Cross-Layer QoS Framework in the IEEE 802.16 Network

Yi-Ting Mai, Chun-Chuan Yang and, Yu-Hsuan Lin
Multimedia and Communications Laboratory
Department of Computer Science and Information Engineering
National Chi Nan University, Taiwan, R.O.C.
s3321906@ncnu.edu.tw, ccyang@csie.ncnu.edu.tw, kahala@vot.url.com.tw

ABSTRACT

To support multimedia service with the Broadband Wireless Access (BWA), the IP layer QoS is one of the keys to success. The IEEE 802.16 technology provides the wide area, high speed and non-light-of-sight wireless network. In the standard, the layer 2 technology of the IEEE 802.16 supports the QoS service with four service types. To improve the performance of the IEEE 802.16 network, we should consider how both the IP layer QoS and 802.16 QoS cooperate. In this paper, we propose a framework of cross-layer QoS support in the IEEE 802.16 network. Two novel mechanisms are proposed in the framework for performance improvement: Fragment Control and Remapping. Fragment Control handles the data frames that belong to the same IP datagram in an atomic manner to reduce useless transmission. Remapping is concerning about the mapping rules from IP QoS to 802.16 QoS and is designed to reduce the impact of traffic burstiness on buffer management. Simulation study has shown that the proposed scheme has higher goodput and throughput than the contrast.

Key words: 802.16, WiMax, PMP, QoS, Cross-Layer

1: INTRODUCTIONS

Broadband Wireless Access (BWA) technology provides an easy, time-saving, and low-cost method for deployment of the next generation (beyond 3G) network infrastructure. Since 1998, IEEE 802.16 working group has launched a standardization process called *Wireless Metropolitan Area Network* (*Wireless MAN*^{TM}) for BWA. The most updated specification of 802.16 (*IEEE Std 802.16-2004*) [1] focuses on fixed location wireless access and supports up to 134 Mbps data rate. Moreover, the standardization of a new 802.16 interface, *802.16e*[2], supports wireless access with high mobility, has also been completed recently. The *WiMax Forum* (*Worldwide Interoperability for Microwave Access*) [3], [4], a wireless industry consortium with about 100 members including major vendors such as *AT&T, Fujitsu, Intel*, and *Siemens Mobile*, is supporting 802.16 technology and promoting its commercial use, which means 802.16 is becoming the most important technology in BWA.

This work was supported in part by the National Science Council, Taiwan, R.O.C., under grant NSC95-2219-E-260-004.

As shown in Figure1, the *PMP* (*Point to Multipoint*) configuration of IEEE 802.16 network consists of a *base station* (*BS*) and a couple of *subscriber stations* (*SS*) that connect to the BS via high-speed wireless link. The BS acts as a gateway to the Internet. Legacy LANs or even more complex subnet systems can connect to the IEEE 802.16 network via SS. An IEEE 802.16 network (including the Legacy LANs that connect to the SS) can cover a large geographical area since the distance between the BS and the SS can be up to 30 miles [1].

In order to provide better QoS service over the 802.16 network, layer 3 (L3) and layer 2(L2) QoS services must be integrated. Therefore, cross-layer mechanisms are designed in the proposed QoS framework, including (1) Mapping from L3 QoS to L2 QoS, (2) the admission control for QoS flow, (3) the fragmentation scheme, and (4) the remapping scheme. Some research works [5]-[9] have been proposed in the literature, in which the dynamic admission and the scheduling scheme are their major focus.

The remainder of the paper is organized as follows. First of all, we present the overall architecture as well as the novel features of the proposed QoS framework in section 2. Key mechanisms in the proposed framework for QoS support in IEEE 802.16 network are presented in section 3. Simulation study for performance evaluation and comparisons is presented in section 4. Finally, section 5 concludes this paper.

