Applying the Policy scheme to the IntServ

Wang-Cheol Song, *Member, KIMICS*

Abstract**—**There is an emergence of Internet applications that have real-time requirements. These applications require IP to support guaranteed capacity, higher priority and lower packet loss rate. To address this, the Internet Engineering Task Force (IETF) is developing a set of protocols and standards for Integrated Services on the Internet. Using RSVP and policies to manage the allocation of network resources in order to provide different levels of service is a topic of great interest to service providers. Currently, it is not possible to dynamically reallocate resources during an application's session. This paper discusses how policies in conjunction with new service class can provide a more enhanced network resource allocation by allowing for this dynamic reallocation.

Index Terms**—**policy based networking, RSVP, Internet QoS, IntServ

I. INTRODUCTION

There has been an increase in distributed applications that involve the exchange of data that is time-sensitive. Applications include video-on-demand, distance education, tele-medicine, tele-conferencing, electronic commerce. Users of these applications expect them to perform at acceptable levels, that is, they expect a high level of quality of service (QoS). Different applications have different QoS requirements. For example, e-mail has a high tolerance for delay and thus has no constraints on delivery. ftp has a high tolerance for delay but is synchronous and thus is time-sensitive but there is flexibility on delivery. Video-on-demand applications have low tolerance for delay.

Currently, IP networks provide best-effort delivery for all applications. At each router, data packets are queued and then forwarded with no distinction made between packets of different applications. This works well for e-mail, ftp and telnet. However, best-effort delivery does not work as well for delay-sensitive applications, such as Internet telephony, video-conferencing and interactive games, since these applications have relatively strict requirements on throughput, loss and delay that must be satisfied. This suggests that the Internet needs to be able to provide different levels of service in order to support the increasing diversity in the types of applications running on the Internet.

The Internet Engineering Task Force (IETF) is developing a set of protocols and standards for Integrated Services on the Internet [1,2]. In the IETF's vision, applications request and reserve resources in the network and at the hosts using the end-to-end Resource ReSerVation Protocol (RSVP) [3,4]. The IETF Integrated Service working group standardized two service classes: the Controlled Load (CL) service [5] and the Guaranteed Service (GS) [6]. The purpose of the CL service is to approximate the service a user would experience from a network that does not have a heavy load or congestion. However, if the network routers become overloaded no scheduling algorithm can be used that emulates a lightly loaded network for the CL flow. The purpose of the GS service class is to provide applications with mathematically provable end-to-end delay (deterministic) at all network elements. It guarantees conformant packets, lossless transmission and an upper bound on end-to-end delay. GS provides a deterministic service that tends to wastes resources, since more bandwidth is usually reserved then absolutely necessary. The work in [14] shows that an average load of only 40% is on a link carrying GS flows. The customer who uses the GS service is paying for the unused resource and another flow's resource reservation may be blocked even though there are unused resources. This is costly to both users and vendors and thus only applications with hard real-time requirements should use GS [13]. CL has the potential to underallocate resources while GL has the potential to over-allocate resources. In either case, it may be desirable to dynamically reallocate resources during an application's session which currently cannot be done. This paper addresses this problem with the introduction of a new service class: Modified Guaranteed Service with Pool (MGSP).

The IETF Resource Allocation Protocol (rap) working group has also standardized various architectural elements for allocating resources based on policies (i.e., rules). The use of policies allows the network elements to determine which flows are entitled to which service [7].

The use of RSVP and policies as an approach to manage the allocation of network resources in order to provide different levels of services is a topic of great interest to Service Providers. However, most of the recent research has focused on defining the framework and functionality of each element for provisioning network resources based on policies [8,9,10]. The primary application of policies has been for determining pricing based on usage [11]. This paper discusses how policies in conjunction with a new service class, Modified Guaranteed Service with Pool (MSGP) can provide a more dynamic network resource allocation.

Section 2 introduces RSVP and an architecture for policy-based RSVP. Section 3 proposes a new service

Manuscript received June 17, 2003. This work was supported by Cheju National University.

Wangcheol Song, Faculty of Commun. & Computer Eng., Cheju National University, Jeju, Jeju, Korea (Email: philo@cheju.ac.kr)

class - MGSP and suggests a mechanism for policy-based QoS service. Section 4 describes briefly related work and section 5 describes conclusions.

