信ネットワークに導入し、packets をシェーピングさせることにより、良好なシェーピング特性が得られ、回線のサービス品質向上および回線使用率の向上が実現できた。今後も PureFlow シリーズを充実させ、帯域制御装置分野でさまざまなニーズに応える装置を開発していきたい。

参考文献

東山, 茂木, 関水, 山口 “バーストラフィックのシェーピング機能に関する一考察”, 電子情報通信学会 NS 研究会, 2004 年 6 月

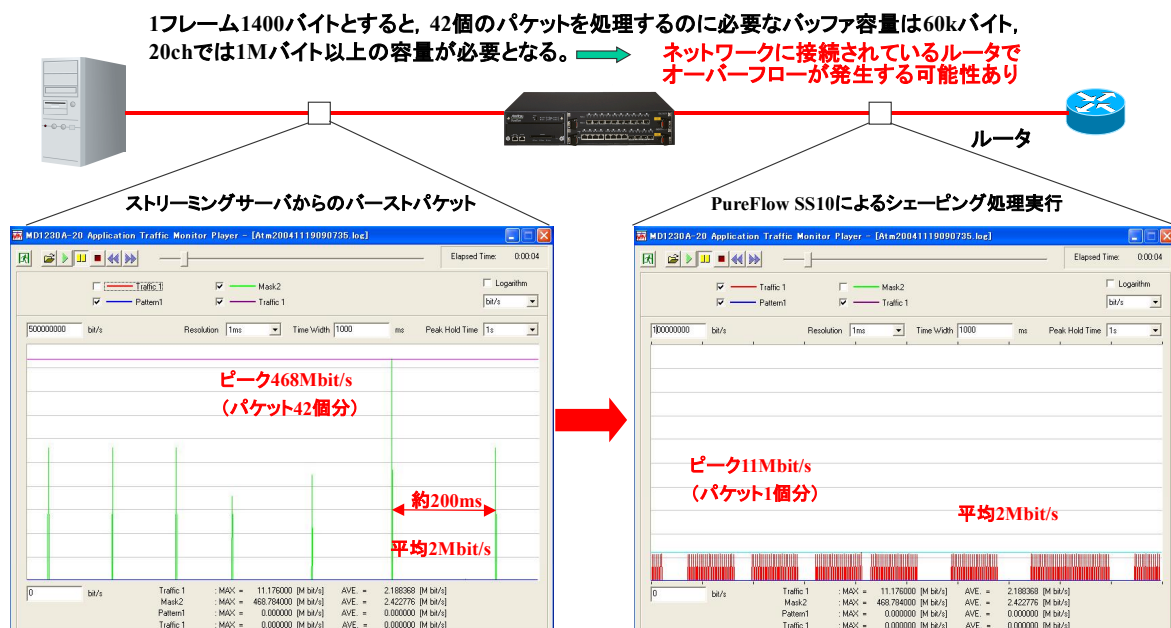
執筆者



茂木正英
ピュアフロー事業推進部
業務管理部



関水宗樹
ピュアフロー事業推進部
開発部



アンリツ製 IPネットワークアナライザ MD1231A1で測定

図 7 バーストラフィックとシェーピングトラフィック波形
Monitor for burst and shaping traffic

表1 主要機能
Specifications

項目		主要機能
制御可能帯域幅		8Kbit/s～800Mbit/s
最大帯域制御キュー数		4,095
最大アクセスコントロール数		32,768
インタフェース	ラインカードスロット数	1
	ポート数	・GbE/10SFP ラインカード MU071001A 1000BASE-X(SFP ポート) :10 ポート ・GbE/8T+4SFP ラインカード MU071002B 1000BASE-X(SFP ポート) :4 ポート 1000BASE-T :8 ポート (最大, 任意の 10 ポート使用可能)
	回線数	5(1 回線が標準, オプションで追加可能)
	対応 SFP(注 1)	・1000BASE-SX(MMF 対応, 最大距離 550m レベル, コネクタ形状 LC) ・1000BASE-LX(SMF 対応, 最大距離 10km レベル, コネクタ形状 LC) ・1000BASE-LE(SMF 対応, 最大距離 40km レベル, コネクタ形状 LC) ・1000BASE-LR(DSF 対応, 最大距離 80km レベル, コネクタ形状 LC)
	コンソールポート	RS-232C(RJ-45)(RJ-45/DB9 ケーブル付属)
	PC カードスロット	Type II ATA 規格準拠
	マネジメントポート	100BASE-TX
帯域制御	分類方法	・IPv4 IP アドレス, プロトコル番号, TCP/UDP ポート番号 ・IPv6 IP アドレス, ネクストヘッダ番号, TCP/UDP ポート番号
	優先制御	3 段階
	帯域割り当て	固定帯域幅設定, 受信許容バースト数, 送信可能バースト数
	帯域測定機能	フローごとの自動帯域測定機能
管理	設定	コンソール/Telnet による CLI(注2)
	管理	コンソール/Telnet による CLI SNMPv1/v2c/v3(注3), EnterpriseMIB(注4)
	装置アラーム出力	電源, ファン, 装置内温度異常の検出
冗長化機能	電源	2 重化, ホットスワップابل対応, ロードシェアリングによる供給
	装置	ASRP(注5)によるホットスタンバイ冗長
電源		90～110VAC, 50/60Hz±1Hz 消費電力 300VA 以下
環境条件	動作温度	0～50℃
	動作湿度	20%RH～80%RH(ただし, 結露なきこと)
寸法		434(W)×452(D)×88.9(H)mm
重量		18kg(全実装時) 11kg(1 ラインカード, 1 電源ユニット挿入時)

注 1: SFP(Small Form Factor Pluggable): 光トランシーバモジュール

注 2: CLI(Command Line Interface): 装置の動作パラメータの表示や設定を行うインタフェース

注 3: SNMP(Simple Network Management Protocol): ネットワークに接続された通信機器をネットワーク経由で監視・制御するためのプロトコル

注 4: MIB(Management Information Base): SNMP で管理されるネットワーク機器が, 自分の状態を外部に知らせるために公開する情報

注 5: ASRP(Advanced Shaper Redundancy Protocol): PureFlow 独自の装置冗長プロトコル

論 文

ストリームシェーパー PureFlow SS10 の開発

「アンリツテクニカル」82号（2006.3）より抜粋

アンリツ株式会社