A Bandwidth-efficient Scheduler for MPLS DiffServ Networks

Ximing Zeng (frankzxm@sce.carleton.ca)
Chung-Horong Lung (chlung@sce.carleton.ca)
Changcheng Huang (huang@sce.carleton.ca)
Department of System and Computer Engineering
Carleton University, K1S 5B6 Canada

Anand Srinivasan
(anand@eion.com)
Eion Incorporation
K1Y 2X5 Canada

Abstract

DiffServ networks support three forwarding classes: EF, AF and BE. Scheduling algorithms based on priority queueing such as Cisco LLQ and MDRR seem to be a natural choice to provide EF service. The main problem with priority queueing is that, due to lack of service isolation, QoS to AF classes may not be guaranteed if EF traffic is bursty. As an alternative, Fair Queueing schedulers can provide service isolation and guarantee QoS to all classes by reserving certain bandwidth. However, over-provisioning required by EF service leads to low bandwidth utilization. In this paper, we propose a bandwidth-efficient scheduler to provide QoS guarantees to both EF and AF classes. Simulation results show that the proposed scheduler outperforms Cisco MDRR in guaranteeing QoS to AF classes and excels fair queueing schedulers in bandwidth utilization.

1. Introduction

Traditional Internet only provides one type of service, namely the best effort service, to all applications. Under the best effort concept, all packets are treated the same in a first come first serve basis and there is no way to differentiate the services to different applications. The traditional Internet works well in supporting non-real time applications such as HTTP or FTP. However, with the increasing demands of supporting the multimedia real time traffic such as Internet Telephony, Video Conferencing, Interactive Game and the mission critical business traffic, reliable services with quality of service guarantee are required.

Best effort service can no longer meet the diverse QoS requirements. DiffServ [1] model was proposed in 1998 as a scalable model to support QoS in the Internet. In DiffServ network, QoS is provided to service classes. Packets from applications with similar QoS requirements are assigned the same service class at the edge of the DiffServ network and aggregated in the core network. DiffServ model meet the QoS requirements of different service classes by providing Per Hop Behaviors (PHB). Three PHBs are currently supported.

Expedited Forwarding (EF) PHB [2] is the key ingredient in DiffServ for providing a low-loss, low-latency, low-jitter, and assured bandwidth service. EF can be implemented using priority queueing with rate limiting on the class. Real time applications with stringent delay requirement such as VoIP, interactively game are especially suitable to be forwarded using EF. Although EF can provide the premium service, only the critical applications should be provided by it since under congestion it is not possible to treat all traffic as high priority traffic. Cisco suggested that the volume of EF traffic should not exceed 33% of the link speed.

Assured Forwarding (AF) PHBs [3] are defined to provide different forwarding assurances. The AFxy PHB defines four AFx classes; namely, AF1, AF2, AF3, and AF4. Each class is assigned a certain amount of buffer space and interface bandwidth to guarantee certain QoS. Within each class AFx, three drop precedence values are defined. Under congestion, the packets marked with high drop precedence will be dropped first. Therefore, packets within the same class AFx may experience similar QoS in delay and jitter but different QoS in loss rate. Usually, packets are marked according to their service agreements with the service provider. Packets exceed the service profile will be marked a high drop precedence and dropped first under congestion. Those non real-time applications such as streaming video can use AF service. AF classes can be implemented using fair queueing scheduler.

Like in the traditional Internet, Best Effort (BE) PHB provides no service guarantee. All packets belong to BE class are treated the same. Cisco suggested that to prevent

Proceedings of the The IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'04)
1526-7539/04 $20.00 © 2004 **IEEE**

service starvation a minimum throughput should be guaranteed to the BE traffic by pre-allocating certain bandwidth to it. This is the reason in OSPF-TE a link cannot declare all its bandwidth to be reservable. Typically, the maximum reservable bandwidth on a link is 75% of the total bandwidth. Therefore, 25% of the bandwidth is pre-allocated to BE service.

Traditional IP DiffServ Network is based on the Service Level Agreements (SLAs). A SLA is the service contract a customer (can be an organization or an upstream service provider) signed with its service provider to guarantee certain PHBs. At the access point (Edge Router) of a DiffServ domain, traffic is policed according to the SLA. Traffic conforms to the SLA is allowed to enter the network regardless of their destinations. With traffic routes and destinations being random and dynamic, it is hard to perform connection admission control and resource reservation on a per connection basis. Under an overloaded network, service disorder may happen and it is difficult to guarantee end-to-end QoS. One way to solve this problem is to provide destination-based SLA so that traffic can be policed according to their destinations. Usually a bandwidth broker is needed as a central control point in each DiffServ domain to manage the bandwidth allocations and perform the admission control.

