Development of 10 Gbits/s Traffic Shaper

Kenji Anzai, Kenji Yoshinaka, Hiroshi Harada, Koichi Ryu, Masaya Suzuki, Atsushi Saegusa, Masanao Kobayashi, Takayuki Sato, Ryota Watanabe

[Summary] We developed the PureFlow GSX-XR as a high-precision bandwidth controller for a cloud service company. It has a processing performance of 10 Gbit/s and a very accurate bandwidth function for controlling network packets. Furthermore, operation cost-cutting is realized by using the GSX-XR traffic monitoring and Web API functions.

1 Introduction

The number of businesses using cloud-based services for their backbone systems is increasing in recent years. Until now, backbone systems have been based on servers configured at each business office. However, data center administrators are gradually promoting a shift to so-called cloud-based services using open Software as a Service (SaaS) and Infrastructure as a Service (IaaS), etc., both as a means for cutting key server and network infrastructure costs, and also to cut operation costs. In particular, more recently, there has been a trend to increased use of cloud services as a means to cut the cost of back-up needed to ensure service continuity in the event of a natural disaster.

With a cloud-based service, the services used by each separate business are aggregated as virtual servers at a data center, increasing the server efficiency. However, since the servers and applications are at one location, the volume of traffic passing over the network tends to increase, driving the deployment of 10-Gbit/s network infrastructure.

We developed the NF7101A PureFlow GSX-XR traffic shaper supporting 10-Gbit/s speeds in this market background. This PureFlow traffic shaper supports flexible and separate bandwidth control for businesses, back offices, applications, etc., for up to 40,000 virtual lines in 8 layered groups. In addition, it prevents TCP communications reduced by using high precision shaping control, and assuring continued stable communications.

Figure 1 Appearance of PureFlow GSX-XR

2 Development Concept

As well as supporting previous business networks, the Pureflow has been developed to support future cloud networks. In addition to high-precision 10-Gbit/s bandwidth control, it supports layered control for cloud services, as well as monitor network traffic and also has a new WebAPI function for coordinated operation with cloud services.

(1) High-accuracy and high-performance bandwidth control

Increases in server performance and network bandwidth are resulting in sudden large increases in traffic volumes at millisecond time spans (microbursts). These microbursts can cause transient traffic congestion on networks, resulting in lost packets and lowering the performance of TCP communications networks. As a result, sometimes the throughput of even 10 GbE data circuits falls far below 10 Gbit/s. Since the PureFlow GSX-XR supports high-precision control of packet send timing at the microsecond level, it increase throughput performance up to maximum 10 Gbit/s as well as average-out microbursts to the ideal packet interval to prevent packet loss, resulting in increased network usage efficiency and optimum cloud network efficiency.

Figure 2 Packet Loss Caused by Microburst Traffic

(2) Layer bandwidth control

When used in a cloud data center, the bandwidth can be controlled for up to 8 layers. Figure 3 shows the traffic in cloud data centers with services, customers, offices and users stratified into layered groups and requiring control of layered bands. Layered bandwidth control supports shaping for up to 8 layers by adding scenarios and filters for each layer. Not only the maximum bandwidth control, but also the communications bandwidth is assured, while scenario and filter settings offer flexible and fine control for data centers.

(3) Support multiple network interfaces

The PureFlow GSX-XR supports various interfaces for 10GBASE-SR/10GBASE-LR/1000BASE-SX/ 1000BASE-LX/1000BASE-T standards and media. Most cloud data center internal networks use the 10 GbE standard but external networks and WAN connections with customers' networks use a variety of standards, including both GbE and 10 GbE. Moreover, users with GbE circuits are planning to change to 10 GbE due to future traffic increases. The PureFlow GSX-XR offers a variety of network interface modules for each speed and media to support the changing customer environment.

(4) Traffic monitoring function

The built in traffic monitoring function is required to optimize asset management. Although network administrators must minimize degraded communications quality caused by traffic increases, wasteful investment in unnecessary infrastructure increases capital and operation costs. As a result, maintaining network quality whilst reduce costs requires a detailed understanding of traffic conditions. Moreover, responsive network troubleshooting requires detailed information. When problem occurs, traffic monitoring can help to confirm whether the total traffic volume is exceeding the network capacity. Traffic monitoring can also be used to identify faulty servers and clients from changes in traffic volumes.