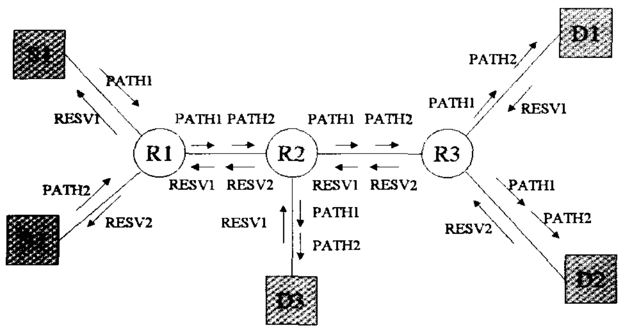

Fig. 1 An architecture for RSVP operation

II. ARCHITECTRE FOR POLICY-BASED RSVP

This section presents an overview of the RSVP protocol and the architecture for policy-based RSVP is described.

A. RSVP operation

RSVP is a signaling protocol designed to enable applications to reserve network resources to satisfy a requested level of QoS requirements. Using RSVP senders, receivers and routers communicate with each other to setup the necessary router state needed to obtain the network resources that satisfy the QoS requirements for their data flows [3].

There are two important RSVP messages: PATH and RESV message. The PATH message is initiated by a sender and is used to set routing state (this is called soft state) on the routers on the path between a sender and a receiver. It is also used to provide receivers with information about the characteristics of the traffic flow (using the Tspec specification) so that a receiver can make the appropriate reservation request using the RESV message. When a PATH message is received by a RSVP enabled router (i.e., a router with a RSVP daemon on it), the router records the source address from where it received the PATH message. If the router detects an error it sends a PATH ERR message to the sender. When the receiver receives a PATH message, it first examines the Tspec specification. This specification has information including the following: (1) The bandwidth needed to maintain that data is being transferred at a specific rate. (2) The extent to which the data rate can exceed the sustainable average for short periods of time. (3) The maximum rate to send data and (4) The size of the smallest packet that should be generated by sending application. Based on this information and the service class (GS or CL) specified by the sender, the receiver makes an appropriate RESV message, and then sends it to the sender to reserve network resources. The RESV message includes a request specification (Rspec) that indicates the type of service required (GS or CL) and may include information such as maximum delay and packet loss probabilities. Each router that receives the RESV message checks to see if sufficient resources are available and tries to reserve the required resources in the router. If there are not enough resources, a RESV ERROR message is sent to the receiver. Routers that receive a RESV ERROR message cancel the reservation for the flow.

The PATH and RESV messages are independent and a PATH message of a session might induce RESV messages to request different amount of resources. As the RESV messages from the receivers traverse upstream to the sender, they are merged by the routers at the merging points. After the reservation, the sender continues to send PATH REFRESH messages to maintain the soft state of the flow. The soft state information in the routers is periodically refreshed by using REFRESH messages. If there are no REFRESH messages for a particular flow then the entry in the soft state associated with that flow is deleted which implies that the resources reserved for that flow are released.

Fig. 1 graphically depicts a network used for illustrating RSVP operation. There are three routers and five host machines, in which S1 and S2 are senders and D1, D2 and D3 are receivers in the same multicast group. PATH messages from S1 and S2 are sent to all of the receivers. D1 and D3 accept the PATH message from S1 and D2 accepts the PATH message from S2. Router R2 merges the RESV messages from D1 and D3 to make one reservation of resources on the flow from S1 on the link between R1 and R2. Router R1 also merges the RESV messages from D1 and D3 to make one reservation of resources on the flow from S1. After the reservation, S1 and S2 send the PATH REFRESH messages periodically in each refresh interval, D1, D2 and D3 send the RESV REFRESH messages in response to the PATH REFRESH messages.

B. Policy-Based RSVP QoS Control

The use of policies allows a sophisticated regulation of access to network resources. A policy is a set of rules. A rule is of the form if condition then action that is used to control which users, applications or hosts should have access to which resources and under what conditions. Thus, policy-based admission control is not only based on the amount of available resources, but also on the application and user. For example, telephony applications have higher priority than games. The policies could specify that the amount of bandwidth available to all games is x. Assume that a flow is requesting x', but that flow is associated with a game and the amount of bandwidth available to games had already been allocated. The flow is not given the requested resources. Thus, even though there may be enough resources the requesting flow does not get its resource reservation request satisfied. This is controlled by the policies.