2: CROSS-LAYER QOS FRAMEWORK

Although the 802.16 standard only defined up to layer 2 specification for the BS and SS, the proposed framework requires the BS and SS to be equipped with some of the layer 3 functionalities, such as IP header processing and layer 3 service class interpretation, for better service support. Since the traffic flows in the 802.16 network are classified as downlink or uplink, we present the framework in the downlink mode and the uplink mode respectively in the following:

2.1: Downlink mode

In the downlink mode, we assume the sender is located outside the 802.16 network and the receiver is located within the 802.16 network as displayed in figure 2. The framework in the downlink mode is illustrated in

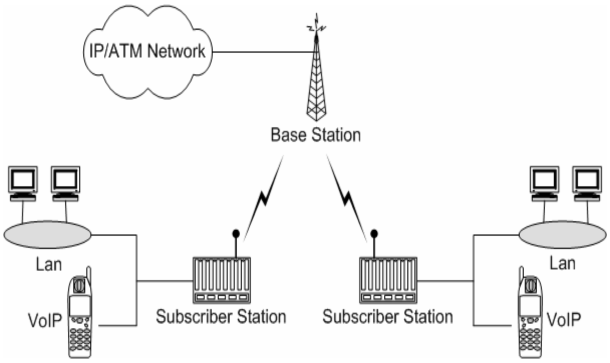

Figure 1.IEEE 802.16 PMP mode

figure 3. Main functional blocks in the proposed QoS framework are briefly explained as follows:

(1) Connection Setup: Since there are mainly two categories of Quality-of-Service framework in Layer 3 (IP layer), *Integrated Service (IntServ)* [10] [11] and *Differentiated Service (DiffServ)* [12], the functional blocks of Classifier and QoS Mapping from L3 to L2 are required at the BS for resource management in Layer 2 admission control. In other words, the BS must be equipped with some of IP layer functionality, such as interpretation of IP header, to have a better support of QoS.

(2) Fragmentation: The size of an IP datagram can be up to 64K bytes, but the size of each slot *(Maximum Transmission Unit, MTU)* in IEEE 802.16, although depending on the channel rate and the time frame length, is much smaller than the size of IP packets. Thus, fragmentation is a required function at the BS. Moreover, considering that all fragments coming from the same IP datagram must be successfully delivered to the destination for reassembly, these fragments should be treated as a whole in the 802.16 network. Therefore, the mechanism of *Fragment Control* is proposed in the framework, which maintains the dependency of the fragments from the same IP packet during Layer 2 operations.

(3) Downlink Scheduler: The scheduler at the BS is responsible for dispatching IEEE 802.16 data frames of different service types at proper times (time slots). Since there are four service types, namely *UGS (Unsolicited Grant Service), rtPS (real-time Polling Service), nrtPS (non-real-time Polling Service), BE (Best-Effort)*, defined in IEEE 802.16, four queues are required in the scheduler.

(4) Queue Monitor: Queue Monitor is used for monitoring the state of each queue in the scheduler and cooperates with the Mapping functional block for better resource management. More specifically, Queue Monitor can change the mapping rule from L3 service type to L2 service type under certain situations to increase the utilization of the queues in the scheduler. The mechanism of changing the mapping rule is called *Remapping* in the paper.

2.2: Uplink mode

We assume the sender is connected to the SS in the uplink mode. As illustrated in figure 4, the operation of the framework in the uplink mode is more complicated than the downlink mode, since the SS must negotiate with the resource manager BS. Major differences of the uplink mode from the downlink mode are explained in the following:

(1) Cooperation of SS and BS: Although BS is the administrator in IEEE 802.16 network, the SS shall negotiate with the BS in the phase of connection setup and uplink scheduler. For example, message *DSA (Dynamic Service Addition), DSC (Dynamic Service Change), DSD (Dynamic Service Deletion)* are used in the admission control. Moreover, the SS must send out *BW_REQ(Bandwidth Request)* messages to the BS for resource allocation and channel access.

(2) Virtual Reassembler: Since an IP packet received at the SS comes from a subnet system (e.g. a legacy LAN) in which the source host locates, the IP packet is probably merely one of the fragments of its original datagram. To mark the fragment control mechanism more effective, a virtual reassembler is added before fragmentation. The virtual reassembler is used of identifying the IP fragments that belong to the same original datagram by virtually reassembling the fragments.