(5) Coordination with cloud services

The PureFlow GSX-XR supports a WebAPI for coordinating operation with cloud infrastructure administration systems to easily automate bandwidth control policy changes.

Figure 4 outlines the PureFlow GSX-XR WebAPI operation. The administration system changes the virtual server configuration according to requests from the operator. Changing the bandwidth control policy via the PureFlow GSX-XR WebAPI assures that the bandwidth control always matches the server configuration.

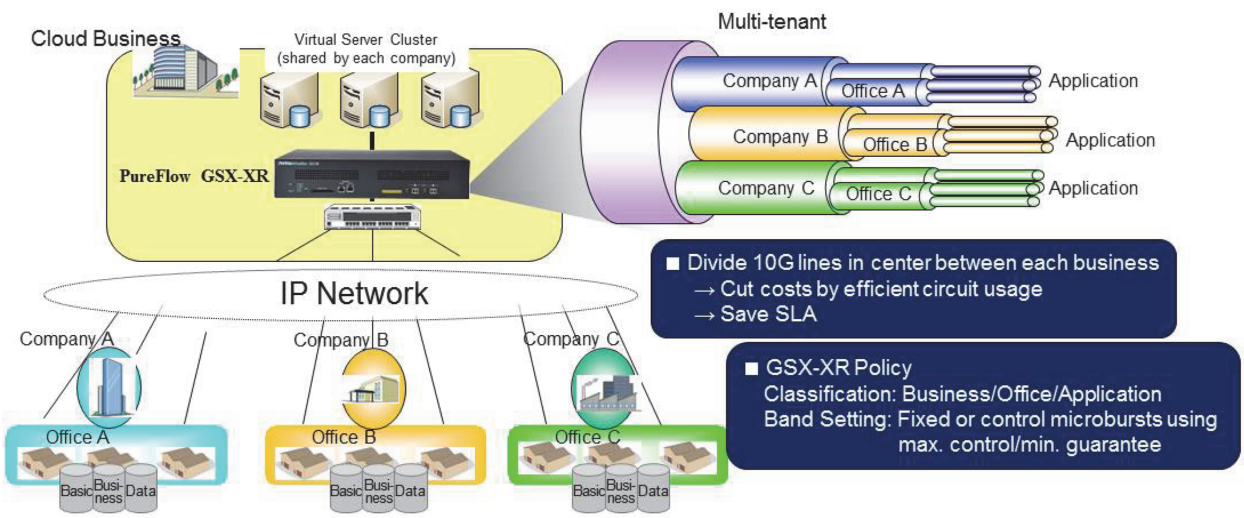

Figure 3 Overview of Cloud Services

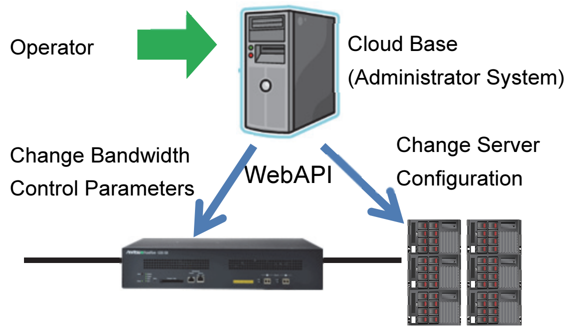

Figure 4 WebAPI

3 Equipment Design

3.1 Composition

Figure 5 shows the block diagram of the PureFlow GSX-XR. It is composed of a management unit, packet controller, power supply unit, and fan unit. To reduce the risk of errors, each unit operates independently. Even if an error occurred in the management unit, network services continue because the packet controller continues without interruption.

3.1.1 High-precision Bandwidth Control Engine

The bandwidth control engine separates received packets into band control units called scenarios. The bandwidth is controlled by controlling the send timing for each scenario. To achieve both high performance and high precision, a multi-core processor CPU is used. Since the packet division processing and schedule processing are performed in parallel, both high processing resolution and high-performance packet transfer are achieved.

