

# 10GbE 対応帯域制御装置 PureFlow GSX-XR の開発

Development of 10 Gbits/s Traffic Shaper

安齋賢二 Kenji Anzai, 吉仲謙二 Kenji Yoshinaka, 原田 浩 Hiroshi Harada, 龍 浩一 Koichi Ryu,  
鈴木雅也 Masaya Suzuki, 三枝 淳 Atsushi Saegusa, 小林正直 Masanao Kobayashi, 佐藤敬幸 Takayuki Sato,  
渡部良太 Ryota Watanabe

**[要 旨]** クラウドサービス事業者向けに高精度の帯域制御装置 **PureFlow GSX-XR** を開発した。本装置は、**10 Gbit/s** の処理性能と高精度な帯域制御機能を実現した。本装置はネットワークでパケットが廃棄されることを抑制するため、ネットワーク管理者はネットワークの利用効率を向上することが可能である。さらに、ネットワーク管理者は、本装置のトラフィック監視機能と **Web API** 機能を利用することにより、ネットワークの運用コストを削減することが可能である。

**[Summary]** We developed highly precise bandwidth controler PureFlow GSX-XR for a cloud service company. This device realized processing performance of 10 Gbit/s and a highly precise bandwidth control function. This device controls that a packet is discarded in a network, and the network administrator can improve the use efficiency of the network. Furthermore, the network administrator can reduce the operative cost of the network by using traffic monitoring function and the Web API function of this device.

## 1 まえがき

昨今、企業の基幹システムをクラウド型サービスを利用する企業が増加している。これまで、企業の基幹システムは、企業単位でサーバ構築や運用を行ってきた。しかし、データセンター事業者が展開する **SaaS (Software as a Service)** や **IaaS (Infrastructure as a Service)** など、いわゆるクラウド型サービスを利用する形態に移行すると推定される。この理由は、クラウド型サービスを導入することにより、基幹システムを担う主要サーバやネットワーク機器等にかかる設備投資を抑制できるとともに、運用コストを低減できるからである。特に近年では、大災害が発生した場合でも事業を継続するためのバックアップシステムを低コストで実現できることから、クラウド型サービスの普及が進んでいる。

クラウド型サービスでは、各企業が個別に所有していたサーバがデータセンターの仮想サーバに集約され、サーバが効率的に配置される。つまり、サーバやアプリケーションが一箇所に集約されていくことからネットワークを通過するトラフィック量は増加傾向にあり、**10 Gbit/s** ネットワークが普及してきている。

こうした背景のもと、**10 Gbit/s** 対応帯域制御装置 **NF7101A PureFlow GSX-XR** を開発した。本装置は企業、拠点、アプリケーション等を最大 **40,000** の仮想回線、**8** 階層のグループに分けることで柔軟な帯域制御を実現することができる。また、**TCP** 通信における帯域低下要因の一つであるパケット損失を、高精度シェーピング制御により低減させ、安定した通信帯域を維持することができる。



図1 PureFlow GSX-XR の外観  
External view of PureFlow GSX-XR

## 2 開発方針

従来型の企業ネットワークに加え、クラウドネットワークで利用することを考慮し開発を行った。本装置は、**10 Gbit/s** の高精度高性能帯域制御に加えて、クラウド型サービスに対応する階層化帯域制御、ネットワークサービスの状態を監視するトラフィック監視機能、クラウド型サービスと連携して動作するための **Web API** 機能を新たに開発した。

### (1) 高精度・高性能帯域制御

サーバの高性能化、ネットワークの高帯域化により、ミリ秒単位で瞬間的にトラフィック量が増大する現象(以後、マイクロバースト)が発生しやすくなっている。このマイクロバーストはネットワークにおいて、トラフィックの輻輳によるパケット損失を引き起こし、**TCP** 通信のパフォーマンスを低下させる。**10GbE** のデータ回線を利用しても、そのスループットは **10 Gbit/s** を大きく下回る場合がある。**PureFlow GSX-XR** はパケットの送出タイミングをマイクロ秒オーダーで制御する高精度制御でありながら、装置スループット性能を **10 Gbit/s**