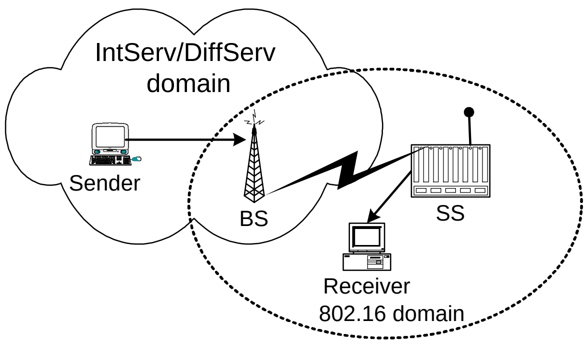

Figure 2.Downlink diagram

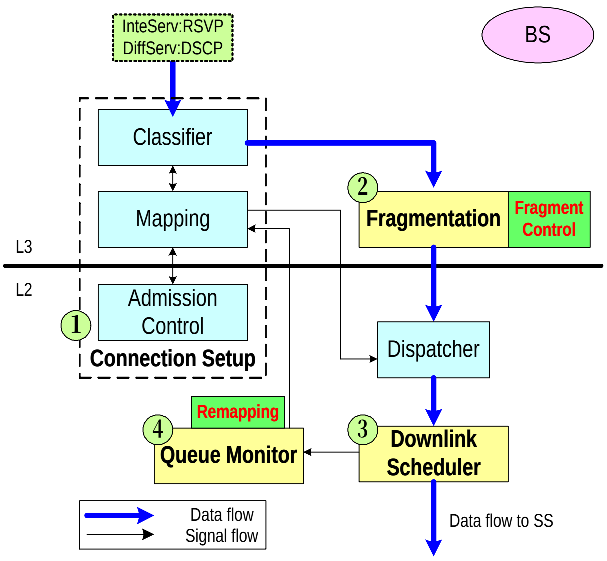

Figure 3.QoS framework in the downlink mode

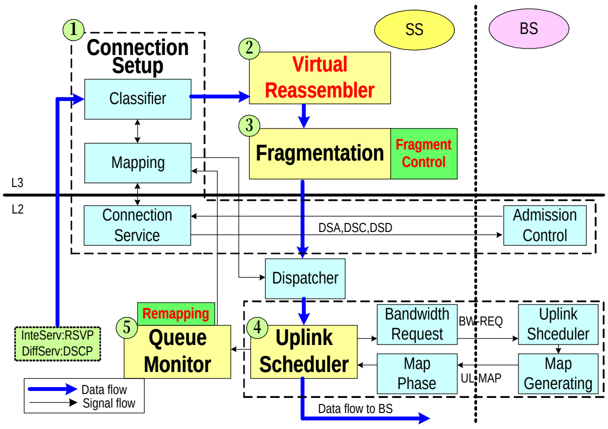

Figure 4.QoS framework in the uplink mode

3: CROSS-LAYER QOS MECHANISMS

3.1: QoS mapping from L3 to L2

There were mainly two QoS frameworks in IP layer: Integrated Service (IntServ) and Differentiated Service (DiffServ), each of them defines different classes of QoS. We adopted a simple mapping rule from IP layer QoS to 802.16 QoS types [6] in our proposed framework as illustrated in figure 5.

	IP QoS	802.16 QoS
IntServ	*Guarantee Service (GC)*	*Unsolicited Grant Service (UGS)*
	Controlled Load (CL)	*Real-time Polling Service (rtPS)*
DiffServ	*Expedited Forwarding (EF)*	
	Assured Forwarding (AF)	*Non-Real-time Polling Service (nrtPS)*
IntServ, DiffServ	*Best Effort (BE)*	*Best Effort (BE)*

Figure 5. Mapping rule from IP QoS to 802.16 QoS

3.2: Admission control

We adopt a simple rate-based admission control scheme, in which the new QoS flow must provide the required bandwidth and the BS check if there is enough capacity for the new flow. The algorithm of the admission control is displayed in Figure 6. For example, a new UGS flow with bandwidth requirement b_{UGS} is accepted when the remaining capacity (i.e. the total capacity of the link B – the current load b_C) is larger than b_{UGS}. Moreover, since the characteristic of the flow in each service type varies, the required bandwidth defined for each service type should be different. More specifically, the *peak rate* for an UGS flow, the *average rate* for an rtPS flow, and the *minimum rate* of an nrtPS flow are used in the admission control respectively.

