The Local and Global Effects of Traffic Shaping in the Internet

Massimiliano Marcon
MPI-SWS
mmarcon@mpi-sws.org

Marcel Dischinger
MPI-SWS
mdischin@mpi-sws.org

Krishna P. Gummadi
MPI-SWS
gummadi@mpi-sws.org

Amin Vahdat
UCSD
vahdat@cs.ucsd.edu

***Abstract*—The Internet is witnessing explosive growth in traffic, in large part due to bulk transfers. Delivering such traffic is expensive for ISPs because they pay other ISPs based on peak utilization. To limit costs, many ISPs are deploying ad-hoc traffic shaping policies that specifically target bulk flows. However, there is relatively little understanding today about the effectiveness of different shaping policies at reducing peak loads and what impact these policies have on the performance of bulk transfers.**

In this paper, we compare several traffic shaping policies with respect to (1) the achieved reduction in peak network traffic and (2) the resulting performance loss for bulk transfers. We identified a practical policy that achieves peak traffic reductions of up to 50% with only limited performance loss for bulk transfers. However, we found that the same policy leads to large performance losses for bulk transfers when deployed by multiple ISPs along a networking path. Our analysis revealed that this is caused by certain TCP characteristics and differences in local peak utilization times.

I. INTRODUCTION

The Internet is witnessing explosive growth in demand for bulk content. Examples of bulk content transfers include downloads of music and movie files [13], distribution of large software and games [33], online backups of personal and commercial data [3], and sharing of huge scientific data repositories [32]. Recent studies of Internet traffic in commercial backbones [20] as well as academic [7] and residential [8] access networks show that such bulk transfers account for a large and rapidly growing fraction of bytes transferred across the Internet.

The bandwidth costs of delivering bulk data are substantial. A recent study [22] reported that average monthly wholesale prices for bandwidth vary from \$30,000 per Gbps/month in Europe and North America to \$90,000 in certain parts of Asia and Latin America. The high cost of wide-area network traffic means that increasingly *economic* rather than *physical* constraints limit the performance of many Internet paths. As charging is based on peak bandwidth utilization (typically the 95^{th} percentile over some time period), ISPs are incentivized to keep their bandwidth usage on inter-AS links much lower than the actual physical capacity.

To control their bandwidth costs, ISPs are deploying a variety of ad-hoc traffic shaping policies today. These policies target specifically bulk transfers, because they consume the vast majority of bytes [7, 24, 29]. However, these shaping policies are often blunt and arbitrary. For example, some ISPs limit the aggregate bandwidth consumed by bulk flows to a fixed value, independently of the current level of link utilization [16]. A few ISPs even resort to blocking entire applications [12]. So far, these policies are not supported by an understanding of their economic benefits relative to their negative impact on the performance of bulk transfers, and thus their negative impact on customer satisfaction.

Against this backdrop, this paper poses and answers two questions:

1. What reduction in peak utilization can an ISP achieve by traffic shaping only bulk flows, and how much do such shaping policies penalize bulk flows? Using traces from the access links of 35 universities, we show that diurnal patterns in bandwidth consumption offer an opportunity to significantly reduce the peak bandwidth consumption by only shaping bulk flows. However, we found that naive traffic shaping policies can dramatically increase the completion time of bulk transfers. More intelligent policies combined with simple queueing techniques have the potential to minimize the impact of traffic shaping on bulk transfers while achieving a near-optimal reduction in peak bandwidth.

2. Assuming most ISPs adopted traffic shaping of bulk transfers to reduce bandwidth costs, how would that affect the performance of bulk transfers traversing multiple inter-ISP links? Given the significant reduction in peak bandwidth usage (and thus in costs) that can be achieved with traffic shaping of only bulk flows, it is very likely that most ISPs would adopt such policies eventually. However, we found that even if ISPs deploy policies that are designed to minimize the *local* performance loss of bulk flows, the *global* performance loss of flows traversing multiple traffic shapers is substantial. In our analysis we found that this is caused by TCP characteristics and differences in local peak utilization times of ISPs.

The rest of this paper is structured as follows: Section II describes real-world traffic shaping policies in use today. Section III discusses the goals of an ideal traffic shaping policy. Section IV compares different traffic shaping policies when traffic traverses only one traffic shaper, while Section V analyzes the effects of multiple shapers active on a network path. Finally, Section VI discusses related work and Section VII concludes the paper.

978-1-4244-8953-4/11/\$26.00

II. TRAFFIC SHAPING POLICIES IN USE TODAY

ISPs today deploy a variety of traffic shaping policies. The main goal of these policies is to reduce network congestion and to distribute bandwidth fairly amongst customers [5]. This is typically achieved by reducing the peak network usage through traffic shaping applied either to single flows or to the aggregate traffic of a user. The reduction in peak network usage also has the side-effect that it reduces inter-AS traffic and thus bandwidth costs. At the same time, ISPs are also concerned to affect as few flows as possible to keep the effect on user traffic low [28].

To the best of our knowledge, there exists no previous study that analyzed the exact benefits of these policies and their impact on targeted flows when deployed in practice. In this section, we present three canonical examples of traffic shaping policies in use today. Most of these policies traffic shape bulk transfers [5, 16]. We investigate the benefits of these policies and compare them with more sophisticated policies in later sections.

1. Traffic shaping bulk applications on a per-flow basis. This policy shapes every flow belonging to bulk transfer applications to some fixed bandwidth. For example, Bell Canada revealed that it throttles traffic from P2P file-sharing applications in its broadband access networks to 256 Kbps per flow [5]. Traffic shaping applies to flows both in the downstream and in the upstream direction. Bell Canada chose to traffic shape only P2P file-sharing applications because it found that a small number of users of these applications were responsible for a disproportionate fraction of the total network traffic.

2. Traffic shaping aggregate traffic. Here, traffic shaping is applied to the aggregate traffic produced by multiple network flows. For example, Comcast handles congestion in its access network by throttling users who consume a large portion of their provisioned access bandwidth over a 5-minute time window [11]. All packets from these users are put in a lower priority traffic class in order to be delayed or dropped before other users' traffic in case of network congestion. Another example of such a policy was deployed at the University of Washington in 2002 [16]. The university started limiting the aggregate bandwidth of all incoming peer-to-peer file-sharing traffic to 20 Mbps to reduce the estimated costs of one million dollars that this type of traffic was causing per year.

3. Traffic shaping only at certain times of the day. This policy is orthogonal to the previous two policies and is typically used in combination with these. An ISP can decide to traffic shape throughout the day or restrict traffic shaping to specific time periods. For example, the University of Washington shapes P2P traffic during the entire day [16], while Bell Canada and Kabel Deutschland announced to only traffic shape during periods of "peak usage", i.e., between 4:30 pm and 2:00 am [5, 28]. Since many ISPs pay for transit bandwidth based on their peak load, shaping only during peak usage appears to be an effective way to reduce bandwidth costs.