1) The IETF Architecture

Fig. 2 shows placement of the IETF policy architectural elements with a monitoring agent. When a router receives a PATH message, the Policy Enforcement Point (PEP) on the router contacts the Policy Decision Point (PDP). This communication is defined by the Common Open Policy Service (COPS).

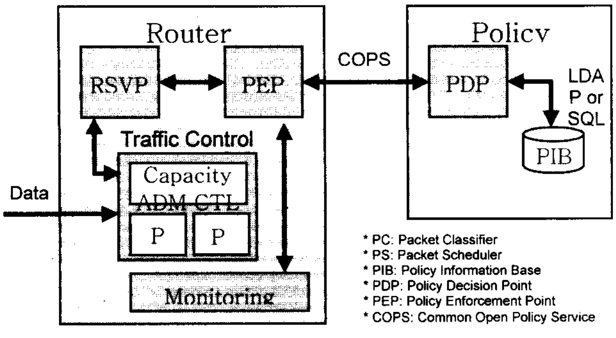

Fig. 2 Placement of IETF policy Elements with the monitoring agent

The PDP may determine the service class(es) that can be associated with the requesting flow. This is based on the policies and the current state of the network. An admission control decision is returned to the requesting PEP through COPS. The PDP has access to a policy database (through SQL or LDAP) containing the policies for the administrative domain. This combination of PDP and policy database is referred to as a policy server. The logical combination of a PEP and PDP is referred to as a policy module.

2) Relationship Between Policy and Service Classes

The use of policies proposed for RSVP is to control the allocation of network resources. A service level agreement (SLA) is defined as a contract specifying expected QoS requirements between a customer and a service provider. Customers and service providers are administrative domains. An end-user that initiates a communication is in an administrative domain. This administrative domain is a customer of an adjacent administrative domain that is on the path from the end-user to the target. Basically, the adjacent upstream administrative domain is a service provider for the adjacent downstream administrative domain. The adjacent administrative domains agree on a SLA. A request for resources is examined by the policy module and the admission control module of an administration domain. The policy module is invoked only at an egress router. The policy module check the SLA and determin the appropriate service class, and the admission control module regulates resource allocation by result from the policy module.

As stated earlier, the IETF Integrated Service working group standardized two service classes: the Controlled Load (CL) service [5] and the Guaranteed Service (GS) [6]. The GS service class often results in a wastage of resources since more bandwidth is usually reserved than needed. The CL service class provides a service level based on an approximation of the service a user would experience from a network that does not have a heavy load or congestion. However, if the network routers become overloaded no scheduling algorithm can be used that can emulate a lightly loaded network for a CL flow. If the flow needs more resources, the QoS of the flow may not be guaranteed. This suggests the need to be able to dynamically reallocate resources during an application's session. This can be supported through a new service class. This service class is defined and a description of how policies can be applied is described in the next section.

III. RSVP BASED QoS SERVICE WITH NEW SERVICE CLASSES

This section describes a new service class, Modified Guaranteed Service with Pool (MGSP) and how it can be incorporated into the RSVP operation. MGSP is similar to GS with the following difference: some portion of the resources reserved but not used by a flow will be returned to a pool of available resources that can be used by other flows. If extra resources are required, then it can use available resources from the pool. This allows for a dynamic allocation of resources i.e. during a session, it is possible to request additional resources from the pool without starting a new session (as must happen for a flow of the GS or CL service class).

A. Modified Guaranteed Service Class with Pool (MGSP)

Flows that request MGSP service are setup in a similar fashion as flows requesting GS service. It is assumed that each RSVP-enabled router has a monitoring agent that monitors the actual resource usage of the flow after completion of the reservation. The monitored value is an attribute of the soft state. Figure 3 describes the procedure for the MGSP service class. Whenever a RESV REFRESH signal is received, the most recently monitored value and the reserved value are compared. If the monitored value is greater than or equal to the reserved value, the reserved value is maintained and the need for resources that exceeds the reserved value is satisfied by taking from the pool (if possible). Otherwise, the reserved value for the flow is set to the monitored value. The amount of the pool resource can be updated on the basis of the monitored and estimated value of pool. One example equation that can be used is the following:

