Distributed measurement system for the end-to-end network performance using a dynamic job management

Takayuki Kushida (kushida@acm.org) and Kunikazu Yoda
IBM Research, Tokyo Research Laboratory

Toshio Nishiyama and Tetsuo Takahashi
NTT Communications, Network Business Division

The network monitoring and the analysis of the end-to-end performance is a critical task for Internet Service Providers (ISPs) that would like to estimate their subscribers' actual network performance to improve their service quality. Although many research studies for the network performance monitoring have been only focused on the Internet backbone, inter-AS connectivity with their monitoring systems, the study of a monitoring infrastructure for the distributed small clients like home appliances and PCs is quite few.

In addition, ISPs would like to know the quality of their subscribers' network in real time. Since their subscribers are distributed and have client PCs with a very short period time of the network connection, we can't directly apply the traditional network measurement methods to the environment. This paper proposes a new approach for the end-to-end network measurement that is focused on the performance for ISP subscribers to the Internet hosts.

Our contributions of the paper are as follows:

1. Design and development of a new approach for the end-to-end network performance monitoring with a dynamic program loading to distributed small clients like home appliances and PCs, and the analysis of the monitoring data is done at a central server.

2. The experiment to evaluate the proposed method on the operational ISP network based on the topology construction with the analysis of several links from clients.
1. Introduction

- Objective
  - Study a distributed measurement for the end-to-end network performance on ISP network using a dynamic job scheduling
- Internet is a best-effort service and a common infrastructure
  - A simple and powerful mechanism has been a critical to deploy the network as a common infrastructure. High speed networks and wireless networks have been deployed without QoS.
  - The Internet currently doesn’t have a QoS guaranteed at all. There have been a lot of efforts to define the set of network standards for the Internet QoS.
- Network performance measurement vs QoS network
  - Instead of providing the full guaranteed QoS function for the operational networks, the network performance should be measured at the end-to-end level.
- ISP (Internet Service Provider) requirements
  - They provide a high-speed broadband service and want to provide a differentiated service for a network quality for their customers. They currently provide only the connection service.
  - In ISP, there are regional networks, the core network and the rest of the Internet. It is difficult to measure the end-to-end network performance including PCs and home appliances with the current network management configuration.

The Internet is a best-effort service using a packet switched network. A simple and this powerful mechanism to forward a packet to a destination node on the Internet is a critical reason to be able to widely deploy the network as a common infrastructure. On the other hand, although there are many research studies for a QoS guaranteed network, the current Internet hasn’t become QoS guaranteed network[1]. Therefore, the end-to-end performance for all packets on the network isn’t assured at all. The Internet is commonly used as a communication infrastructure because there is the advantage for the scalability of its network scalability to a global size. Many studies for the end-to-end performance measurement have been done within a network research community recently. Many applications on the Internet need the estimation of the performance among end-to-end hosts because they provide the better service to end users[2][3][4]. On the other hand, there are a lot of Internet Service Providers (ISPs) to provide the connection service for the Internet with a broadband using high speed links, but ISPs can only offer the connection service without a better performance service to their subscribers. ISPs can only monitor the performance of their backbone networks with the network management system, but it is hard to get an enough information of the network end-point which is located at subscriber’s house or small office. Because there are regional networks that split their subscribers from the backbone network of ISP. As a results, ISPs’s subscribers logically connect to the ISP network through the regional networks that provide the connection service at a regional area such as a prefecture and the access link service using the optical fiber and the high speed DSL to their subscribers. ISPs would like to provide a better service than just the simple network connectivity. ISPs have several issues and requirements for the current end-to-end network monitoring. In general, there are two different kinds of monitoring methods which are an active and a passive probing for the end-to-end network performance. A newly developed network monitoring program is usually introduced from the network research community to be public as a published research paper. The activity for the development of a new monitoring program will be continued in future. In addition to introduce a new monitoring program, pre-setting parameters for the monitoring program should be changed based on the actual network status and ISPs’ experiences on the actual operational network. There is also a requirement that the monitoring data at many sites should be sent to the central site and stored for the analysis of the future statistical trend for the ISP network. The analysis and browsing tasks for a network monitoring should be done by the central site. This activity of monitoring and the time synchronization for monitoring data should be also managed by the central site. Because a precise time at many clients are usually not managed by their client owners. In this results, the design approach of a network monitoring infrastructure would be that the network monitoring activities should be done at distributed network end-points and the data analysis should be done at the central site.

