Packet Scheduling with Buffer Management for Fair Bandwidth Sharing and Delay Differentiation

Dennis Ippoliti and Xiaobo Zhou*
Department of Computer Science
University of Colorado at Colorado Springs
Colorado Springs, CO 80918

Liqiang Zhang
Dept. of Computer and Information Sciences
Indiana University South Bend
South Bend, IN 46615

Abstract— **Packet delay and bandwidth are two important metrics for measuring quality of service (QoS) of Internet services. Traditionally, packet delay differentiation and fair bandwidth sharing are studied separately. In this paper, we first propose a generalized model for providing fair bandwidth sharing with delay differentiation, namely FBS-DD. It essentially aims to provide multi-dimensional proportional differentiation with respect to both QoS metrics at the same time. We design packet scheduling schemes that take both packet delay and packet size into considerations, without assuming admission control. Furthermore, we propose a control-theoretic buffer management scheme. The packet scheduling with buffer management approach provides delay and bandwidth differentiation in an integrated way, while existing approaches consider delay and loss rate differentiation as orthogonal issues. It enhances the flexibility of network resource management and multi-dimensional QoS provisioning. It is capable of self-adapting to varying workloads from different classes, which automatically builds a firewall around aggressive clients and hence protects network resources from saturation. Simulation results by the use of trace files demonstrate that the approach can provide predictable fair bandwidth sharing with delay differentiation at various situations. The control-theoretic buffer management scheme improves the controllability.**

I. INTRODUCTION

Differentiated Services (DiffServ) is one of the recent major efforts to meet the demand of provisioning different levels of quality of service (QoS) on the Internet. It aims to provide differentiated services between classes of aggregated traffic flows within a router, rather than offer QoS guarantees to individual flows. To receive different levels of QoS, packets are assigned with different service types or traffic classes at the network edges. DiffServ-compatible routers in the network core perform stateless prioritized packet forwarding or dropping, called "per-hop behaviors" (PHBs), to the classified packets. Due to its stateless processing, the DiffServ architecture exhibits good scalability. It is an active research topic.

The proportional differentiation model [2] is a popular DiffServ model. It aims to provide per-class service quality level in proportion to the pre-specified differentiation parameters of the classes, independent of those class workloads. Delay and bandwidth are two important QoS metrics considered in the model. The representative algorithms for proportional delay differentiation (PDD) consider lossless and work-conserving packet scheduling [2], [3], [4], [5], [7]. When the overall workload of classes exceeds the link bandwidth capacity, the algorithms for bandwidth differentiation aim to enforce that the ratio of the loss rates of two classes be proportional to the ratio of their differentiation parameters [1], [5]. However, most of those algorithms consider delay differentiation and bandwidth differentiation as orthogonal issues.

While the proportional differentiation model is popular due to its proportionality fairness to clients, it is insufficient and might be unfair from the perspective of the network resource providers. It is because the model does not consider another important issue, i.e., fair bandwidth sharing. Fair bandwidth sharing is a classic issue. Its short-term behaviors were originally studied as fair queueing. While those PDD algorithms can ensure that experienced delay of different classes be proportional, there is no assumption nor guarantee on the fair bandwidth sharing, be in short term or in long term.

Consider two traffic classes (Class-1 and Class-2) with the pre-specified differentiation parameters 2 and 1, respectively. Consider the scenario that Class-1's workload is 80% of the link capacity and Class-2's workload is 5% of the link capacity. According to the proportional delay differentiation model, the ratio of the average packet delay of Class-1 to that of Class-2 would be 1 to 2. However, the workload of Class-1 is 16 times of that of Class-2 while their differentiation parameter ratio is only 2 to 1. The scenario illustrates that the current workload-independent differentiation model can be very unfair to some network traffic. Even worse, some aggressive or malicious clients can utilize this unfairness and weak controllability to saturate network resources.