In MPLS DiffServ Network [4], things are different. MPLS network is a connection-oriented network that facilitates traffic engineering. Resource reservation protocols can be use to reserve the bandwidth along each LSP from the ingress node to the egress node. Connection admission control can be done in a distributed way as each router can make its local decision according to its own resource reservation status. OSPF-TE and Constraint Based Routing can be used to route the traffic with QoS requirements. In this paper, we will focus on designing a scheduler that can facilitate the traffic engineering in the MPLS DiffServ network.

In the following sections, we will discuss in detail how bandwidth can be allocated to support EF, AF and BE PHBs under different schedulers. The rest of the report is organized as following: Section 2 reviews Cisco's solutions to support DiffServ. Section 3 discusses the fair-queueing based solutions. In section 4 we propose a new scheduler, the WFQ-P (Weighted Fair Queueing with Priority) scheduler. Section 5 studies the performance of WFQ-P scheduler using simulations. Section 6 summarizes the paper.

2. Cisco solutions

Cisco proposed two scheduling algorithms to support DiffServ PHBs, Modified Deficit Round Robin (MDRR) and Low Latency Queueing (LLQ). Their structures are similar. Under both schedulers, EF is implemented using priority queueing. AF and BE are provisioned by fair queueing. The major difference is that LLQ uses weighted fair queueing and MDRR uses Deficit Round Robin to implement fair queueing.

Since EF traffic is given the high priority, it is guaranteed small delay and jitter. A policer is needed to regulate EF traffic to prevent the service starvation of other low priority service classes. If EF traffic exceeds certain rate limit, it will be dropped before enters the network. AF and BE traffic share the "rest" of bandwidth that is not used by EF class. However, Cisco does not specify how much is the "rest" of bandwidth that is allocable to AF and BE classes after EF is policed at certain rate and given the high priority. If the EF traffic has a constant bit rate r, the rest of bandwidth is clearly $R-r$, where R is the link speed. However, if EF traffic has a variable bit rate with r_{peak} and $r_{avg.}$ as the peak and average rate respectively, it is unspecified how much the rest bandwidth should be. Obviously, it is not an economic way to reserve according to the peak rate of EF traffic and count the reservable bandwidth to other service classes as $R-r_{peak}$ since most of the time EF traffic is not generated at its peak rate. The waste of bandwidth becomes dramatic when the EF traffic is bursty. To maintain a high link utilization ratio, generally the rest of bandwidth that is available should be counted as: $R-r_{reserve}$, where $r_{reserve}$ has a value between r_{peak} and $r_{avg.}$ However, declaring any bandwidth ($R-r_{reserve}$) larger than $R-r_{peak}$ to be allocable may affect the QoS guarantees to AF classes. Since when a burst of EF traffic comes, the priority scheduler will keep serving the EF traffic, the rest of service classes cannot be served as their expected rate: $R-r_{reserve.}$ As a result, if the QoS is estimated based on $R-r_{reserve}$, it may not be always guaranteed.

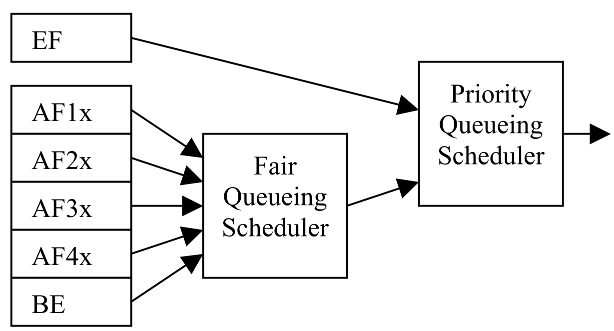

Figure 1. Cisco MDRR or LLQ scheduler

3. Solutions based on fair queueing

As discussed in the previous section, the disadvantage of using priority queueing is that, due to lack of service isolation, the service to lower priority classes cannot be guaranteed when the high priority traffic is bursty. Fair queueing schedulers are capable of providing service isolation among service classes. By engineering the bandwidth properly, each service class can be guaranteed its QoS independently to other service classes.

