Flow-based Reservation Marking in MPLS Networks

Nianbo Liu[a], Jiannong Cao[b], Ming Liu[a], Jiazhi Zeng[a]
[a] Department of Computer Science, University of Electronic Science and Technology, Chengdu, China
[b] Department of Computing, Hong Kong Polytechnic University, Hongkong, China
E-mail: liunb@uestc.edu.cn, csjcao@comp.polyu.edu.hk, liuming@uestc.edu.cn, jzzeng@uestc.edu.cn

Abstract— **Marking in DiffServ at the edge of the network often follows a demand side policy. It meters a traffic stream and marks its packets according to some predefined traffic parameters as Committed Information Rate, Committed Burst Size, Excess Burst Size and so on. Such marking based on traffic characteristics is irrespective to network dynamics, which causes collision and QoS degradation in DiffServ. This paper proposes flow-based Reservation Marking as a supply side marking at the network edge, which marks stream packets reserved or unreserved according to flow-specific reservation in a distributed resource reservation environment. When congestion occurs, anticipant per flow QoS is secured by protecting reserved packets on core routers without any per flow or per trunk management. It provides a simple, scalable and adaptive mechanism of implementing quantitative end-to-end QoS by mapping per flow in IntServ into per class in DiffServ. A "once reserve, no more manage" framework is constructed to eliminate flow state, avoiding unexpected collision and flow management simultaneously on core routers. Performance evaluation reveals that it affords controllable and quantitative QoS, keeps networks core-stateless and achieves high link utilization at the same time.**

I. Introduction

In next generation high-speed networks, it is very important to provide QoS guarantees to a wide range of applications in a scalable manner. An important requirement is to prevent congestion collapse, keep congestion levels low, and guarantee fairness. DiffServ develops a class-orientated framework, including differentiated traffic marking and Per Hop Behaviors, to remedy congestion problems when that occurred. IntServ, as a per flow orientation framework, develops distributed admission control and resource reservation to obviate the occurrence of congestion. However, many technical QoS solutions are not ideal, no matter DiffServ-like ones based on congestion control or IntServ-like ones based on congestion avoidance. Admitting congestion means less QoS guarantee while avoiding congestion entirely (or close entirely) can be rather expensive and unscalable.

What motivate us to develop our approach lies in an interesting question: how to combine flow-based resource reservation in IntServ with class-based packet forwarding in DiffServ. Marking in DiffServ at the edge of the network often follows a demand side policy. It meters a traffic stream and marks its packets according to some predefined traffic parameters as Committed Information Rate, Committed Burst Size, Excess Burst Size and so on. Such marking based on traffic characteristics is irrespective to network dynamics, which causes collision and QoS degradation in DiffServ. This paper proposes flow-based Reservation Marking as a supply side marking at the network edge, which marks stream packets reserved or unreserved according to flow-specific reservation in a distributed resource reservation environment. When congestion occurs, anticipant per flow QoS is secured by protecting reserved packets on core routers without any per flow or per trunk management. It provides a simple, scalable and adaptive mechanism of implementing quantitative end-to-end QoS by mapping per flow in IntServ into per class in DiffServ. By using Reservation Marking to coordinate actions of resource reservation and data forwarding, our approach achieves controllable and quantitative QoS of individual streams in a scalable manner.

II. Related Work

The stateless DiffServ possesses excellent scaling properties, but mismatch between traffic forecast and actual load including distribution among DiffServ classes is unavoidable in dynamic network environments. Many approaches focus on implementing adaptive mechanisms on DiffServ framework to improve QoS on different aspects, especially on end-to-end QoS assurance earlier [1]-[3] and bandwidth fairness or delay guarantee later [4]-[7]. But these adaptive mechanisms, often running in some control theory based feedback models, are far from overall solutions due to network dynamics. In particular, it seems all but impossible to construct a single model that, on the one hand, represents all aspects of congestion control accurately and, on the other hand, is simple enough to be useful.

Alternatively, approaches based on traffic trunk to bundle and implement QoS of a number of flows have also been considered, which are known as RSVP-TE and DS-TE [8]-[10]. Such form of aggregation simplifies the allocation of network resources and promotes the deployment of QoS frameworks notably, whereas the scalability problem still remains. Traffic trunks are not only used for data transmission but also for resource management, which generates the spending for maintaining "soft state" and obstructs the aggregation of trunks with different QoS constrains.

