

高精度トラフィックシェーパ PureFlow GS1 の開発

Development of High Accuracy Traffic Shaper PureFlow GS1

瀧山 雅朋 Masatomo Takiyama, 安齋 賢二 Kenji Anzai, 西 正博 Masahiro Nishi

[要 旨] 企業ネットワーク向けにトラフィックシェーパ PureFlow GS1 を開発した。IP 電話、業務データなどを指定し帯域確保を行うことが可能で、ネットワークを安定して利用することを可能にする。本装置は、設定誤差 1% 以内、2 階層の平滑化(シェーピング)を実現し、さらに、リバースフラグメント対応やアプリケーション層の認識など高度な機能を実現するとともに、Web 対応などユーザフレンドリな機器とした。企業に導入可能なように、装置構成を根本から見直し、低価格、低消費電力、省スペースを実現した。

[Summary] We have developed the PureFlow GS1 Traffic Shaper for enterprise networks, bringing stability to network connections by allowing the user to secure the necessary bandwidth for VoIP or business data requirements. The setting error is $\leq 1\%$ and the bandwidth control uses two-stage shaping. Many other advanced features, including reverse fragment support, application layer recognition, and web compatibility, enhance operability and flexibility. The hardware/software architecture has been fully re-designed to offer the enterprise market low cost and power consumption in a small footprint.

1 まえがき

企業ネットワークにおいて、従来の専用線や ATM 回線に変わり、より高速で低コストの新型 WAN(Wide Area Network) サービスである IP-VPN(Internet Protocol-Virtual Private Network) や広域イーサネット(以下 VPN サービス)により拠点間を接続する形態が普及してきた。高速な接続サービスの普及に伴い、従来は別々であった電話系通信とデータ系通信を IP ネットワークに集約する企業が増加している(図2)。また、業務効率化のためにネットワーク通信を利用した業務管理アプリケーションなどを採用する企業もあり、データ系通信の中でも通信品質の確保が必要なトラフィックが増加している。

こうした背景のもと、PureFlow シリーズの企業ネットワーク向け製品としてトラフィックシェーパ PureFlow GS1 を開発した。本装置は、重要なトラフィックを回線帯域の不足や通信遅延から守るために、回線帯域を拠点やユーザまたはアプリケーションごとに分割し、必要な帯域の確保やトラフィックの優先制御が可能な帯域制御装置である。

本装置は、従来製品である映像配信サービスにおけるパーストトラフィックの平滑化(シェーピング)を目的として開発されたストリームシェーパ PureFlow SS10¹⁾ をベースとし、企業ネットワークのニーズを考慮し、階層化帯域制御、アプリケーション層認識、Web 監視などの機能を追加するとともに、低価格、低消費電力、省スペースを実現した。図1に PureFlow GS1 の外観を示す。



図1 PureFlow GS1 の外観
External view of PureFlow GS1

2 開発方針

本製品の開発に当たっては、企業ネットワークへの導入を目指し以下の点を考慮した。

(1) 高精度帯域制御

帯域設定値に対する設定誤差を 1% 以下とし、回線帯域を極限まで活用可能にする。

また、パケット間隔を調整する平滑化においてもマイクロ秒の高精度シェーピングを実現し、ネットワーク上のパケット中継装置やエンド端末に瞬間的な負荷を軽減する。図3に平滑化のイメージを示す。

(2) 階層化帯域制御

多様化した企業ネットワークのトラフィックを考慮し、階層化帯域制御を実装する。この機能により、拠点ごとの通信帯域を制御すると同時に、アプリケーションごとの通信帯域を制御することが可能となる。

階層化帯域制御は、物理回線において仮想パイプ (VP: Virtual Pipe) と仮想パイプ内の仮想チャネル (VC: Virtual Channel) を定義し、階層的な帯域制御を提供する。平滑化のための最大帯域だけでなく、通信帯域を確保する保証帯域や、仮想パイプや仮想チャネルの構成も設定可能とし、企業ネットワークのトラフィックを考慮した柔軟性の高い帯域制御を提供する。

