Fast Admission Control for Short TCP Flows

Yingxin Jiang and Aaron Striegel
Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556 USA
{yjiang3,striegel}@nd.edu

Abstract—Over the last decade, numerous admission control schemes have been studied to allocate network resources. Although per-flow control schemes can provide guaranteed QoS, such schemes face scalability issues in large networks due to the tremendous number of flows present. While aggregation-based approaches such as Differentiated Services relieve the storage of state in the core, admission control of flows, especially short-lived flows, is still a serious bottleneck. To that end, we propose an admission control scheme, Fast Admission for Short Flows (FASF), that enables accelerated admission control at the edge rather than via centralized or in-path mechanisms. FASF not only reduces the burden on admission control by largely distributing the dominant resource requests (i.e. short-lived flows), but also improves flow completion time and hence network goodput.

Index Terms—Admission control, Quality of Service (QoS), DiffServ, flow completion time

I. INTRODUCTION

Over the last decade, numerous admission control schemes have been studied to allocate network resources [1]–[8]. One of the earlier large scale effort, Integrated Services (IntServ) [1], drew heavily from the connection-oriented voice system, i.e. an end-to-end path is reserved with most if not all routers participating in policing the resource request. Although IntServ provides the capacity for guaranteed QoS, it suffers from scalability problems as each router must process per-flow resource reservation requests and maintain per-flow states.

To improve scalability, approaches such as Differentiated Services (DiffServ) streamlined network core operations to operate on aggregates (classes) of flows rather than individual flows. While DiffServ improves scalability in the core, resource negotiations still take place on a per-flow basis at each domain, typically in a bandwidth broker (BB) type of environment [2]. Meanwhile, although admission process latency can be amortized by long-term flows over the long transmission time, the setup latency significantly affects the performance of short-term flows. Hence, as the majority of the flows on the Internet are short-lived [9], the bottleneck of admission control was simply shifted in DiffServ as opposed to being relieved in such an environment. Conversely, the approach of relegating short-lived to entirely best effort (i.e. no QoS) is not feasible either as e-commerce or other priority-based (ex. subscriptions, text messaging, etc.) flows still demand QoS. To that end, this paper poses the following questions: Does there exist a tipping point under which well-behaved short-lived flows can be streamlined for admission control? Can that tipping point be exploited to delegate admission control to the edge and improve scaling while maintaining network stability?

In this paper, we propose an admission control scheme, Fast Admission for Short Flows (FASF), where edge routers can provide instant admission for short-lived flows. Flows which are not accepted by edge routers through fast admission will follow the traditional resource request mechanism. The key contributions of this paper include:

• Reduced processing load of the admission control agent (BB): With FASF, edge routers do not send resource reservation requests for fast admitted flows to the BB. Thus the number of requests processed by the BB decreases significantly.

• Decreased amount of per-flow states: In FASF, edge routers aggregate packet policing for all flows admitted through the fast admission procedure. Thus, per-flow states of such flows are avoided and the burden of edge routers for flow state storage is relieved.

• Reduced flow completion time (FCT) for short-lived flows: By reducing the admission process latency for fast admitted short-lived flows, FASF achieves a lower FCT for such flows. As the flow completion time indicates when users can finish their transactions, such as downloading a web page, reading/sending an email, etc., it is a critical performance metric for users [10], [11].

While we recognize that this work at first glance appears to be incremental amongst the large body of existing QoS research, we believe this work represents a significant improvement towards achieving practical QoS
in the reality of today’s Internet. Critically, we show that it is possible to safely ignore the QoS requests of short-lived traffic subject to a derived tipping point without relegating short-lived traffic to best effort. The net result is a removal of one of the chief objections towards per-flow QoS, namely that short-lived traffic imposes untenable overhead in terms of both delay and scalability. Moreover, we believe that future work may be able to extend the core concepts of this work to an end-to-end perspective, providing end-to-end admission/reservation without requiring per-flow end-to-end negotiation, i.e., removal of much of the per-flow control plane messaging.

The remainder of this paper is organized as follows. Section II comments on related work. Section III presents an overview of the proposed admission control scheme, Fast Admission for Short Flows (FASF), and explains how fast admission can be done at edge routers. Next, Section IV presents and discusses the simulation results and Section V concludes the paper.