Note that we do not intend to deny the merit of the proportional differentiation model. Essentially, it considers the single-dimensional QoS provisioning with either delay or bandwidth. It needs the support of admission control schemes that shape the traffic according to the service level agreements or some adaptive schemes that promote the differentiation parameters dynamically according to the workload conditions. Generally, the pre-specified differentiation parameters are used by the network operators to control the quality spacings between the multiple classes. They are often associated to the differentiated pricing, say proportionally. But the model is insufficient when multi-dimensional QoS should be considered.

---

[1] Corresponding author: Xiaobo Zhou, email: zbo@cs.uccs.edu.

1-4244-1251-X/07/$25.00 2007 IEEE.

Given that both bandwidth and delay are important QoS metrics, we propose a generalized model, namely FBS-DD, for providing fair bandwidth sharing with delay differentiation at the same time. It is to ensure that the ratio of the average delay of two classes normalized by their experienced bandwidth be proportional to the pre-specified differentiation parameters. It aims to provide multi-dimensional proportional differentiation with respect to both QoS metrics. One uniqueness is that the delay differentiation and loss rate differentiation are integrated with traffic policing capabilities for providing better controllability to network operators and more fairness to clients.

We design packet scheduling algorithms for FBS-DD provisioning. Two VPS (various packet size) schemes take both packet delay and packet size into scheduling considerations. We study the impact of packet size distributions on FBS-DD provisioning. Simulation results by the use of IP trace files show that the schemes are capable of self-adapting to varying workloads of different classes. They automatically build a firewall around aggressive clients and protect network resources from saturation. We further study the performance controllability. When the overall workload of the classes is below the link capacity, the FBS-DD model actually is to achieve the proportional delay differentiation weighted by the throughput of the classes. When the overall workload of classes is beyond the link capacity so that there will be packet loss, the FBS-DD model is to achieve the proportional delay differentiation weighted by the experienced bandwidth ratio of the classes. This is however a non-trivial issue. We propose a control-theoretic buffer management scheme. It enhances the controllability of network resource management.

Our work is to address the integration of traffic policing with proportional differentiation. In Section II, we review existing packet scheduling algorithms for PDD provisioning. Section III presents the FBS-DD model with packet scheduling and buffer management schemes. Section IV focuses on the performance evaluation. Section V concludes the paper.

II. RELATED WORK

Fair bandwidth sharing was initially studied as fair queueing, which aims to allow each flow passing through a network device to have a fair share of network resources. There are classic mechanisms for achieving the short-term per-flow fairing sharing, see PGPS [9] for an example. In the context of DiffServ, the QoS provisioning is concerned with per-class behaviors. The FBS-DD model considers the long-term fair bandwidth sharing with delay differentiation.

Delay differentiation in packet networks is an active research topic. The PDD model is to provide differentiated delay services among traffic classes [2], [3]. A class is assigned a delay differentiation parameter. The packet scheduler of a router aims to keep the ratio of average delay of a higher priority class to that of a lower priority class equal to the pre-specified value. The existing PDD algorithms can be classified into three categories, rate based, time-dependent priority based, and Little's law based. Rate based packet scheduling algorithms adjust service rate allocations of classes dynamically to meet the proportional delay differentiation constraints; see BPR [2] and JoBS [5] for representatives. Time-dependent priority based packet scheduling algorithms adjust the priority of a backlogged class according to the experienced delay of its head-of-line packet; see WTP [3] and AWTP [4] for representatives. Little's law based packet scheduling algorithms correlate the average queue length to the average arrival rate and the average queueing delay of packets. They control the actual delay ratio between two different classes by equalizing their normalized queue lengths with the pre-specified delay differentiation parameters; see PAD, HPD [3], MDP [7], and LAD [10] for representatives.

There are a number of interesting differentiated buffer management and packet dropping schemes for loss rate differentiation. PLR droppers in [1] aim to provide proportional loss rates to different traffic classes according to their differentiation weights. JoBS in [5] extends the proportional loss rate model by providing both absolute loss and delay guarantees and proportional differentiations. HPPD in [11] aims to reduce the retransmission cost of the dropped packets for congestion mitigation by hop-count based differentiated dropping.