The bandwidth control engine packet classification processing not only uses the common IP address and UDP/TCP port number, but also uses the cloud environment QinQ (multi-VLAN) VLAN ID for packet classification, permitting bandwidth control for complex network configurations used by cloud environments, such as identification of customers and services by multi-VLAN.

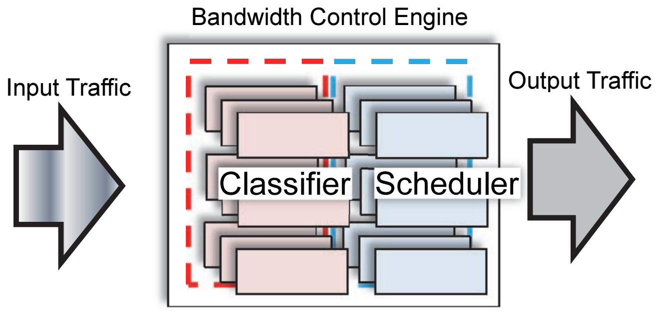

Figure 6 multi-core processor
–Bandwidth Control Engine–

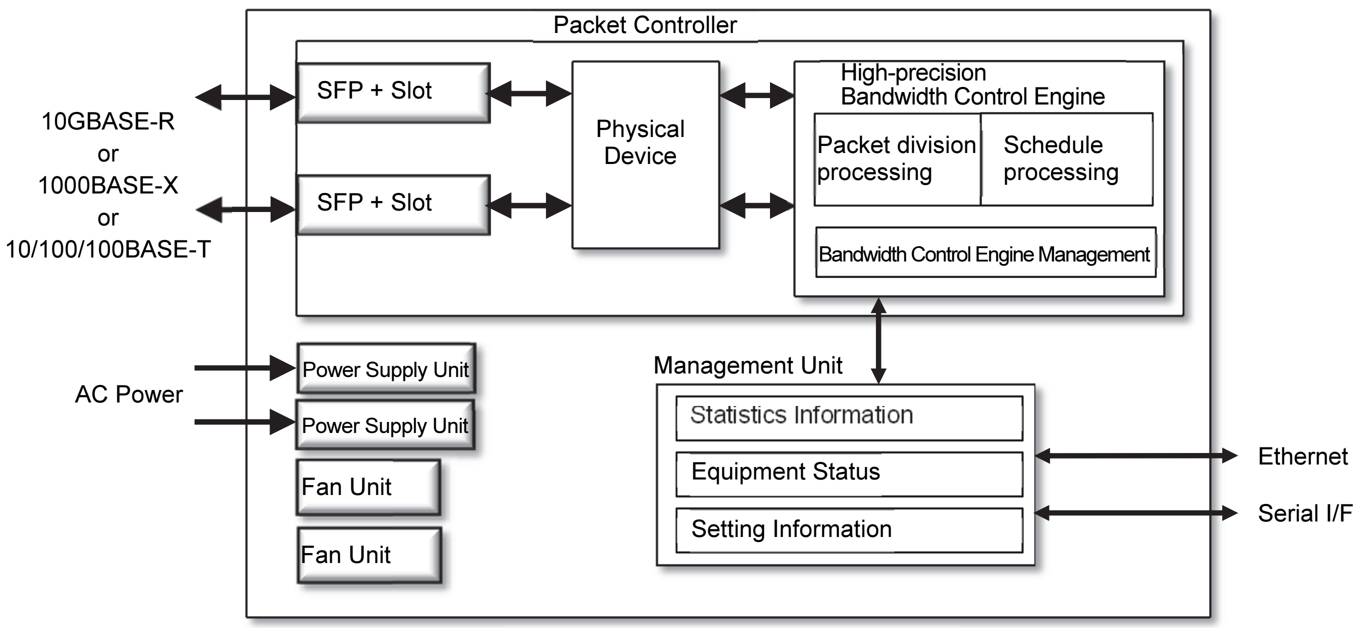

Figure 5 System Block Diagram

The bandwidth control engine scheduled processing controls the packets send timing at a resolution of 1 ppm per packet. This high-accuracy bandwidth control is very effective for microburst traffic and achieves optimum traffic flows with the ideal packet interval.