図4 に階層化帯域制御の使用例を示す。

(3) リバースフラグメント対応

データサイズがネットワークの最大伝送サイズ (MTU:

Maximum Transmission Unit) を超える場合、データを細分化 (フラグメント) して送信するが、一部サーバにおいては細分化したデータの送出順が通常と逆順となる (リバースフラグメント)。本装置では、リバースフラグメントにも対応する。

(4) アプリケーション層の認識

IP 電話 (Voice over IP: VoIP) やテレビ会議システムで使用されるプロトコルの SIP (Session Initiation Protocol) や H.323、また、ファイル転送で使用されるプロトコルの FTP があるが、これらをアプリケーション層レベルで解析し自動的に分類可能とし、さらに SIP と H.323 については自動的に帯域の割り当てを行うことを可能とする。

これにより、コンテンツごとのトラフィック制御が容易に行えるようになり、通信品質の向上が計られる。また、設定を自動化することで電話増設、移動による設定変更が不要となり、運用コストの大幅な削減が可能となる。

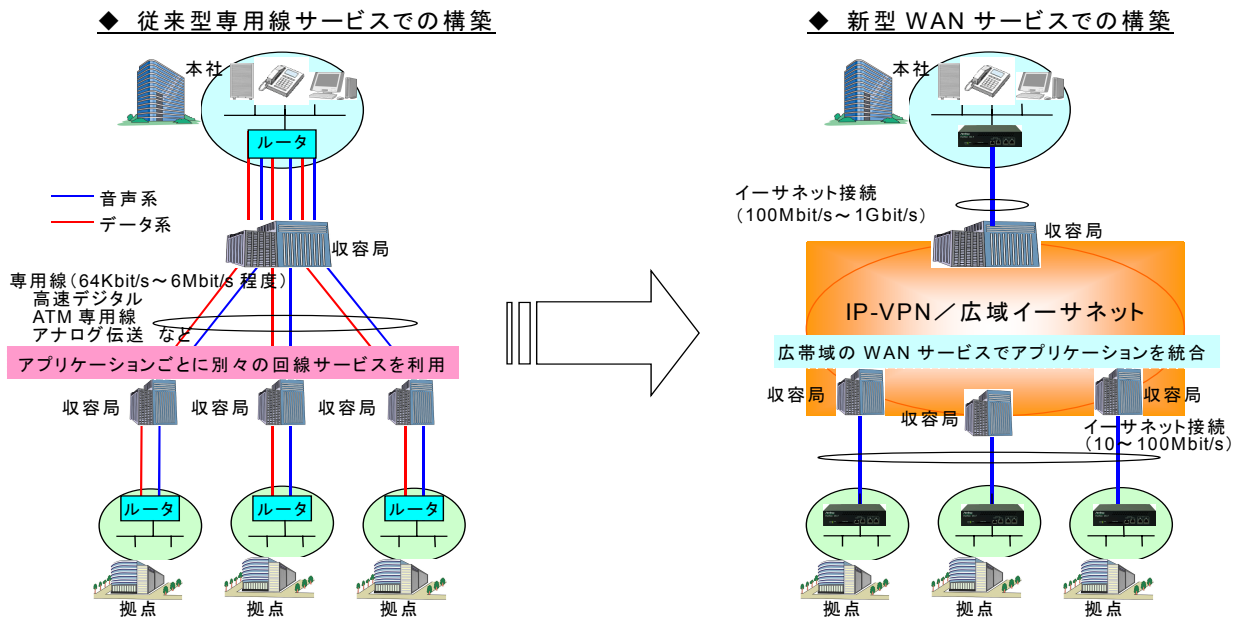


図2 専用線サービスと新型 WAN サービス
Leased line and new WAN services



図3 平滑化イメージ
Shaping image

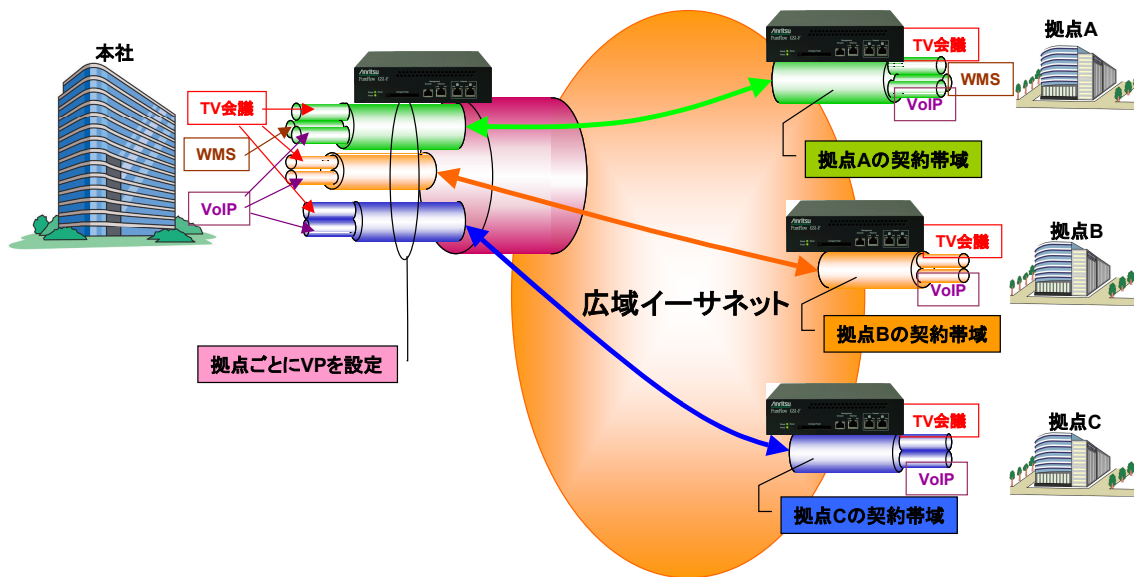


図4 階層化制御使用例
Hierarchy control

(5) Web ブラウザによる監視機能

Web ブラウザを使用したユーザフレンドリな GUI を提供する。遠隔のパソコンのブラウザから装置を監視可能とし、これにより装置の設定が正しく行われているか、異常な通信が発生していないか、視覚的に分かりやすくする。

