Development of PureFlow Profiler Traffic Monitoring and Analysis Software

Takayuki Sato, Kenji Anzai, Tomohiro Wakita, Masato Aketo

[Summary] The recent spread of cloud services and remote working is driving explosive increases in network traffic. Moreover, networks are more complex and fault troubleshooting is becoming more difficult. Since work productivity drops remarkably as more network faults occur, businesses face increasing challenges in how to maintain network quality. To solve this problem, we have developed this PureFlow Profiler software described here to visualize traffic con-ditions, and quickly discover and recover faults.

1 Introduction

With the migration of business trunk systems from previous on-premises services to the cloud, system configurations have become very complex with a mixture of on-premises and cloud servers. Additionally, the rapid spread of remote working, online meetings, etc., is increasing network loads, causing faults such as inability to connect to the network and slow servers which impact work efficiency. Typical causes of these faults are listed below.

(1) Line Congestion Faults

As parts of a network handle larger traffic volumes, some segments become congested, causing instability and delay between clients and servers, packet loss, and retransmission. Additionally, if this traffic concentration occurs periodically, it can be very difficult to characterize the cause of the fault and recover the normal condition.

(2) Equipment and Aging Faults

Communications can be dropped by faults occurring in network equipment such as switches and routers, as well as in servers, resulting in interrupted network services. Additionally, aging prior to the occurrence of an equipment fault can cause unstable communications.

Troubleshooting fault causes and locations requires visualizing the network communication conditions. 24/7/365 network monitoring and analysis permits rapid responses to fault occurrences and minimizes network downtime.

We developed the PureFlow Profiler (PFP) software to monitor network traffic and help solve these challenges.

2 Development Concept

The PFP software is composed of server and client parts (Figure 1). The server software collects statistical data from up to 255 bandwidth controllers (PureFlow nodes) installed on the network and saves the data to a database for consolidated management. The client software connects to the server software to visualize information stored in the database.

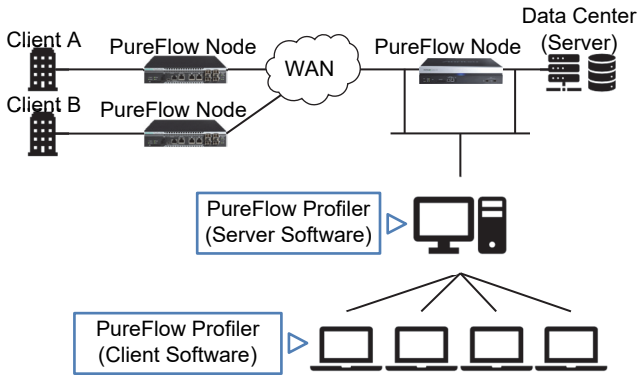

Figure 1 System Configuration

Using PFP supports individual simultaneous monitoring of multiple clients and servers as well as multiple PureFlow nodes on the network to efficiently monitor and analyze the overall system traffic conditions. To help troubleshoot fault locations and faults, PFP has been designed bearing the following points in mind to facilitate fast recovery.

2.1 Real-time Data Visualization and Analysis

To help the system administrator troubleshoot fault causes and locations as quickly as possible, PFP collects statistical data in real time from PureFlow nodes. It has functions for visualizing and analyzing real-time changes not only in traffic volumes but also in latency, packet loss, etc., to help administrators understand network communications conditions.

2.2 Visualization and Analysis of Long-Term Data

Comparing normal and abnormal traffic conditions when a fault occurs requires collection of long-term statistical data. The PFP software collects long-term time-based data to visualize and analyze changes in traffic by day, week, month, and year.

2.3 Simple Graph Displays

To compare previous and current traffic conditions, the software makes it easy to specify the graph display range. In concrete terms, a 1-click operation designates the fault analysis time range between 00:00 to 23:59 yesterday and 00:00 to the current time today.

2.4 Fast Error Detection

When a fault is detected by normal monitoring and analysis, the system administrator is notified immediately about the occurrence time, location, fault condition, etc., using two network management protocols: SNMP (Simple Network Management Protocol), and SYSLOG (System Logging Protocol).