Those dropping schemes are able to achieve different differentiation objectives. But the schemes, exception of JoBS [5], consider delay differentiation and bandwidth differentiation as orthogonal issues. There are two significant differences between JoBS and our work. JoBS's goal is to support both absolute and relative DiffServ. It executes an optimization on every packet arrival. Ours is to provide multi-dimensional QoS with respect to both bandwidth sharing and delay differentiation. Packet scheduling is lightweight. Second, JoBS is rate based, which assumes a fluid-flow interpretation of traffic. It hence needs the support of scheduling algorithms that closely approximate the fluid-flow schedulers with rate guarantees. Our work follows a practical per-packet scheduling discipline. Furthermore, it adopts a unique control-theoretic buffer management scheme for the controllability improvement.

III. MODELING AND ALGORITHMS

We consider a lossless and work-conserving packet scheduler that serves N queues, one for each class. For two classes i and j, the FBS-DD model is to maintain the multi-dimensional QoS spacing of two classes with respect to their delay ratio (D_i/D_j) normalized by their bandwidth sharing ratio (B_i/B_j) be proportional to their pre-specified differentiation parameters δ_i and δ_j. That is,

$$\frac{D_i(T,T+t)}{D_j(T,T+t)} \cdot \frac{B_j(T,T+t)}{B_i(T,T+t)} = \frac{\delta_i}{\delta_j}, 1 \le i,j \le N \quad (1)$$

for time intervals $(T, T+t)$ where t is the monitoring timescale. Note that the lower average delay or higher bandwidth sharing represents higher QoS. FBS-DD is essentially a fair tradeoff between PDD provisioning and FBS provisioning. When two classes experience the same bandwidth, the model is reduced to PDD model. When two classes experience the same average delay, the model is reduced to FBS model.

1-4244-1251-X/07/$25.00 2007 IEEE.

A. Packet Scheduling Schemes

1) VPS-TWP: Throughput normalized waiting time priority scheduling: We consider the general case, that is, packets from a class have various sizes and different classes may have different packet size distributions. We revisit the time dependent priority scheduling discipline and design the VPS-TWP scheme, tailored from WTP [3]. The time dependent priority scheduling was initially studied in queueing systems. It was later studied by Dovrolis *et al.* for PDD provisioning [3]. We describe VPS-TWP as the throughput normalized waiting time priority scheduling for FBS-DD provisioning by further taking the packet size into considerations.

At the beginning of scheduling, $TWP_i = \infty$ for $1 \leq i \leq N$. Suppose that class i is backlogged at time t, $s_i(t)$ is the size of the packet at the head of the class i at t, and $w_i(t)$ is the the waiting time of the packet at the head of the class i at t. We define the throughput normalized head waiting time of class i at t as

$$TWP_i(t) = \frac{w_i(t)}{\delta_i s_i(t)}. \tag{2}$$

Every time a packet is to be transmitted, the VPS-TWP scheduler selects the backlogged class j with the maximum throughput normalized head waiting time,

$$j = arg \quad max_{i \in G(t)} TWP_i(t), \tag{3}$$

where $G(t)$ is the set of backlogged classes at time t. Tie is broken by the use of the differentiation priority. The throughput of class j is increased by the size of the transmitted packet. Its throughput normalized head waiting time will be minimized as its packet delay will not increase any more. VPS-TWP attempts to minimize the differences between the bandwidth normalized waiting times of successively departing packets. It essentially aims to achieve instantaneous FBS-DD.

2) VPS-TAD: Throughput normalized average delay scheduling: While VPS-TWP focuses at the instantaneous behavior, we propose the VPS-TAD scheme which focuses on the long-term behavior. (1) can be rewritten as

$$\frac{D_i(T, T+t)}{\delta_i B_i(T, T+t)} = \frac{D_j(T, T+t)}{\delta_j B_j(T, T+t)}. \tag{4}$$

That is, the throughput normalized average delay (TAD) factor must be equal in all classes, i.e., $TAD_i = TAD_j$. Note that given a same interval, bandwidth sharing ratio of two classes is the same as the throughput ratio. The VPS-TAD scheme, tailored from PAD [3], aims to equalize the throughput normalized average delays among all classes so as to achieve the FBS-DD goal.