(6) IPv6 対応

次世代ネットワークへの導入も考慮し、IPv6 パケットも識別・制御可能とする。

(7) 通信のバイパス機能

停電による装置電源断時やソフトウェアの異常時においても通信状態を確保するためにバイパス機能を実装する。

図5 にバイパス機能イメージを示す。

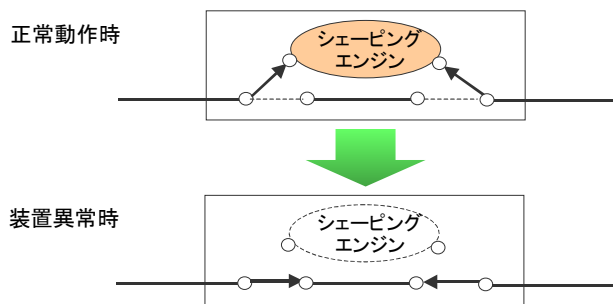


図5 バイパス機能
Bypass function

(8) 低消費電力・省スペース

オフィス環境に適応するために、以下の目標値を定めた。

部品点数を削減し、軽量化、低消費電力化を図るとともに、19 インチラックに 2 台並べての設置が可能となるような省スペース化を実現する。

消費電力 30VA 以下

質量 2kg 以下

寸法 幅(W)215mm 以下、高さ(H)60mm 以下

(9) 低価格

企業ネットワークへ容易に導入可能な価格を実現する。

3 構成

本装置のブロック図を図6 に示す。

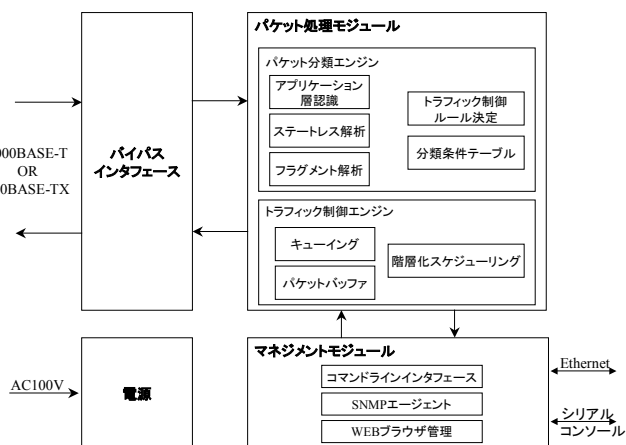


図6 ブロック図
Block diagram

3.1 パケット処理モジュール

パケット処理モジュールは、パケット分類エンジンとトラフィック制御エンジンを有する。パケット分類エンジンは、主にパケットフローの分類とトラフィック制御ルールの決定を行う。トラフィック制御エンジンは、パケットフローごとにパケットをキューイングし、キューからのパケット送出速度を制御する。

(1) パケット分類エンジン

パケット分類エンジンは、パケットの識別情報を検出する解析部とトラフィック制御ルールを決定するトラフィック制御ルール決定部を有する。

(a) ステートレス解析部

ステートレス解析部は、受信したパケットのヘッダ情報を解析し、以下の識別情報を抽出する。

- ・ L2 ヘッダ情報 (CoS, VLANID)
- ・ L3 ヘッダ情報 (IP アドレス, プロトコル番号など)
- ・ L4 ヘッダ情報 (TCP/UDP ポート番号)

(b) フラグメント解析部

フラグメント解析部は、UDP ヘッダ情報が付加されていないフラグメントパケットから UDP ポート番号を抽出する。本装置は、リバースフラグメントにも対応した。

(c) アプリケーション層認識部

アプリケーション層認識部は、IP 電話などのアプリケーションが送出した制御パケットやデータパケットを解析し、通信の接続状態を認識することにより、アプリケーション層における識別情報を抽出する。また、アプリケーションが使用する帯域をトラフィック制御ルール決定部に通知し、自動帯域割り当てに使用することが可能である。

(d) トラフィック制御ルール決定部

ユーザがあらかじめ設定したパケット分類条件に対応するトラフィック制御パラメータを決定し、トラフィック制御エンジンに通知する。パケット分類条件に一致する識別情報を持つパケットをパケットフローとして認識する。さらに、パケットフローに対してトラフィック制御パラメータ値を決定し、トラフィック制御エンジンに通知する。主なトラフィック制御パラメータを以下に列挙する。

- ・ 保証帯域
- ・ 最大帯域
- ・ バッファサイズ

(2) トラフィック制御エンジン

トラフィック制御エンジンは、キューイング部と階層化スケジューリング部を有する。キューイング部は、受信パケットをパケットバッファに格納するとともに、パケットフローごとにパケットをキューイングする。階層化スケジューリング部は、主に、仮想パイプ単位の送出帯域や仮想チャネル単位の送出帯域を制御する。

3.2 マネジメントモジュール

マネジメントモジュールは、設定情報、統計情報、装置ステータスを管理するためのユーザインタフェースを提供する。シリアルコンソールと TELNET (コンピュータ遠隔操作の標準方式) によるコマンドラインインタフェース、SNMP (Simple Network Management System)、および SYSLOG (システムメッセージの転送・記録システム) によるネットワーク監視インタフェース、さらに Web によるグラフィカルな監視インタフェースを有する。

3.3 電源

AC100V から装置内部で使用する DC 電圧を生成する。企業ユースであるが、長寿命、高調波対応とした。

3.4 バイパスインタフェース

停電による電源断時やソフトウェア異常時にバイパス接続状態の切り替えを行い、通信状態を確保する。

バイパス接続状態となった場合、本装置に接続された装置同士がクロスケーブルで接続された状態と同等になる。

4 設計の要点

設計の要点を以下に述べる。

4.1 高精度な帯域制御

独自のアルゴリズムを採用したトラフィック制御エンジンを新規に開発し、送信帯域と送信パケット長から次回送信するパケットのタイミングを計算することで、仮想パイプや仮想チャネルの送信帯域を設定値の 1% 以下の誤差で制御可能とした。帯域誤差を小さくすることで、使用できるネットワーク帯域を最大限に有効活用できる。