3 Basic Functions

In addition to traffic volumes and flow counts (minimum traffic recognized by PureFlow node), PFP also collects long-term measurement data indicating communications quality, such as Round Trip Time (RTT), packet loss rate, etc. The PFP software has the following basic functions for analyzing changes in communications quality not only to review line bands and bandwidth control policy but also for early fault discovery. The key PFP specifications are listed in Table 5.

3.1 Traffic Monitoring

(1) Visualizing Traffic Volume

By visualizing average Tx traffic (Figure 2) and bit rates (Figure 3) on the network, PFP facilitates monitoring and analysis of large amounts of communications data for the line band. Average Tx traffic volumes can be separated by traffic type using a stacked line graph to understand both changes in specific types of traffic and overall traffic. The maximum bit rate per minute can be displayed by measuring Tx traffic bit rates every second. Comparisons showing large differences with average Tx traffic rates can help understand traffic flows when network resources are limited or not.

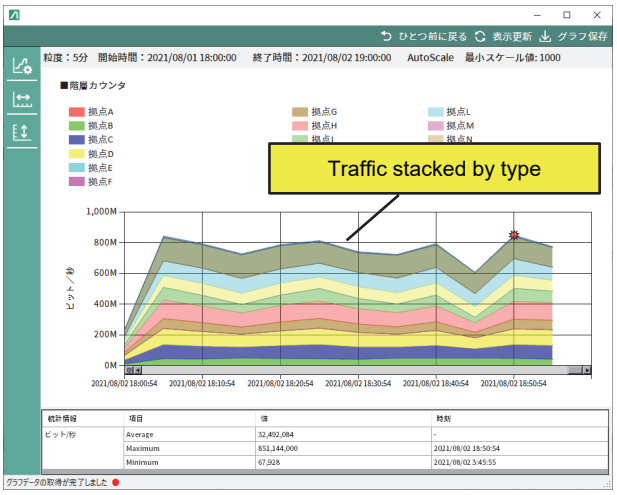

Figure 2 Average Tx Traffic

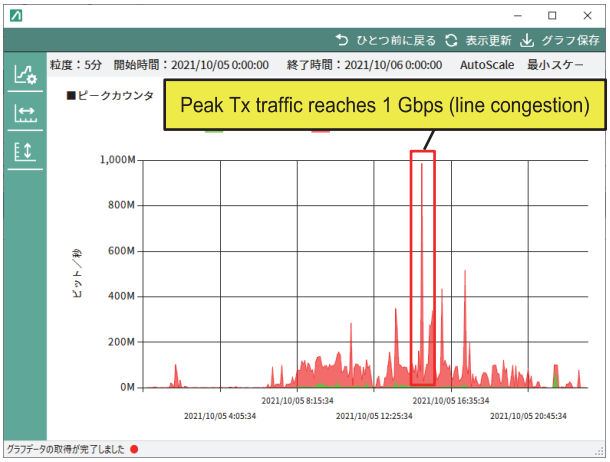

Figure 3 Peak Rate

Flow count graphs can indicate time ranges with most users to confirm predicted performance limits and update facilities reaching these limits (Figure 4). Moreover, statistical data can be displayed along with markers on graphs to facilitate easy understanding of maximum and minimum values in the graph display range.