While the above policies are simple to understand, they raise several questions:

1) How effective are the different traffic shaping policies at reducing network congestion and peak network usage?
2) What is the impact of traffic shaping policies on the performance of the targeted network flows?
3) Are there policies that achieve similar or better reduction in bandwidth costs, while penalizing traffic less?

To answer these questions, we first need to define the precise goals of traffic shaping, as well as the metrics with which we evaluate the impact of traffic shaping policies on network traffic.

III. GOALS AND POTENTIAL OF TRAFFIC SHAPING

In this section, we identify three goals for traffic shaping policies as deployed by ISPs: minimizing the peak network traffic, minimizing the number of flows targeted by traffic shaping, and minimizing the negative impact on these flows. We argue that traffic shaping policies should be designed around these goals, and quantify the potential of such policies through an analysis of real-world network traces.

A. Network traces

In our analysis of traffic shaping performance, we use publicly available NetFlow records collected at the access links of 35 different universities and research institutions. The records contain incoming and outgoing traffic between these universities and the Abilene backbone [1]. The NetFlow records were collected during a 1-week period starting on January 1st 2007, and contain durations and sizes of TCP flows. The NetFlow data has two limitations: (1) long flows are broken down into shorter flows (with a maximum duration of 30 minutes), and (2) flows' packets are sampled with a 1% rate. To recover long flows from the NetFlow data, we combine successive flows between the same TCP endpoints into longer flows using the technique employed in [21]. To account for the sampling rate, we multiply packet and byte counts by 100. While this approach is not reliable when applied to small flows, it was shown to be accurate for large bulk flows [30], which are the object of the traffic shaping policies considered in this paper.

B. Goals and potential

We identify the following three goals as the main practical objectives for an ISP that deploys traffic shaping.

Goal 1: Minimizing the peak network traffic. The main motivation for ISPs to deploy traffic shaping is often network congestion [5, 28]. With traffic shaping, ISPs can lower the risk of congestion by reducing the peak network usage. At the same time, lowering the peak network usage also reduces bandwidth costs for ISPs since they are often charged based on the near-peak utilization (e.g., 95^{th} percentile traffic load) of their links. This creates an incentive for ISPs to keep the peak network usage as low as possible to minimize bandwidth costs. Using our traces, we quantify the maximum peak reduction in

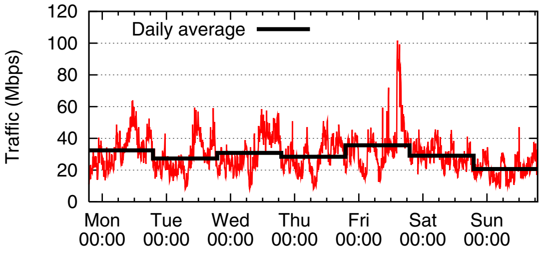


Figure 1: Downstream network traffic at Ohio State University: The traffic shows diurnal variations with large peak-to-average ratios.

network traffic that ISPs can achieve with an optimal traffic shaping policy.

Figure 1 plots the network traffic in one of our traces (collected at the Ohio State university). The traffic exhibits strong diurnal variations, with traffic peaking around noon and dropping in the early morning. As a result of these variations, the daily traffic peak is considerably higher than the average daily traffic. Intuitively, the lower bound for any realistic peak reduction scheme is the average daily traffic, because this is the minimum traffic level that can assure that all traffic will eventually be delivered within the day[1].

Averaging across all access link traces, the daily peak is 2.6 times larger than the average traffic load, while the daily 95^{th} percentile is 1.7 times larger than the average traffic load. These results suggest that traffic shaping has the potential to reduce ISPs' peak load by a factor of two.

Goal 2: Minimizing the number of traffic shaped flows. While ISPs have an economic incentive to reduce the peak network usage as much as possible, they are also concerned with affecting as few flows as possible to keep the effect on user traffic low. As a consequence, most ISPs today target either users that are responsible for a disproportional large fraction of traffic (so-called "heavy-hitters"), or applications known to be bandwidth-hungry (e.g., file-sharing applications). Using our traces, we quantify the minimal fraction of bulk flows that need to be shaped to achieve a near-optimal reduction in peak load.

Typically, an ISP would use deep packet inspection to identify flows belonging to bandwidth-intensive applications. However, since our traces do not contain information about application-level protocols, we identify bandwidth-intensive flows based on their size, i.e. the number of transferred bytes.

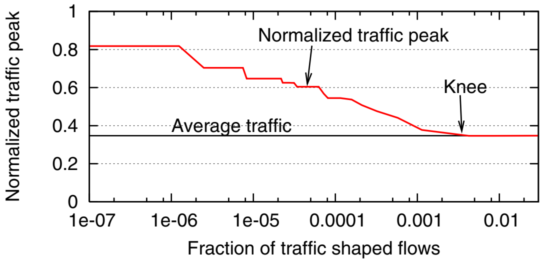


Figure 2: Tradeoff between maximum achievable peak reduction and fraction of traffic shaped flows: Intuitively, shaping more flows lowers the peak. However, the peak cannot be lower than the average traffic rate without dropping flows. At this point, shaping more flows has no further benefits.

We sorted all flows in each of our trace by decreasing size. We then selected all flows larger than a certain size T for traffic shaping and computed the theoretical maximum peak reduction achievable. For this analysis, we assume that flows can be arbitrarily throttled, as long as they complete within the trace's time-frame of one week. We then repeated this for decreasing values of T, thus selecting more and more flows. Figure 2 plots the results for one of our traces. After selecting only 0.4% of the largest flows, the traffic peak reaches the average traffic load and no further reduction is possible (the "knee" in the figure). In this trace, this translates to flows that are larger than 10 MB. Across all traces, traffic shaping less than 4% of the flows is always sufficient to achieve the maximum peak reduction, and in 30 of our 35 traces traffic shaping less than 1% of the flows also suffices. This result suggests that ISPs can considerably reduce their peak while shaping a very small fraction of flows.

Goal 3: Minimizing the delay that traffic shaped flows incur. We found that ISPs have to shape only a small fraction of flows to achieve an optimal reduction in peak network usage. Note that this optimal reduction can be achieved without dropping any flows. Instead, in our analysis, we ensured that all shaped flows complete within the time-frame of the trace. However, even if only a small fraction of flows are affected by traffic shaping, ISPs should try to limit the delay incurred by these flows in order to minimally penalize the applications or users generating the bulk flows. With respect to this goal, focusing on bulk flows has the advantage that these flows, being large, have completion times on the order of minutes, hours or even days. Therefore, they can endure considerable absolute delays without severe damage to their performance. For example, the bulk flows in our trace take on average 3.5 minutes to complete when they are not traffic shaped, suggesting that they can be delayed by seconds without negative effects for applications.