また、この独自のアルゴリズムは、マイクロ秒で送出タイミングを管理することにより、瞬間的なバーストのあるトラフィックをマイクロ秒の単位で平滑化することが可能となった。

以下に、平滑化した映像トラフィックの例を示す。平滑化前のトラフィック(図7)は、ピークレートが 109Mbit/s であり、瞬間的なバーストは 9 パケット分に達しているが、平滑化後のトラフィック(図8)では、1ms の間に 1 パケットが転送されるような平滑なトラフィックになった。

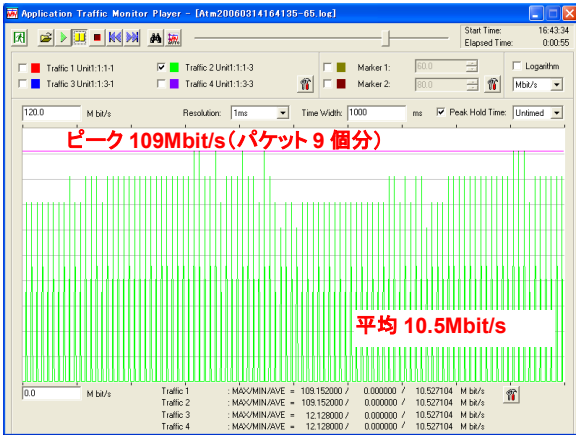


図 7 平滑化前のトラフィックパターン
Traffic pattern before shaping

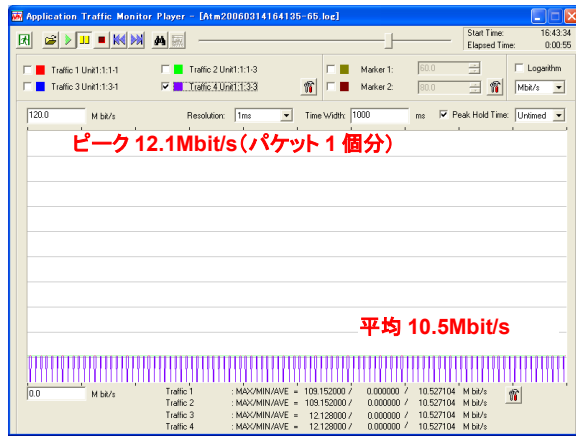


図 8 平滑化後のトラフィックパターン
Traffic pattern after shaping

4.2 階層化帯域制御の対応

パケット分類エンジンおよびトラフィック制御エンジンに対し、仮想パイプおよび仮想チャネルの概念を導入し、階層化帯域制御を実現した。

パケット分類エンジンは、仮想パイプ分類フィルタと仮想チャネル分類フィルタを有し、受信したパケットに対し、まず、仮想パイプ分類フィルタでトラフィックを分類し、次に、2階層目の仮想チャネル分類フィルタでトラフィックを細く分類する。

トラフィック制御エンジンは、仮想チャネルごとにパケットをキューイングし、仮想パイプと仮想チャネルの送出帯域を階層的に制御する。トラフィックの平滑化は、仮想パイプ単位のトラフィックのバーストと仮想チャネル単位のトラフィックのバーストを同時に制御可能とした。また、仮想パイプの帯域を仮想チャネルごとに分配する帯域分配機能を実装することにより、仮想チャネルの帯域保証を実現した。

4.3 リバースフラグメントへの対応

リバースフラグメントに対して、正しくレイヤ 4 フィルタを適応させるため、リバースフラグメント処理機能をトラフィック制御エンジンに実装した。

ストリーミングデータ等の UDP (User Datagram Protocol) を使用したアプリケーションでは、最大ネットワーク伝送サイズ (MTU) を超える大きなデータを細分化 (フラグメント) してネットワーク上に送出する。フラグメントパケットには、パケット分類に必要な情報が先頭パケットにしか含まれていないため、フラグメントパケットの処理は (1) 先頭パケットに含まれる UDP ヘッダから、処理に必要なポート番号等の情報を抽出、処理方法を決定し、(2) 継続パケットのフラグメント ID をトレースすることにより、一つのパケットフローと認識し、レイヤ 4 フィルタを実現する。