A lot of scheduling algorithms have been proposed in the literature [5]-[8]. They all follow the General Processor Sharing (GPS) concept. QoS under fair queueing schedulers can be guaranteed by allocating certain bandwidth. Specifically, if the traffic source is leaky bucket constrained with parameters (σ, ρ) at the edge of the network, the worst-case end-to-end delay on a route with n nodes can be guaranteed as shown in [7]:

$$D \leq \frac{\sigma}{r} + \sum_{i=1}^{n} C_i^s \qquad (1)$$

D is the end-to-end delay. σ is the leaky bucket depth. r is the minimum guaranteed rate of all n nodes and should be larger than the leaky bucket rate ρ. C_i^s is the latency of the scheduler of node i. Interested reader may refer to [7] for the details on latency of different type of schedulers. In brief, latency is the worst case waiting time a session has to wait to begin its service at the assigned rate. The importance of equation (1) is, although each node may use different type of schedulers to implement fair queueing, the end-to-end delay bound still has a simple form by introducing the latency.

The real packets delays are usually much smaller than the worst-case delay given in equation (1) and delay violation to some extent is tolerable by real applications. Therefore, allocating bandwidth according to equation (1) is too conservative and causes low link utilization. In practice, bandwidth allocation may be based on measurement [11]-[12] or other methods such as effective bandwidth [13]-[16]. However, equation (1) does give us some insight on the relation between delay guarantee and bandwidth allocation. For service class with the strict delay requirement such as EF class, to allow some burstness, the assigned rate r can be much larger than its average rate ρ. As a result, if static bandwidth allocation is used in fair queuing schedulers, significant over-provisioning to bursty EF traffic may occur.

To achieve higher link utilization, instead of assigning bandwidth statically, Dynamic Weighted Fair Queueing (DWFQ) scheduler is proposed in [10]. To minimize the over-provisioning to EF, bandwidth allocation is adjustable according to the current QoS. If EF packets currently receive good QoS, its bandwidth allocation will decrease to allow other service classes to use the bandwidth. On the other hand, if EF traffic has a bad QoS, its bandwidth allocation will increase. In particular, an average queue length of EF packets is calculated using exponential average of the instant queue length. The bandwidth allocation changes linearly as the average queue length increases or decreases between a minimum and maximum queue length thresholds. To maintain a small delay to EF packets, typically the maximum average queue length threshold is set to be 2 packets.

A major problem with this kind of dynamic approach is how often the QoS should be measured and bandwidth allocations be adjusted accordingly. If the adjustment cannot keep up with the traffic dynamics, good QoS cannot be provided. Obviously, better QoS can be expected by more frequent bandwidth-reallocations. But monitoring QoS and changing the bandwidth allocations too often in a high-speed core router may lead to dramatic computational overhead. Another problem with this dynamic approach is that connection admission control becomes difficult as the bandwidth allocations change in time.

4. A bandwidth-efficient scheduler

Unlike the traditional IP DiffServ network where the routes are dynamic and traffic destinations are not under control, MPLS DiffServ network is a connection–oriented network. Resource reservation protocols such as RSVP, RSVP-TE, RSVP with MPLS extension [4] can be used to reserve bandwidth for each LSP along its path. While receiving a new LSP setup request with its flow specifications and desired service, extra bandwidth required to guarantee the QoS can be calculated. Constraint based routing can be used to find a route. With the signaling or resource reservation protocols, the router along the route makes resource reservation and accepts the connection if there is enough bandwidth. After each successful LSP setup or tear down, the new bandwidth utilization information is broadcasted so that constraint based routing can always have the up-to-date information. Therefore, MPLS makes it possible to perform the end-to-end connection admission control and bandwidth management on a per LSP basis. Bandwidth allocations in each router can be done in a more accurate sense.

Both schedulers based on priority queueing and fair queueing have their drawbacks in using bandwidth efficiently. The former lacks of the protection to AF service classes. As discussed in section 2, to minimize the service impact bursty EF traffic has on AF traffic, the amount of bandwidth allocable to AF classes are limited. The latter needs over-provisioning to provide bursty EF traffic low delay service. As a result, they both are inefficient in terms of bandwidth utilization. The dynamic weighted fair queueing solves the problem but it is computationally more expensive. Furthermore, dynamic bandwidth allocations make it difficult for DWFQ to be used in any connection admission control scheme to provide hard QoS guarantees.