Researches about QoS mapping between different specifications, requirements, services and frameworks mainly focus on multi-domains on heterogeneous networks [11, 12], which treat these domains separately in geography and come down

978-1-4244-2075-9/08/$25.00 ©2008 IEEE

to IntServ over DiffServ [13] finally. Few of them aims at the mapping between multiple QoS frameworks of single stream on homogeneous networks. Developing RSVP-TE and DS-TE, we present our contribution for handling this mapping form per flow reservation into per class packet forwarding in a connection-based network. A "once reserve, no more manage" framework is constructed to eliminate flow state, avoiding unexpected collision and flow management simultaneously on core routers.

III. Flow-based Reservation Marking Framework

As RSVP-TE and DS-TE mainly proposed to deal with TCP flows and multimedia streams in MPLS networks, we also build Reservation Marking (RM) framework in such environment. Although RSVP-TE and MPLS given as examples, RM doesn't involve any details of message protocols and networks. The details of RSVP-TE, DiffServ and MPLS are discussed in [8, 14, 15].

A. Resource Reservation

RM is a core-stateless framework, in which per flow states of streams are only installed on edge routers. Except that, the Label Switch Path (LSP) establishment in RM is very similar to RSVP-TE. At first the ingress router transmits an PATH message downstream and the egress router responses an RESV message upstream in Fig. 1. Then the RESV message establishes LSP independently at every hop by distributing label bindings and requesting resource reservations via Admission Control and Police Control modules. At last the flow-specific reservation from the Resv message is stored in per flow state on the ingress node. After the whole procedure, routers have a LSP installed for transmission instead of traffic trunk in RSVP-TE.

Admission Control module on core routers supports measurement-based admission control (MBAC) schemes as Measured Sum, Hoeffding Bounds and so on, which rely on instantaneous traffic measurement and require no prior knowledge about the traffic. It doesn't support Simple Sum or other parameter based schemes require flow state info on core routers. In this study we use a MSPK algorithm that employs flow peak rather than token rate to calculate measured sum.

B. Marking and Forwarding

Traditional DiffServ traffic conditioner uses a meter to determine the compliance to traffic parameters, and then marks incoming packets appropriately. For instance a Time Sliding Window Two Color Marker (TSW2CM) [16] marks stream packets in-profile or out-of-profile according to its Committed Information Rate.

Instead we propose a new traffic conditioner performing Reservation Marking without traffic meter on edge router in Fig. 2. It introduces TSW2CM to mark stream packets reserved or unreserved respectively as the in/out profile in DiffServ. The only marking threshold is not Committed Information Rate but the flow reservation of the stream, which is obtained from per flow state previously installed on the ingress node. Therefore Reservation Marking becomes a supply side marking and represents the throughput rate negotiated by user and service provider on a per-flow level.

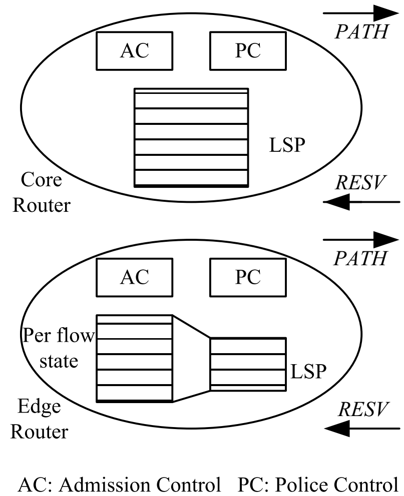

Fig. 1. Core router and edge router in RM

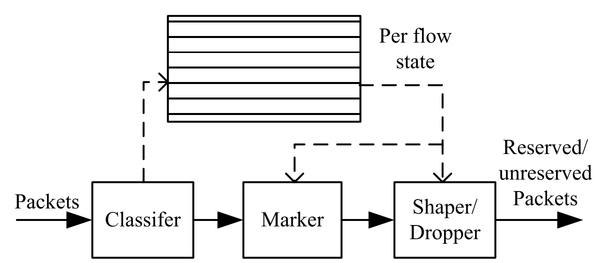

Fig. 2. Traffic conditioner on edge router

RM provides three service types including Guaranteed Service (GS), Best Effort Service (BE) and Controllable Service (CS). Different QoS service levels associated with different reservations of stream are shown in Fig. 3. We assume that the reservation value R specifies packet classification and indicates service differentiation. A R_1 marking with all packets reserved indicates GS; a R_0 marking with no packet reserved indicates BE service; a R_i marking with quantitative packets reserved indicates CS. The reserved/unreserved classification claims different Per Hop Behaviors without any flow management on core routers. Once the queue becomes congested, core routers drop the unreserved packets with much higher probability than reserved packets. Consequently reserved packets are doubly protected by early distributed resource reservation and Per Hop Behavior on core routers during packets forwarding period. In this way Reservation Marking successfully maps flow reservation into reserved/unreserved traffic and seamlessly connects LSP establishment based on RSVP-TE with data forwarding based on DiffServ.