2. Background (I)

ISP provides the high speed service for their customers

- For example, the ISP service called OCN provides ADSL from 1Mbps to 50Mbps and FTTH with 10Mps/a dedicate 100Mbps/shared 100Mps as access links. Bandwidth and performance estimation is more important than before.
- ISP subscribers must connect to ISP via a regional IP network that is different company. It’s difficult to estimate the end-to-end performance on the network with subscribers PC continually.

Related studies of the end-to-end network performance measurement:

- PlantLab is a distributed test bed for providing a distributed environment, and it supports a virtualized service called “slice”. In the PlanetLab, the triangle inequality of roundtrip times among Planet nodes was investigated with end-to-end measurement[9]. It showed how PlanetLab has a similarity with the whole Internet.
- The study of the network measurement on the PlanetLab basically uses the Planet nodes that are located at sites in research organizations. On the other hand, this paper for a distributed measurement system focuses on the end-to-end network performance from many clients that are located at residential or business premises in ISP.
- The approach of the distributed environment to collect the end-to-end network performance data has been reported as NET@HOME [10]. It has a capability of a passive packet monitoring at a distributed host, and its monitored data is sent to the central server to analyze the network performance.
- SPAND (Shared Passive Network Performance Discovery) is the mechanism to collect network information passively. The shared data has been analyzed.
- Tool called “pathchar” can investigate the end-to-end performance on the path, but can’t get the precise value if value is too large.
- The bandwidth investigation of “packet-pair and “packet-tailgating” has been studied.

2. Background and previous studies (I)

This section describes the related studies of the paper, and both a distributed network measurement and a dynamic task management that we propose as a background. The PlanetLab is a distributed test bed to provide a distributed environment with the overlay network infrastructure on the Internet[8], and it supports a virtualized service called “slice”. In the PlanetLab, the triangle inequality of roundtrip times among Planet nodes was investigated with end-to-end measurement[9]. It showed how PlanetLab has a similarity with the whole Internet.

The network tomography is the study to understand the Internet inside as a performance characteristics [13]. The path quality tool called “pathchar” is well-known to investigate the end-to-end performance on the path. This tool can precisely estimate the delay of the path, but can’t get the precise value if value is too large [14]. In addition, there have also been studies to investigate the bandwidth with “packet-pair” and “packet-tailgating”. In those studies, the method of the packet-tailgating method can give us precise results with only a few packets [3][15]. ISP can monitor their backbone network with the network management system, and will upgrade the network link if the bandwidth of the links is insufficient to carry their subscribers traffic to the Internet.

There are many menus for the broadband service that is provided by ISP. For example, Open Computer Network (OCN) by NTT Communications provides ADSL with 1/1.5/8/12/24/40/47/50 Mbps and FTTH with 10/100/shared 100Mps as access link for their subscribers [5]. A typical ISP is using a regional network operator to connect their subscribers [6]. ISP subscribers are logically connected using PPPoE (PPP over Ethernet) to get the connection service from ISP [7]. On the other hand, regional networks are interconnected to the ISP backbone network at several locations. Different kinds of broadband speeds for their service and the logical connection from subscribers are hard to support ISP subscribers for the end-to-network performance even if they have a performance issue on the network.

2. Background (II)

- Related studies Distributed resource management by Grid Computing
  - Grid has been defined as to use a heterogeneous computer resource efficiently on distributed sites.
  - Agent-based middleware by PC Grid has been studied. The mobile agent runs on the PC Grid platform to process the CPU intensive application.