II. RELATED WORK

Various admission control approaches have been proposed to satisfy the QoS requirements in IP networks [1]–[8]. In the probe-based distributed admission control approach [3]–[5], flows probe the network and admission decisions are made based on the QoS of the probe packets. In the centralized admission control approach [2], [7], [8], new connections must be approved through an agent, called Bandwidth Broker (BB) or Network Broker (NB). Since most admission control schemes do not make a distinction between short-lived and long-lived flows, both short and long flows have similar admission process latencies. While the admission process latency for long-term flows can be neglected due to long data transmission time, the setup latency significantly impacts the performance of short-term flows.

To reduce the transmission time of short-lived flows, several mechanisms have been proposed to differentiate short-lived flows from long-lived flows and give short flows higher priority to transmit packets [12]–[14]. In [12], the packet priority is set according to the application type or the state of a TCP connection. In [13], after edge routers mark packets belonging to short or long flows, the RIO queues in core routers treat short-lived and long-lived flows differently. The work [14] presented a packet level stateless, threshold-based scheduling mechanism, RuN2C. Notably, these mechanisms focus on reducing the transmission time for short-lived flows without traffic policing for individual flows.

III. FAST ADMISSION FOR SHORT FLOWS

Specifically, we focus on intra-domain admission control to assess the feasibility of the approach with future work to examine end-to-end admission, a considerably more difficult problem. Before presenting the proposed FASF scheme, we briefly introduce the key concepts of the BB admission control scheme in one DiffServ domain. In such a scheme, the Bandwidth Broker (BB), a centralized agent, is in charge of processing resource reservation requests and per-flow information is stored at the BB or edge routers. When the resource reservation request for a new flow arrives at the first edge router (ER) of the domain, the ER sends the resource request to the BB. The BB then processes the request and sends the admission decision to the ER. Finally, the ER forwards the admission decision back to the source of the new flow.

Figure 1 shows the model of the proposed FASF scheme in a single DiffServ domain. When the first edge router (ER) of the domain receives the resource reservation request of a new flow, the ER makes a decision for if the new flow can be admitted instantly through the fast admission control procedure which we will explain shortly. If this flow is short-lived and can be admitted instantly, the ER admits the flow and the source may start data transmission immediately. Otherwise, the ER will follow the traditional admission control scheme via communicating with the BB. By reducing the amount of resource reservation requests which need to be processed by the BB, the FASF scheme not only significantly reduces the amount of stored per-flow states and the processing load of the BB, but also decreases the FCT for traffic substantially.

A. Fast Admission Based on Flow Throughput

In this and next subsections, we describe the mechanisms which enable FASF. For the purpose of simplicity,
we assume that the BB in one domain has employed appropriate mechanisms to ascertain the definition of short flows, \( S_{sf} \) (size of a short flow), and the edge-to-edge bandwidth that can be utilized for fast admission by an edge router, \( BW_f \). We also assume that a flow informs the network about its size together with the resource reservation request in the connection setup phase\(^1\). Section III-C details the solution for the event that fast admitted flows estimate or report their sizes mistakenly (such as reporting a long-lived flow being short). In addition, we assume edge routers gather the information of end-to-end RTTs through active probing or passive monitoring [6].

Conceptually, FASF attempts to bound the instantaneous throughput of short-lived flows (transfer less than \( S_{sf} \) bytes) on an edge-to-edge basis to less than \( BW_f \). As these flows appear and disappear rapidly, it is easier to view them as an aggregated number bounded by \( MAX_f \) for a given edge-to-edge path rather than individual flow states. In short, \( MAX_f \) represents the maximum number of active short-lived flows which are admitted through fast admission on the edge-to-edge path without sacrificing the QoS of themselves and other traffic (i.e. the tipping point). When a new TCP flow arrives at an edge router (see Figure 1), the edge router can quickly make a decision about whether the new flow can be instantly admitted. If the new flow is a short TCP flow and the number of active fast admitted flows on the edge-to-edge basis is less than \( MAX_f \), the flow is admitted through fast admission. Otherwise, the traditional admission control procedure is followed.