We developed a simple scheduler that not only uses the bandwidth more efficiently but also fits well in the MPLS DiffServ architecture. The basic idea is simple: the bandwidth over-provisioned to EF class can be used by BE class. Below is the architecture of our scheduler:

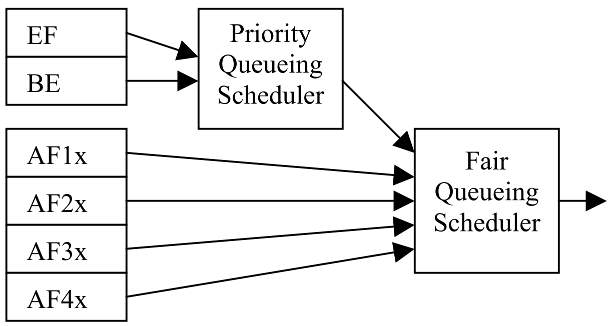

Figure 2. WFQ-P scheduler

The scheduler is a hierarchical scheduler. In the higher level, it is a fair queueing scheduler. Each AF class has its own bandwidth reservation so that hard QoS can be provided. EF and BE classes share the same bandwidth allocation. In the second level, EF and BE classes are served by a priority queueing scheduler. EF packets always have the priority to use the allocated bandwidth. BE packets get served only if there are no EF packets waiting. We call this scheduler the Weighted Fair Queueing scheduler with Priority. For notation convenience, we use the name WFQ-P to refer to the scheduler in the rest of the paper.

4.1 Bandwidth Allocations under WFQ-P

Bandwidth to EF and BE classes is allocated as following:

$$r = \max(r_{avg}^{EF} + r^{BE}, r_{res}^{EF}) \qquad (2)$$

Where r_{avg}^{EF} is the average rate of EF traffic; r_{res}^{EF} is the bandwidth needed to guarantee QoS to EF class under fair queuing concept. If hard QoS such as the delay bound is required, r_{res}^{EF} can be calculated using equation (1). In the case that the QoS is guaranteed in a probability sense, methods based on effective bandwidth are applicable. To provide a low delay service, r_{res}^{EF} is generally much larger than r_{avg}^{EF} if the traffic is bursty. r^{BE} is bandwidth required to guarantee the minimum throughput of BE class.

Compared with fair queueing under which the bandwidths of EF and BE classes are allocated separately, our approach saves bandwidth as following:

$$r_{gain} = r_{res}^{EF} + r^{BE} - \max(r_{avg}^{EF} + r^{BE}, r_{res}^{EF}) \qquad (3)$$

Specifically, if EF traffic is not so bursty in that r_{avg}^{EF} + $r^{BE} > r_{res}^{EF}$, we have:

$$r = r_{avg}^{EF} + r^{BE}, r_{gain} = r_{res}^{EF} - r_{avg}^{EF} \qquad (4)$$

Equation (4) shows the reservation can be made based on the average rates of EF and BE traffic. There is no over-provisioning, or in other words, waste of bandwidth.

Under the case that EF traffic is extremely bursty, $r_{avg}^{EF} + r^{BE} <= r_{res}^{EF}$. Over-provisioning is unavoidable. We have:

$$r = r_{res}^{EF}, r_{gain} = r^{BE} \qquad (5)$$

Equation (5) means the reservation should be based only on the need of EF traffic. No extra bandwidth needs to be reserved for BE traffic. In fact, all the BE traffic can be transferred during the intervals when the EF traffic is low. All the bandwidth needed by BE can be saved.

Considering the fact that r^{BE} is usually about 25% of the total link speed, the results of equation (4) and (5) tell us either there will be no bandwidth wasted or we can save all the bandwidth required by BE traffic if EF traffic is extremely bursty. The gain is significant.

4.2 QoS under WFQ-P

We have shown the advantage of saving bandwidth under WFQ-P, a natural question is how the QoS can be guaranteed.

All QoS requirements under fair queueing schedulers can be mapped to bandwidth requirements. It is beyond the scope of this paper to answer the question as how much bandwidth is needed. Methods such as effective bandwidth have been studied extensively to solve this problem. It is important that under WFQ-P QoS to both EF and AF classes can be guaranteed by assigning certain bandwidth.

AF services under WFQ-P are provided by fair queueing, as a result, their QoS are guaranteed by satisfying the bandwidth requirements. EF and BE traffic share the same amount of bandwidth under WFQ-P. According to equation (2), QoS to EF traffic is always guaranteed by allocating the bandwidth equal or large than r_{res}^{EF}. If EF traffic is not very bursty and r_{avg}^{EF} + $r^{BE} > r_{res}^{EF}$, since EF traffic has the priority in using a larger amount of bandwidth $r_{avg}^{EF} + r^{BE}$, QoS to EF class will be better than required. There is no QoS guarantees to BE traffic except for a minimum throughput. However, there is no explicit definition in what time period the throughput should be measured. Since applications using BE service can tolerate a long delay, it is the throughput guarantee under a relatively long time period that matters.