In summary, we found that a traffic shaping policy should not only minimize the peak network traffic, but also affect as few flows as possible and minimize its impact on the shaped flows. In the next section, we compare how well different traffic shaping policies perform relative to these goals.

IV. LOCAL PERFORMANCE OF TRAFFIC SHAPING POLICIES

In this section we analyze how different traffic shaping policies perform based on the three metrics from Section III: the peak reduction, the fraction of shaped flows, and the delay that shaped flows incur. As we only consider a single traffic shaper in the network path here, we call this the local performance of traffic shaping policies. In Section V, we

---

[1] A higher peak reduction is only possible if traffic is dropped from the network, e.g., by blocking certain applications' traffic. However, blocking is a very intrusive form of traffic shaping and ISPs that previously deployed it had to deal with very negative media coverage about this practice [31].

analyze the effect of multiple traffic shapers in the networking path.

A. Selecting flows for traffic shaping

ISPs target only a subset of flows for traffic shaping, typically flows from bandwidth-intensive applications. Doing so, ISPs achieve very good peak reductions while keeping the number of affected flows low. In the following, we call flows that are subject to traffic shaping "low-priority traffic" and the remaining flows "best-effort traffic".

To identify flows from bandwidth-intensive applications, ISPs often employ deep packet inspection (DPI), which is widely available in routers [9] or provided by special DPI equipment [27]. Additionally, today's networking equipment allows ISPs to collect statistics on flow sizes at line speed, which can be used to mark large flows for traffic shaping [9, 19]. In practice, flow classification is implemented at ISPs' ingress routers. Flows are marked as low-priority or best-effort by setting the DSCP field in the IP header[2]. The traffic shaping equipment then selects the packets to traffic shape just based on the value of the DCSP field.

As our traces do not contain information to identify application protocols, we rely on flow sizes instead, i.e., flows that are larger than a certain "flow size threshold" are shaped. Picking the right flow size threshold is nontrivial, because a higher threshold will affect fewer flows, but at the same time will give ISPs fewer bytes to traffic shape, and thus limit its ability to decrease peak usage. To select the right threshold for each trace, we use the analysis from Section III-B and pick the threshold that results in the maximum potential for peak reduction with the minimum fraction of flows being shaped.

In all traffic shaping policies in this section, unless explicitly stated otherwise, we keep a running counter of the bytes sent by each network flow, and use its value to classify the flow. For example, if the flow size threshold is 10 MB, a 20 MB flow will send the first 10 MB as best-effort traffic. After that, the flow is classified as low-priority traffic and the remaining 10 MB of the flow are traffic shaped. This technique can also be used by ISPs to deploy a protocol-agnostic traffic shaping policy that targets all flows larger than certain flow size threshold. Modern traffic shaping equipment can sample packets to keep accurate per-flow byte counts even on high-speed links [9], and some recent sampling techniques enable identification of large flows with high accuracy and low memory requirements [14].

B. Selecting aggregate bandwidth limits

Some traffic shaping policies (e.g., as used by the University of Washington [16]) shape low-priority flows only when the traffic rate exceeds a certain "bandwidth limit". This limit can refer to the aggregate traffic (best-effort + low-priority traffic) or to the low-priority traffic only. For example, an ISP could traffic shape only when the total traffic rate exceeds 20 Mbps or when the low-priority traffic alone exceeds 20 Mbps.

The bandwidth limit determines the total reduction in traffic peak. As we showed in Section III, the average traffic rate is the minimum value that enables delivery of all traffic. Therefore, in all policies that use a bandwidth limit, we set the bandwidth limit to the average traffic rate of the previous day plus 5% to account for small increases in demand. We found that this approach works well in practice because the average rate is quite stable across days. In fact, in our 35 one week traces, we found only two days were this was not the case, i.e., the average traffic varied considerably from one day to the next. If there is a sudden increase in daily average traffic, too many low-priority flows may compete for too little bandwidth, thus incurring large delays or even starvation. To overcome this problem, ISPs can monitor the number of low-priority flows and the overall traffic in their network and increase the bandwidth limit if they detect a significant difference from the previous day.

C. Traffic shaping policies

We now describe the traffic shaping policies we evaluate. All of the traffic shaping policies described here can be implemented using well-known elements like token buckets, class-based rate limiting, and strict priority queuing, available in today's networking equipment [10, 23]. To design the traffic shaping policies we start from the real-world examples from Section II and develop more complex policies that attempt to reduce the peak traffic while minimize the delay incurred by the traffic shaped flows. Note that all of the traffic shaping policies presented here shape only flows classified as low-priority; best-effort traffic is never shaped.

Per-flow bandwidth limit (PBL). With PBL, each low-priority flow is shaped to a fixed maximum bandwidth. Traffic shapers use a dedicated queue for each low-priority flow, and dequeue packets according to a token bucket algorithm. In our simulations, we limit the bandwidth consumed by each low-priority flow to 250 Kbps.

We also evaluate a variant of this policy called **PBL-PEAK**, where low-priority flows are shaped only between 9 am and 3 pm local time. This period corresponds to 6 hours centered around the peak utilization in our traces at about noon. Both PBL and PBL-PEAK require routers to allocate a new queue for each new low-priority flow, thus potentially limiting the practicality of these two policies.

Low-priority bandwidth limit (LBL). In this policy, the aggregate bandwidth consumed by all low-priority flows is bounded by a bandwidth limit. Traffic shapers deploy two queues: one for best-effort traffic and one for low-priority traffic. A token bucket applied to the low-priority queue limits the low-priority traffic rate to the desired bandwidth limit. The bandwidth limit is determined based on the average bandwidth consumed by low-priority traffic on the previous day, as described before. No bandwidth limit is applied to the best-effort traffic. This policy can also be used to approximate PBL by using a dynamic bandwidth limit proportional to the number of low-priority flows.

---

[2] The DSCP field allows up to 64 different traffic classes.

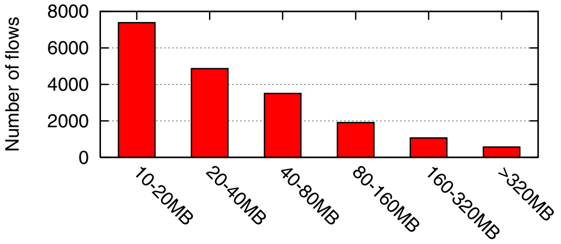


Figure 3: Distribution of flows in the Ohio State trace: Number of flows larger than 10 MB falling within different flow size ranges.