$$(RB_{GS\_i} - MB_i)_{max} + MB_{pool,\,avg}. \qquad (1)$$

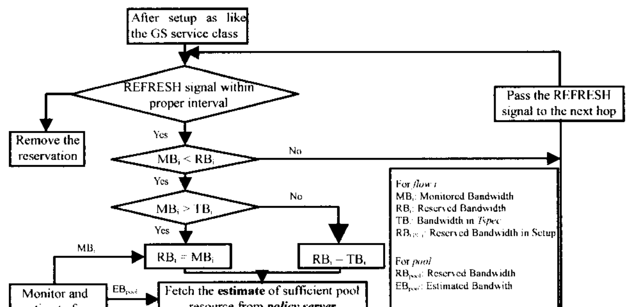

Fig. 3 the procedure for the MGSP

For a MGSP flow *i*, RB_{GS_i} represents the initial resources allocated to flow *i*. MB_i represents the current monitored value for flow *i*. $MB_{pool\ ,avg}$ represents the monitored current average value of the pool. $(RB_{GS\_i} - MB_i)_{max}$ represents the maximum difference among all flows between monitored amount and the initially reserved amount. Since a flow of type MGSP is setup as a flow of type GS, then the initial allocation of resources is based on worst case needs. $(RB_{GS\_i} - MB_i)_{max}$ represents the maximum difference for all flows between the reserved resource amount and the monitored resource usage. This difference is added to the pool. When the worst-case delay occurs and a flow needs resources, the flow can use some of the resources from the pool of available resources.

B. Policy-Based RSVP QoS Service

This section describes possible applications of policies in the RSVP operation with the new service class. There are two types of policies: static and operational. Static policies are applied when a device or process is configured. This often initializes variables to be used during the lifetime of the device or a process or at least is not changed very often. Basically, in this case the condition is always true and the action is the assignment. Operational policies describe the action to take place when some condition arises during the operation of the system.

1) Static Policies and Service classes

Using the MGSP service class and existing IETF service classes, we can clearly define a set of service classes. Policies can be used to set the priorities that the different classes have in being able to reserve resources. Assigning priorities to service classes is an example of static policies. They are static in the sense that the priority assignment is considered part of the router setup. The classes with possible priorities are defined below.

* The Guaranteed Service (GS) : This is the service class of RFC 2212. As it uses more resources than what is actually used, its cost is very high. The level of service is the highest (priority = 1).
* The Modified Guaranteed Service with Pool (MGSP) : This is the service class proposed in this paper. It is supposed to provide the same quality as GS, but efficient resource usage is considered (priority = 1).
* Prioritized Controlled Load Service (PCL) : This is the service class to prioritize CL. It has precedence to use the available resources over CL only when the PCL flow needs more resources than reserved and there are unused resources available to PCL and CL flows (priority = 1).
* The Controlled Load Service(CL) : This is the service class of RFC 2211 (priority = 2).
* Best-Effort : This represents current Internet traffic flow (priority=3).

This prioritization for the above classes gives GS, MGSP and PCL the same level of priority and make all of them have equal privilege over unused resources. In case of PCL and CL, by assigning higher priority to PCL, implies that the flow of type PCL gets a precedence of unused resources over flow of type CL when both PCL and CL flows need more resources than what has been reserved.

2) Static Policies and the Pool

Policies can also be used to manage the buffer pool. Examples of possible policies include the following: (1) Policies can be used to state the percentage of the pool made available to each of the service classes. Alternatively, this could be stated in absolute numbers. (2) Policies can be used to determine how much of the excess resources assigned to a flow is returned to the available pool (Equation 1 represents only one way to do this).

It is expected that these policies may change more often then the policies in Section 3.2.1. For example, an analysis of past history of usage of the resources by flows of each type of service class may suggest that the percentages should change.