まで向上、高精度・高性能帯域制御を実現した。マイクロバーストを理想的なパケット間隔に平滑化し、パケット損失を防止することができる。この結果、ネットワークの使用効率が向上し、結果、クラウドネットワークの最適化をすることができる。



図2 マイクロバーストが及ぼすパケット損失  
Packets loss caused by micro-burst traffic

(2) 階層化帯域制御

クラウドデータセンターでの利用を想定し、帯域制御の階層を8階層とした。クラウドデータセンターのトラフィックは、図3のように、サービス、顧客や拠点、ユーザといったグループが階層化されているため、通信帯域も階層的に制御することが必要となる。階層化帯域制御は、階層ごとにシナリオとフィルタを追加することで、8階層までの階層化シェーピングを提供する。平滑化のために最大帯域だけでなく、通信帯域を確保する保障帯域や、シナリオやフィルタの構成も設定可能とし、データセンター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。

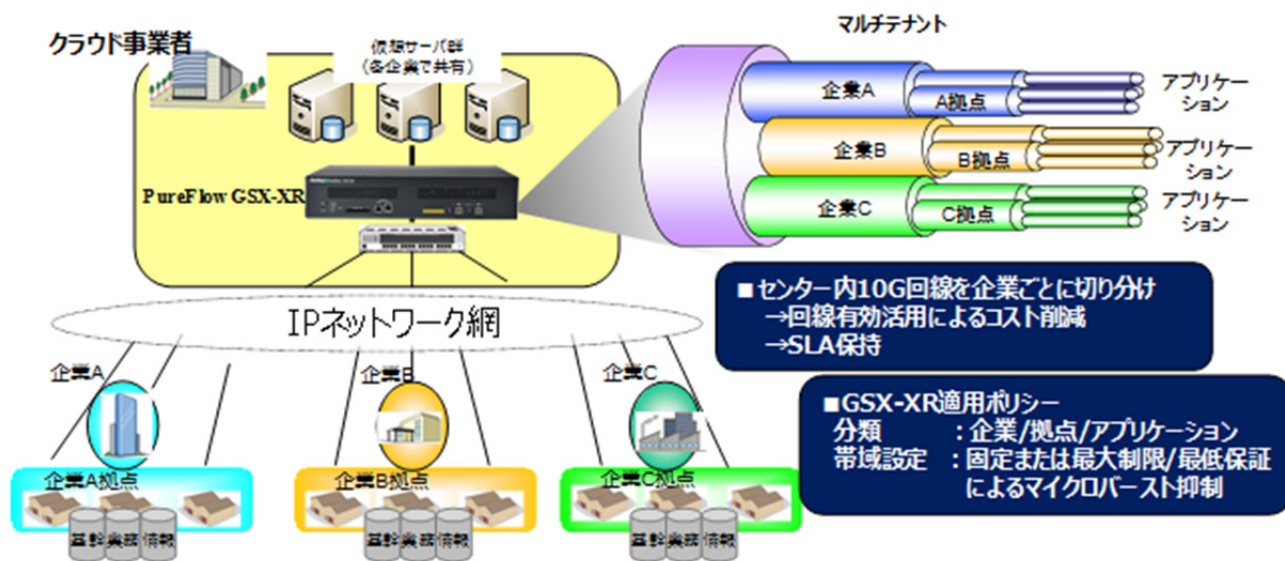


図3 クラウド型サービスの概略図  
Overview of cloud services

(3) 多様なネットワークインタフェースのサポート

10GBASE-SR / 10GBASE-LR / 1000BASE-SX / 1000BASE-LX / 1000BASE-T の各種回線スピード、メディアに対応した多様なインタフェース回線のサポートを行った。クラウドデータセンターの内部ネットワークは 10GbE が主流であるが、外部、顧客と接続する WAN 回線は GbE であったり、10GbE であったり様々な形態がある。また現在は GbE の WAN 回線を使っているユーザでも、今後はトラフィック増加により 10GbE に変更されていく。このような変化する顧客環境に対応するため、モジュールの差し替えで各種スピード、メディアに対応可能なネットワークインタフェースを提供した。