Aggregate bandwidth limit (ABL). When the aggregate traffic (best-effort + low-priority traffic) approaches the bandwidth limit, low-priority flows are shaped to keep the aggregate traffic below the limit. Note that best-effort traffic is never shaped. Therefore, if the best-effort traffic exceeds the bandwidth limit, this policy cannot guarantee that the aggregate traffic stays below the bandwidth limit. However, in such cases the traffic shaper throttles the low-priority traffic to zero bandwidth until the best-effort traffic falls below the bandwidth limit.

To implement this policy, traffic shapers deploy two queues: a high-priority queue for the best-effort traffic and a low priority queue for the low-priority traffic. Both queues share a single token bucket, which generates tokens at a rate corresponding to the aggregate bandwidth limit. Each time a packet from either queue is forwarded, tokens are consumed. However, best-effort packets are always granted access to the link, even if there are not enough tokens left. This is unlike an ordinary token bucket and can cause the token count to occasionally become negative, thus precluding low-priority packets from using the link. As long as the total traffic rate is below the bandwidth limit, there are always enough tokens to forward both best-effort and low-priority traffic. But, as the total traffic level exceeds the bandwidth limit, low-priority flows are shaped.

Aggregate bandwidth limit with shortest-flow first scheduling (ABL-SFF). This policy is like ABL, but additionally optimizes the usage of the bandwidth available to the low-priority flows. Unlike PBL or LBL, in ABL low-priority traffic is not guaranteed a minimum bandwidth allocation, but all low-priority flows compete for the bandwidth that best-effort traffic is not using. Thus, when the total traffic reaches the bandwidth limit, the bandwidth available to low-priority flows becomes so low that some of these flows get substantially delayed or even stalled.

We gained an insight on how to lessen this problem by looking at the flow-size distribution in our traces. Figure 3 shows the number of low-priority flows that fall into different size ranges in one of our traces. The distribution of flow sizes is heavily skewed with roughly 85% of low-priority flows having size between 10 MB and 100 MB. Under such skewed distributions, it is well-known that giving priority to small flows reduces the mean completion time [18, 26]. Therefore, in the ABL-SFF policy, when selecting a low-priority packet to send over the link, the traffic shaper always chooses the packet from the flow with the smallest size. This effectively replaces the usual FIFO queueing with shortest-flow-first queueing. To implement this policy, the traffic shaper needs to allocate a separate queue for each low-priority flow. Also, the shaper needs a priori knowledge of the size of each flow to select the next low priority packet. This makes this policy not directly applicable to general network flows, whose size cannot be known, but gives an useful lower-bound on the minimum delay that low-priority flows incur with the ABL policy.

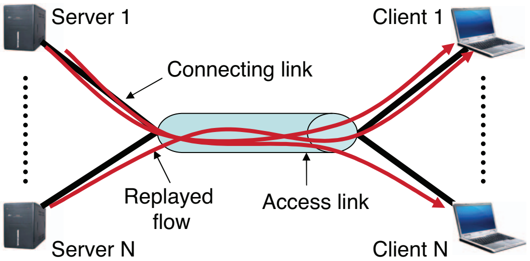


Figure 4: Simulation topology: All replayed TCP flows cross a shared access link where traffic shaping takes place.

Aggregate bandwidth limit with strict priority queuing (ABL-PQ). This policy is a practical version of ABL-SFF and can be implemented by ISPs with today's equipment. It approximates the shortest flow first scheduling of ABL-SFF as follows. First, unlike ABL-SFF, it does not assume a priory knowledge of flow sizes, but instead keeps a running count of the bytes sent by each active network flow and uses this value as an estimate of the flow size. Second, ABL-PQ does not use a separate queue for each low-priority flow, but instead uses a fixed, small number of low-priority packet queues. Each queue accommodates packets of low-priority flows whose size fall in a given range. When the traffic shaper has bandwidth to send low-priority traffic, it schedules the low-priority queues giving *strict priority* to the queues that accommodate smaller flows.

To balance the load of the low-priority queues, we selected contiguous ranges of exponentially increasing width. This is motivated by the typical skewness of the flow size distribution in the Internet. For our traces, where flows larger than 10 MB are classified as low-priority traffic, the first low-priority queue contains packets of flows that have transferred between 10 MB and 20 MB, the second queue contains packets of flows that have transferred between 20 MB and 40 MB, and so on. As opposed to ABL-SFF, this policy uses a limited number of queues (we use 6 in our experiments) and can be easily supported by today's networking equipment.

D. *Comparison methodology*

We used trace-driven simulations to study the behavior of flows under various traffic shaping mechanisms. We conducted our analysis using the ns-2 simulator and the NetFlow traces from Section III. During a simulation, we replayed all TCP flows in a trace using the ns-2 implementation of TCP-Reno. A flow is replayed by having the TCP sender send as many bytes as specified in the flow's NetFlow record.

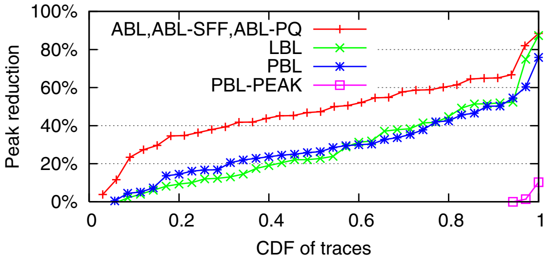


Figure 5: Reduction in peak with different traffic shaping policies: Traffic shaping policies based on aggregate bandwidth limits (ABL) achieve considerable reductions in peak traffic.

To analyze traffic shaping over an access link we used the simulation topology shown in Figure 4. Servers act as TCP senders and clients function as TCP receivers. The access link in our simulations has a capacity of 1 Gbps. For each flow, we set the capacity of the link connecting each server to match the average bandwidth computed from the flow's NetFlow record. This ensures that each replayed flow completes in a time similar to the original flow in the NetFlow trace. We set the length of all packet queues in the simulation to the bandwidth delay product of the corresponding link, using an RTT of 160 ms. Finally, we randomly picked the RTT of each flow from a distribution of latency measurements collected using King [17]. We found that the aggregate traffic generated by our replayed flows matches the original trace very well.

To compare different traffic shaping policies, we focused on the three metrics from Section III: the achieved peak reduction, the fraction of shaped flows, and the delay shaped flows incur.

E. Results

We now present the results of the comparison of the different traffic shaping policies.