3.3: Fragment Control

As mentioned in section 2.1, since fragmentation is always necessary for an IP packet to be transmitted via the 802.16 link, the objective of proposed Fragment Control is to provide a grouping mechanism so that the fragments of the same IP packet are treated as a whole during Layer 2 processing. We assume that all fragments of the same IP packet are put into the layer 2 buffer in an atomic manner such that fragments coming different IP packets are not interleaved in the buffer. Therefore, one bit of a flag field in the header of the 802.16 MAC frame is enough for grouping the fragments. The reserved bit (*Rsv 1*) in the header of the 802.16 MAC frame (Figure 7) is used for fragment grouping.

The fragments coming from the same IP packet are marked with the same value ('0' or '1') alternately in the flag field of the MAC frames and put into the 802.16 queue as illustrated in Figure 8. Layer 2 buffer operations are designed to treat the fragments with the same marking as a group. Therefore, in the case of congestion, the fragments of the same group should be removed all together for saving unnecessary frame transmissions in the congestion control mechanism such as *Drop Tail* or *Random Early Detection* (*RED*).

(1) Since the sender connected the SS directly, the uplink mode was easy to identify the data traffic belonged to the same sender in the SS. We added the Virtual reassembler to mark the all data frames with the same mark before put on the L2 buffer.

3.4: Remapping

The proposed remapping scheme is concerning with integrated buffer management of rtPS and nrtPS queues to achieve better buffer utilization and reduce frame dropping. Since the framework adopts static mapping rules from L3 QoS classes to 802.16 service types, there are cases that the rtPS queue overflows due to bursty traffic condition while the nrtPS queue still can accept more data frames. To better utilize buffers in the queues, a remapping rule is designed for L3 higher priority CL and EF packets to use nrtPS buffers when the rtPS queue is going to be full.

To support the remapping scheme, buffer utilization of rtPS and nrtPS queues must be monitored. Moreover, two threshold parameters, *Upper-Bound* and *Lower-Bound* as displayed in Figure 9, are defined for the queues. Rules in the remapping scheme are explained as follows:

(1) When buffer utilization of the rtPS queue exceeds its Upper-Bound, the queue monitor notifies the Mapping module in the framework triggering new remapping rules that map CL, EF, and AF packets to nrtPS as illustrated in Figure 10.

(2) In the case of remapping being operated, if buffer utilization of rtPS queue is lower than Lower-Bound, the mapping rules are restored back to the original ones as shown in Figure 5.

(3) In the case of the nrtPS queue exceeds its Upper-Bound (i.e. the nrtPS is going to be full soon), the original mapping rules are restored only when buffer utilization of rtPS queue is lower than the middle line of Upper-Bound and Lower-Bound to reduce oscillations of rule application.