(4) トラフィック監視機能

トラフィック監視機能を実装した。トラフィックの監視は、アセットマネジメントの最適化のために必須である。ネットワーク管理者は、トラフィック量の増加に伴う通信品質の劣化は避けねばならないが、とはいえ、必要以上にネットワーク設備の設備投資をすることは設備コストと運用コストを増大させる。このため、コストを抑えつつネットワークを維持していくためには、トラフィックの状態を的確に把握しなければならない。また、ネットワークサービスで問題が発生した場合、早急に問題を切り分けるためにも重要な情報となる。ネットワークサービスで問題が発生した場合、トラフィックを監視していれば、総トラフィック量が通信容量を超過していないかを確認することが可能である。また、異常な動作をしているサーバやクライアントを、トラフィック量の変化から特定することにも使用することができる。

(5) クラウド型サービスとの連携

本装置は、クラウド基盤の管理システム(以後、管理システム)を連携動作するための WebAPI に対応した。この機能により、帯域制御ポリシーの変更を容易に自動化することができる。

図 4 は、本装置を WebAPI で操作するときの概要図である。管理システムは、オペレータからの要請に従い仮想サーバの構成を変更する。このとき、本装置に対して WebAPI 経由で帯域制御ポリシーを変更することで、サーバの構成にあわせた帯域制御を、常に提供することが可能になる。

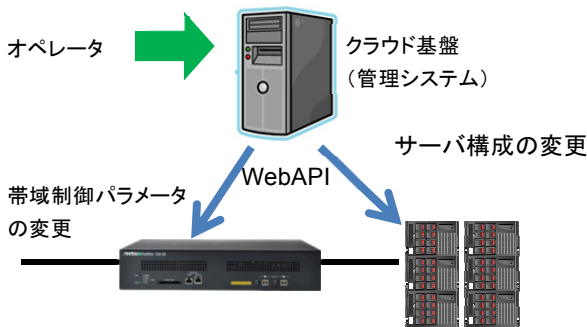


図 4 Web API

3 装置設計

3.1 装置構成

本装置のブロック図を図 5 に示す。本装置は、装置管理部、パケット処理部、電源ユニット、ファンユニットで構成している。障害発生時のリスクを最低限にするため、それぞれのユニットは独立して