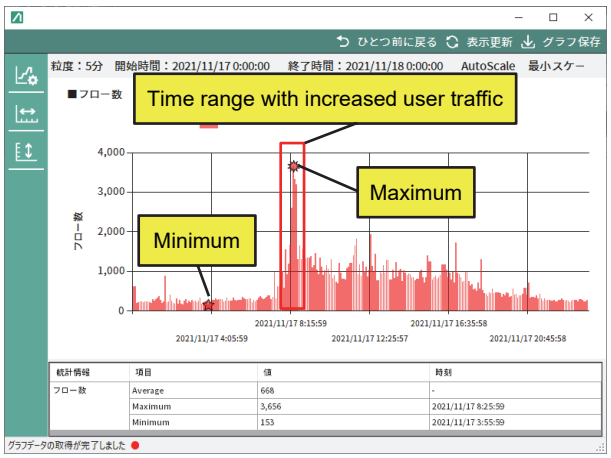

Figure 4 Flow Count

(2) Visualizing Top Traffic Users

On most networks, a very small number of users occupy a very large proportion of the network resources, which can sometimes reduce service quality for other users. PFP can rank these super-users according to their traffic volumes to help identify potential problems. The IP addresses and application port numbers for the top 25 users of all traffic passing via the PureFlow nodes are plotted as a pie chart to visualize top users and clarify traffic-usage conditions. The pie chart in Figure 5 indicating that the top two users take 73% of the bandwidth is the key to understanding the importance of network optimization by limiting communications using band shaping.

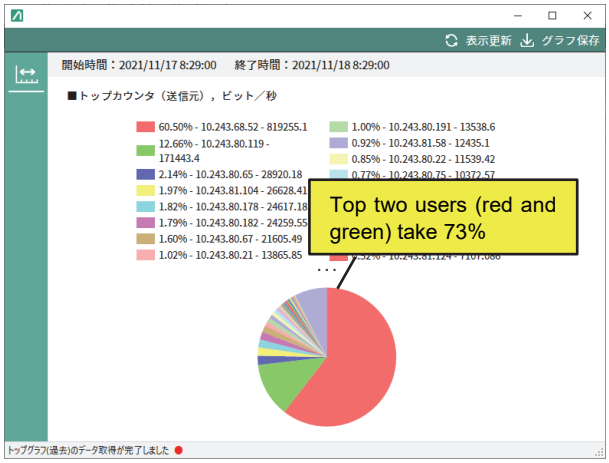

Figure 5 Top 25 Users

3.2 Traffic Analysis

PFP collects latency, packet loss, and retransmit data measured by the PureFlow nodes (Figure 6) to visualize traffic analysis data and understand network and server performance conditions.

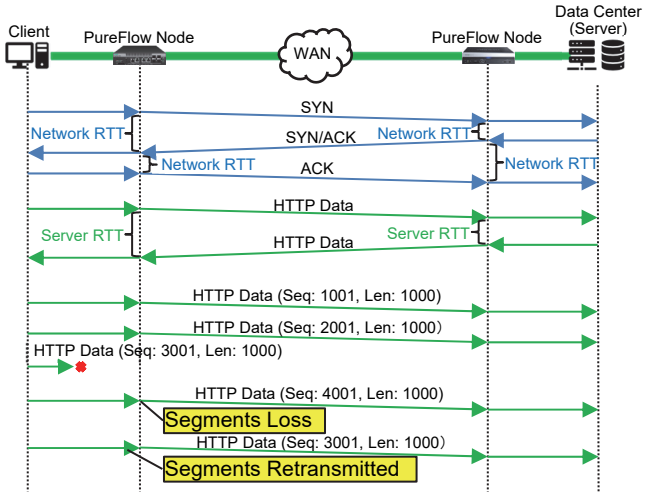

Figure 6 Traffic Analysis Points

(1) Visualizing TCP Latency

To troubleshoot whether a fault is caused by the network or the server, both the network latency (Network RTT) and server latency (Server RTT) can be visualized. Network RTT displays the round-trip delay of the network itself without passage via the application, while Server RTT displays the application start-up time. Network RTT is the time from when the client sends the TCP connection request (SYN) packet until it receives the send request response (ACK or RST) packet from the other side. On the other hand, Server RTT is the time from when the client sends the first data packet until it receives the data packet from the other side.

PFP displays these Network RTT and Server RTT times as graphs and histograms to help understand the randomness and dependency trends for a specific RTT.