1526-7539/04 $20.00 © 2004 IEEE

Because EF traffic is regulated at the edge of DiffServ network, the time EF traffic can be sent at a rate higher than its average rate is limited. BE traffic under WFQ-P will not be starved. According to equation (2), BE traffic use more bandwidth when EF traffic rate is lower than its average rate. BE traffic will be guaranteed the minimum throughput r^{BE} in average. Under the case that EF traffic is extremely bursty and the bandwidth is allocated as $r_{res}^{EF} > r_{avg}^{EF} + r^{BE}$, the BE traffic will have a bigger average throughput than r^{BE}.

The short-term service impact caused by bursty EF traffic is unavoidable under priority queueing that is used in MDRR/LLQ and WFQ-P. The only difference is that under LLQ/MDRR both AF and BE classes are affected while in WFQ-P only BE traffic is affected and more important AF traffic is well protected.

In summary, WFQ-P guarantees QoS to AF and EF classes by fair queueing. Throughput of BE traffic is guaranteed in the long run by making use of the bandwidth over-provisioned to EF class. The QoS requirements of all DiffServ classes can be met.

In the following section, QoS under WFQ-P is tested by simulation and compared to QoS under Cisco MDRR and DWFQ (Dynamic Weighted Fair Qeueuing).

5. Simulation study

The simulations are based on a model running on OPNET simulator. We implement the "single bottleneck" topology to compare the performance of different schedulers on a single core router. The network topology is shown in Figure 3. There are 4 source nodes generating traffic to the server. The link between the scheduler and server are the only bottleneck link. All Other links have enough bandwidth. The schedulers under test are implemented in the scheduler node to assign the bandwidth on the bottleneck link.

EF traffic is generated by source 1 using aggregated ON_OFF traffic. Each ON_OFF traffic source emulates a voice connection. It generates packets at a peak rate of 64 kbps. The lengths of ON and OFF period are exponentially distributed with average length of 35% and 65% of the total connection time respectively. The traffic generated by a single ON-OFF source is bursty. The aggregate traffic of N ON_OFF sources become less bursty as N goes larger and tends to Gaussian traffic when N becomes very large according to the Large Number Theory. In our simulation, to simulate bursty EF traffic, at most 7 ON_OFF sources are used. Source 2-4 generate Possion traffic for AF and EF classes. In our simulations all the packets have the same size as an ATM cell (53bytes). The bottleneck link has a speed of sending 2000 cells per second. Since the objective of the simulations is to compare the performances of different schedulers, the absolute values of link speed or packet size or traffic sources rates will not affect our results as long as they are set to the same for all schedulers.

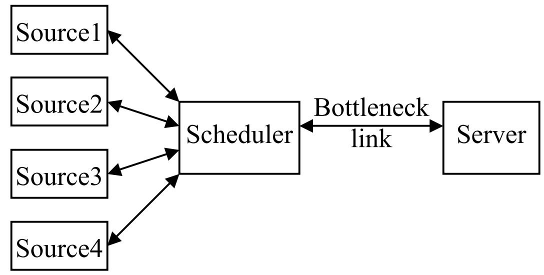

Figure 3. Simulation network topology

5.1 Experiment 1

The objective of this experiment is to test the QoS to AF classes when the EF traffic is bursty under Cisco MDRR. In the experiment, EF traffic is generated using N ON_OFF traffic sources. We fix the allocable bandwidth to AF and BE classes to be 80% of the link rates as N varies from 4 to 7. Both BE and AF traffic are Poisson traffic and they share the rest of bandwidth (80% in the simulations) that is not used by EF traffic. By changing the N we change the over-provisioning ratio to EF traffic and can compare the QoS of AF traffic under different EF traffic loads. The parameters of the experiment are shown in table 1. With a buffer size limit set to 20 packets, AF traffic will have a delay bound of 0.05 seconds if 80% bandwidth is always guaranteed to AF and BE classes as under a ideal GPS scheduler. Therefore, we use the delay bound 0.05 second as the benchmark to measure the QoS to AF classes under Cisco MDRR.

Table 1. Parameters in experiment 1

DS Class	Traffic Characteristics	Bandwidth Allocation	Buffer Size
EF	4-7 ON_OFF sources	Priority to use all bandwidth	20 cells
BE	Exp (average 800 cells/sec)	50% of the rest	100 cells
AF1x	Exp (average 400 cells/sec)	25% of the rest	20 cells
AF2x	Exp (average 400 cells/sec)	25% of the rest	20 cells

The results are shown in table 2. Since AF1 and AF2 are identical, we only show the results of the packet delays of AF1 traffic. Packets with delays longer than 0.05 seconds are considered to have violated the delay bound and received a bad QoS.