Let $G(t)$ is the set of backlogged classes at time t, $L_i(t)$ be the sequence of class i packets that were transmitted during the interval $(T, T+t)$, d_i^m be the delay of the *m*th packet in $L_i(t)$, and s_i^m be the size of the *m*th packet in $L_i(t)$. Assuming that there was at least one packet transmitted from class i during interval $(T, T+t)$, the throughput normalized average delay of class i at t is

$$TAD_i(t) = \frac{D_i(T, T+t)}{\delta_i B_i(T, T+t)} = \frac{1}{\delta_i \sum_{m=1}^{|L_i(t)|} s_i^m} \frac{\sum_{m=1}^{|L_i(t)|} d_i^m}{\mid L_i(t) \mid} \tag{5}$$

where $\mid L_i(t) \mid$ is the number of packets in $L_i(t)$.

At the beginning of scheduling, $TAD_i = \infty$ for $1 \leq i \leq N$. Suppose that a packet is to be transmitted at time t. VPS-TAD selects the backlogged class j with the maximum bandwidth normalized average delay,

$$j = arg \quad max_{i \in G(t)} TAD_i(t). \tag{6}$$

Tie is broken by the differentiation priority. The rationale of VPS-TAD is that each time a packet from class j is transmitted, its throughput normalized average delay decreases. This is because its throughput increases by the size of the transmitted packet. The delay of that transmitted packet will not increase any more, and thus the increase to the average packet delay will be minimized. VPS-TAD therefore attempts to minimize the differences between the throughput normalized average delay of classes. It essentially aims to achieve FBS-DD in the long term. It, however, needs to maintain the state information about the current throughput and average delay per each class.

B. PID Control-theoretic Buffer Management

When the overall workload is greater than the link capacity, packet loss is inevitable and loss rate becomes the dominant QoS metric. The proposed packet scheduling schemes, however, have no control over the loss rate differentiation between classes. We propose a control-theoretic buffer management scheme, to be integrated with the packet scheduling schemes, for the FBS-DD provisioning and proportional loss rate differentiation at the same time. One nice feature of the buffer management based approach is that the packets will be dropped from the tail due to the buffer overflow. This avoids the packet pushout issue and facilitates the packet ordering.

The buffer management is to dynamically allocate the buffer space into a number of virtual mini-buffers, one mini-buffer for one class. The size of a mini-buffer directly affects a class's loss rate. We propose to use a proportional integral derivative (PID) feedback controller to adjust the buffer allocation. Let l_i be the loss rate of class i. The goal is to ensure that the observed relative loss rate l_i be proportional to the pre-specified QoS parameter δ_i, that is, $l_i/l_j = \delta_i/\delta_j$. Let L_i be the relative loss rate ratio of class i, that is, $L_i = \frac{l_i}{l_1 + l_2 + \cdots + l_n}$. Let L_i^d be the desired relative loss rate ratio of class i, that is, $L_i^d = \frac{\delta_i}{\delta_1 + \delta_2 + \cdots + \delta_n}$. During the k_{th} sampling period, the relative error is calculated as difference between the desired value and the observed value, that is,

$$e_i(k) = L_i^d(k) - L_i(k). \tag{7}$$

One property of the model is the sum of the relative errors is always zero since $\sum_{i=1}^{n} e_i(k) = \sum_{i=1}^{n} (L_i^d(k) - L_i(k)) = 0$. This important property makes it feasible for us to adaptively adjust the buffer allocation for a class independent of the adjustments of other classes while maintaining a constant overall buffer size [6].