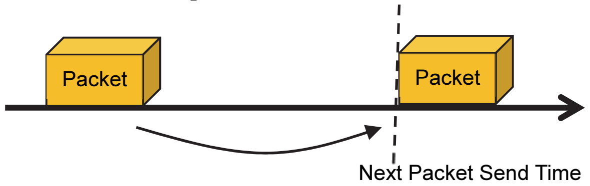

Figure 7 Packet Scheduling

These high processing speeds are achieved by distributed scheduling processing using multiple cores. Received packets are distributed between multiple cores, and packet identification and scheduling processing are performed. Arrange the send timing between cores, the send timing will not overlap even if the packet send timing overlap in multiple scenarios.

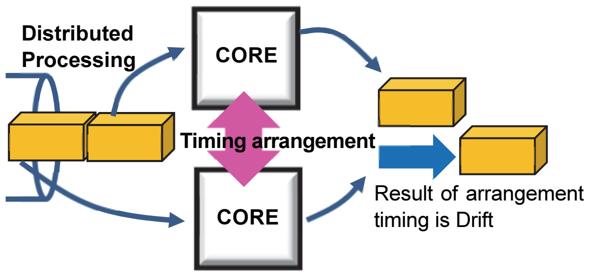

Figure 8 Pallarel scheduling by Multi-core processor

The send timing of the next packet for the scenario with sending delayed by this arrangement is computed in result there is no effect due to sending delay and delays do not accumulate.

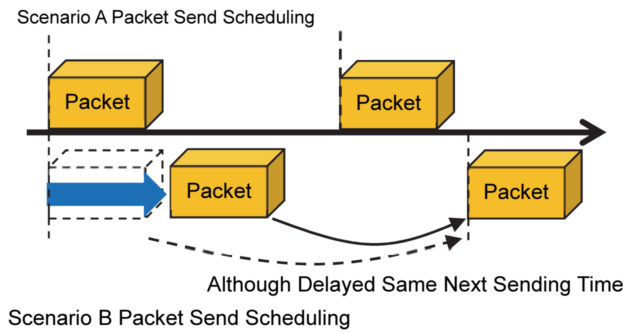

Figure 9 Conflict of Packet scheduling

3.1.2 Phy Layer (Network Interface)

10GBASE-SR/10GBASE-LR are used by the SFP+ module connected via the XAUI interface. To enable connection of 1000BASE-SX/1000BASE-LX/1000BASE-T modules using the same connector, the physical layer is designed to support both XAUI and SGMII connections. Switching between XAUI and SGMII is performed under software control and the circuit speed can be switched to match the installed modules.

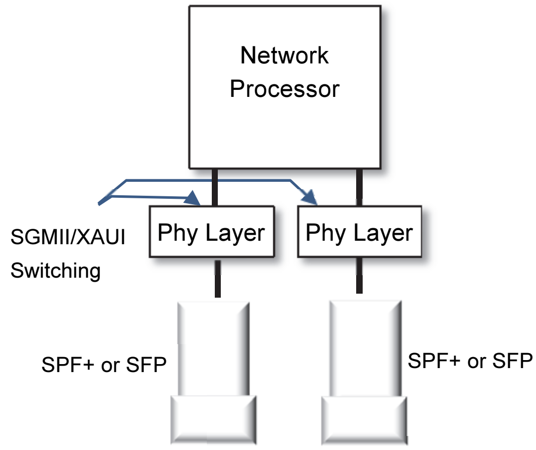

Figure 10 Network Interface

3.1.3 Management Unit

The management unit processes user interface functions for displaying, changing, and notifying setting information as well as managing the equipment status. In addition to common network administration protocols such as serial control, telnet, SNMP, etc., it also supports HTTP supported by the WebAPI.