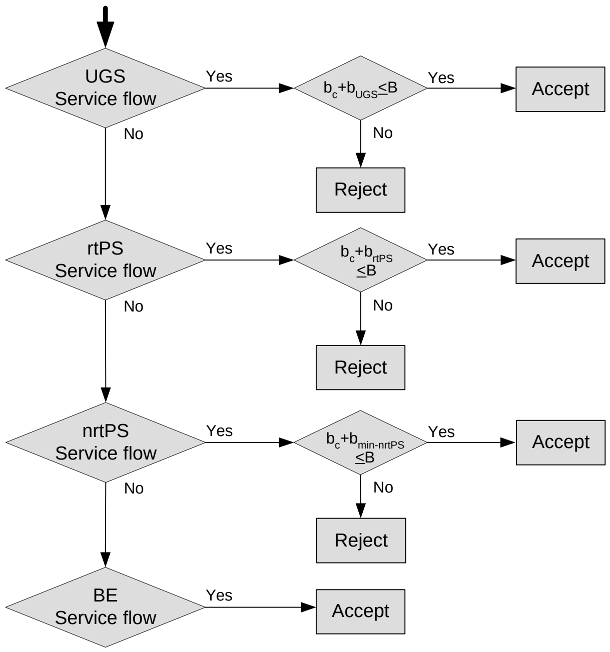

Figure 6. Admission control rule

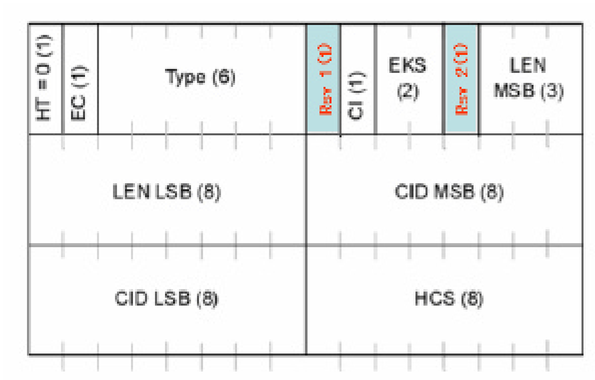

Figure 7.802.16 MAC header format

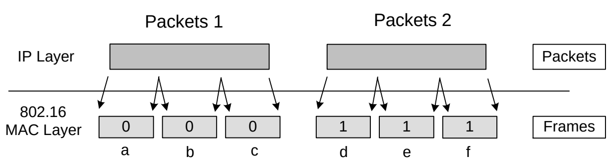

Figure 8.Marking L2 frames according to L3 packets

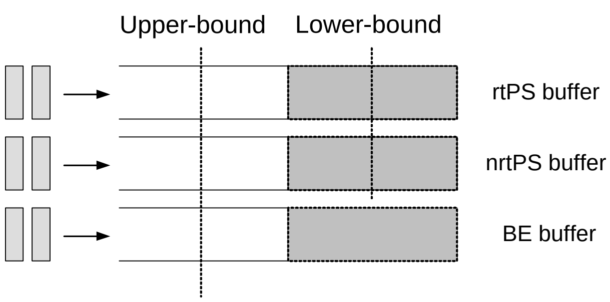

Figure 9.The two thresholds in the L2 buffer

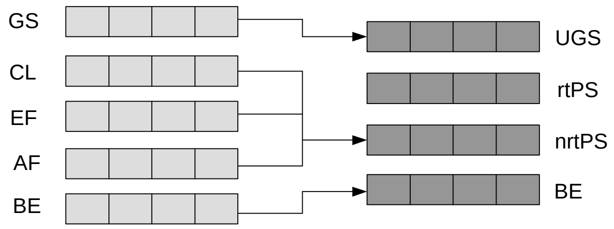

Figure 10. The new rtPS buffer mapping rule

4: PERFORMANCE EVALUATION

4.1:Simulation parameters and performance criteria

Simulation study was conducted to evaluate the proposed Fragment Control and Remapping mechanisms. Two performance criteria were defined for comparison: the *goodput* and *throughput*. The goodput was defined as the data rate arriving at the destination and can be successfully reassembled. The throughput was defined as the data rate arriving at the destination.

The scheme of *DropTail* (DT) as the congestion control mechanism was adopted in the simulation. Cases of DT with and without Fragment Control and Remapping were evaluated. Since there are in total six service types in the two IP QoS frameworks, we generated six types of flows with inter-arrival time and duration exponentially distributed. Parameters in the simulation for Fragment Control and Remapping were displayed in Table 1 and Table 2 respectively.

4.2:Simulation results

As shown in Figures 11-14, goodput in the case of DT with Fragment Control is higher than the contrast. Moreover, as the input load increases, it causes a higher probability of buffer overflow. Therefore, the proposed Fragment Control scheme can achieve even more performance gain over the contrast in heavier loads, especially for nrtPS traffic flows. The figures demonstrate that the proposed Fragment Control mechanism can improve the goodput up to 15% under very heavy input load.

Figure 15 displays the throughput of the Remapping scheme and the contrast. Since the scheme allows rtPS flows to use the buffers of the nrtPS queue under congested condition, the overall throughput is increased. However, as the input load reaches 100% implying that all queues are close to saturation, thus the Remapping scheme does not present any benefit in the case.