- Summary of the background
  - There are a lot of related monitoring activities for this paper. But there is few study that has been focused on the distributed network measurement with dynamic program loading and the job scheduling.
  - Therefore, this paper is focused on studying the system to :
    - Have the capability to execute any application program for the network performance monitoring distributed end nodes
    - Adopt the dynamic job management more than SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs.
    - Utilize the jobs management and schedule applications for Grid Computing as an infrastructure base for the distributed measurement and the submission of the analysis data.

Grid Computing has been defined as to use a heterogeneous computer resource efficiently on distributed sites. The research works for the job scheduling and the CPU resource management for the Grid Computing have been studied [17][18][19]. In addition, the agent-based middleware for PC Grid has been studied [20]. In this system, the mobile agent runs on the platform of PC Grid to process an CPU intensive applications. They have focused on building the distributed environment to run CPU intensive applications on distributed PCs. In stead of focusing on leveraging the computational power, our study in the paper will run an application program for the network performance monitoring on distributed end nodes like PCs that are currently using ISP subscribers on their networks. Job management and scheduling technology of Grid Computing could be related to our study. For an application and service on distributed computers, there is a good example to collect the computational power of clients all over the world. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.

In summary of the section, there are a lot of related activities for this paper, but none of them is focused on the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP. The project name is SETI@HOME [11][12]. In SETI@HOME, the captured radio signal from outside of the earth can be analyzed at approximately 2.4 million PCs. They can break up data from an observation into frequency bands that are essentially independent of one another. In this work, the distributed computer resource can be used to empower the computational power for the science application. SETI@HOME divides the data into many pieces for PCs, and the study of the network monitoring needs to gather the monitoring data from many distributed PCs on ISP.
3. Distributed Measurement (System overview)

Figure shows the overview of the distributed measurement system for the end-to-end network performance using a dynamic job management. The figure shows an architecture for the end-to-end measurement system on the ISP network. The monitoring for the network performance is done for the target host in figure. All monitoring results from clients are stored into the database server through the management server. Those monitoring results will be also restored from the database for the analysis task of the performance.

According to the above figure, the design points of the distributed measurement system for the end-to-end network performance are as follows:

1. Since the captured data should be stored into a single data repository and the central management server for all clients should be used for the system.
2. There are two components such client middleware that has been installed in the client and monitoring applications that are automatically stored in the client.
3. Three components such as the management, the database and the portal server are provided at the central site.

The management server is the management function for all clients and operator can control client middleware and monitoring applications here. All monitoring applications that will be stored in clients. The portal server has the function to analyze monitoring data and to show it to the operator.
3. Distributed Measurement (Architectural points)

These items are design points of the distributed measurement.

1. One central server and many distributed clients for the measurement
2. Geographic analysis and service menu for subscribers
3. Dynamic scheduling with congestion avoidance for measurement and results submission
4. Time difference capturing among all clients and server
5. Software and measurement parameters updates for clients from server side
6. Standard protocol (http) among clients and server
7. Central database to keep all measurement data

1. Distributed clients and one server: The study focuses the end-to-end network performance monitoring for the server from many clients. Since the network performance studies for inter-servers, the backbone networks (intra-AS) and inter-AS have been reported, there has been few study to investigate the end-to-end network performance from many clients with a limited bandwidth to the server on ISP. Since the major traffic on the Internet is WWW between a client-server interaction, the network performance for the server from many clients should be investigated to understand the network performance on the network [21].

2. Geographical analysis and service menus: Since ISPs know the geographical location (prefecture, area or state) of their subscribers with assigned IP addresses and also service menus of the broadband connection that they are providing, the analysis with the geographical location for the network performance and the menu of broadband services like FTTH, fiber or mobile could be done to understand the current network topology and its broadband link service as a snapshot. Because ISP subscribers dynamically connect to the ISP network and the network topology including ISP subscribers are changing, the analysis of the end-to-end network performance could be exhibited for the service quality of ISP at the performance perspective. Therefore, it can show an different aspect of the end-to-end network performance from previous studies of the network performance.