ところが、サーバ用 OS として普及している Linux の一部のバージョンでは、UDP をフラグメントするときに通常の逆順で送出する (リバースフラグメント)。フラグメントパケットが逆順で出力されると、前述の方法では、UDP ポート番号を持った先頭パケットが最後に到着するため、UDP ポート番号を正しく抽出できない。図 9 にフラグメントパケットにヘッダ情報の概要を示す。

リバースフラグメント処理機能は、先に到着した、後続パケットをパケットバッファ内に滞留させておき、フラグメント先頭パケットが到着した時に、フラグメントパケットの関連づけを行い、元の UDP ポート番号を再現させる。

リバースフラグメントに対応したことにより、サーバの OS に依存せず、レイヤ 4 フィルタを適切に動作させることが可能になった。

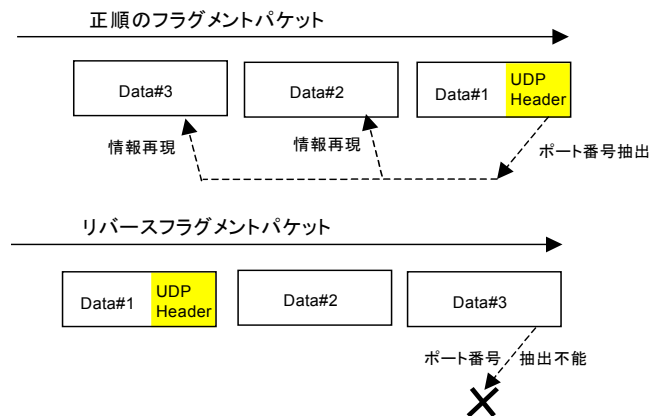


図 9 リバースフラグメントパケット例
Reverse fragment packets

4.4 アプリケーション層認識の対応

IP 電話や TV 会議システムにおいて、音声の途絶や映像の乱れ

を防止し通信品質を確保するには、ネットワークでのパケットロスやパケット遅延を防止する必要がある。この場合、帯域制御装置を使用し、帯域を確保することが最も効果的な処置である。ただし、ネットワーク管理者は、IP 電話が接続している音声や映像のパケットフローを分類し、これらのパケットフローに対し必要帯域を割り当てるといった設定を入力しなければならない。従来の帯域制御装置では、このように変化するUDPポート番号や使用帯域を設定することが困難であった。本装置は、アプリケーションレイヤの認識により、パケットフローの自動設定、割当帯域の自動設定を実現し、音声や映像の通信品質を確保することを可能とした。

アプリケーションレイヤ認識部の構造を図10に示す。アプリケーションレイヤ認識部は、プロトコル解析部とセッション管理テーブルを有する。プロトコル解析部は、端末装置やサーバ装置が送信した制御パケットからアプリケーションごとのイベント情報を抽出し、アプリケーションの接続状態を推定する。セッション管理テーブルは、アプリケーションの接続情報を保持する。以下に、アプリケーション層認識部の設計に際して考慮した点を記述する。