To derive \( MAX_f \), an edge router first computes the predicted FCT (Flow Completion Time) and the throughput for a short-lived flow with no packet loss. As FASF does fast admission only for short-lived TCP flows, we assume such TCP flows do not exit the slow-start phase. Thus, the FCT of a short flow with size \( S_{sf} \) can be computed as:

\[
FCT_s = C + D_{Tx} = 1.5 \times RTT + \log_2 \left( \frac{S_{sf}}{MSS} \right) \times RTT \tag{1}
\]

where \( C \) is the time for connection establishment with the three-way handshake, \( D_{Tx} \) represents the data transmission time, \( MSS \) is the maximum segment size, and \( RTT \) is the estimated end-to-end round trip time\(^2\). Besides the computation of Equation 1, the FCT for short-lived flows can also be gathered at the edge router through passive monitoring.

With \( S_{sf} \) and \( FCT_s \), the throughput of a short-lived flow, \( T_s \), can be computed as:

\[
T_s = \frac{S_{sf} \times (MSS + H)}{FCT_s \times MSS} \tag{2}
\]

\( MAX_f \) can be expressed as:

\[
MAX_f = \frac{BW_f}{T_s} = \frac{BW_f \times FCT_s \times MSS}{S_{sf} \times (MSS + H)}
\]

Example 1: Here, we use a simple example to illustrate the above computations. Suppose \( S_{sf} = 10KB \), \( RTT = 50ms \), \( MSS = 1460B \), \( H = 40B \) and \( BW_f = 10Mb/s \). With Equation 1, \( FCT_s \), the flow completion time of a short-lived flow, is computed as 233.5ms. Then, Equation 2 derives \( MAX_f \) equal to 29. This means that there should be at most 29 active fast admitted short flows existing on the edge-to-edge path with \( BW_f \), the allocated bandwidth for fast admission, being 10Mb/s. If a new short-lived TCP flow enters a domain and the number of existing short-lived flows on the path is less than 29, the edge/ingress router can accept the new flow through the fast admission procedure. Otherwise, the ingress router will follow the traditional admission control scheme by sending the resource reservation request of this new flow to the BB.

B. Fast Admission Based on Modeling

In an alternative sense, we consider a similar fast admission scheme where edge routers make fast admission decisions based on the rate of fast admitted flows. While the fast admission scheme in the previous section is straightforward, this scheme utilizes the modeling of router buffers [11], [15] and exhibits a stronger theoretical foundation. Later, the examples and simulation results will indicate that these two schemes have similar predictability for the fast admission rate. Since short TCP flows are assumed not to exit the slow-start phase (similar to [11] and [15]) and there is no per-flow packet policing for fast admitted short TCP flows, each fast admitted short flow will arrive at an ingress router in \( n \) bursts of size \( X_i \) with \( R \) as the remainder:

\[
X_i = 2, 4, ..., 2^{n-1}, R
\]

If the arrival of short TCP flows follows a Poisson distribution, the router buffer can be then modeled as an \( \text{M}^N/\text{M}/1 \) queue [11] or an \( \text{M}/\text{G}/1 \) queue [15] with a FIFO service discipline. Then, for a certain link load \( \rho \) and predefined distribution of short TCP flows, the buffer
queue length can be computed. In [11], the expected queue length, \( E(Q) \), is expressed as:

\[
E(Q) = \frac{\rho}{1 - \rho} \frac{E(X) + E(X^2)}{2E(X)}
\]

(3)

where \( E(X) \) and \( E(X^2) \) are the first and second moments of the burst size.