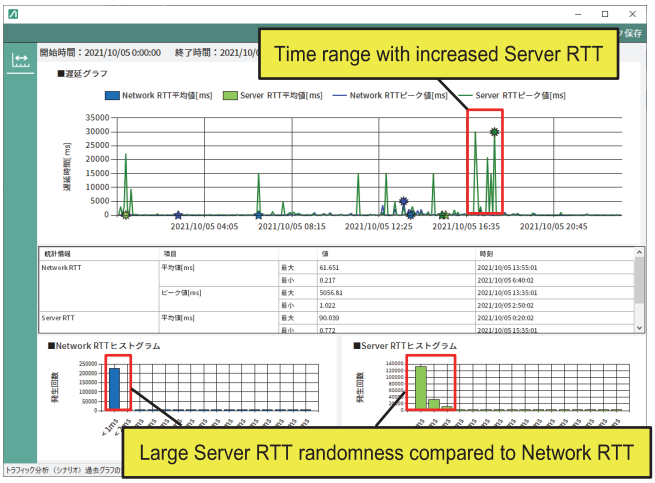

Figure 7 Latency

These latency graphs can help identify whether the network or the server is causing degraded communications quality to take countermeasures such as re-evaluating the communications line and server equipment. Additionally, analyzing the histograms helps understand unusual network or server behavior indicating a potential future problem.

(2) Visualizing TCP Packet Loss and Retransmission

Confirming the communications quality status requires referencing the TCP sequence number of each packet and then visualizing the packet loss and retransmit byte occurrence counts. Packet loss is evaluated as occurring when the TCP header sequence number is bigger than the sequence number of the next packet to be received, and retransmission is evaluated as occurring when the sequence number is smaller.

PFP calculates the packet loss and retransmission rates using these packet loss and retransmission data to display the rates as graphs (Figure 8), helping understand time ranges with degraded communications quality.

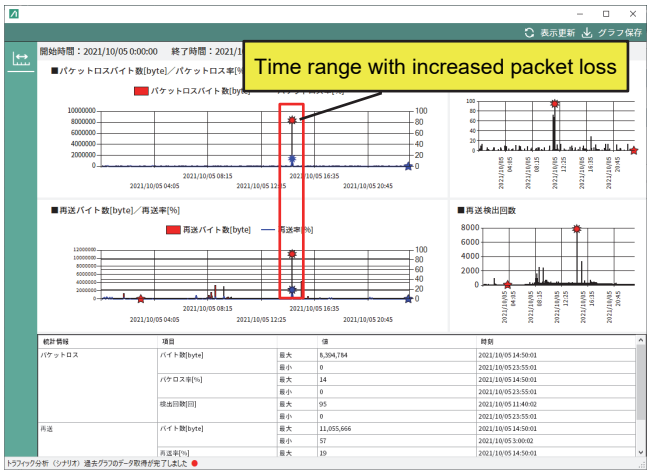

Figure 8 Packet Loss and retransmission

Additionally, comparing packet loss and retransmission rates for each PureFlow node can identify network segments with high packet loss and retransmission occurrences.

3.3 Alarm Notifications

PFP can set sending of both SNMP Trap and Syslog event notifications when normal daily monitoring and analysis detects events exceeding pre-set hi and lo threshold values and continuous detection counts (Table 1). This PFP alarm notification function helps understand intermittent network abnormalities by detecting excessive or too little Tx traffic volumes and flows.

Table 1 Threshold Settings

Threshold Setting	Hi Limit	Lo Limit	Continuous Detection Count
Tx traffic [bps]	○	○	○
Flow count [flows]	○	○	○
Network RTT [ms]	○	—	○
Server RTT [ms]	○	—	○
Packet loss rate [%]	○	—	○

4 Features

4.1 Visualizing Overall Data

Since fault analysis using separate graphs requires more time, PFP also has a function to batch display and print graphs (Figure 9) to show correlations between different data in the same time range and help system administrators with overall analysis of trends in network and server operating conditions.