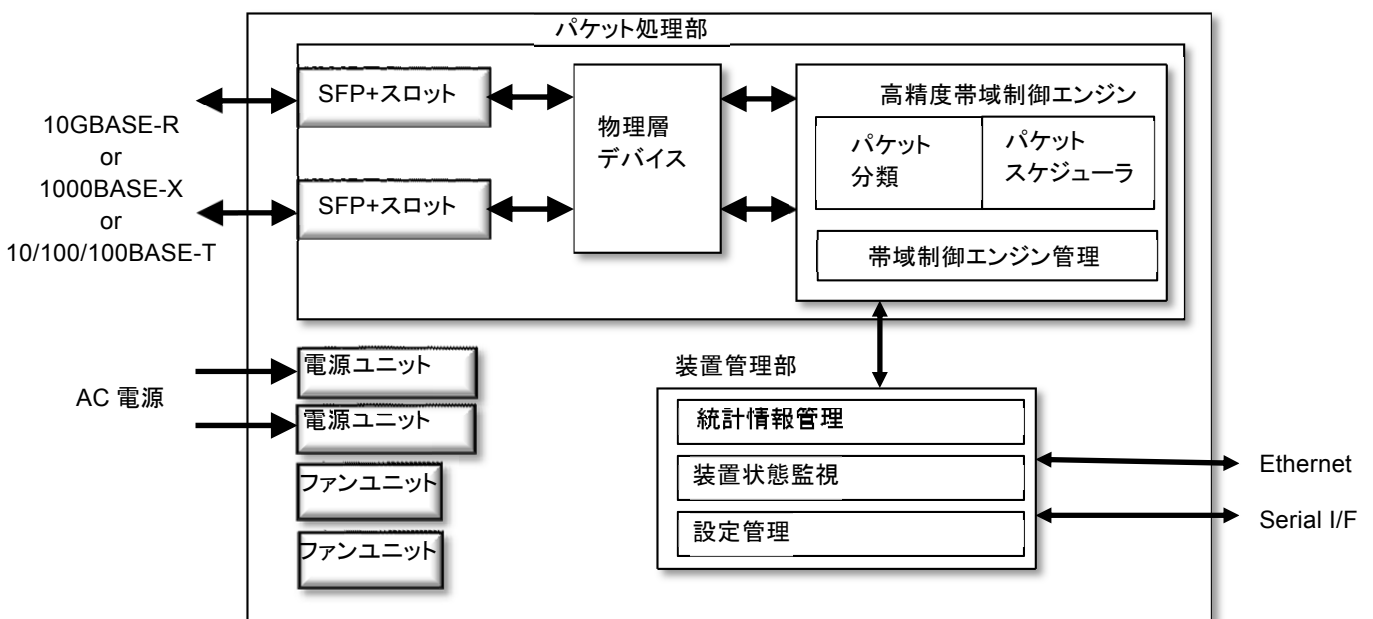


図 5 システムブロック図  
System Block Diagram

動作する。万が一、装置管理部に障害が発生しても、パケット処理部は動作を継続することができるので、ネットワークサービスを継続させることができる。

3.1.1 高精度帯域制御エンジン

帯域制御エンジンは、受信したパケットをシナリオとよぶ帯域制御単位に分類し、それぞれのシナリオごとに送出タイミングを制御することで帯域を制御している。高性能と高精度を両立させるため、多数の CPU コアをもつマルチコアプロセッサを使用した。パケット分類処理とスケジュール処理を並列で処理することにより、処理分解能を高めつつ、高性能なパケット転送を実現した。

帯域制御エンジンのパケット分類処理は、一般的な IP アドレス、UDP/TCP ポート番号に加え、クラウド環境で多用され QinQ(多重 VLAN)の VLAN ID もパケット分類に利用可能である。顧客、サービスを多重 VLAN で区別するなど、クラウド環境で複雑なネットワーク構成でも帯域制御装置を利用可能としている。

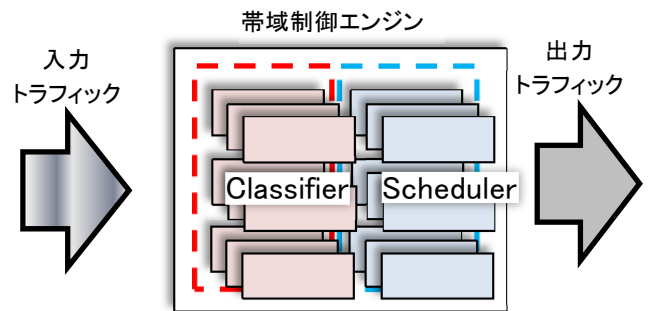


図 6 マルチコアプロセッサ 帯域制御エンジン  
System Block Diagram

帯域制御エンジンのスケジュール処理は、パケットごとに 100 万分の 1 秒の分解能でパケットの送出タイミングを制御している。この高精度な帯域制御は、マイクロバースをもつトラフィックに対して効果を発揮し、理想的なパケット間隔を持つ滑らかなトラフィックに最適化することができる。

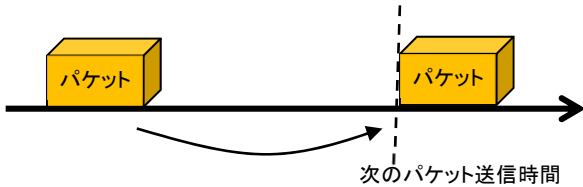


図 7 パケットスケジューリング  
Packet scheduling

これら処理を高速で実現するため、各パケット毎のスケジューリングを複数のコアで分散処理を行う。受信したパケットは複数のコアに分散され、パケット識別、スケジューリング処理を行う。複数シナリオのパケット送出時間が重なる場合はコア間で送出タイミングを調整し、送出タイミングが重ならないようにした。