1) Peak reduction: We start by presenting the overall peak reductions attained by the different policies across all our traces, shown in Figure 5. Since ABL, ABL-SFF and ABL-PQ all cap the traffic at the same limit, we report only one line for all of them. The ABL policies achieve a considerably higher peak reduction than LBL. This is because LBL does not take into account the level of best-effort traffic when computing the low-priority traffic cap. PBL performs similarly to LBL, while PBL-PEAK is by far the worst-performing policy, causing in 90% of the cases an *increase* in traffic peak (these correspond to points that lie on the negative side of the y-axis in the figure, and are not shown).

To better understand the differences in peak reduction among the different policies, we show in Figure 6 time plots of the traffic in an example trace. Flows smaller than 10 MB are marked as best-effort traffic. Figure 6(a) shows the original traffic trace without traffic shaping. Compared to the original trace, the ABL policies (Figure 6(b)) considerably reduce peak bandwidth (-64%). LBL (Figure 6(c)) achieves lower, but still substantial reductions (-51%).

Comparing LBL and ABL, we observe that ABL achieves a much smoother peak as the total amount of traffic is capped to a constant daily threshold (note that best-effort traffic can still occasionally exceed the threshold). The advantage of LBL is that it guarantees a minimum amount of bandwidth to low-priority traffic, and thus avoids stalling low-priority flows. However, the total traffic still shows diurnal patterns and the peak reduction is thus not as large as with ABL.

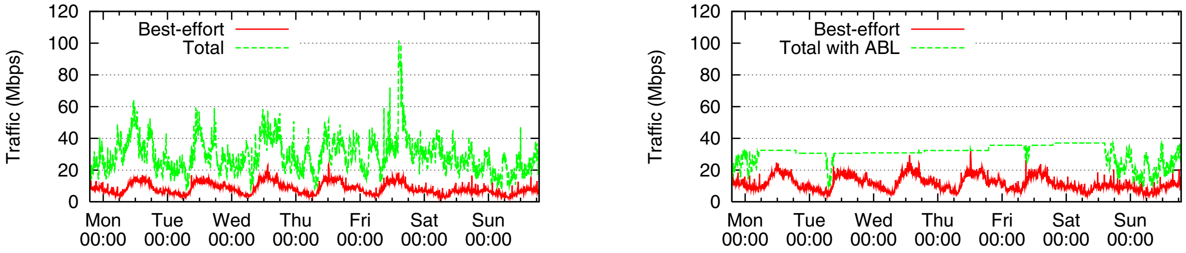
Policy	Flows delayed by >5%	Average peak reduction
ABL	80%	48%
PBL	71%	29%
LBL	61%	28%
ABL-PQ	51%	48%
ABL-SFF	32%	48%
PBL-PEAK	24%	-87%

Table I: Fraction of low-priority flows delayed by more than 5% and average peak reduction: Among the practical policies that maximize peak reduction, ABL-PQ delays the fewest flows.

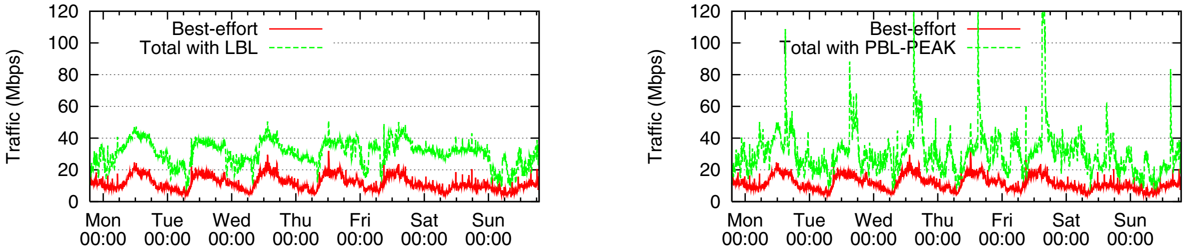
Finally, Figure 6(d) shows that PBL-PEAK is largely ineffective at reducing traffic peak. In fact, PBL-PEAK increases the traffic peak by 11% in this case. To understand this counterintuitive result, consider the following example. During the traffic shaping period (9 am to 3 pm), each low-priority flow is throttled to 250 Kbps. This small per-flow bandwidth makes it hard for low-priority flows to complete. As a result, the number of active low-priority flows increases during the traffic shaping period. At the end of the traffic shaping period all these flows are given full bandwidth again, which they promptly consume. This causes the traffic spikes that are visible in Figure 6(d) on each day at 3 pm, i.e., the end of the traffic shaping period. These spikes can be considerably higher than the original traffic peak. This phenomenon does not occur with PBL because traffic shaping occurs throughout the day (not shown).

2) Number of delayed low-priority flows: Since in our analysis all traffic shaping policies use the same flow size threshold, the flows that are treated as low-priority by each traffic shaping policy are the same. However, depending on the policy, some of these flows may incur only moderate delay. We regard a low-priority flow as delayed if its completion time increases by more than 5% compared to when no traffic shaping is in place. Table I reports, across all traces, the fraction of low-priority flows that are delayed by more than 5% with each traffic shaping policy and the achieved average peak reduction. ABL affects the most flows, followed by PBL, which only gives 250 Kbps to each flow. Compared to ABL, ABL-SFF and ABL-PQ greatly reduce the number of delayed flows. PBL-PEAK delays very few flows because it only rate limits for 6 hours a day, but it also significantly increases the peak usage as pointed out above. Interestingly, although LBL always allocates a minimum amount of bandwidth to low-priority flows, it delays more flows than ABL-PQ and ABL-SFF, which do not provide such a guarantee. The reason is that both ABL-PQ and ABL-SFF give priority to smaller flows, thus shifting the bulk of the delay to a few large flows.

3) Delay of low-priority flows: Figure 7 plots the CDFs of relative and absolute delays of low-priority flows for different

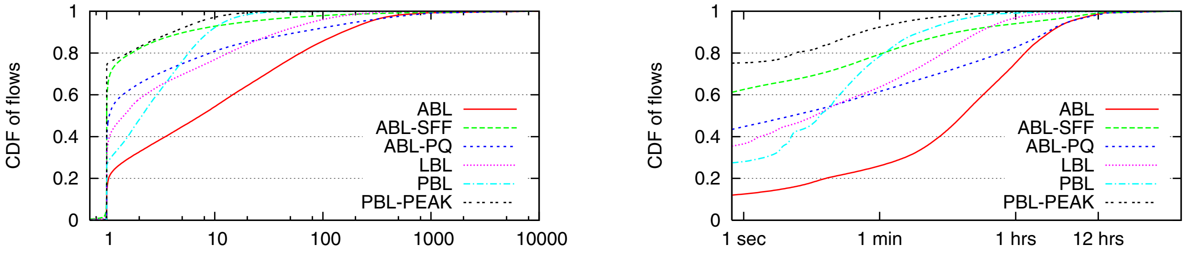


(a) Original traffic (b) Shaped with ABL



(c) Shaped with LBL (d) Shaped with PBL-PEAK

Figure 6: Traffic in the Ohio State trace with different traffic shaping policies: Each plot shows best-effort traffic as well as the total amount of traffic (best-effort + low-priority traffic).