5: CONCLUSIONS

As the most promising Wireless-MAN technology, IEEE 802.16 provides broadband, wide coverage, and QoS support to meet the demand of the next generation BWA (Broadband Wireless Access) network. To achieve the better QoS service in the IEEE 802.16 network, we proposed a cross-layer QoS framework integrating L3 and L2 QoS in the IEEE 802.16 network. Main functional blocks in the framework include: QoS mapping from L3 to L2, Admission control, Fragment Control, and Remapping. Fragment Control handles the data frames from the same IP datagram as a group in L2 operations to reduce useless transmission. Remapping is designed for more flexible use of L2 buffers by changing the mapping rules from IP QoS to L2 service type under congested situation of the rtPS queue. Simulation results have demonstrated that the proposed framework as well as the associated mechanisms can achieve the better performance in terms of the goodput and throughput in the heavy input load.

REFERENCES

[1] IEEE Std 802.16-2004, "IEEE Standard for Local and metropolitan area networks--Part 16: Air Interface for Fixed Broadband Wireless Access Systems," Oct. 2004.-

[2] IEEE Std 802.16e-2005, "IEEE Standard for Local and metropolitan area networks--Part 16: Air Interface for Fixed Broadband Wireless Access Systems--Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands," Feb. 2006.

[3] "Business Case Models for Fixed Broadband Wireless Access based on WiMAX Technology and the 802.16 Standard," WiMax Forum, 10 Oct. 2004.

[4] S. J. Vaughan-Nichols, "Achieving Wireless BroadBand with WiMax," IEEE Computer, pp.10-13, Jun. 2004.

[5] H.S. Alavi, M. Mojdeh, and N. Yazdani, "A Quality of Service Architecture for IEEE 802.16 Standards," Proceedings of 2005 Asia-Pacific Conference on Communications, pp.249-253, Oct. 2005.

[6] J. Chen, W. Jiao, and Q. Guo, "Providing integrated QoS control for IEEE 802.16 broadband wireless access systems," Proceedings of the IEEE 62nd Vehicular Technology Conference (VTC 2005-Fall), vol. 2, pp. 1254-1258, Sept. 2005.

[7] H. Wang; W. Li, and D.P. Agrawal, "Dynamic Admission Control and QoS for 802.16 wireless MAN," Proceedings of Wireless Telecommunications Symposium(WTS 2005), pp. 60-66, April 2005.

[8] J. Chen, W. Jiao, and H. Wang, "A service flow management strategy for IEEE 802.16 broadband wireless access systems in TDD mode," Proceedings of IEEE International Conference on Communications(ICC 2005), vol. 5, pp. 3422-3426, May 2005.

[9] K. Wongthavarwat, and A. Ganz, "Packet Scheduling for QoS Support in IEEE 802.16 Broadband Wireless Access Systems," International Journal of Communication Systems, vol. 16, pp.81-96, Feb. 2003.

[10] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," IETF RFC 1633, Jun. 1994.

[11] J. Wroclawski, "The Use of RSVP with IETF Integrated Services," IETF RFC 2210, Sept. 1997.

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

Table 1. Simulation parameters with Fragment Control

Input	Service type	Mean	Variation
UGS=3Mbps rtPS=4Mbps nrtPS=3Mbps BE=2Mbps (QoS Traffic load with 100% input rate)	GS (UGS)	3072 Kbps	
	CL (rtPS)	2048 Kbps	512 Kbps
	BE (BE)	1024 Kbps	256 Kbps
	EF (rtPS)	2048 Kbps	512 Kbps
	AF (nrtPS)	3072 Kbps	768 Kbps
	BE (BE)	1024 Kbps	256 Kbps
Total Bandwidth (Downlink)		10240 Kbps = 10Mbps	
MAC Frame Size		5 ms, 100 slots	
Simulation Time		100000 ms = 100 sec	
L2 Buffer Size		50Kb = 100 slots	
Packet size		1216Bytes	

Table 2. Simulation parameters with Remapping

Input	Service type	Mean	Variation
UGS=3Mbps rtPS=3Mbps nrtPS=2Mbps BE=2Mbps (Traffic load with 100% input rate)	GS (UGS)	3072 Kbps	
	CL (rtPS)	2048 Kbps	512 Kbps
	BE (BE)	1024 Kbps	256 Kbps
	EF (rtPS)	1024 Kbps	256 Kbps
	AF (nrtPS)	2048 Kbps	512 Kbps
	BE (BE)	1024 Kbps	256 Kbps
Total Bandwidth (Downlink)		10240 Kbps = 10Mbps	
MAC Frame Size		5 ms, 100 slots	
Simulation Time		100000 ms = 100 sec	
L2 Buffer Size		50Kb = 100 slots	
Packet size		1216Bytes	
Upper-bound		80%	
Low-bound		40%	