3. Scheduling with congestion avoidance: A central management server is a key to control many monitoring clients at the distributed sites. Because those monitoring programs should be synchronized to have a different timing for the measurement among clients to avoid the traffic congestion with probing packets into the network paths. For example, if many clients send monitoring probing packets or get them at the same time, the link paths are suffered from them and the performance degradation will be happened at the path on the network. The server workload should be also controlled by the number of client accesses at the same time. The time to send and receive data for the server should be slightly split at each client to avoid the congestion for the server access. The client control should be done by the central server remotely.

4. Time synchronization: The time synchronization among clients are needed to manage the scheduling for the measurement and the congestion avoidance to generate the measurement data for the target host. The simple method to synchronize the time among all clients can be used to get the time information from the server using SNTP (Simple Network Time Protocol). The synchronized time is only used for client middleware and its monitoring applications to avoid the unexpected behaviors for other applications in clients.

5. Software and parameters updates: The dynamic program loading feature for the client application and their parameters from the server to clients can be used to update the parameters in the monitoring program when they are needed to change the monitoring parameters. There is also a set of initial parameters that monitoring programs need before they are started. For example, those parameters for the monitoring program are target host, starting time, ending time, packet length, periodical time of monitoring and other options that the monitoring program is needed to set. They are decided for the network performance that depends on the networking environment that the ISPs would like to estimate. The operator can put those parameters appropriately before starting the monitoring. The functions of the dynamic program loading is also used to upgrade or replace the monitoring program that should be improved in future. Since a new better monitoring program for the end-to-end network performance will be invented and released to the public, the monitoring infrastructure in this paper should support the dynamic program loading for the end-to-end network monitoring.

6. Standard protocol: The monitoring system in this paper should employ the standard protocol for the Internet. In this system, the HTTP is used for a interaction between client and server. Because the TCP port 80 for the HTTP from the client to the server is usually opened for the Internet without blocking the firewall. The most transparent packets on the Internet are HTTP transactions. Therefore, all activities between client and server should be done by a normal HTTP transaction. The access direction for those HTTP transactions should be initiated by the client to the server because there is usually broadband routers that also acts as the firewall to prevent an illegal access to the client from the other Internet hosts. The unidirectional access for the HTTP port to the server from the client needs to pass the firewall on the client side. This type of access is a pull model from the client for the server. All HTTP transactions in this system are initiated by the client.

7. Database: The monitoring data at the monitoring application in the client is collected at the central server, and it is stored into the central database. It will be analyzed with parameters that the operator sets up. The advantage of the database for the network monitoring and the analysis of the data is that the monitoring data can be retrieved later and displayed at different aspects with given parameters.
4. System Configuration

Figure shows the system configuration of the distributed network measurement system. In figure, agent programs are the network monitoring programs that can investigate the end-to-end network performance for the target host after they are downloaded to User PCs from the server. There are three kinds of servers such as the management server, the database server and the portal server in figure. The management server manages the access control from User PC, authenticate User PCs with the unique identification for client middle, retrieve the agent programs to User PCs from database server, to support to upload the measurement results from clients to store the DB server. The DB server is the central repository for all data to store agent programs, their parameters and the measurement results that have been sent by User PCs. The portal server provides a user interface for the network performance administrator who would like to manage agent programs and their parameters to investigate the network performance with this distributed network performance system. The administrator can also analyze the measurement results to investigate issues for the end-to-end network performance on the ISP network. Client middleware in figure is installed on User PC, downloads the agent programs from the management server if it's available to execute on the User PC. It also runs the agent program that has the function of the active monitoring program that generates many probing packets for the target host. The measurement results of the agent program can be sent to the database server through the management server using the same link between server and user PC.

The target host in this figure is the Web server. In addition, any kinds of servers than the Web server could be used for the target host if they have an enough capability to upload and download the measurement data for the agent program of the network performance monitoring and generate the ICMP echo reply for the ICMP echo request for the agent program for the path discovery on the network.