With Equation 3, we can derive the link load, \( \rho \), for a predefined queue length. Thus, for a known \( BW_f \), if \( \rho \) is the traffic load on \( BW_f \), the rate of fast admitted flows for that edge-to-edge path can be computed as:

\[
\text{rate} = \frac{BW_f \times \rho}{MSS + H} \times \frac{MSS}{S_{sf}}
\]

(4)

Example 2: Here, we use another example to explain how fast admission based on modeling can be done and compare this mechanism to the one described in Section III-A. Suppose the network has the similar parameters as in Example 1 (\( BW_f = 10Mb/s \), \( S_{sf} = 10KB \), \( MSS = 1460B \) and \( H = 40B \)) and \( \rho = 0.9 \). The rate of fast admitted flows on that edge-to-edge path can be computed by Equation 4 yielding approximately 112. This means that for every second, the edge router can admit 112 short TCP flows through the fast admission process on that edge-to-edge path. We can transfer the result of Example 1 into the rate of fast admitted flows which is equal to \( MAX_f/FCT_s = 124 \). As the method used in Example 1 tends to consume all \( BW_f \) for fast admission rather than \( \rho \times BW_f \), the rate of fast admitted flows in Example 1 is slightly larger than the rate in Example 2 (\( 124 \times 0.9 = 111.6 \)).

C. Bounding Misbehavior Users

To prevent fast admitted (short) flows from consuming more bandwidth than allocated, there can be an aggregate policing for all fast admitted flows on each edge-to-edge path. Thus, fast admitted flows will not affect the QoS of other traffic even if an edge router temporarily admits too many short flows.

Since there is no per-flow policing for fast admitted flows, each fast admitted flow will naturally consume as much bandwidth as it can. Thus, if a long-lived flow fakes as a short-lived flow to avoid long admission process latency and packet policing, this long-lived flow will consume a large amount of bandwidth and severely affect the performance of other fast admitted flows. For such situation, the fast admission procedure can have a metric recording the start time of a flow in the fast admission list. If a flow stays in the list over a certain period, it will be removed from the list and resigned to normal resource negotiation.

For our simulation studies, the simulations were conducted using the ns-2 simulator. The goal of the studies is to examine the performance of FASF. We compared four approaches, the Best Effort (no admission control approach), RuN2C [14], the DiffServ Bandwidth Broker (BB), and FASF. In RuN2C, routers classify packets according to the current TCP sequence number, then put packets belonging to short-lived TCP flows into the priority queue. In FASF, the edge routers make fast admission decisions based on flow throughput (Section III-A).

The performance metrics used to evaluate the approaches are number of admitted flows, number of finished flows, and flow completion time (FCT), etc. The simulations were performed in a simple network topology (Figure 2) to isolate the characteristics of the proposed FASF scheme. The remaining information concerning the simulation setup is listed below:

- Traffic was generated to work with a sizable number of flows and to ensure complete utilization of the links in the network. A typical core link had hundreds of simultaneous active flows present.
- The settings for FASF was \( BW_f = 10Mb/s \).
- The maximum (cap) link utilization of the BB scheme was 90%.
- A token bucket policing mechanism was used on flows accepted through the traditional BB admission control scheme such that any packets out of the profile were immediately dropped rather than remarked.
- Flows were composed of 10% UDP traffic and 90% TCP traffic. The short/long-term flows followed an exponential distribution with an average length of 5 and 50 seconds for UDP flows and an average size of 10KB (the default value of \( S_{sf} \)) and 5MB for TCP flows. The average size of short-term TCP flows could be changed to 15KB, 20KB, and 25KB to examine the performance of FASF for different short flow compositions (Section IV-B).
- The bandwidth QoS requirements of flows fol-
TABLE I

<table>
<thead>
<tr>
<th>Approach</th>
<th>Core Link Util. (%)</th>
<th>Num. of Admitted Flows</th>
<th>Num. of Finished Flows</th>
<th>FCT for Short Flows (s)</th>
<th># of Requests Processed by BB</th>
<th>Admission Process Latency (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Effort</td>
<td>98.703</td>
<td>11142</td>
<td>4486</td>
<td>8.218</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>RuN2C</td>
<td>98.634</td>
<td>11142</td>
<td>7125</td>
<td>5.585</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>BB</td>
<td>80.772</td>
<td>3308</td>
<td>3203</td>
<td>0.353</td>
<td>3308</td>
<td>0.025</td>
</tr>
<tr>
<td>FASF</td>
<td>87.196</td>
<td>10085</td>
<td>9983</td>
<td>0.196</td>
<td>282</td>
<td>0.01</td>
</tr>
</tbody>
</table>