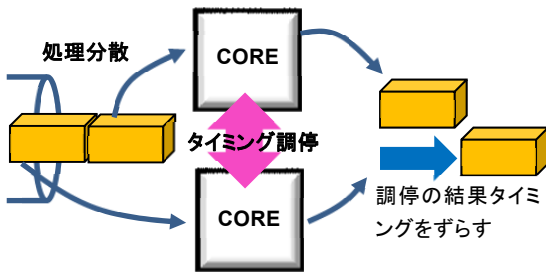


図 8 マルチコアプロセッサによるパラレルスケジューリング  
Parallel scheduling by Multi-core processor

この調整により送出を遅らされたシナリオの次のパケットの送出可能時間は、送出が遅延したことによる影響を受けないように算出し、遅延が蓄積しないようにしている。

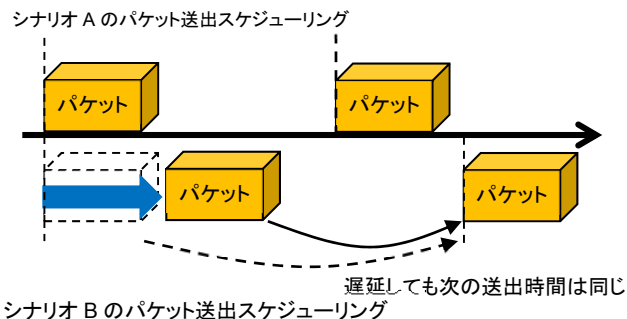


図 9 パケットスケジューリングの競合  
Conflict of Packet scheduling

### 3.1.2 物理層デバイス（ネットワークインタフェース）

10GBASE-SR / 10GBASE-LR は SFP+モジュールを採用した。SFP+を XAUI インタフェースで接続。同一コネクタに 1000BASE-SX / 1000BASE-LX / 1000BASE-T モジュールも接続できるようにするため、物理層は XAUI と SGMII の両方が接続可能となるように設計した。ソフトウェア制御により XAUI / SGMII の切り替えが可能であり、装着したモジュールに合わせて回線スピードを切り替えることができる。

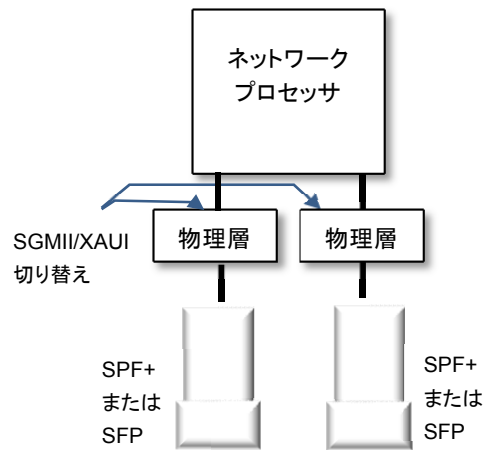


図 10 ネットワークインタフェース  
Network Interface

### 3.1.3 装置管理部

装置管理部は、設定情報の表示や変更や通知、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機能を処理する。シリアルコンソール、TELNET、SNMP など一般的なネットワーク管理プロトコルに加え、Web API に対応した HTTP にも対応する。

### 3.1.4 帯域制御エンジン管理部

帯域制御エンジン管理部は、帯域制御エンジンの設定変更や監視、統計情報の収集を行う。最大で 40000 シナリオ分の統計情報を定期的に収集する必要から、帯域制御エンジン管理部も高性能 CPU で構成すると同時に、収集処理も最適化し、高速で統計情報の収集を行っている。

### 3.1.5 電源ユニット・ファンユニット

本装置は企業のデータセンターなど、ネットワークの中でもトラフィックが集中する部分に設置される事を想定している。そのため、電源ユニットのみならずファンも活線挿抜を可能にし、寿命のある可動部分をサービス中に容易に交換できるようにしている。

## 4 高精度帯域制御の効果

広帯域ネットワークにおける高精度帯域制御の効果を示す。クラウドネットワークを想定した擬似ネットワークエミュレータを使用し検証・評価を行った。擬似ネットワークエミュレータは転送レート 2 Gbit/s をピークレートとし、それを超えるトラフィックのポリシングを行う。擬似ネットワークエミュレータを挟み、データセンター側に HTTP サーバを、クライアント側に HTTP クライアントを配置する。

本検証では、1 MByte のファイルをサーバ側からクライアントに繰り返しダウンロードし、合計で 1000 MByte のファイルをダウンロードした場合の転送レートと転送時間を測定した。