The buffer size allocated to a class is adjusted in proportion to the error between the desired relative loss rate ratio and the observed one. Specifically, the operation of the PID controller is described as follows:

$$s_i(k+1) = s_i(0)+G_P e_i(k)+G_I \sum_{j=0}^{k-1} e_i(j)+G_D \Delta e_i(k). \quad (8)$$

$s_i(k+1)$ denotes the buffer size allocated to class i in the new sampling period. $s_i(0)$ denotes the initial buffer size allocated. The three terms added to $s_i(0)$ denote proportional, integral, and derivative components, respectively. Setting a large proportional feedback gain (G_P) typically leads to faster response at the cost of increasing system instability. The integral controller (G_I) can eliminate the steady-state error and avoid over-reactions to measurement noises. The derivative control (G_D) considers the change of errors in adjusting the buffer size allocation and hence responds fast to errors. The derivative error with class i is calculated as

$$\Delta e_i(k) = e_i(k) - e_i(k-1). \quad (9)$$

IV. PERFORMANCE EVALUATION

We developed a simulator to study the performance of the packet scheduling schemes and the feedback control based buffer management. Due to the space limitation, we report the two-class experimental results. Note that the number of classes for DiffServ often varies from 2 to 3 [3], [5], [12]. For the packet size distribution of two classes, we used two Bell Labs-I trace files adopted from the National Laboratory for Applied Network Research [8]. The characteristics and the packet size distributions for the traces are illustrated in Figure 1. Without loss of generality, let Class-1 be the high priority class and Class -2 be the low priority class.

A. Performance of the packet scheduling algorithms

The first set of experiments is to study the impact of the packet scheduling schemes on FBS-DD provisioning when the overall workload is within the link capacity. We considered a lossless model. Figure 2 shows the performance of the packet scheduling schemes. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) is set to 1:2 and their workload ratio is set to 1:3. Figure 2(a) shows the achieved FBS-DD ratio with its 95th and 5th percentiles when the overall workload changes from 55% to 100% of the link capacity. Figure 2(b) shows the achieved delay ratio with its 95th and 5th percentiles. The results show that the scheduling schemes can achieve the goal of providing fair bandwidth sharing with delay differentiation when the overall workload is greater than 60%. But the variance, as demonstrated by the 95th and 5th percentiles, is a nontrivial issue. It is due to the variance of the packet size distributions and the inter-arrivals. When the workload is light, there is a feasibility issue with the packet scheduling for service differentiation provisioning [3].

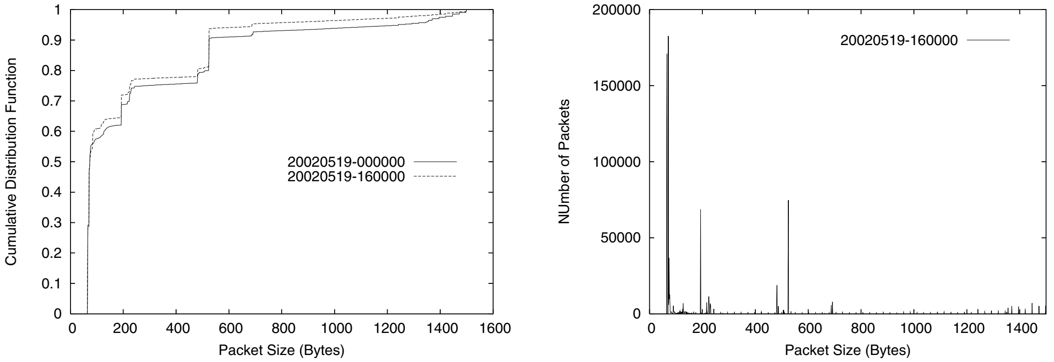
We next change the ratio of $\delta_1 : \delta_2$ to 1:3. We fix the overall workload to 80% and vary the Class-1's workload from 10% to 90% of the overall workload. Figure 3(a) shows the achieved FBS-DD ratio with its 95th and 5th percentiles. It shows the FBS-DD ratio can be achieved as expected. But the variance is high when Class-1's workload deviates from the middle value 50%. This is due to the fact that there are too few or too many packets from Class-1, limiting the capability of the packet scheduling schemes. Figure 3(b) shows the achieved delay ratio with its 95th and 5th percentiles. We can see that the proposed VPS scheduling schemes can achieve the fair bandwidth sharing with delay differentiation when the workload percentage of the classes changes dynamically. When the Class-1 contributes 75% of the overall workload, the delay ratio of two classes becomes 1. This demonstrates the benefit of the FBS-DD model that can make adaptive tradeoff between delay differentiation and fair bandwidth sharing.