(a) Relative delay (b) Absolute delay

Figure 7: CDFs of relative and absolute delays for low-priority flows across all our experiments: The relative delay is the ratio of the completion time of the traffic shaped flow to its completion time with no traffic shaping. With the exception of ABL and ABL-PQ, few low-priority flows get delayed by more than 1 hours, and almost none is delayed by more than 12 hours.

policies across all our experiments. ABL causes the largest delays while both ABL-SFF and PBL-PEAK lead to very low delays. However, as mentioned above, PBL-PEAK significantly increases peak usage and has therefore little practical utility. With ABL, about half of low-priority flows take 10 times longer or more to complete compared to when they are not traffic shaped. With ABL-PQ, only 20% of low-priority flows take 10 times longer or more to complete. Regarding the absolute delay of flows (Figure 7(b)), we observed that at most 20% of low-priority flows are delayed by more than 1 hour for all policies, and almost no flow is delayed by more than 12 hours.

Although delay is the main performance metric for many bulk flows like file downloads, real-time bulk applications like *video on demand (VoD)* have additional requirements, such as steady throughput to ensure a smooth playback. Ideally, we would have liked to directly quantify how our traffic shaping policies affect such applications. However, our traces don't contain application-level information, making such analysis hard. To arrive at a rough estimate, we gathered data on the sizes of over 1.2 M YouTube videos selected at random and estimated the fraction of the video flows that would be affected by traffic shaping. In 60% of our experiments, at least 94% of these videos would have never been traffic shaped because they are smaller than the flow size threshold. In the remaining 40% of our experiments, when using ABL-PQ, an average of 52% traffic shaped videos would have completed in the first low-priority queue, thus limiting their performance loss. These results suggest that YouTube videos are unlikely to be severely affected by our ABL-PQ policy.

F. Summary

We compared the performance of five traffic shaping policies with respect to our goals of peak reduction, minimum number of delayed flows, and minimum increase in completion time. We found that the ABL policies result in the best peak

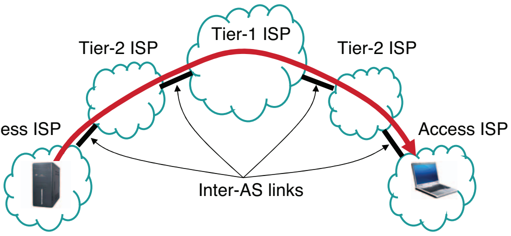


Figure 8: Flow traversing multiple ISPs: It is likely that a transfer between a server and a client traverses multiple traffic-shaping ISPs.

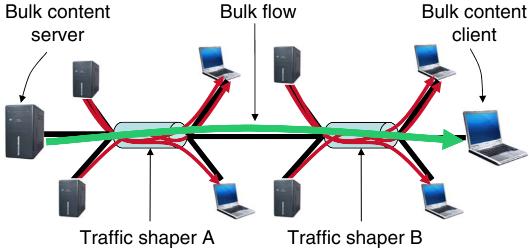


Figure 9: Simulation topology for analyzing the performance of a flow passing two traffic shapers: A long-running bulk TCP flow transfers data from a server to a client and traverses two traffic shapers that act independently.

reduction (almost 50% in half of our traces). In addition, ABL-SFF keeps the delay incurred by low-priority flows to a minimum. However, it might not be possible to implement ABL-SFF in practice as it requires a distinct router queue for each low-priority flow. A more practical alternative to ABL-SFF is ABL-PQ, which achieves both high peak reduction and moderate delay of low-priority flows.

Finally, although we used traces from academic networks in our experiments, the characteristics of the traces that are relevant to our analysis (such as diurnal variations and skewness in flow size distribution) are consistent with what was observed in several previous studies of commercial Internet traffic [2, 4, 29]. This suggests that our results would also apply to commercial Internet traffic.

V. THE GLOBAL EFFECTS OF LOCAL TRAFFIC SHAPING

In this section, we focus on the impact that widespread deployment of traffic shaping has on the end-to-end performance of bulk flows in the Internet. Economic incentives are likely to drive ISPs to deploy traffic shaping at their network boundaries. This is especially true for access ISPs that pay for transit. Since most Internet flows traverse at least two access ISPs, they are likely to be traffic shaped at two or more inter-AS links (Figure 8).

For the analysis, we assume that each ISP implements the ABL-PQ policy from Section IV, because this policy enables maximum peak reduction with low impact on network flows.

A. Analysis methodology

Our analysis is based on trace-driven simulation experiments conducted using ns-2. Figure 9 shows the topology we used in our analysis: it consists of two traffic shaped links connected to each other. We used our university traces to simulate the local traffic traversing each of the shapers using the same methodology as in the previous section. In addition to the flows from the traces, we simulated a week-long bulk TCP flow that traverses both traffic shaped links. We analyzed the performance of this week-long bulk flow to understand the impact of multiple traffic shapers. We focused on a single long-running bulk flow because small flows are left largely unaffected by the ABL-PQ policy.

Although our long-running flow is active throughout the simulated week, we focus solely on the performance achieved from Tuesday to Thursday. The reason is that there is often sufficient available bandwidth to serve all traffic around weekends, and as a consequence our traffic shapers are mostly active during the central days of the week.

As a measure of a bulk flow's performance, we count the number of bytes the bulk flow was able to send from Tuesday to Thursday. To quantify the impact of multiple traffic shapers on a flow, we define a metric called *end-to-end performance loss*. End-to-end performance loss is defined as the relative decrease in performance of a bulk flow traversing multiple traffic shapers compared to the minimum performance the bulk flow achieves when it traverses either of the two traffic shapers separately. More formally, consider a flow that transfers B_1 and B_2 bytes when it separately traverses traffic shapers S_1 and S_2, respectively. If the same flow transfers G bytes when it simultaneously traverses S_1 and S_2, the end-to-end performance loss of the flow is: $(min(B_1, B_2) - G)/min(B_1, B_2)$.

B. The impact of multiple traffic shapers on end-to-end performance

To study the effects of multiple traffic shapers, we used traces from 15 of our 35 university access links. We simulated traffic shaping on these links and analyzed the performance of bulk flows over all possible (105) combinations of the 15 traffic shaped links. The universities are spread across the four U.S. time zones. When replaying the traces in our simulation we eliminated the differences in local time by time-shifting all traces to the Eastern Standard Time. We discuss the impact of time zone differences in Section V-C.