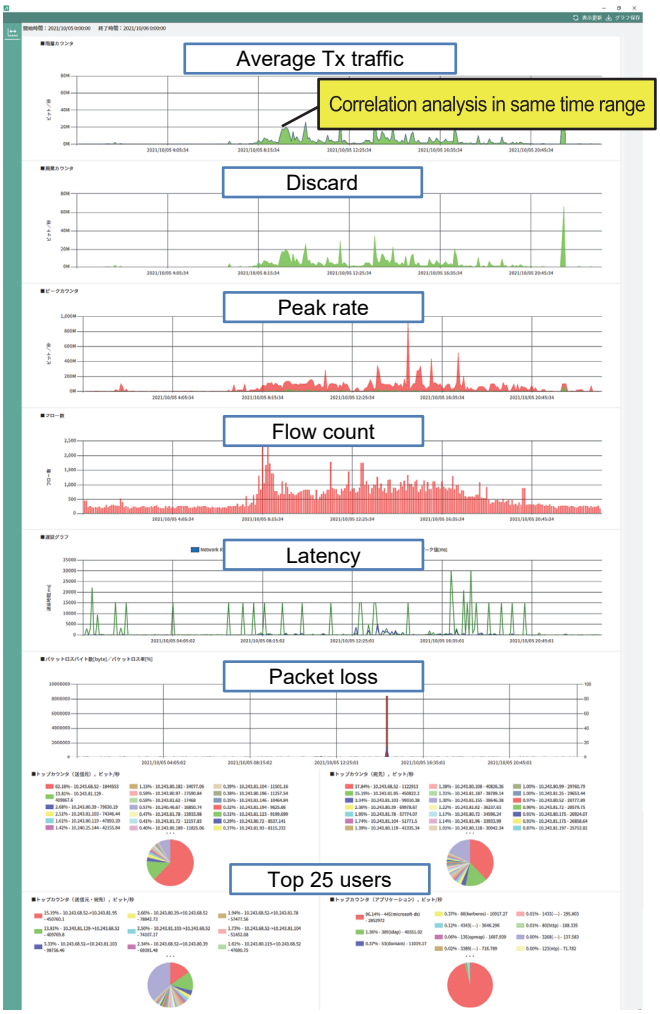

Figure 9 Overall Data Graphs

Table 2 lists the types of overall data graphs.

Table 2 Overall Data Graph Types

Data Type	Graph Category
Monitoring data	Average TX traffic, Congestion, Bit rate, Flow count
Top data	Top
Analysis data	Latency, Packet loss

Analyzing the overall data graphs can help evaluate line congestion conditions as well as correlations, such as latency and packet loss occurrence due to users' usage conditions, helping diagnose faults related to deteriorated communications quality. For example, monitoring and traffic-analysis data can be compared for time ranges with slow speeds and dropped server connections. In time ranges with high packet loss and high retransmission levels, information on faults such as unoccupied line bands can be investigated to support

appropriate countermeasures. Moreover, confirming by matching with top user data can help understand which users are taking excess line bandwidth before implementing measures such as bandwidth shaping policies.

4.2 Visualizing RTT Segment Bottlenecks

The locations of all bottlenecks in the entire system must be understood when determining fault locations. Although measured data collected from each PureFlow node can be analyzed for each measurement point, it is not possible to analyze which segments in the overall network have bottlenecks. As a result, PFP extracts data from the database about up and down flows using the RTT information for each measured point and calculates RTT differences between adjacent nodes as RTT segments for display as a ladder diagram. Indicating the maximum value for each RTT segment offers an intuitive indication of which network segments have large latency.

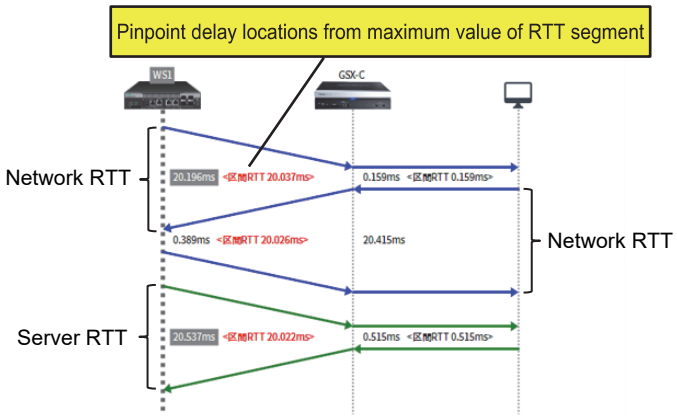

Figure 10 Ladder Structure

4.3 Active Measurement

Network measurement methods are broadly divided into active and passive methods. The passive method measures packets travelling on the network between the client and server. On the other hand, the active method measures by sending packets spontaneously to the target device.