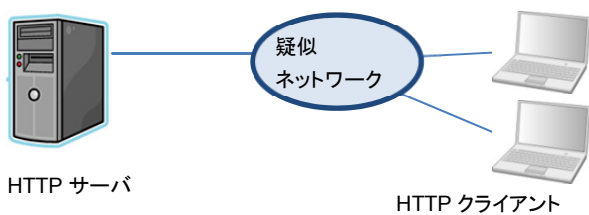


図 11 試験構成1(帯域制御なし)  
Test network 1 (Without Traffic shaping)



図 12 試験構成 2(帯域制御あり)  
Test network 2 (With Traffic shaping)

表1 転送レートと合計ダウンロード時間  
transfer rate and total downloading time

項目	平均転送レート	最大転送レート	合計ダウンロード時間
帯域制御なし	1.2 Gbit/s	1.9 Gbit/s	28.7 秒
帯域制御あり	2.0 Gbit/s	2.0 Gbit/s	16.9 秒

表 1 は、実測した擬似ネットワークの転送レートと全てのファイルダウンロードが終了するまでの接続時間である。帯域制御が無い場合と帯域制御を実施した場合を比較すると、帯域制御を実施した方が、平均転送レートも高く、また、全てのファイルダウンロードが終了するまでの時間も短い。

また、図 13 と図 14 は、帯域制御が無い場合と帯域制御を実施した場合のトラフィック状態を 1 ms の分解能でサンプリングシグナチャ化したものである。赤のグラフは HTTP サーバの送信レートである。また、緑のグラフは、擬似ネットワークを通過し、HTTP クライアントが受信できた受信レートである。

帯域制御が無い場合(図 13)は、擬似ネットワークの実効レートは 2 Gbit/s であるにもかかわらず、平均レートは、1.2 Gbit/s となった。これは、トラフィックがバーストしていることによるパケット廃棄が原因である。このバーストのため、擬似ネットワークにおいてパケット廃棄が発生し、TCP の輻輳制御が動作する。このため、転送レートが安定せず、全体的な実効帯域が抑圧された。

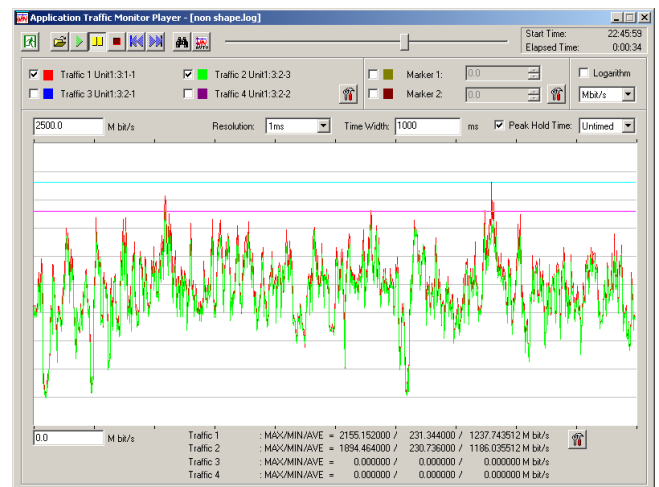


図 13 ダウンロードのトラフィック(帯域制御なし)  
Traffic of download (Without Traffic shaping)

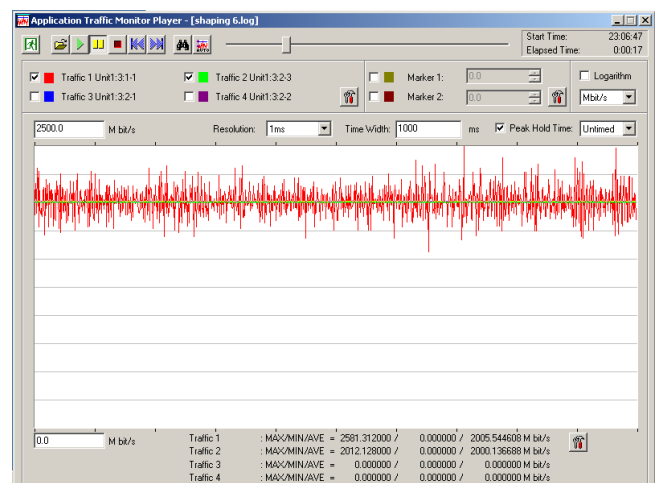


図 14 ダウンロードのトラフィック(帯域制御あり)  
Traffic of download (With Traffic shaping)