Figure 10 shows the relative end-to-end performance loss experienced by flows traversing each pair of traffic shaped links. The loss in end-to-end performance is significant. In almost 80% of the cases flows crossing two shapers sent 40% less data than what was sent when crossing only a single traffic shaper. In 50% of the simulations, the loss in performance is larger than 60%. While we do not show the data here, the performance continues to decrease with each additional traffic shaper in the path.

The considerable performance loss for flows traversing multiple traffic shapers can be mainly attributed to two factors. First, at any time t of the simulation, a flow traversing

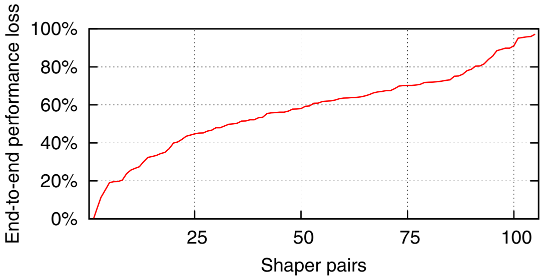


Figure 10: The impact of two traffic shapers in the path of bulk flows: In most of our simulations the end-to-end performance loss for flows is substantial.

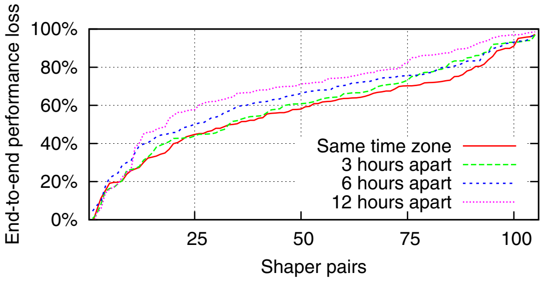


Figure 11: Impact of time zone offset on performance: Compared to two traffic shapers located in the same time zone, topologies with 3, 6, and 12 hours time difference between the shapers lose performance only moderately.

two traffic shapers S_1 and S_2 is limited to using only the *minimum bandwidth allowed by each traffic shaper at time* t. However, the sum of these minimum bandwidths over the entire simulation time can be lower than the total bandwidth available at either of the two traffic shapers during the same time.

The second limiting factor is that TCP congestion control may prevent the flow from fully using the full bandwidth available at any time t. In fact, prior work [15] has shown that a long TCP flow traversing two or more bandwidth bottlenecks (e.g., caused by congestion or in our case traffic shapers) suffers a severe drop in throughput when competing with shorter flows traversing only a single bottleneck (or shaper). We quantify the relative contribution of these two factors in detail in a longer tech report [25].

C. Impact of traffic shaping across time zones

We now quantify the performance loss caused by traffic shapers located in different time zones. The diurnal patterns of these traffic shapers will be offset. As a result, since bulk flows can only use the minimum end-to-end bandwidth available at any time t, we expect an additional reduction in the end-to-end bandwidth available to bulk flows, with consequent additional performance loss. In Figure 11, we show the end-to-end performance loss when the two shapers are in the same time zone, and when they are 3, 6, and 12 hours apart. We also show the average performance loss across all link pairs in Table II.

Interestingly, while the performance loss does increase with the time zone difference between traffic shapers, the additional performance loss is rather small. While two shapers in the same time zone cause an average 58% loss in performance, spacing them by 3 hours decreases performance by only an additional 2%. A time shift of 12 hours results in an average total performance loss of 69%.

Time zone difference	Median total perf. loss	Avg. total perf. loss
0 hrs	60%	58%
3 hrs	62%	60%
6 hrs	68%	64%
12 hrs	72%	69%

Table II: Median and average performance loss for different time zone offsets: Compared to the performance loss caused by adding a second traffic shaper to the path, increasing the time difference of these shapers has only moderate impact.

To understand why most of the performance loss occurs even when the two traffic shapers are in the same time zone, we took a closer look at the behavior of our week-long bulk TCP flow when it traverses only a single traffic shaper.

We found that the flow's throughput shows extreme diurnal patterns: the throughput is very high during a short time period, typically between midnight and early morning, and it's low during the rest of the day. The reason for such extreme diurnal patterns is that, as the total bandwidth available to low-priority flows decreases due to diurnal patterns in best-effort traffic, low-priority flows find it hard to complete and their number increases. As this happens, there are increasingly more low-priority flows competing for less bandwidth, causing the *per-flow* available bandwidth to decrease even faster, resulting in extremely pronounced diurnal patterns in per-flow throughput. In fact, we found that more than 90% of the flow's bytes are transferred during short peak periods, which account for less than 10% of the flow's duration.

When our long-running bulk flow traverses two traffic shapers, its throughput will have characteristic peak periods at each one of these shapers. If the peak periods coincide, the flow will be able to transfer a considerable amount of data. However, since the peak periods are short, a marginal misalignment of peak periods is likely to occur even when the two traffic shapers are in the same time zone. In this case, the end-to-end throughput drops considerably. At this point, most of the damage has been done and increasing the time zone difference does not result in a large additional throughput reduction. A more detailed explanation of this phenomenon can be found in our longer technical report [25].

To avoid the negative global effects that arise when transiting through multiple traffic shapers, one could break up TCP transfers along an end-to-end path into multiple smaller TCP transfers, each spanning a path segment containing only one traffic shaper. In this case, the data would have to be stored and forwarded at intermediate routers along the path using techniques similar to the ones presented in [22]. Our longer technical report [25] presents a detailed investigation of the effectiveness of store-and-forward techniques in this scenario and quantifies the required storage.

VI. RELATED WORK

Today's networking equipment enables ISPs to deploy complex traffic shaping policies at line speeds [9, 27]. This equipment typically supports deep packet inspection, which allows ISPs to identify particular applications, and a wide range of traffic shaping and queue management techniques. The queueing and shaping techniques used in this paper were first developed in the context of QoS and DiffServ [6], and originally used to provide flow delay guarantees that are better than best-effort.

To the best of our knowledge, our paper is the first study of the local and global effects of traffic shaping in the Internet. The only related work we are aware of is from Laoutaris et al. [22], who quantified how much additional "delay tolerant" data (i.e., data that can tolerate delivery delays of hours or days) ISPs could send for free by exploiting 95^{th} percentile billing and diurnal traffic patterns. They present and evaluate simple end-to-end scheduling policies as well as "store-and-forward" techniques that use storage deployed in the network. They show that it is possible to transfer multiple TBytes during off-peak times with no additional costs.