The prototype system on those servers in this figure has been developed on the Java servlet (WebSphere Application Server) for the management and the portal server, the database server (DB2) is the database that is enable to retrieve the monitoring data with SQL commands. On the other hand, the system on the client side in the figure has been developed as a native program on Windows 2000/XP operating system because the majority of ISP subscribers is currently using those Windows operating systems. Therefore, the agent programs from a dynamically loaded on the client middleware are based on the Windows program. In addition, the agent programs have a capability to run any kinds of Windows application program such as a CPU intensive program for high performance computing applications. Therefore, this prototype system can be used for an application platform for the distributed computing environment than only the end-to-end network performance.
5. Measurement method and network topology

- Measurement vs simulation
  - Performance model and simulation results are a powerful tool to analyze the actual operating network. It needs to validate the simulation results and the actual network behavior.
  - The study has been focused on the analysis of the measurement results on the operational network.

- Advantage of the distributed measurement
  - ISP subscribers in the same region or area can be used for the aggregated result of the analysis data. It will be precise data than each different subscribers data.
    - E.g. Three clients are connected to the same POI. They will exhibit the same characteristics of the path and can be analyzed at the bottleneck of the same link.
  - Measurement in this paper is roundtrip time, packet losses, path discovery to the target host.
    - Those methods are very simple, but it’s powerful to investigate the end-to-end performance on the Internet.
    - Ping, Traceroute and http throughput have been used for getting the measurement results.

Figure shows the graph of the network topology that could be constructed by the path discovery of the traceroute agent on the end-to-end network measurement. In this figure, the root host would be the target node, those intermediate nodes are routers and leaf nodes are Client PCs. The intermediate nodes and links are categorized into the ISP network, regional networks and access routers that ISP subscribers have installed. The tree structure from the root node can be constructed in the figure. Those links of the graph are connected among routers and have the delay and the packet loss as a property of a link between routers.
6. Topology of ISP, regional networks and clients

Measurement results could be shown with the topology.

Topology of ISP, regional networks and clients

The prototype system for the distributed measurement with a dynamic job management runs for a preliminary test on the ISP networking environment. This evaluation for the prototype system has been experimentally conducted at the ISP and its subscribers from Nov. 14, 2004 to Jan. 14, 2005. The total participants of this experiment are sixteen clients which are located at different prefectures in Japan, and the number of clients is increased. The major six clients are shown in this section.

Figure shows the actual ISP’s topology that has been analyzed with the traceroute agent of the prototype system. In figure, the root node is the target host that the monitoring programs on all six clients have estimated the network performance on the path. ISP provides the target host for the experiment. In figure, the IP addresses for those routers are shown as a IP+digits with hiding a real IP address although the traceroute agent of the path discovery could capture them and all results can be stored at the database.

The topology construction of the experiment can be enhanced with those clients, different target hosts on the Internet and the hosts on the regional network. If some jobs with different agent parameter for the traceroute agent will be executed at clients, a big scale of the topology from ISP subscribers to the Internet servers can be constructed.

The system finds subgraphs with the path discovery method, and the vertices in subgraphs are examined whether they have the shared vertices. Those subgraphs are overlapped with shared vertices. Therefore, the constructed graph will be the part of the whole graph that is the network including the ISP network, the regional networks and access links.

The ISP’s network topology in this experiment is that the regional network assigns their own IP addresses to each subscriber and the subscribe connects ISP using PPP over Ethernet. Therefore, the global IP address is assigned by PPP. The protocol construction for transferred packets is IP (regional network)/PPP/IP (global network). The topology construction like figure is a critical task for the end-to-end performance measurement.

7. Results (I)

Figure shows the download bandwidth with CDF for clients. There are six measurement data sets for clients at different locations. Two higher links are FTTH 100Mbps at access lines, and other four lines are DSL 12Mbps at access lines. Two higher links have gentle slopes and exhibit that there is a cumulative distribution for the downloaded bandwidth. The gentle slope in figure shows the bottleneck with the traffic congestion on path. In this figure, there is no common trend for the bandwidth measurement. Therefore, three shared links from the target host are ISP routers which are a normal state without any congestion at all.
7. Results (II)

The relative roundtrip time drastically increases the first hop and the link characteristics shows the low bandwidth for the capacity. Those first links for clients are access lines of FTTH (100Mbps) and DSL (up to 50Mbps) that are relatively slower than other backbone links on the regional networks and the ISP network. Therefore, the relative roundtrip time for the first hop is increased. In addition, most of the relative roundtrip time reaches the relative value of 1 until the fifth hops from client. It means that the routers of regional networks has the slower link or the longer distance link than the ISP backbone network.