While both VPS-TAD and VPS-TWP schemes can achieve FBS-DD provisioning from the long-term perspective, they have different behaviors. Figure 4 shows the FBS-DD ratios achieved by the scheduling schemes in different sampling intervals. Interestingly, in short sampling intervals, VPS-TAD does not perform well for FBS-DD provisioning. Figure 4(a) shows that its performance improves as the sampling interval increases. This is explained by the fact that VPS-TAD takes into account the average of a number of packets in the interval. It aims to minimize the differences between the normalized average class delays and thus its performance improves as the sampling interval increases. On the other hand, Figure 4(b) shows that VPS-TWP achieves desirable FBS-DD ratios when the sampling interval is short and the performance deteriorates as the interval increases. This is due to the fact that VPS-TWP attempts to minimize the differences between the normalized head waiting times. Essentially, it aims to achieve the instantaneous FBS-DD provisioning.

B. Performance of the PID control based buffer management

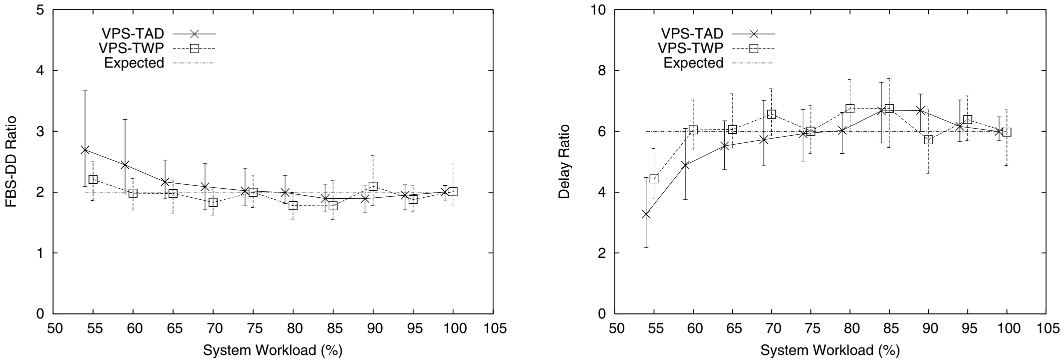
Previous experimental results have shown that the packet scheduling schemes can achieve the fair bandwidth sharing and delay differentiation at the same time. But when the overall workload is beyond the link capacity, there will be packet loss and the packet scheduling schemes have no control over the loss rate differentiation between classes. Figure 5 depicts the impact of the PID control-theoretic buffer management on loss rate differentiation and the controllability of FBS-DD provisioning. The overall workload is set to 150% of the link capacity. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) is set to $1 : 3$. Figure 5(a) shows the impact of the PID control-theoretic buffer management on the proportional loss rate differentiation. It shows that with the buffer management, the loss rate ratio of two classes is fairly proportional to the differentiation weight ratio as the percentage of the Class-1's workload changes from 10% to 90% of the overall workload. On the other hand, without the buffer management, both classes experience almost the same loss rate.

Figures 5(b) shows the achieved FBS-DD ratio by the use



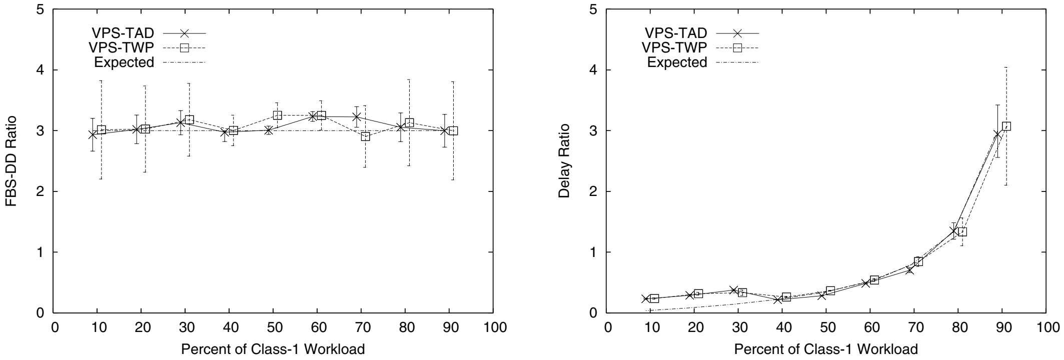
(a) The cumulative distribution. (b) Number of packets for one trace.