From the view of congestion theory, RM combines congestion avoidance mechanism with congestion control one and brings forth powerful QoS exceeding IntServ and DiffServ. Controlled-load Service in IntServ, Premium Service

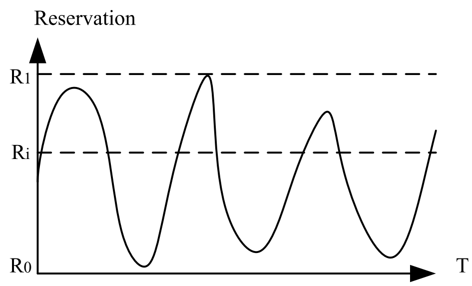

Fig. 3. Reservation Marking of different reservations

and Assured Service in DiffServ are some coarse QoS level descriptions, which are some approximating differentiations of delay from different loaded networks. The QoS details of specific stream in these services is not only transparent to end users, but also to service providers to some extent. This blur on QoS also causes difficulties in pricing. Services in RM, in the form of a mixture of reserved packets and unreserved ones on a user-requested ratio, enforce performance guarantees at a any-grained level. As a consequence, RM afford controllable QoS to both end users and service providers in a quantitative manner. QoS pricing also benefits from a greater level of accuracy and a finer level of granularity.

C. Implementation Discussion

A simple RM example in MPLS networks is given in the two previous subsections. It doesn't mean that the deployment of RM is limited in MPLS networks. Using the existent QoS techniques or modules in IntServ and DiffServ, RM fits any networks in theory, from IP networks to Service Unit networks [17], which had already developed necessary mechanisms to support IntServ and DiffServ. Similarly RM also has a favorable compatibility because these LSP establishments and classified packets can be directly reused in any IntServ or DiffServ domains on heterogeneous networks without extra interaction and mapping.

The scalability of RM is greatly enhanced in comparison with RSVP-TE or other trunk-based frameworks. Unlike TE tunnels, LSPs in RM framework is only used for transmission and dispenses with any resource management function. As a result, RM supports large scale aggregation like label merging in original MPLS networks. Label Merging is the capability of forwarding two different packets belonging to the same FEC, but arriving with different labels, with the same outgoing label. For example a merging of LSPs with common egress node is feasible in RM.

Another problem is the performance-complexity tradeoff of marking. Reservation Marking represents a simple supply side marking, which may easily cooperate with the traditional demand side DiffServ markings. Such multi-factor marking (or multiple makers) could offer more differentiated services and improve the fairness in bandwidth allocation of unreserved packets. However, we don't discuss such compound marking in this study and insist on a practical view: "We need QoS at the lowest layer, below IP, and it only needs to be simple - a two level (one bit) scheme. The lowest level QoS mechanism needs to be exceedingly cheap to implement and deploy, encouraging innovative use with minimal inconvenience." [18]. We prefer one-bit Reservation Marking because the powerful QoS is already achieved in a simple way.

As outlined in section II, many approaches appeared as promising IP QoS solutions. In order to evaluate RM, we investigate the resource allocation and forwarding management parts of these frameworks in the following tables. (The approaches implementing adaptive mechanisms on DiffServ are attributed to Enhanced DiffServ.)

TABLE I
RESOURCE ALLOCATION IN DIFFERENT QOS FRAMEWORKS

Framework	Type	Aggregation	Reserve	Effect
DiffServ	edge AC	flow	no	not accurate
Enhanced DiffServ	edge AC with probe feedback	flow	no	statistically accurate
IntServ	distributed AC	flow	yes	accurate
RSVP-TE	distributed AC	flow/trunk	yes	accurate
DS-TE	distributed AC	flow/trunk	yes	accurate
RM	distributed AC	flow/trunk	yes	accurate

TABLE II
FORWARDING MANAGEMENT IN DIFFERENT QOS FRAMEWORKS

Framework	Unit	Maintain	QoS
DiffServ	class		not guaranteed
Enhanced DiffServ	class		statistically guaranteed
IntServ	flow	flow	guaranteed
RSVP-TE	trunk	trunk	nearly guaranteed
DS-TE	class over trunk	trunk	nearly guaranteed (better than RSVP-TE)
RM	class		controllable quantitative guaranteed