3) Operations of Policy Module

The arrival of a RESV message at an egress router results in the PEP contacting the PDP. The PDP determines the service class(es) that the requesting flow can be assigned to. If there is more than one service class then an ordered list of service classes is returned where the order represents the preferences. The admission control module is then contacted. The admission control module assigns the flow available resources on the basis of the results from the policy module and the current availability of resources. Policies can be used to regulate this setup in a number of ways. For example, there may be a policy that states that if the requesting flow has higher priority then other existing flows and there are not currently enough resources for the requesting flow then pre-empt lower priority existing flows until there are enough resources. Another policy could state that if there are not enough resources then the flow is blocked until there are resources. The behavior of the admission control module differs depending on the actual policies used. The first policy described above allows the policy module to guarantee a specific entity/user some level of QoS (as defined by the SLA) by allowing the pre-emption of resources assigned to other flows. The second policy described above is simpler and basically only ensures that flows are not admitted until there are sufficient resources. This policy does not interfere with existing flows that have resources assigned to them. If more than one flow is blocked waiting for resources then policies can be used to determine which flow next gets resources that are freed up. In this paper, only the second policy is considered.

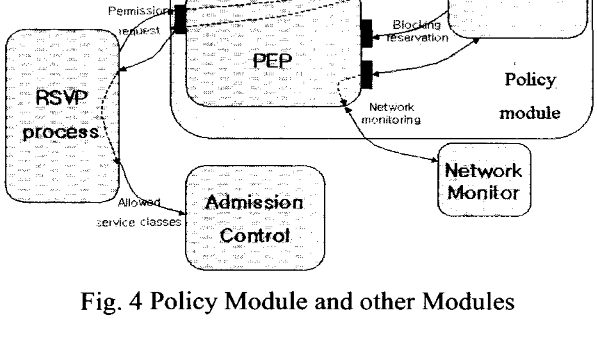

Fig. 4 Policy Module and other Modules

If a flow requests more resources then allowed by the SLA, policies can be used to determine the action to be applied which could be that the flow is treated as best-effort or that the flow is assigned a lower level service class or is not to be allowed admission. Again, policies have an impact on the behavior of applications.

For the PDP to be able to make decisions it needs current information about resource usage. Logically, it is assumed that it receives this information from the network monitor through the PEP. Physically, the network monitor may consist of a number of monitoring processes including agent processes that monitor resource usage at each router.

In order to provide the level of QoS control described in this paper, the RESV messages must certain policy information. This is represented in the following policy data structure (this represents only a subset of the possible fields).

```
Policy data structure {
   token or session id : identification of session, sender
                          and receiver
   flow id
   sender's list of addresses & ports
   receiver's   list of addresses & ports
   requested service class
   service class list which can be reserved
   reserved service class list
   service style {WF, SE, FF}
                          // : Part of the RSVP standard
   receiver's current resource reservation status for each
        class
}
```

The above description of policy usage primarily applies to the egress router (ER). The following are some of the interface methods assumed for the PEP at the egress router.

Service_class_list request_permission *(token,flow id, sender_addr, receiver_addr, requested service class)*
Boolean blocking_reservation (*required_resource*)
resource_monitor(*service_class_list, measured_resource, available_resource*)

The request_permission interface method is used by the RSVP process to get permission for the reservation from the PEP. The PEP sends a service request to the PDP. The PDP identifies the requesting flow (through flow id) and the sender and receiver information. The PDP determines if the requested service class (if this is not nil) is allowed for the flow. If the flow is allowed to use more resources than indicated in the requested service class, then a list of available service classes (service_class_list) is returned for the flow. This assumes that the resources are available. If the requested resources (as indicated by the requested service class) are not available then the service_class_list is returned with permissible service classes for the requesting flow. Otherwise, the blocking_reservation interface method is called. The blocking_reservation interface method allows the PDP to request the blocking of new reservations of flows that are requesting more resources than can be provided (this is indicated by the resource amount passed to this method). Using this method, the policy module can guarantee some level of QoS for flows already assigned resources. The resource_monitor interface method is used by the PDP to request resource usage information from the PEP which uses the network monitoring module to get this information. In the inner domain routers, each flow has already gotten permission for the proper service class from PEP in ER. Thus, the blocking_reservation and resource_monitor methods are needed in the inner domain routers.

IV. RELATED WORK

As discussed earlier, research on policy based RSVP for QoS is led by IETF's rap working group [7]. This working group discusses and organizes the basic concepts and framework. Based on this effort, Cisco, KT, Nortel and the other network-related companies have announced many policy-based QoS service and/or products with RSVP.