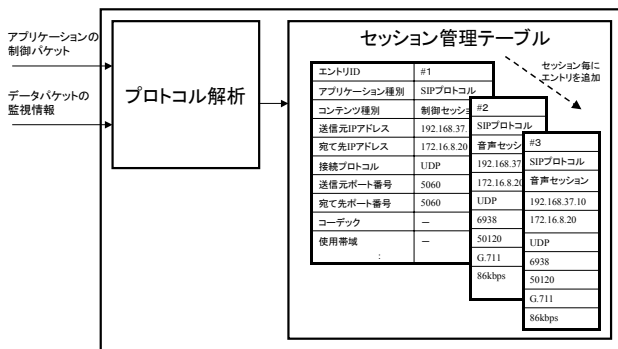


図10 アプリケーション層認識
Application layer recognition

- (1) システム構成や機種の違いを考慮したアプリケーション認識
- システム構成が異なる場合、制御セッションの接続先が異なる。図11の例はSIPサーバを使用したシステム構成の一例である。図中では、SIPサーバは制御セッションだけを端末装置と接続している。別のSIPサーバでは、制御セッションだけでなく音声セッションも接続し、音声データを中継する装置もある。一方、端末装置の設定によっては、SIPサーバを介さずに制御セッションも端末同士(peer to peer)で接続する場合もある。このように、端末が接続するセッションの数や種類はシステムによって異なっている。

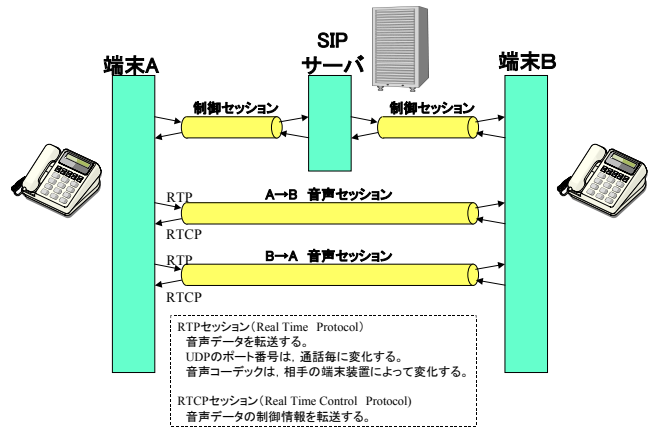


図11 IP電話の接続例
VoIP Connection

図10のセッション管理テーブルは、管理単位を端末の接続単位ではなくセッション単位とし、接続するセッションの数や種類の変化に柔軟に対応可能なように設計した。

また、制御パケット中のイベント情報の順番が異なったり、端末独自の情報が追加されていたりなど、機種によっては同一のイベント情報を通知する場合でも制御パケットのフォーマットが異なったり、端末独自のイベント情報を含む場合がある。

図10のプロトコル解析部は、制御パケット中のイベント情報を個別に抽出し、接続先や帯域などの帯域制御に必要なイベント情報のみを使用するように設計した。これにより、端末独自のイベント情報を含む制御パケットを受信した場合や、イベント情報の順番が異なる制御パケットを受信した場合でも、異常パケットと判断することなく、アプリケーション層認識を可能としている。

- (2) 柔軟性の高い帯域認識

端末装置やサーバ装置から通知される帯域情報は、通常、コーデックの情報速度で通知される。たとえば、G.711の音声コーデックの場合、符号化された音声の情報速度である64Kbit/sで通知される。しかし、IPネットワーク上では、EthernetヘッダやIPヘッダなど付加され、転送帯域は約85Kbit/sを必要とする。

プロトコル解析部は、コーデックの情報速度に加え、EthernetヘッダやIPヘッダなどの付加情報やパケット送信速度を考慮し、IPネットワーク上での転送帯域を算出可能とした。また、シグナリングプロトコルが情報レートを通知しない場合は、G.711などコーデック種別から転送帯域を算出可能とした。

(3) 端末装置の異常処理への対応

アプリケーションの異常などにより、切断要求などの制御パケットを通知できずに通信が強制終了される場合もある。

プロトコル解析部は、上記のような異常状態を考慮し、制御セッションだけでなく、制御セッション以外のパケットの流れも監視し、異常なシーケンスによる切断を検出可能とした。

4.5 Web ブラウザによる監視機能

Web ブラウザによる監視機能を実現するためのアプリケーションとして Java Web Start を採用した。Java Web Start とは Web を利用して Java アプリケーションをエンドユーザに配布する仕組みである。これにより、ブラウザだけでは実現が困難であったリアルタイムでのグラフ表示など、ユーザフレンドリな機能を実現可能とした。