一方、帯域制御を実施した場合（図 14）、図 13 で見られたマイクロバーストが抑えられ、かつ、転送レートが安定している。マイクロバーストによるパケット廃棄が無くなった結果、擬似ネットワークにおけるパケット廃棄が無くなり、転送レートが安定していることが分かる。この結果、平均の転送レートが約 2 倍になり、効率的なデータ転送が可能となった。

## 5 トラフィック監視

最適なネットワークを維持していくために必要となるトラフィックの監視機能としてシナリオカウンタと、トップカウンタと、ピークレートモニタを実装した。ここでは、これらをグラフ化した表示例と利用方法を説明する。なお、これらのグラフ表示には NF7201A モニタリングマネージャ 2 を使用した。これは、PureFlow GSX-XR のトラフィック流量を収集する監視ソフトウェアである。トラフィック流量をグラフ化するとともに、定期的にトラフィックレポートを自動生成できる。

### 5.1 シナリオカウンタによるトラフィック流量の監視

トラフィックごとの平均帯域を監視し、1ヶ月、四半期、半年、1年などの周期でトラフィック量をグラフ化し、トラフィック量が増加傾向にあるかどうかを判断する。

図 15 は、本装置のシナリオカウンタをつかってトラフィック流量をグラフ化した例である。PureFlow GSX は、シナリオ単位でトラフィック流量を監視できる。この例では、トラフィック種別ごとにシナリオを割り当てた。トラフィック種別ごとにトラフィック流量を色分けし、積み上げることによって特定のトラフィック種別が占める割合と総トラフィックの変化を可視化する。このグラフにより、トラフィックの peak 時間帯は 13:00~14:00 であり、そのほかにもバースト状に peak 時間が存在すること、また、peak 時間帯のトラフィック種別の内訳が確認できる。

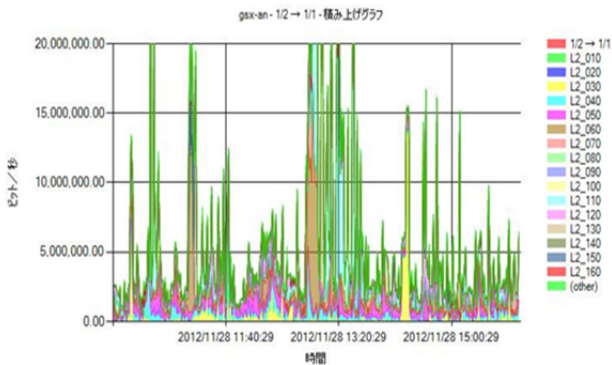


図 15 シナリオグラフ  
Scenario Graph

### 5.2 トップカウンタの監視

ネットワークによっては、極めて少数のネットワーク機器だけがネットワークリソースを占有している場合があり、ネットワークトラフィックを最適化するには、これらの機器を特定する必要がある。

図 16 は、本装置のトップカウンタを使って、ネットワークリソースを占有している 20 台のネットワーク機器を IP アドレスごとに表示した例である。IP アドレスのほか、トラフィック量が多いアプリケーションポート番号もトラフィックが多い順に表示することができる。

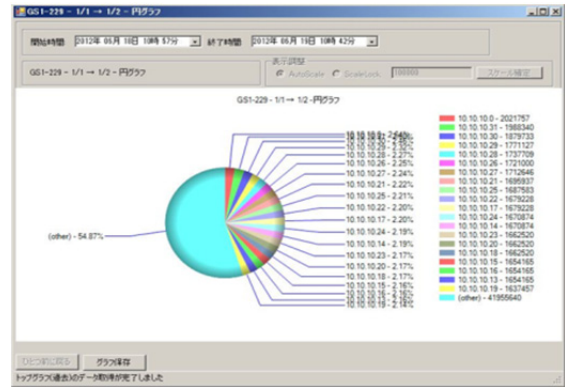


図 16 トップグラフ  
Top Graph

### 5.3 ピークレートと平均レートの監視

ピークレートと平均レートを比較することで、ネットワークリソースの空き情報を可視化することができる。ここで、ピークレートとは、1秒間隔でレート値を測定し、この中から 1 分間 (60 回分) の最大値を選択したものである。また、平均レートとは、1 分間の平均レートである。ピークレートと平均レートを比較したとき、この差分は大きければ、ネットワークリソースに余裕があるとみなす。逆に、差分がほとんど無い場合は、ネットワークリソースに逼迫するようなトラフィックが流れ続けているとみなす。