Since the active method must be used to monitor and analyze a system even when communication requests are not sent from the client to the server, PFP can send packets periodically from a PureFlow node to any server to measure the RTT, packet loss, and retransmission rates. Active measurement generates either Hypertext Transfer Protocol Secure (HTTPS) or Internet Control Message Protocol (ICMP) traffic at 1-minute intervals sent to any server to measure and analyze traffic. Figure 11 shows an example of generated ICMP traffic. PHP visualizes the time from when an ICMP Echo Request is first sent until the ICMP Echo Reply is received from the other side as the Network RTT.

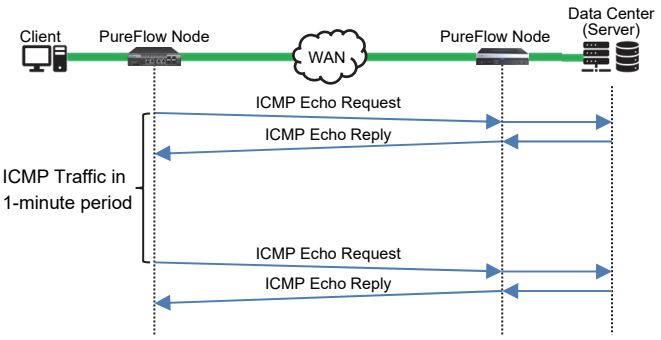

Figure 11 Active Measurement (ICMP Traffic)

4.4 Faster Graph Displays

When confirming the operation status of clients and servers, although analysis of prior events over the long term is possible, real-time analysis in the short term requires processing time to extract and plot all data from long-term graphs possibly covering years-long spans. Consequently, PFP accelerates graph display processing times by supporting selection of the data granularity for the graph display range (Table 3).

Although PFP collect data from PureFlow nodes into a real-time database using a collection period, the collected data can be sampled at different granularities of 5 minutes, 1 hour, 3 hours, and 1 day to create a database with the optimum granularity. Averages for Tx traffic volumes for monitoring data, maximum values for flow counts, and top 25 positions calculated from adding values for each IP address and application port number for top data are extracted from the relevant database for the relevant time period to aggregate data for each granularity.

Table 3 Graph Display Range and Usage Data

Display Range	Date Use
<24 hours	Real-time data
24 h to <7 days	5-minute granularity
7 days to <1 month	1-hour granularity
1 month to <1 year	3-hour granularity
≥1 year	1-day granularity

4.5 Optimizing Database

The size of the database containing aggregated data is optimized to satisfy long-term analysis requirements. Since storage capacity is limited, separate save periods can be set

for each data granularity to prevent capacity becoming too small, and data outside the save period is overwrite to limit data aggregation.

Table 4 Database Save Intervals

Monitored Data	Save Interval	Default
Real time	1 to 60 days	1 day
5-minute granularity	1 to 300 weeks	4 weeks
1-hour granularity	1 to 24 months	1 month
3-hour granularity	1 to 5 years	1 year
1-day granularity	3, 6, and 9 years	3 years

5 Supporting Silent Fault Detection

Sometimes, part of a network may dropout due to a network or server fault, server bug, etc. In this case, the equipment often does not recognize the fault and even normal monitoring does not detect the problem. These problems are called silent faults and are often noticed by users as an inability to connect. These silent faults are most effectively detected by monitoring communications for each network route.