3.1.4 Bandwidth Control Engine Management Unit

The bandwidth control engine management unit changes and monitors the bandwidth control engine settings, as well as collects statistical information. To collect statistics up to 40,000 scenarios periodically, the bandwidth control engine management unit also uses a high-performance CPU to optimize processing and collection of statistical data at high speed.

3.1.5 Power Supply Unit and Fan Unit

The PureFlow GSX-XR is assumed to be located in somewhere like a data center where it will be collecting network traffic continuously. Consequently, it has easily replaceable power supply and fan units to help assure long life without interrupting network services while being serviced.

4 Effect of High Precision Bandwidth Control

This paragraph explains the effectiveness of precision bandwidth control on wideband networks. Inspection and evaluation are performed using a network emulator emulating the cloud network. The network emulator has a peak transmission rate of 2 Gbit/s and polishing the traffic exceeding this rate. It is located between the HTTP server of the data center side and the HTTP client of the client side.

In the test, a 1-MB file is downloaded repeatedly from the server side to the client side to measure the time to transfer 1 GB and calculate the transmission rate.

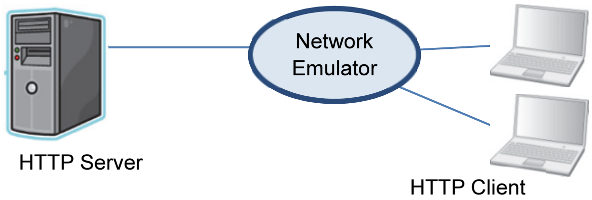

Figure 11 Test Network (without Shaping)

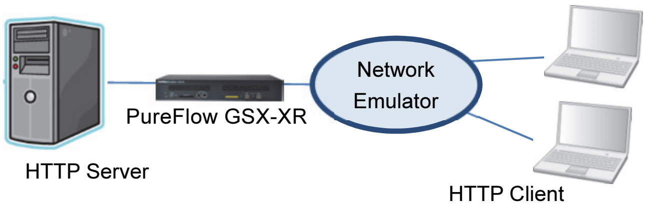

Figure 12 Test Network (with Shaping)

Table 1 Transfer Rate and Total Download Times

Bandwidth Control	Avg. Transfer Rate	Max. Transfer Rate	Total DL Time
Without	1.2 Gbit/s	1.9 Gbit/s	28.7 s
With	2.0 Gbit/s	2.0 Gbit/s	16.9 s

Table 1 lists the measured network emulator transmission rates and connection times until file download is completed. The average transfer rate is faster and the time until all files are downloaded is shorter with bandwidth control (traffic shaping) than without it.

Additionally, Figures 13 and 14 show the traffic status with and without bandwidth control sampled at a resolution of 1 ms. The red graph is the HTTP server transmission rate. The green graph is the HTTP clients receive rate via the network emulator.

Without bandwidth control (Figure 13), the average rate is only 1.2 Gbit/s although the effective rate of the network emulator is 2 Gbit/s.

The cause of this low rate is packet loss due to burst traffic. Packets are lost at the network emulator as a result of these bursts and TCP collision control is executed, resulting in an unstable transmission rate and suppressing the overall effective bandwidth.

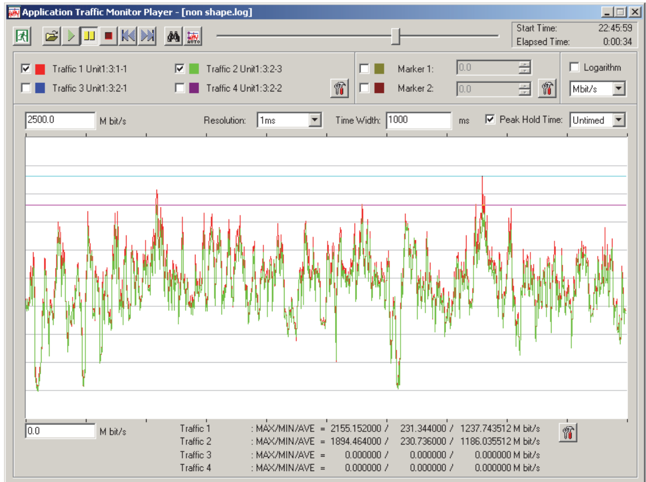

Figure 13 Download Traffic (without Shaping)