In TABLE I, we list the type of admission control, aggregation scale, reservation, and the general effect of resource allocation. In TABLE II, we focus on management unit, maintaining objective, and the general QoS level. When we draw a comparison between RM and other frameworks, we find that RM framework develops a special "once reserve, no more manage" method. In resource allocation, it employs accurate distributed admission control instead of edge admission control or probe feedback mechanism. On the aspect of forwarding management, it adopts concise class-based packet forwarding to smooth the fussy flow or trunk states on core routers. By mapping the quantitative flow reservation into differentiated classes, RM provides quantitative, controllable, and guaranteed QoS in a scalable manner. Some early researches proposed the combination of IntServ and DiffServ and described the blueprint of providing good QoS while maintaining the scalability of the networks [19]. DS-TE was deemed to the most promising scheme to this aim. Bandwidth constraints models are used in DS-TE for admission control of traffic trunks by enforcing different bandwidth constraints for different classes of traffic so that QoS degradation can be minimized. Such bandwidth constraints models, to some extent, provide some congestion control rules within the congestion-avoiding traffic trunks. Thus DS-TE offers a "class over trunk" structure and achieves better QoS than other trunk-based frameworks. On the contrary, RM performs a complete mapping from per flow to per class, which discards traffic trunk and eliminates the maintaining spending. In comparison to DS-TE, we believe that RM has a further step as a middle course and accomplishes the combination of IntServ and DiffServ. To a single stream, RM behaves stateful on a per-flow level and stateless on a per-class level simultaneously in the networks.

IV. SIMULATION

As the discussions in section III, the main objectives of RM are to map distributed reservation into reserved/unreserved traffic and to protect reserved traffic by minimizing its loss in comparison to that of unreserved traffic. The purpose of the simulation in this study is to demonstrate the QoS and link utilization of RM. It doesn't relate to the performance evaluation of RM and other QoS frameworks, which should be the next research agenda.

A. Topology and Simulation

The topology used in simulations are shown in Fig. 4. There is a S node that generate flows destined for a R node through the network of two ER nodes and one CR node. All of these flows are reserved and marked prior to entering the ER and CR network. So the links within ERs use RED queue management and the links out of ERs use DropTail. The bottleneck of this network lies the CR-ER2 link labeled 5M bandwidth. For the purpose of distributed reservation, we assume the MSPK scheme to be adopted in AC model on CR and the maximum allocable bandwidth to be 0.95 of total. Once accepted and reserved, data flows entering the network are marked into two priorities according to their reservations by TSW2CM on the ER1 node. The packet drop rate of RED queues within ERs is set to 0.20 for unreserved traffic, and 0.05 for reserved traffic as a rigid protection.

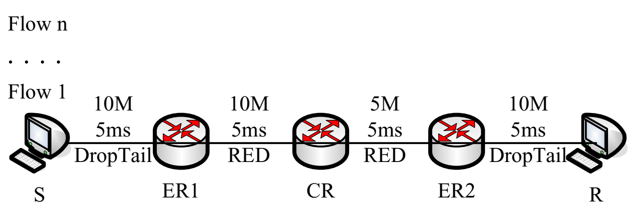

Fig. 4. Topology employed in simulations

Since we mainly focus on the MBAC and TSW2CM, which are the connection-level resource allocation and marking, the Exponential process is a good approximation for our purpose. The traffic generator settings on the S node are as follows: the individual flow uses 64K exponential traffic; the flow arrival distribution is exponential with an average of 400ms; the flow lifetime also has an exponential distribution with an average of 300s. In order to validate the support for different services of RM, all generated flows randomly claim GS, CS or BE service. A GS flow requests a full 64K reservation and a BE flow requests zero. Although a CS flow could request any value within the 64K limit, we assume that all GS flows have a uniform 32K reservation request. We applies a different police in treating the admission control of BE flows, while it always approved in RSVP-TE. In the simulation, BE flows participate in MBAC with a zero reservation request, which could avoid obvious congestion when estimate bandwidth exceeds the max limit. After a 3000s simulation, we get the traffic status in Fig. 5, which describes the changes of GS, CS and BE flow numbers during the simulation.