Fig. 1. Packet size distributions of two Bell Labs-I IP traces.



(a) FBS-DD ratio of Class-2 to Class-1. (b) Delay ratio of Class-2 to Class-1.

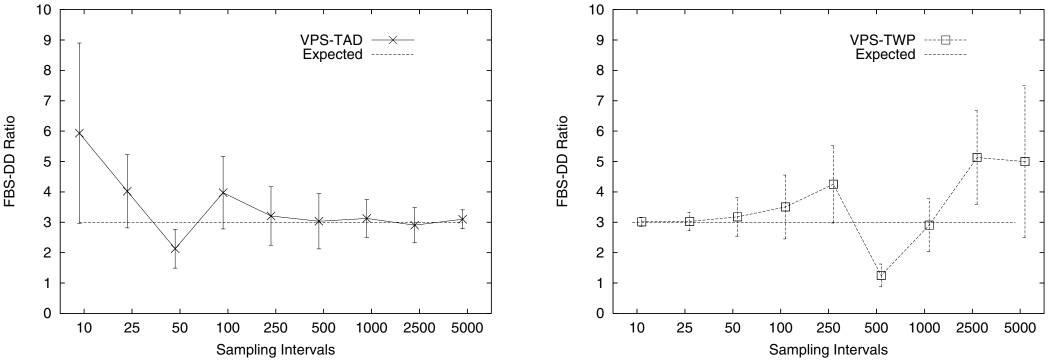
Fig. 2. The performance of VPS packet scheduling schemes when overall workload changes from 55% to 100%.



(a) FBS-DD ratio of Class-2 to Class-1. (b) Delay ratio of Class-1 to Class-2.

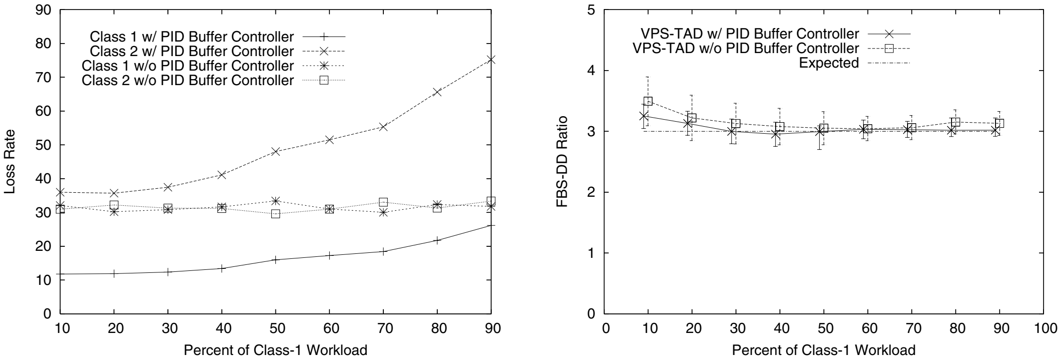
Fig. 3. The performance of VPS packet scheduling schemes when Class-1's workload changes from 10% to 90% of the overall workload.

of VPS-TAD packet scheduling scheme with and without the PID controller based buffermanagement. The results show that with the PID controller based buffer management, the VPS-TAD scheme is able to achieve more consistent and desirable FBS-DD ratios with respect to both the mean and the variance. Without the feedback control based buffer management, the variance of the FBS-DD ratio is higher. One reason is that if there is no feedback based buffer management, a class with some bursty traffic can saturate the buffer easily, leaving little buffer space for another class. The VPS-TAD scheduling schemes aims to minimize the normalized average delays. Its capability is limited by the availability of packets from certain classes for scheduling. This benefits the low-priority but high-workload class. Therefore, the buffer management should be integrated with packet scheduling for controllable FBS-DD provisioning. The integrated approach is capable of self-adapting to varying workloads from different classes, which automatically builds a firewall around aggressive clients and hence protects network resources from saturation.