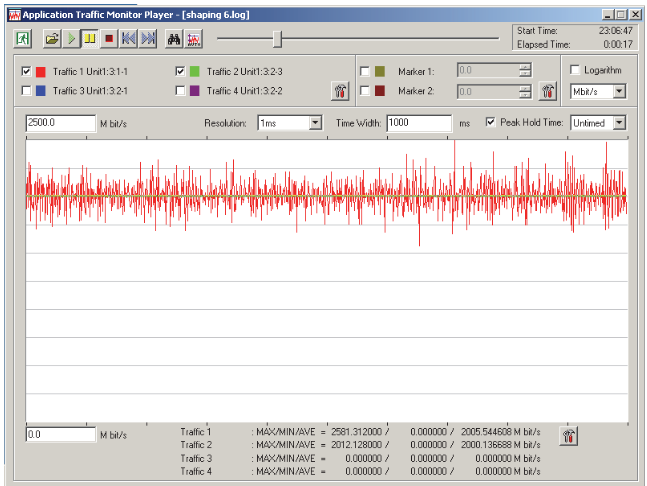

Figure 14 Download Traffic (with Shaping)

On the other hand, with bandwidth control (Figure 14), the microbursts seen in Figure 13 are suppressed and the transfer rate is stable. Since there is no packet loss due to microbursts, there is no packet loss in the network emulator and the transfer rate is stable, resulting in a nearly doubled average transfer rate and efficient data transmission.

5 Traffic Monitoring

The PureFlow GSX-XR incorporates scenario counter, top counter, and peak rate monitor functions for monitoring traffic to maintain optimum network conditions. This section explains some examples and methods using graphs obtained with the NF7201A Monitoring Manager 2 software for collecting data on the PureFlow GSX-XR traffic. In addition to graphing the traffic, traffic reports can be also generated automatically at regular periods with this software.

5.1 Monitoring Traffic Flow rate using Scenario Counter

This scenario counter monitors the average bandwidth for each type of traffic and plots traffic graphs for each month, quarter, half year, and year, etc., to evaluate the increasing tendency of traffic.

Figure 15 shows an example of the graph of traffic flow rate using this scenario counter. The PureFlow GSX-XR can monitor traffic flow rate in scenario units. For example, a scenario can be allocated to each traffic type and the volume of traffic can be color-coded to support visual identification of changes in total traffic and ratios of specific traffic types. Using these graphs, we can see the traffic peak time band is 13:00 to 14:00 as well as the presence of burst peaks and traffic types in the peak time band.

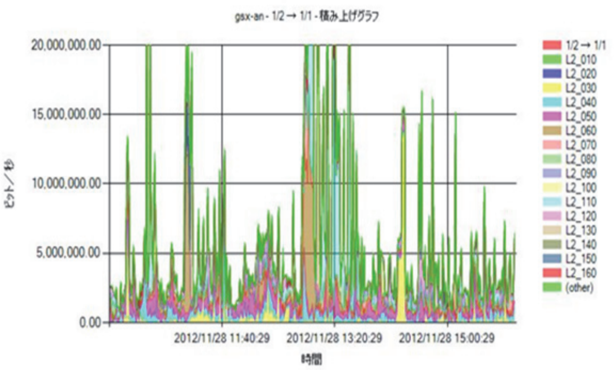

Figure 15 Scenario Graph

5.2 Top Count Monitoring

Some networks are occupied by just a few appliances and these appliances must be specified to optimize network traffic.

Figure 16 shows an example of using the PureFlow GSX-XR top counter function to display the IP addresses of 20 network appliances occupying the network resources. In addition to the IP addresses, it can also rank the application ports with highest traffic volumes in order.