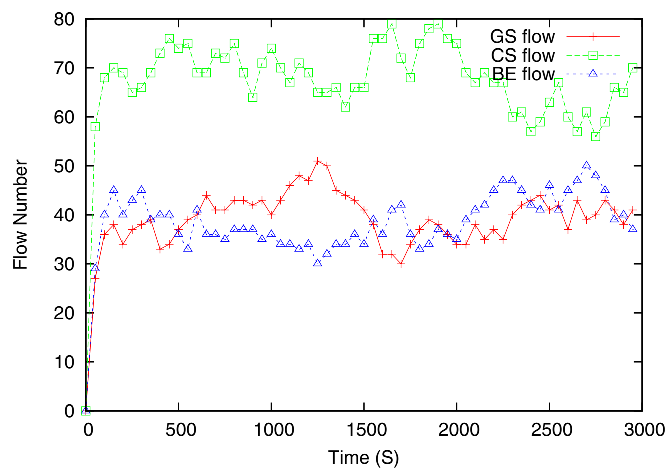

Fig. 5. Flows generated in simulations

B. Results

In order to examine the QoS of GS, CS or BE service, packet transmission information of the CR-ER2 link is traced and collected in TABLE III and TABLE IV. The column *ldrops* in these tables corresponds to the queue overflow, and the column *edrops* lists early drops due to the implemented RED mechanism. We find that reserved traffic is completely protected in RED queue, in which no packet dropped in comparison with a total 104495 packets dropped in unreserved traffic. But here are several reserved packets dropped due to the queue overflow. Indeed, the appearance of queue overflow is uncertain and inevitable on a MBAC environment. As the loss rate is limited to 0.0010% we concluded that reserved traffic is very close to entirely protected. From the view of per flow QoS, GS flows got 64k bandwidth guarantee and CS flows had a 32k minimum guarantee in the simulation. Consequently, we proved quantitative guaranteed QoS acquired in GS and CS service. The average loss rate of unreserved traffic is about 0.84% which shows that distributed MBAC works well and the whole congestion level is controlled on a low level.

The link utilization in Fig. 6 presents the simulation results examining the influences of incoming traffic characteristics to the utilization efficiency of the CR-ER2 link bandwidth. The estimate utilization is often beyond the 5M limit because of the existence of CS and BE flows, which is very helpful to increase actual utilization. The average actual utilization keeps a little above 91% during the most time of simulation, which proves

TABLE III
PACKET LOSS OF RESERVED TRAFFIC

Time	TotPkts	TxPkts	ldrops	edrops	DropRate %
500	48083	48083	0	0	0
1000	95999	95998	1	0	0.0010
1500	143957	143955	2	0	0.0014
2000	191889	191886	3	0	0.0016
2500	240030	240027	3	0	0.0012
3000	288055	288052	3	0	0.0010

TABLE IV
PACKET LOSS OF UNRESERVED TRAFFIC

Time	TotPkts	TxPkts	ldrops	edrops	DropRate %
500	2123801	2106586	1167	16048	0.81
1000	4398062	4357365	2604	38093	0.93
1500	6664373	6603432	3946	56995	0.91
2000	8925393	8844822	5311	75260	0.90
2500	11167608	11071527	6437	89644	0.86
3000	13416506	13304199	7812	104495	0.84

the high efficiency in exploiting network resources of RM. With MSPK algorithm and TSW2CM, we showed that RM achieved controllable quantitative guaranteed QoS and high link utilization. Studying the applicability of RM in different conditions more deeply is a topic of future work

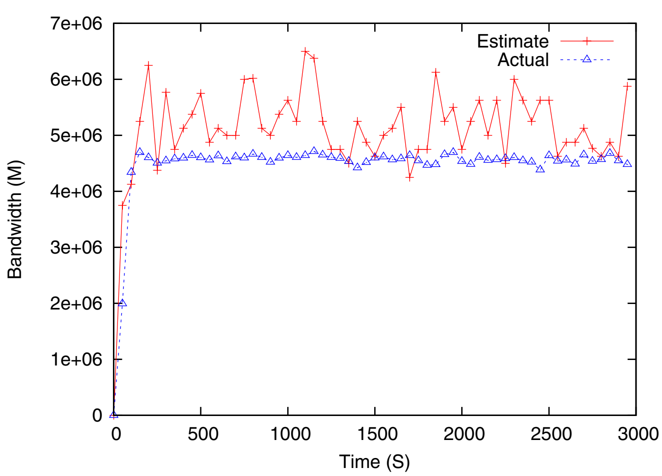

Fig. 6. Link utilization of bottleneck

V. CONCLUSION

We constructed a scalable RM framework to provide powerful QoS in the networks. The basic idea of RM is simple: per flow QoS can be achieved in a stateless way, if flow-specific reservation is quantitatively mapped into stream packets associated with corresponding Per Hop Behaviors.