The non-communicating condition can be detected using the PFP alarm notification function. Active measurement from a PureFlow node to any server and monitoring the result with PFP can detect non-communicating routes. Figure 12 shows an example of flow counts generated when there are no communications.

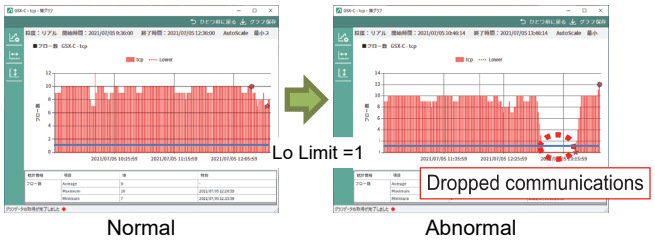

Figure 12 Flow Counts

Setting the flow count lo limit to 1 detects flow counts of 0 (= no communications) and issues an alarm notification. Using this method to monitor each route to each server helps clarify partial network faults.

6 Conclusions

This article has described PFP functions and features with the main focus on TCP communications latency and packet loss. Using PFP for monitoring and analysis helps early fault discovery and recovery. As accelerating rollout of 5G networks results in new applications and services, new challenges are expected in monitoring and analyzing 5G communications quality.

We hope PFP will play a role in improving network system communications quality even further by supporting future analysis of applications, such as video using the Real-time Transport Protocol (RTP) and planned upgrades, such as automating predictive fault detection.

References

1) WHITE PAPER Information and Communications in Japan, Ministry of Internal Affairs and Communications, 2021

Authors


Takayuki Sato
Solution Development Dept.
Infrastructure Resilience Division
Environmental Measurement Company


Kenji Anzai
Solution Development Dept.
Infrastructure Resilience Division
Environmental Measurement Company


Tomohiro Wakita
Solution Development Dept.
Infrastructure Resilience Division
Environmental Measurement Company


Masato Aketo
Solution Development Dept.
Infrastructure Resilience Division
Environmental Measurement Company

Table 5 Specifications

	Item				PureFlow Profiler	Remarks
1	Node Management				○	255 nodes max
2	Data Monitoring	Network port counter			○	—
		Scenario counter (layer, congestion)			○	—
		Bit rate			○	1-minute peaks
		Flow count			○	User comms count
3	Top Usage	Top counter	Tx source, destination, applications		○	Top 25
4	Traffic Analysis	Scenario analysis (protocol totals per scenario)	TCP	Network RTT	○	Mean/max/min/dist
				Server RTT	○	Mean/max/min/dist
				Packet loss	○	Loss rate/bytes/times
				Retransmission	○	Retransmit rate/bytes/times
		Flow analysis (Top total per flow)	Top TCP	TCP flow with highest latency	○	Top 100
			Top ICMP	ICMP flow with highest latency	○	Top 100
			Segment delay (RTT segment)		○	Ladder plot
5	Graphs	Real-time graph			○	—
		Elapsed-time graph			○	—
		Overall data graphs			○	Set of graphs
6	Reports	Periodic report (CSV/HTML)			○	—
		Manual report (CSV/HTML)			○	—
7	Graph/Report Types	Stacked line, line, pie, bar			○	—
		Line/bar/histogram			○	Network RTT/ Server RTT graphs
		Flow list/ladder chart			○	Segment analysis
8	Alarms (Syslog Comms/ SNMP Trap Tx/ Console Display)	Traffic rate			○	Hi/lo limits
		Flow count			○	Hi/lo values
		TCP Scenario analysis	Network RTT/Server RTT (ms)		○	Hi value
			Packet loss (%)		○	Lo value
9	Database Save Period	Monitor analysis			○	9 years max
		Top data			○	1 year max
		Traffic analysis data			○	1 year max
10	Backup	Online backup/restore			○	—
11	Scenario Comment	Import/export			○	Batch operation
12	Radius Authentica-	Local authentication/Radius authentication			○	—
13	Connection model	NF7101C PureFlow GSX			○	—
		NF7501A PureFlow WS1			○	—

Publicly available