For example, the ACCA 12Mbps (Fukuoka) link has lower relative roundtrip until the fifth hop count, and it shows that those links between the second hop count and the fifth hop count have a constrained bandwidth or a longer distance links. On the other hand, in figure the ISP network beyond the fifth hop count doesn’t have any constrains for their link bandwidths because the ISP backbone network has usually an enough network resources with a good condition.

Figure shows the relative roundtrip time for the hop count from those clients to the target host. Since the roundtrip time to the target host is 1 at Y-axis in the figure, those values at Y-axis are the relative roundtrip times at the intermediate routers. Since the roundtrip times in the figure are the minimum value among measured values for a short captured period of the time, this figure shows the link characteristics until intermediate routers that have generated the reply packets to clients.

Figure shows the relation between the loss rate and the hop count from clients to the target host. The loss rate is increased at the first or the second hop count, and other hops are almost flat. This is a different characteristic from figure. Packet losses are accumulated at connected routers because those probing packets are sent from the client. Both those routers of both the first and second hop are located at the regional network. Therefore, those losses are based on the traffic congestion in the router at all times. Routers at other hop counts are relatively stable for the packet loss, and those routers beyond the third hop are not directly related to the packet loss on the path. The irregular peak at the second hop in ACCA 12Mbps (Chiba) has been shown in the figure, and it isn’t directly related to traffic congestion on the path because other hop counts on the same path have few packet loss. This peak needs to be investigated with the further experiment.
8. Discussions and conclusion

Discussions:
Here is the discussions of the issues and the future direction for the proposed distributed measurement system. The number of clients in this experiment can be increased. This experiment could be only used for the evaluation of the proposed system. If it will be deployed in a real operational ISP network and the number of clients for the network performance monitoring will be increased, some paths are shared by some clients that the monitoring agent is running on. This situation takes advantage of the distributed network measurement. When the specific link is congested by the overflowing traffic, all monitoring agents that the same link is shared can find the congested link along with the path precisely. Therefore, if the number of clients is increased, the congested link can be identified by the proposed system in the paper. This work will be done by a large number of client installations.

The deployment for the client middleware installations is critical for the proposed system. If a lot of clients is installed at the ISP network, the current network status including regional networks and access link is also clearly understood at the central ISP operation site. Increasing the number of clients also needs the business aspect for the deployment with the proposed mechanism. The question is that "are ISP's subscribers willing to install the measurement client for their PCs or is there any incentive for the client installation to deploy those monitoring clients?". This is also the challenge and an interesting work, but it's beyond this technical scope of the study.

Our direction of the further study for the distributed network monitoring could be thought as two items such as automated measurement for the end-to-end network monitoring to detect a performance anomaly and building a large scale common platform for the distributed network measurement.

An automated distributed measurement should be required to reduce an operational workload for the network management of ISP. The operation for the network needs to update their parameters based on analyzing the results data. For example, the monitoring period for the network performance is ten minutes at the normal network performance. If the network performance degradation is occurred at the link, the monitoring period for the paths that are related to the link will be shorter than one at the normal network performance.

Conclusion:
The paper proposes the distributed network measurement system with a dynamic job management on distributed clients, and the design and development for the prototype system has been done. We had the experiment to evaluate the proposed system on the operational ISP network. The preliminary experiment with the prototype system shows the effectiveness of the distributed measurement for the management of the ISP network operation and a customer management for ISPs service. Therefore, the proposed system can be useful for the ISP network to manage their subscribers for the perspective of the network performance.

* Acknowledgement
The authors would like to thank R. Shimomura, H. Hayashi, T. Yuguchi and M. Nakamura of NTT Communications and M. Haley, D. Saha and P. Pradhan of IBM for valuable discussions and kindly suggestions to study this distributed measurement system for the end-to-end network performance.