図12に仮想チャネルごとのトラフィックを10秒おきに表示した表示例を示す。

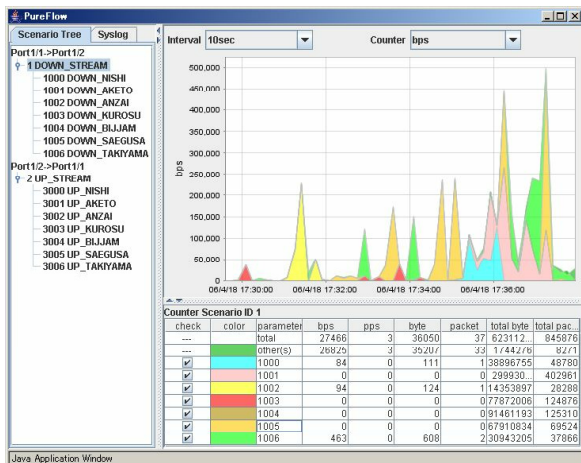


図12 Web によるトラフィック量の表示例
Web browser traffic

4.6 低価格、低消費電力、省スペース

パケット処理モジュールおよびマネジメントモジュールのハードウェアデバイスを1チップ化し、デバイス数を削減することにより低価格、低消費電力、省スペースを実現した。

5 機能

表1に本装置の主要諸元を示す。

6 むすび

今後、企業ネットワークにおいて、データと音声・映像の統合がますます進み、帯域制御の必要性がより一層高まると考えられる。より使いやすく、高性能で、低価格な装置を提供できるよう顧客ニーズに適應した製品の開発を行いたい。

参考文献

- 1) 茂木, 関水: “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, (Mar. 2006)

執筆者



瀧山 雅朋
アンリツネットワークス株式会社
システム開発センター
第1開発部



安齋 賢二
IPネットワーク事業推進部
第1プロジェクトチーム



西 正博
IPネットワーク事業推進部
第1プロジェクトチーム

表 1 主要緒元
Specifications

PureFlow®GS1 シリーズ		GS1-F	GS1-FB	GS1-G	GS1-GB
型名		PF7000A	PF7001A	PF7010A	PF7011A
制御可能帯域幅		10Kbit/s~100Mbit/s		10Kbit/s~1Gbit/s	
帯域制御設定単位		1Kbit/s		1Kbit/s	
仮想パイプ数		512		1,024	
仮想チャンネル数		512		1,024	
フィルタルール数		8,192		8,192	
インタフェース	ネットワークポート(*1)	100BASE-TX×2		1000BASE-T×2	
	コンソールポート	RS232C (RJ-45) (RJ-45/DB9 ケーブル付属)			
	コンパクトフラッシュスロット	CFA 規格 3.3V 対応			
	管理用イーサネットポート(*1)	100BASE-TX			
帯域制御	制御可能トラフィック	VLAN タグフレーム, IPv4 パケット, IPv6 パケット			
	管理可能な項目	レイヤ 2	VLAN ID, CoS		
		レイヤ 3	IP アドレス, プロトコル番号, ToS (IPv4), トラフィッククラス (IPv6)		
		レイヤ 4	TCP/UDP, ポート番号		
		レイヤ7(*2)	H.323(*3), SIP(*3), FTP		
帯域設定	仮想パイプの保証帯域, 仮想チャンネルの最大帯域および最低保証帯域				
運用管理	設定	コンソール/Telnet による CLI			
	管理	コンソール/Telnet による CLI, SNMPv1/v2c/v3, EnterpriseMIB, SYSLOG			
	その他	SNTP			
障害対策	バイパス機能	—	対応	—	対応
電源		AC100V, 50/60Hz(消費電力:30VA 以下)			
環境条件	動作温度	0~40℃			
	動作湿度	20%~80%(ただし, 結露なきこと)			
寸法		212(W)×375.5(D)×55(H)mm			
質量		2kg			
オプション		ラックマウントキット(1 台取付用・2 台取付用), CF カード			

*1: Auto-MDIX 機能に対応

*2: 別途, ソフトウェアオプション(GS1 ライセンスキー1【型名:PU700301A】)が必要です。

*3: Auto Scenario(自動帯域割当て機能)に対応