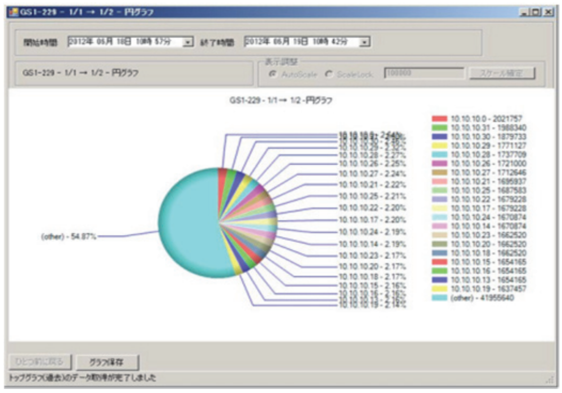

Figure 16 Top Graph

5.3 Monitoring Peak and Average Rates

Comparing peak and average rates can help visualize network resource availability. Here, peak rate is the maximum value measured by every 1second within 60seconds (60 times). The average rate is the average for the 1-minute period. When the difference between the peak rate and the average range is large, it seems like there are surplus network resources but conversely, when there is little or no difference, traffic which is tight in a network resource is continuing flowing.

Figures 17 and 18 show graphs of the peak band and average rate obtained using the PureFlow GSX-XR peak rate monitor function. Figure 17 is the peak graph without network congestion. When the traffic load is low, the peak rate increases compared to the average rate. In contrast to this, Figure 18 shows the overloaded network condition when there is almost no difference between the average and peak rates. Using this type of overlapping comparison permits visualization of the network load for network resources.

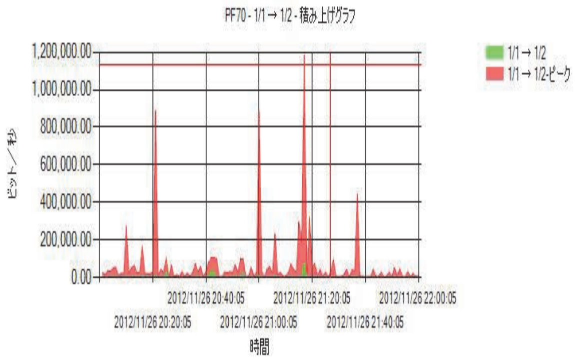

Figure 17 Peak Graph when Network Not Congested

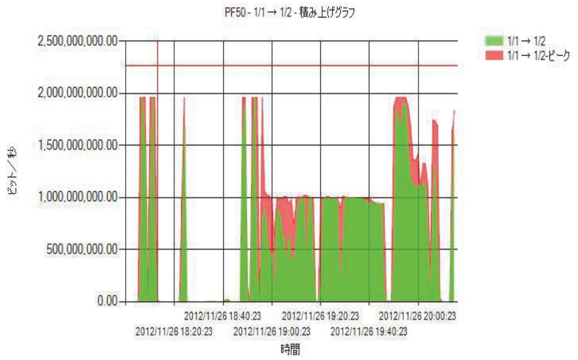

Figure 18 Peak Graph when Network Congested

6 Conclusion

We have developed the PureFlow GSX-XR for shaping traffic on 10-Gbit/s networks. Its predecessor PureFlow GS1 with high-performance, high-accuracy bandwidth control functions has taken the top share in Japan's traffic shaping market. The PureFlow GSX-XR builds on the PureFlow GS1 features with 10times better performance and continues to drive Anritsu Networks' lead in the traffic shaping market.

The cloud network market is changing rapidly with technology innovations. As well as a key device in the cloud business revolution, the PureFlow GSX-XR will help to drive developments of future technologies such as Software Defined Networks, offering collective network control to meet new market needs.

Authors


Kenji Anzai
Anritsu Networks Ltd.
Development Division


Kenji Yoshinaka
Anritsu Networks Ltd.
Development Division


Hiroshi Harada
Anritsu Networks Ltd.
Development Division


Koichi Ryu
Anritsu Networks Ltd.
Development Division


Masaya Suzuki
Anritsu Networks Ltd.
Development Division


Atsushi Saegusa
Anritsu Networks Ltd.
Development Division


Masanao Kobayashi
Anritsu Networks Ltd.
Development Division


Takayuki Sato
Anritsu Networks Ltd.
Development Division


Ryota Watanabe
Anritsu Networks Ltd.
Development Division