There are two main differences between our work and theirs. First, while [22] aims to send additional (delay-tolerant) data without increasing bandwidth costs for ISPs , our work reduces the peak bandwidth usage of ISPs for today's traffic with only moderate impact (i.e., delay) on shaped flows.

Second, their analysis is based on data limited to aggregate load on network links and ignores issues related to transport protocols. On the contrary, we use flow-level NetFlow data to simulate the behavior of individual TCP flows and perform a more realistic analysis. Our more detailed analysis enables us to identify global effects of traffic shaping that are related to TCP characteristics, which would have escaped an analysis based on traffic aggregates only.

VII. CONCLUSIONS

In this paper, we conducted a systematic analysis of traffic shaping. Even though traffic shaping is widely deployed today, most deployed policies are ad-hoc without a clear understanding of their effectiveness at reducing peak network load and their impact on network traffic.

We compared several traffic shaping policies with respect to the achieved reduction in peak network traffic and the resulting performance loss for shaped network flows. We identified a practical policy that achieves peak traffic reduction of up to 50% with only limited performance loss for bulk transfers. However, we also found that the same policy leads to large performance losses for network flows when deployed by multiple ISPs along a networking path. We were able to show that this performance loss can be attributed to certain characteristics of TCP and differences in local peak utilization times.

REFERENCES

[1] Abilene Backbone Network. http://abilene.internet2.edu.

[2] S. Agarwal, A. Nucci, and S. Bhattacharrya. Measuring the Shared Fate of IGP Engineering and Interdomain Traffic. In *Proc. of ICNP*, 2005.

[3] Amazon Simple Storage Service. http://aws.amazon.com/s3/.

[4] N. B. Azzouna and F. Guillemin. Analysis of ADSL Traffic on an IP Backbone Link. In *Proc. of IEEE GLOBECOM*, 2003.

[5] Comments of Bell Aliant Regional Communications, Limited Partnership and Bell Canada. http://www.crtc.gc.ca/PartVII/eng/2008/8646/c12_200815400.htm#2b.

[6] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. An Architecture for Differentiated Service. RFC 2475 (Informational), Dec. 1998. Updated by RFC 3260.

[7] N. Brownlee and kc claffy. Understanding Internet Traffic Streams: Dragonflies and Tortoises. *IEEE Communications Magazine*, Oct 2002.

[8] K. Cho, K. Fukuda, H. Esaki, and A. Kato. The Impact and Implications of the Growth in Residential User-to-user Traffic. In *Proc. of SIGCOMM*, 2006.

[9] Cisco IOS Classification. http://www.cisco.com/en/US/docs/ios/12_2/qos/configuration/guide/qcfclass.html.

[10] Cisco IOS Congestion Management. http://www.cisco.com/en/US/docs/ios/12_2/qos/configuration/guide/qcfconmg_ps1835_TSD_Products_Configuration_Guide_Chapter.html.

[11] Comcast: Description of planned network management practices. http://downloads.comcast.net/docs/Attachment_B_Future_Practices.pdf.

[12] M. Dischinger, A. Mislove, A. Haeberlen, and K. P. Gummadi. Detecting BitTorrent Blocking. In *Proc. of IMC*, 2008.

[13] P. Elmer-DeWitt. iTunes store: 5 billion songs; 50,000 movies per day, June 19th, 2008. http://tech.fortune.cnn.com/2008/06/19/itunes-store-5-billion-songs-50000-movies-per-day.

[14] C. Estan and G. Varghese. New Directions in Traffic Measurement and Accounting: Focusing on the Elephants, Ignoring the Mice. *ACM Transactions on Computer Systems*, August 2003.

[15] S. Floyd. Connections with multiple congested gateways in packet-switched networks part 1: One-way traffic. *ACM Computer Communication Review*, 21:30–47, 1991.

[16] S. Friederich. Bandwidth restrictions save almost $1 million, Oct 22nd, 2002. http://thedaily.washington.edu/2002/10/22/bandwidth-restrictions-save-almost-1-million/.

[17] K. P. Gummadi, S. Saroiu, and S. D. Gribble. King: estimating latency between arbitrary internet end hosts. In *Proc. of IMW*, 2002.

[18] L. Guo and I. Mitta. Scheduling flows with unknown sizes: An approximate analysis. In *Proc. of SIGMETRICS*, 2002.

[19] ipoque GmbH. OpenDPI, 2009. http://www.opendpi.org.

[20] T. Karagiannis, A. Broido, N. Brownlee, K. Claffy, and M. Faloutsos. Is P2P Dying or just Hiding? In *Proc. of IEEE Globecom*, 2004.

[21] T. Karagiannis, A. Broido, M. Faloutsos, and K. claffy. Transport Layer Identification of P2P Traffic. In *Proc. of IMC*, 2004.

[22] N. Laoutaris, G. Smaragdakis, P. Rodriguez, and R. Sundaram. Delay Tolerant Bulk Data Transfers in the Internet. In *Proc. of SIGMETRICS*, 2009.

[23] Linux Advanced Routing & Traffic Control HOWTO. http://lartc.org/lartc.html.

[24] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On dominant characteristics of residential broadband internet traffic. In *Proc. of IMC*, 2009.

[25] M. Marcon, M. Dischinger, K. Gummadi, and A. Vahdat. The local and global effects of traffic shaping in the internet. Technical Report MPI-SWS-2010-004, MPI-SWS, 2010. http://www.mpi-sws.org/~mmarcon/TrafficShaping-TECHREPORT.pdf.

[26] I. Matta and L. Guo. Differentiated predictive fair service for tcp flows. In *Proc. of ICNP*, 2000.

[27] Blue Coat PacketShaper. http://bluecoat.com/products/packetshaper.

[28] J. Röttgers. Internetanbieter bremst Tauschbörsen aus. Focus Online, Mar 6th, 2008. http://www.focus.de/digital/internet/kabel-deutschland_aid_264070.html.

[29] A. Shaikh, J. Rexford, and K. G. Shin. Load-Sensitive Routing of Long-Lived IP Flows. In *Proc. of SIGCOMM*, 1999.

[30] S. Shalunov and B. Teitelbaum. TCP Use and Performance on Internet2. In *Proc. of ACM SIGCOMM Internet Measurement Workshop*, 2001.

[31] R. Topolski. Comcast is using Sandvine to manage P2P connections, May 2007. http://www.dslreports.com/forum/r18323368-Comcast-is-using-Sandvine-to-manage-P2P-Connections.

[32] Universities Prepare for Data Deluge from CERN Collider, May 2007. http://www.hpcwire.com/hpc/1572567.html.

[33] Valve Corp. Steam. http://store.steampowered.com.