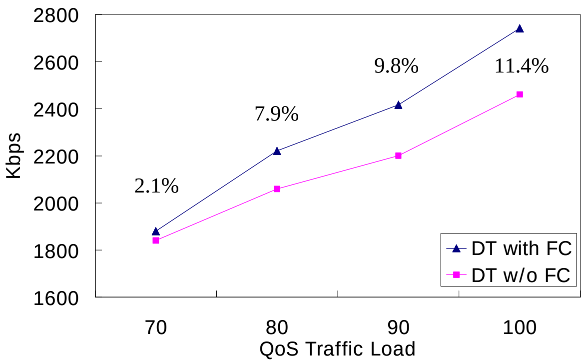

Figure 11. The goodput with UGS

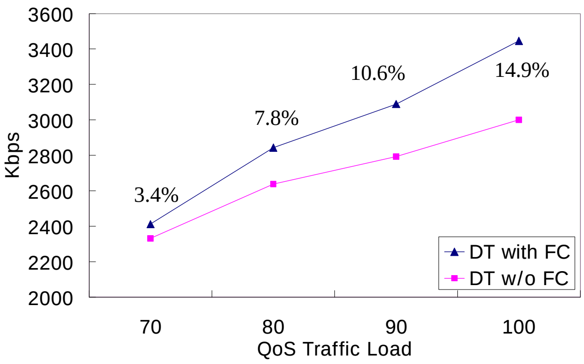

Figure 12. The goodput with rtPS

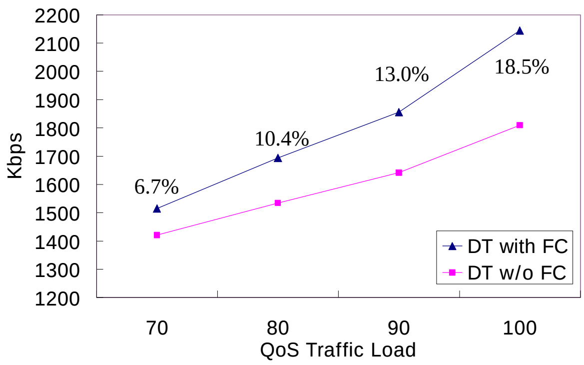

Figure 13. The goodput with nrtPS

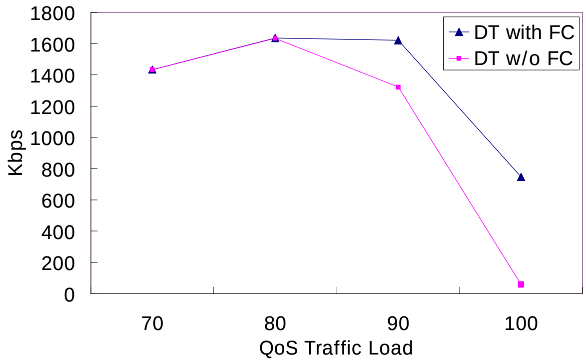

Figure 14. The goodput with BE

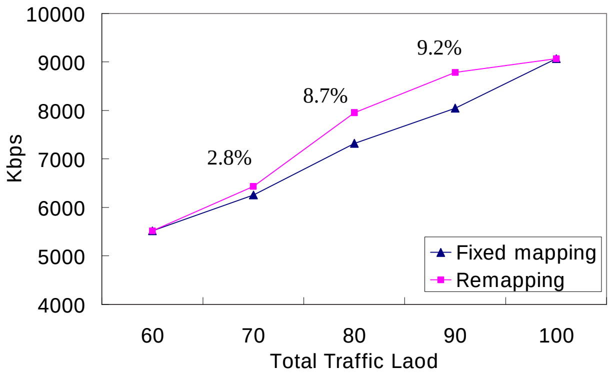

Figure 15. Total throughput with Remapping