Research on improvement of IETF's service classes and development for a new service has been carried out with a view focused mainly on the data link layer and/or ip layer – this has led to work on RSVP scalability [15]. We can see some effort for adopting various applications to the guaranteed service [16]. An extensive search of the literature indicates no work where a new class of service is proposed, and that work related to policy in RSVP based QoS control have been focused on pricing and/or measuring usage[11,17].

V. CONCLUSIONS

In this paper, in order to provide policy based QoS service with RSVP appropriately, we have proposed a new service class, MGSP and have considered a mechanism for the QoS service with the differentiated services – GS, MGSP, PCL and CL.

Led by IETF's rap working group, there has been a good deal of research on policy-based QoS with RSVP. However, these research results have been mainly focused on concepts, protocols and frameworks. Litle work has been done on the applicability of policies to RSVP. Therefore, this paper has discussed how the service could be provided practically, proposed a new service class, MGSP, and described a mechanism in which policies lets users/applications actually control the QoS with differentiated services. This policy-based QoS control mechanism is based on the SLA, so it needs to discriminate communication entities/users. In this paper we simply use "token" as an field of the policy data structure, but further study is needed for token to be tied up with an authentication system. The proposed MGSP service class may be supposed not to provide theoretically the complete lossless transmission, but we predict it to practically service it.

REFERENCES

[1] R. Braden, D. Clark, and S. Shenker, Integrated services in the Internet architecture: An overview, *RFC 1633*, July 1994.

[2] Black, D., Blake, S., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", *RFC 2475*, December 1998.

[3] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification, *RFC 2205*, September 1997.

[4] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, RSVP: A new resource ReSerVation Protocol, *IEEE Network*, pages 8-18, September 1993.

[5] J. Wroclawski, Specification of the controlled-load network element service, *RFC 2211*, September 1997.

[6] S. Shenker, C. Partridge, R. Guerin, Specification of guaranteed quality of service, *RFC 2212*, September 1997.

[7] http://www.ietf.org/html.charters/rap-charter.html

[8] Yavatkar, R., Pendarakis, D. and R. Guerin, "A Framework for Policy Based Admission Control", *RFC 2753*, January 2000.

[9] Boyle, J., Cohen, R., Durham, D., Herzog, S., Raja, R. and A. Sastry, "COPS Usage for RSVP", *RFC 2749*, January 2000.

[10] Herzog, S., "RSVP Extensions for Policy Control", *RFC 2750*, January 2000.

[11] M. Karsten, J. Schmitt, and R. Steinmetz, "Generalizing RSVP's Traffic and Policy Control Interface", *Proceedings of the 7th International Conference on Parallel and Distributed Systems Workshops (ICPADS'00), Iwate, Japan*, pages 249-254. IEEE, Piscatay Way, NJ, USA, July 2000

[12] L-N. Hamer, B. Gage, "Framework for session set-up with media authorization", *Internet- Draft, draft-ietf-rap-session-auth-00.txt*, April 2001.

[13] R. Guerin, V. Peris, "Quality-of-service in packet networks: basic mechanisms and directions", *Computer Networks*, Vol. 31, pages 169-189, 1999.

[14] L. Georgiadis, R. Guerin, V. Peris and R. Rajan, "Efficient support of delay and rate guarantees in an Internet", in:*Proc. SIGCOMM, San Francisco, CA*, August 1996, pp. 106-116.

[15] Stoica and H. Zhang, "*Providing guaranteed service without per flow management*", ACM Computer Communication Review, vol. 29, pp. 81-94, Oct. 1999.

[16] F. Borgonovo, A. Capone, L. Fratta, M. Marchese, C. Petrioli, "*PCP: A bandwidth guaranteed service for IP networks*", to be presented at IEEE ICC'99, also available at http://cerbero.elet.polimi.it/ntw/ip/qos.html

[17] M. S. Borella, V. Upadhyay, I. Sidhu, "*Pricing Framework for a Differential Services Internet*", European Transactions on Telecommunications, Vol. 10(2), March/April 1999.

Wang-Cheol Song
He received his B.S., M.S and Ph.D. degrees in Electronic Engineering from Yonsei University in 1989, 1991 and 1995 respectively. Since 1996 he has been working for Cheju National University and during a year in 2001 he was a postdoctoral fellow in Department of Computer Science at University of Western Ontario , Canada. His research interests include Internet QoS, PBN(Policy based Networking) and network management.