1526-7539/04 $20.00 © 2004 IEEE

From the results we can see that priority queueing cannot guarantee the QoS to low priority traffic if the high priority traffic is bursty. For example, under 5 ON_OFF sources, although the bandwidth is over-provisioned (152.4%) to EF classes, 2% AF1 packets still violate their delay bound. This is because due to lack of service isolation, 20% of total bandwidth to AF1 class cannot be guaranteed when burst of EF traffic comes. It is hard to provided QoS guarantees to AF classes and at the same time maintain high link utilization when EF traffic is bursty.

Table 2. Simulation result of experiment 1

N	Over-provisioning ratio (EF bandwidth / EF average rate)	AF delay violation ratio
4	190.5%	0.7%
5	152.4%	2%
6	127.0%	6%
7	108.8%	8%

5.2 Experiment 2

The objective of experiment 2 is to test the performance of WFQ-P. As discussed before, the QoS requirements of each DiffServ classes are different. For EF class, low delay service is expected; For AF classes, the delay requirements are not as stringent as EF class but the delays should still be bounded; For BE class, service starvation should be avoided by guaranteeing a minimum throughput. We will evaluate the performance of WFQ-P from the above three perspectives.

The performance of WFQ-P will be compared with the performance of Cisco MDRR and DWFQ. To be fair, under all schedulers, the input traffic, the buffer allocations and queue management schemes are set to the same. Table 3 and 4 show the parameters settings.

In DWFQ, the bandwidths assigned to EF and BE classes are adjustable according to the average queue length of EF class. To provide EF packets small delays, all the bandwidth of BE class will be assigned to EF class if EF average queue length reaches 2 packets. The average queue length is calculated with the exponential weight set to 0.1 to keep up with the traffic dynamics.

QoS to EF traffic in terms of packets delay CDFs is shown in Figure 4. As expected, Cisco MDRR provides the smallest delays to EF class by giving it the priority to use all the bandwidth. As we shall see later, this approach affects QoS to other classes. Although not as small as under MDRR, EF delays under WFQ-P are guaranteed to meet the stringent delay requirement by over provisioning. It is hard to provide QoS guarantees under DWFQ since the bandwidth allocations are dynamic. To keep the delay small, under DWFQ the maximum average queue length threshold should be small and frequent bandwidth adjustments are required. Compared with DWFQ, WFQ-P provides QoS guarantee without any extra computational overheads introduced by the bandwidth reallocations.

Table 3. Parameters in experiment 2

DS Class	Traffic Characteristics	QM	Buffer Size
EF	7 ON_OFF sources	Drop tail	20 cells
BE	Exp (average 800 packets/sec)	Drop tail	100 cells
AF1x	Exp (average 400 packets/sec)	Drop tail	20 cells
AF2x	Exp (average 400 packets/sec)	Drop tail	20 cells

Table 4. Bandwidth allocations in experiment 2

DS Class	MDRR	DWFQ	WFQ-P
EF	100% with high priority	20%-60%	60% with high priority
BE	50% of the rest of bandwidth	40%-0%	60% with low priority
AF1x	25% of the rest of bandwidth	20%	20%
AF2x	25% of the rest of bandwidth	20%	20%

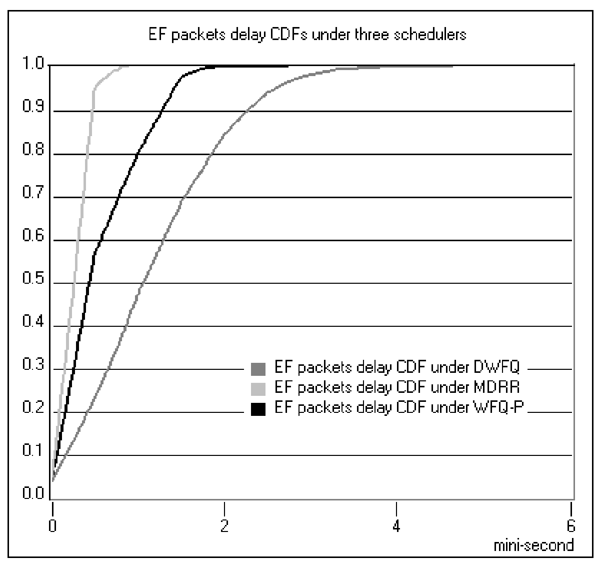

Figure 4. EF QoS under WFQ-P, Cisco MDRR and DWFQ

Figure 5 shows the AF1 packets delay CDFs under three schedulers. The results show that both DWFQ and WFQ-P guaranteed the delay bound (0.05 second) for

AF1 class by allocating the bandwidth using fair queueing. In comparison, there are about 8% of AF1 packets violated their delay bound under Cisco MDRR due to lack of service isolation. Therefore, compared to MDRR, both DWFQ and WFQ-P provide service isolation and guarantee hard QoS to AF classes.