図 17 と図 18 は、本装置のピークレートモニタ機能をつかってピーク帯域と平均レートをグラフ化したものである。図 17 はネットワークが輻輳していないときのピークグラフである。トラフィック負荷が低い場合であり、平均レートに比べピークレートが大きく上回っている。これに比べ、図 18 は、平均レートとピークレートの差がほとんどなく、ネットワークが過負荷状態になっていることが分かる。このように重ねて比較することで、ネットワークリソースに対するネットワーク負荷の状況を可視化することができる。

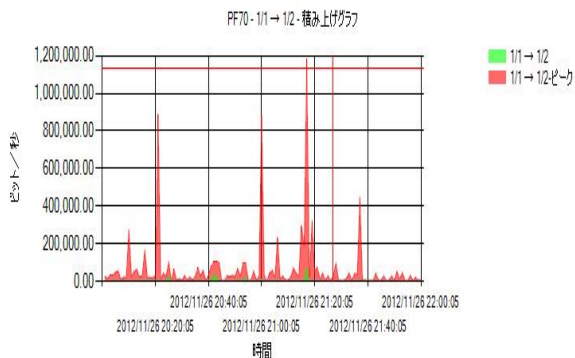


図 17 ネットワークが輻輳していないときのピークグラフ  
Peak graph when networks are not congested

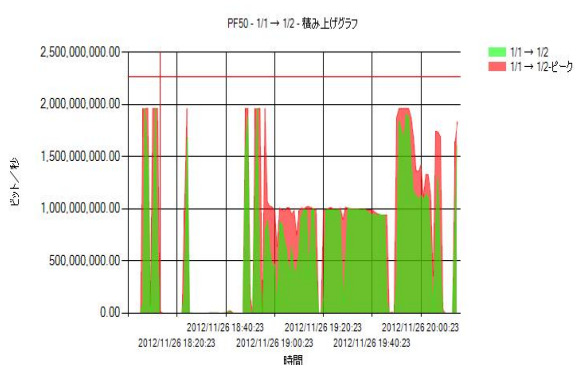


図 18 ネットワークが輻輳しているときのピークグラフ  
Peak graph when networks are congested

## 6 むすび

今回 10 Gbit/s 帯域制御装置 PureFlow GSX-XR を開発した。既存製品である PureFlow GS1 は高性能・高精度で国内帯域制御市場でトップシェアを獲得してきた。PureFlow GSX-XR は PureFlow GS1 の特長である高精度を継承し、性能は 10 倍を達成したことにより、帯域制御装置のリーディングカンパニーとして引き続き帯域制御市場をけん引し続ける。

クラウド業界は目覚ましい進歩を遂げ、日進月歩で技術革新し続けている。今後も PureFlow GSX-XR はクラウドビジネスを更に進化させるキーデバイスになるとともに、市場要請に応え、ソフトウェアによるネットワークの一括制御 (Software Defined Network) などの新技術対応開発などを進めながら進化し続けていきたい。

## 執筆者



安齋 賢二  
アンリツネットワークス(株)  
開発部



吉仲 謙二  
アンリツネットワークス(株)  
開発部



原田 浩  
アンリツネットワークス(株)  
開発部



龍 浩一  
アンリツネットワークス(株)  
開発部



鈴木 雅也  
アンリツネットワークス(株)  
開発部



三枝 淳  
アンリツネットワークス(株)  
開発部



小林 正直  
アンリツネットワークス(株)  
開発部



佐藤 敬幸  
アンリツネットワークス(株)  
開発部



渡部 良太  
アンリツネットワークス(株)  
開発部

告知