We proposed Reservation Marking, a supply side marking method on the edge of the network, which marks stream packets according to its per flow reservation. RM adopts this mapping in a distributed resource reservation environment, discriminates reserved packets from unreserved ones, and secures anticipant QoS by protecting reserved packets on core routers. We also studied the implementation of RM, and evaluated its resource allocation and forwarding management in comparison to other QoS frameworks. The simulations demonstrated the powerful QoS of RM through the support to three types of service known as Guaranteed Service, Controllable Service and Best Effort Service. With the simulation results, we showed that RM behaved a good performance both in QoS and link utilization.

ACKNOWLEDGMENT

The authors would like to thank Shenyi Lin for proofreading the paper. This research is supported in part by Natural Science Foundation for Young Scholar of UESTC (L08010601JX0746 and 747), China NSF grant (60703114).

REFERENCES

[1] C. Cetinkaya, V. Kanodia, E. Knightly, *Scalable Services via Egress Admission Control*. IEEE Transactions on Multimedia, 2001, pp. 69-81.

[2] J. Yang, J. Ye, S. Papavassiliou, *Enhancing end-to-end QoS granularity in DiffServ networks via service vector and explicit endpoint admission control*. Proceedings of IEE Communications, 2004, pp. 77-81.

[3] Y. Yin, G. Poo, *End-to-end QoS guarantees for a network based on Latency-Rate Max-Min service curve*. Proceedings of 2005 IEEE International Conference on Communications, pp. 255-259.

[4] C.-K. Tham, T. Hui, *Reinforcement learning-based dynamic bandwidth provisioning for quality of service in differentiated services networks*. Computer Communications, vol 28, pp. 1741-1751, 2005.

[5] X. Zhou, C.-Z. Xu, *Harmonic proportional bandwidth allocation and scheduling for service differentiation on streaming servers*. Parallel and Distributed Systems, IEEE Transactions on Volume 15, Issue 9, pp. 835-848, 2004

[6] G. Miao, Z. Niu, *Bandwidth Management for Mixed Unicast and Multicast Multimedia Flows with Perception Based QoS Differentiation*. Proceedings of 2006 IEEE International Conference on Communications.

[7] P. Kulkarni, M. Nazeeruddin, S. McClean, *Building a Controlled Delay Assured Forwarding Class in Differentiated Services Networks*. Proceedings of 2006 SIGCOMM workshop on Internet network management, pp. 11-16.

[8] D. Awduche, L. Berger, et al., *RSVP-TE: Extensions to RSVP for LSP Tunnels*. RFC 3209, December 2001.

[9] F. Le Faucheur, W. Lai, *Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering*. RFC 3564, July 2003.

[10] S. Herrerła, M. Veiga, M. Rodrłguez, A. Surez, C. Lpez, *Edge-to-edge proactive congestion control for aggregated traffic*. Computer Communications, Volume 29, Issue 7, April 2006.

[11] Z. Mammer, *Framework for parameter mapping to provide end-to-end QoS guarantees in IntServ/DiffServ architectures*. Computer Communications, Volume 28, Issue 9, June 2005.

[12] A. Pereira, E. Monteiro, *Admission Control in IntServ to DiffServ mapping*. Proceedings of the International conference on Networking and Service, 2006.

[13] Y. Bener, P. Ford, R. Yavatkar, *A Framework for Integrated Services Operation over Diffserv Networks*. RFC 2998, November 2000.

[14] S. Blake, D. Black, et al., *An Architecture for Differentiated Services*. RFC 2475, December 1998.

[15] E. Rosen, A. Viswanathan, et al., *Multiprotocol Label Switching Architecture*. RFC 3031, January 2001.

[16] D. Clark, W. Fang, *Explicit allocation of best effort packet delivery service*. IEEE/ACM Transactions on Networking, Volume 6, Issue 4, August 1998.

[17] J. Zeng, J. Xu, Y. Wu, et al., *Service unit based network architecture*. Proceedings of the Fourth International Conference on Parallel and Distributed Computing, Applications and Technologies, 2003.

[18] J. Crowcroft, S. Hand, R. Mortier, et al., *QoSs Downfall: At the bottom, or not at all!*. Proceedings of SIGCOMM Workshop on Revisiting IP QoS (RIPQOS'03), August 2003.

[19] I. Stoica, H. Zhang, *Providing Guaranteed Services Without Per Flow Management*. Proceedings of ACM SIGCOMM' 99, September 1999.