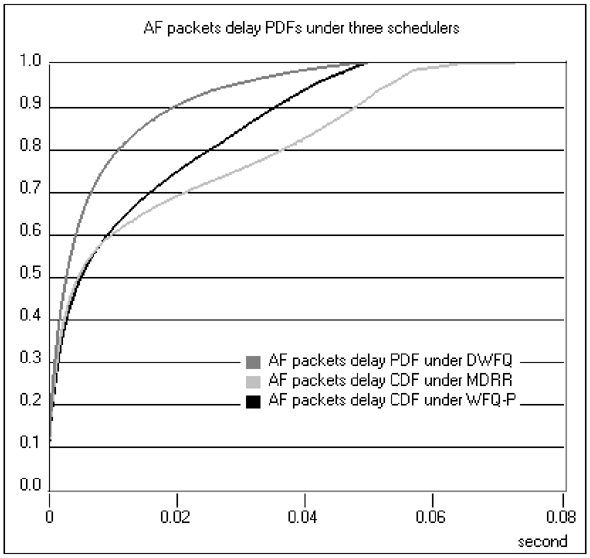

Figure 5. AF QoS under WFQ-P, Cisco MDRR and DWFQ

Figure 6 shows the BE class throughputs for one second time period with 90% confidence interval. The throughputs under all schedulers are close. BE class under Cisco MDRR has a slightly bigger throughput since it is given the same priority as AF classes. When burst of EF traffic arrives, QoS of both BE and AF classes are affected. Under WFQ-P or DWFQ, when burst of EF traffic arrives, AF classes are protected. BE class is the only class that has their QoS affected. This is the reason under the same traffic conditions BE traffic under MDRR has a better throughput. It is important to understand that the slightly better throughput of BE class under MDRR is achieved at the cost of scarifying the QoS to more important AF classes.

Intuitively, One may think that BE throughput under WFQ-P will be lower than under DWFQ since there is no bandwidth guarantee to BE class under WFQ-P while bandwidth is dynamically assigned to BE class under DWFQ. However, the simulation result is quite opposite. BE throughput under WFQ-P is actually higher. This is because of the hierarchical structure of WFQ-P. When bandwidth assigned to EF class is not used, it is always assigned to BE class first. Although no bandwidth is guaranteed to BE class, total bandwidth is guaranteed to BE and EF classes. In DWFQ, things are different. If EF traffic cannot use up its assigned bandwidth, the spare bandwidth will be reallocated among all service classes. Unfortunately, this is not a rare case under DWFQ. The bandwidth requirements of the next phase are predicted based on the current average queue length of EF class. The prediction may not be accurate. As a result, the chance that EF class cannot use up its allocated bandwidth is high. At the same time, when large bandwidth is allocated and not used by EF class, BE class is assigned a small bandwidth and has to compete the spare bandwidth with AF classes. In average, BE class may get a lower bandwidth and AF classes may get a higher bandwidth under DWFQ compared with WFQ-P. Consequently the throughput of BE class under DWFQ can be lower than under WFQ-P. This also explains in Figure 5, under DWFQ most of the AF packets have smaller delays than under WFQ-P although they are all guaranteed the same delay bound.

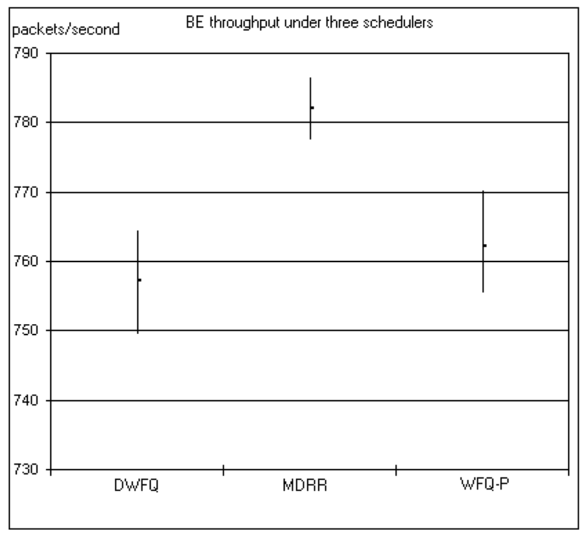

Figure 6. Throughput of BE traffic under WFQ-P, Cisco MDRR and DWFQ

In conclusion, under the same traffic and link utilization ratio, WFQ-P scheduler is best among the three schedulers to satisfy the QoS requirements of all EF, AF and BE classes.