(a) FBS-DD ratio due to VPS-TAD. (b) FBS-DD ratio due to VPS-TWP.

Fig. 4. The behaviors of VPS scheduling schemes for FBS-DD provisioning in different sampling intervals.



(a) Loss rate due to the buffer management. (b) FBS-DD ratio due to VPS-TAD.

Fig. 5. The impact of the control-theoretic buffer management on FBS-DD provisioning.

V. CONCLUSION

In this paper, we investigated the problem of multi-dimensional QoS differentiation with respect to both packet delay and bandwidth sharing. We proposed a generalized model, FBS-DD, for providing fair bandwidth sharing with delay differentiation at the same time. One uniqueness is that the delay differentiation is integrated with traffic policing capabilities for providing better controllability to network operators and more fairness to clients. We designed packet scheduling schemes that take both packet delay and packet size into considerations, without assuming admission control. We conducted the performance evaluation of the schemes with a wide range of sensitivity analysis. Experiment results by the use of Internet trace files have shown that the packet scheduling schemes are able to achieve the FBS-DD provisioning at different workload conditions. Results have also demonstrated the significance of the feedback control based buffer management on the performance controllability at overload conditions. The integrated approach can significantly enhance the flexibility of network resource management and enable the multi-dimensional QoS provisioning.

ACKNOWLEDGMENT

The authors want to thank the support from the Center for Homeland Security at the U. of Colorado at Colorado Springs.

REFERENCES

[1] C. Dovrolis and P. Ramanathann. Proportional differentiated services, part ii: Loss rate differentiation and packet dropping. In *Proc. of the Int'l Workshop on Quality of Service*, 2000.
[2] C. Dovrolis, D. Stiliadis, and P. Ramanathan. Proportional differentiated services: Delay differentiation and packet scheduling. In *Proc. ACM SIGCOMM*, pages 109–120, 1999.
[3] C. Dovrolis, D. Stiliadis, and P. Ramanathan. Proportional differentiated services: Delay differentiation and packet scheduling. *IEEE/ACM Trans. on Networking*, 10(1):12–26, 2002.
[4] M. K. H. Leung, J. C. S. Lui, and D. K. Y. Yau. Adaptive proportional delay differentiated services: Characterization and performance evaluation. *IEEE/ACM Trans. on Networking*, 9(6):908–817, 2001.
[5] J. Liebeherr and N. Christin. JoBS: Joint buffer management and scheduling for differentiated services. In *Proc. Int'l Workshop on Quality of Service*, pages 404–418, 2001.
[6] Y. Lu, T. F. Abdelzaher, and A. Saxena. Design, implementation, and evaluation of differentiated caching services. *IEEE Trans. on Parallel and Distributed Systems*, 15(5):440–452, 2004.
[7] T. Nandagopal, N. Venkitaraman, R. Sivakumar, and V. Bharghavan. Delay differentiation and adaptation in core stateless networks. In *Proceedings of IEEE INFOCOM*, pages 421–430, 2000.
[8] NLANR. PMA: Special traces archive. http://pma.nlanr.net/Special/.
[9] A. K. Parehk and R. G. Gallager. A generalized processor sharing approach to flow control in Integrated Sevices networks: the single-node case. *IEEE/ACM Trans. on Networking*, 1(3):344–357, 1993.
[10] J. Wei, C.-Z. Xu, and X. Zhou. A robust packet scheduling algorithm for proportional delay differentiation services. In *Proc. of Globecom*, 2004.
[11] X. Zhou, D. Ippoliti, and T. Boult. HPPD: A hop-count based probabilistic packet dropper. In *Proc. IEEE ICC*, pages 116-121, 2006.
[12] X. Zhou, J. Wei, and C.-Z. Xu. Quality-of-service differentiation on the internet: A taxonomy. *Journal of Network and Computer Applications, Elsevier*, 30(1):354–383, 2007.