A. Comparison of Best Effort, RuN2C, BB, and FASF

Table I shows the simulation results for the four approaches. As traffic was generated to ensure a complete utilization of the core link in the network, for the best effort and RuN2C schemes, the monitoring TCP flow almost stalls (not shown) and the core link utilization is close to 100%. For the admission control approaches (BB and FASF), the monitoring TCP flow receives its requested bandwidth (around 100KB/s, not shown) and the link utilization is stable. The more interesting differences of the schemes emerge in the third column of Table I, namely the number of admitted flows. The best effort and RuN2C schemes have the largest amount of admitted flows since each flow is allowed to enter the network. For the BB and FASF approaches, only a fraction of all traffic is admitted to ensure QoS. As a large amount of short-lived flows are accepted through the fast admission procedure in FASF, FASF admits significantly more flows than BB does. The forth column gives the number of finished flows among the admitted flows shown in the third column. In the best effort and RuN2C schemes (no admission control approaches), the extremely high link utilization results in packet losses and TCP backoffs. Thus, only a small percentage of flows finish their data transmission. In the BB and FASF schemes, however, most admitted flows finish their data transmission.

Table I presents the average FCT of short-lived flows for the four approaches as well. Since TCP flows in the schemes without admission control approaches endure large packet losses, the best effort approach receives an extremely high FCT. Compared to best effort, RuN2C receives a shorter FCT as RuN2C puts the data transmission of short-lived TCP flows at a higher priority. The BB and FASF approaches reduce the average FCT substantially as the network is not congested. In addition, FASF has a shorter FCT than BB does due to the short admission process latency (shown in the last column) and no traffic policing for individual fast admitted flows.

Finally, the second column from the right shows the number of requests that are processed by the bandwidth broker. In the BB approach, every resource reservation request has to be processed. In the FASF approach, however, only a small amount of requests needs to be processed by the bandwidth broker since most flows are instantly admitted by edge routers. Thus, compared to other approaches, the FASF approach receives the shortest FCT and reduces the processing load of the bandwidth broker substantially.

B. Different Compositions for Short Flows in FASF

To ensure the fast admission schemes (presented in Section III-A and Section III-B) show similar predictability for fast admission rate under various conditions, we ran simulations where the composition of short flows was changed. As presented in the simulation setup, the size of short-term TCP flows followed an exponential distribution with the average being 10KB, 15KB, 20KB, or 25KB.

Figure 3 shows the rate of fast admitted flows in one edge when fast admission is based on flow throughput (Section III-A) and based on modeling (Section III-B). For a predefined $BW_f$ (equal to $10\text{Mb/s}$), with the increase in the size of short-term TCP flows, the rate of fast admitted flows decreases in both schemes as short-term TCP flows need to transmit more data. For the fast admission based on flow throughput scheme, the simulation results match closely with the theoretical results. Finally, the theoretical results of the fast admission based on modeling scheme have less amount of fast admitted flows comparing to the former scheme. This is because the former scheme tends to consume all allocated bandwidth ($BW_f$), while the latter scheme has the limitation of $\rho = 0.9$ on bandwidth utilization (similar to the comparison between Example 1 and Example 2).

Figure 4 shows the average FCT of short-lived flows for the four approaches. In all approaches, with the increase in the average size of short-term TCP flows, the FCT keeps increasing as well since short flows have...
more data to transmit. Similar to the results in Table I, the FCT in the best effort, RuN2C, BB, and FASF approaches follows a decreasing order for each short file definition. Moreover, FASF offers the improved performance to user while preserving network stability and reducing the amount of control messaging interactions.

V. SUMMARY

In this paper we proposed an intra-domain admission control scheme, Fast Admission for Short Flows (FASF). By reducing the amount of resource reservation requests that need to be processed by the centralized admission control agent, the FASF scheme not only reduces the burden on admission control, but also decreases flow completion time. The ability to relegate admission control to the edge of a large portion of flows while preserving network stability has significant implications for network scaling and QoS deployment that merit future exploration. There are also significant future research opportunities unexplored by this work with regards the general applicability of this work to end-to-end admission control with both participating and non-participating domains. Our immediate future work includes generalizing the derivations of the key parameters ($BW_f$ and $S_{sf}$) in FASF under varying network topologies and conditions.

REFERENCES