6. Conclusion and future research areas

In this paper, we proposed a new scheduler, namely WFQ-P, to support DiffServ in a MPLS core router. Through both theoretical analysis and simulation study, we demonstrated the proposed WFQ-P scheduler has the following advantages:

1) It can improve bandwidth utilization under bursty EF traffic by allowing EF and BE traffic to share the same bandwidth.

2) QoS guarantees to all EF, AF and BE classes can be provided by a simple bandwidth allocation scheme.
3) WFQ-P can be used easily to perform bandwidth management and connection admission control to facilitate MPLS traffic engineering.

Currently we only studied the performance of WFQ-P in a single DiffServ node case. Its performance in a MPLS DiffServ network will be studied in the future. It is highly expected that with the bandwidth allocation methods such as effective bandwidth, WFQ-P can improve the network resource utilization ratio while providing end-to-end QoS guarantees to all DiffServ classes.

7. References

[1] Blake, S., Black, D., Carlson, M., Davies, E., Zhang, Z., Weiss, W., "An Architecture for Differentiated Services", *IETF RFC 2475,* 1998.

[2] Jacobson, V., Nichols, K., Poduri, K., "An Expedited Forwarding PHB", *IETF RFC 259,* 1999.

[3] Heinanen, J., Baker, F., Weiss, W., Wroclawski, J., "Assured Forwarding PHB Group", *IETF RFC 2597*, 1999.

[4] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., Heinanen, J., "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", *IETF RFC 3270*, 2002.

[5] Parekh, A.K., Gallager, R.G., "A generalized processor sharing approach to flow control in integrated service networks: the single node case", *IEEE/ACM Transactions on Networking*, 1993, pp. 344 – 357.

[6] Parekh, A.K., Gallager, R.G., "A generalized processor sharing approach to flow control in integrated services networks: the multiple node case", *IEEE/ACM Transactions on Networking*, 1994, Vol.2, Issue2, pp.137 – 150.

[7] Stiliadis, D., Varma, A., "Rate-proportional servers: a design methodology for fair queueing algorithms", *IEEE/ACM Transactions on Networking*, 1998, Vol.6, Issue:2, pp.164 – 174.

[8] Stiliadis, D., Varma, A., "Efficient fair queueing algorithms for packet-switched networks". *IEEE/ACM Transactions on Networking,* 1998, Vol.6, Issue:2, pp.175 – 185.

[9] Mao, J., Moh, W.M., Wei, B., "PQWRR scheduling algorithm in supporting of DiffServ", *IEEE International Conference on Communications,* 2001, Vol.3, pp.679 – 684.

[10] Wang, H., Shen, C., Shin, K.G., "Adaptive-weighted packet scheduling for premium service", *IEEE International Conference on Communications,* 2001, Vol.6, pp.1846 – 1850.

[11] Oottamakorn, C., Bushmitch, D., "A Diffserv measurement-based admission control utilizing effective envelopes and service curves", *IEEE International Conference on Communications*, 2001, Vol.4, pp.1187 – 1195.

[12] Bianchi, G., Blefari-Melazzi, N., "Admission control over assured forwarding PHBs: a way to provide service accuracy in a DiffServ framework", *Global Telecommunications Conference*, 2001, Vol.4, pp.2561-2565.

[13] Kelly, F., "Notes on effective bandwidths", *Stochastic Networks: Theory and Applications*, Clarendon Press, Oxford. 1996, pp.141-168.

[14] Norros, I., "On the use of fractional Brownian motion in the theory of connectionless networks", *IEEE Journal on Selected Areas in Communications*, 1995, Vol.13, No.6, pp. 953-962.

[15] Elwalid, A., Mitra, D., "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks", *IEEE/ACM Transactions on Networking*, 1993, Vol. 1, No.3, pp. 329-343.

[16] Chang, C.S., Thomas, J.A., "Effective bandwidth in high-speed digital networks", *IEEE Journal on Selected Areas in Communications*, 1995, Vol. 13, No.6, pp. 1091-1010.