Admission Control of a Stateless Network: Effective Bandwidth Approach

1Marlene Angulo, 2Deni Torres-Roman, 3Marco Turrubiartes, 4David Muñoz-Rodríguez
UABC1,2, CINVESTAV Gdl.1,2,3, ITESO7, ITCM Mt.4
mangulo@gdl.cinvestav.mx, dtorres@gdl.cinvestav.mx, mturrubiartes@ieee.org,
dmunoz@itesm.mx

Abstract
This paper describes an admission control general framework on a packet-based Internet environment. Each application is accepted based on its effective bandwidth, bandwidth requirement and the network resources availability. This approach is applied on stateless networks, thus, the resource reservation is not required. Moreover, the core router will not distribute periodically the QoS state information. So, the network available resources will be estimated by the incoming device by means of a Local Resource Measurement. The main advantage of our proposed framework is the system scalability, because there are no extra-mechanisms to be implemented at the core routers rather than the incoming ones. Additionally, the admission mechanism avoids the QoS degradation of previously accepted applications on the network path.

Index Terms.- Admission control, Effective Bandwidth, Stateless networks.

1. Introduction

Admission control on Connection Oriented networks computes the available bandwidth by comparing the utilized bandwidth of previously admitted flows to a link capacity. Connectionless networks have multiple shared links. Consequently, the available bandwidth can not be calculated, must be estimated.

There are several Quality of Service(QoS) IP based networking frameworks, such as integrated Services (IntServ), Differentiated Services (DiffServ), and Multiprotocol Label Switching (MPLS). A trade-off between granularity and scalability is applied in each framework. That is, a trade-off between the application’s requirement fulfillment and system scalability. The main concern of IntServ and MPLS networks is the fulfillment of per-flow requirements rather than DiffServ which focuses on a common type of service shared by a group of applications.

An option to decrease the scalability problem improving the response time is pre-computation scheme for QoS routing [1]. Routing consists of two basic functions: distribution of the network state, and computation of appropriate paths. Routing with QoS implies the necessity of consider the application's requirements and the availability of network resources [2].

On IntServ and MPLS frameworks the establishment of a route before sending the information is required, which does not occur with DiffServ. The acceptance of a new flow connection requires a comparison of application requirements versus network resource availability, as the constraint-based algorithms presented by Korkmaz and Krunz in [3]. However, there is a localized approach The Local Resource Measurement proposed by Nelakuditi et.al. [4]. It is implemented without knowledge of the available resources in the whole network, working only with the flow blocking probability.

Previous work reported by Kelly [5] proposed the use of effective bandwidth to accept or reject a connection considering static flows. It is important to remark that the use of the effective bandwidth term is applied to non-aggregated applications. So, the effective bandwidth of each flow is accurately modeled by conventional distributions such as Pareto and short exponentials.

The paper approach is applied on stateless networks, thus, the resource reservation is not required. Moreover, the core router will not distribute
periodically the QoS state information. So, the available resources in the network will be estimated by the incoming device by means of a Local Resource Measurement.

The remainder of this paper is organized as follows. Section two presents the effective bandwidth concept. The following section describes two methodologies to estimate the available bandwidth of a path. Section four describes admission control adjusted by effective bandwidth. Finally, conclusions and future work are presented.

2. Effective Bandwidth

Effective bandwidth is a concept to provide a measure of resource usage which adequately represents the trade-off between sources of different types, considering their statistical characteristics and QoS requirements. The definition of the effective bandwidth associated with a source is [S]:

\[
\alpha(s,t) = \frac{1}{st} \log \left[ \frac{1}{1 + p(1 - \exp(st \alpha(s,t)))} \right] 
\]

Where \( X[0,t] \) is the amount of work that arrives from a source in the interval \([0,t]\). Assume that \( X[0,t] \) has stationary increments.

There are different kinds of traffic sources: voice, video, WWW. Consequently, the effective bandwidth should be computed according to the traffic characteristics.

Traffic models (previously reported in literature) could be used to compute the effective bandwidth of individual sources. They are summarized by the following table:

<table>
<thead>
<tr>
<th>Traffic</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice</td>
<td>On-Off source / two states markovian chain [6].</td>
</tr>
<tr>
<td>Video</td>
<td>Four-states Markov-Modulated transform-expand sample (TES) traffic model with TES durations [7]. Markov-Modulated Gamma-based [8].</td>
</tr>
<tr>
<td>WWW</td>
<td>On-Off source with Poisson session arrivals and paretian distributions [9],[10].</td>
</tr>
</tbody>
</table>

The effective bandwidth of an on-off source model (for an individual application) is given by the following equation:

\[
\alpha(s,t) = \frac{1}{st} \log \left[ \frac{1}{1 + p(1 - \exp(st \alpha(s,t)))} \right] 
\]

Where \( \alpha(s,t) \) is the effective bandwidth of the on state and \( p \) is the proportion of time spent in the on state. The mean and peak of the source are represented by \( M \) and \( h \) respectively, so, \( \alpha(s,t) = h \) and \( p=M/h \). Then the available bandwidth is represented by:

\[
\alpha_{st}(s,t) = \frac{1}{st} \log \left[ \frac{1}{1 + \frac{h}{M} (1 - \exp(st \alpha(s,t)))} \right] 
\]

Figures 1 and 2 present graphical examples (obtained by Matlab) of the On/Off source effective bandwidth described by Eq. 3, where \( M \) and \( h \) are set to 1 and 100 workload units respectively.

Since \( X[0,t] \) is the amount of work (load) produced by the source over an interval of length \( t \). Therefore, if the time interval is extended then the effective bandwidth is increased. This effect is described by figure 1.

Figure 2 shows S-axis evolution of the On/Off Source Effective Bandwidth. When \( s \) is approaching zero, then, the effective bandwidth becomes the mean load produced by the source. If \( s \) value is moving toward one, then \( h \) is approaching the resource capacity. Consequently, the effective bandwidth is increasing.

![Figure 1. Time evolution of the On-Off source Effective Bandwidth](image)
3. Path available bandwidth

Idealized proportional routing model assume that all paths between a source and a destination are disjoint. However, IP networks paths have shared links.

The path's available bandwidth concept is similar to the concept of virtual capacity [4] since both are modeling a direct virtual link with a certain amount of path's capacity. However, the latest concept is also related to the corresponding blocking probability observed by the source. The proposed framework makes use of available bandwidth estimation by a local mechanism. There are some available bandwidth measurement methods used as local mechanisms such as: Self-Loading Periodic Streams (SLoPS) [11], and Trains of Packet Pairs (TOPP) [12].

The SLoPS methodology involves monitoring variations in the one-way delays of the probing packets. If the stream rate R is higher than the path's available bandwidth A, the stream causes a temporary overload in the queue of the tight link. Consequently One-way delay of probing packets will be increased (caused by queuing delay). On the other hand, if the stream rate R is lower than the available bandwidth A, the probing packets will go through the path avoiding accumulation at the tight link (without increasing One-way delay).

TOPP sends many packet pairs at gradually increasing rates from the source to the sink assume a packet pair is sent from the source with initial dispersion $\Delta_0$. The probing packets have a size of S bytes; thus, the offered rate of the packet pair is $T_{O}=S/\Delta_0$. If $T_O$ is greater than the end-to-end available bandwidth A, then, the second probing packet will be queued behind the first probing packet, and the measured rate at the receiver will be $T_R < T_O$. On the other hand, if $T_O < A$, TOPP assumes that the packet pair will arrive at the receiver with the same rate it had at the sender (i.e., $T_R = T_O$).

Since the available bandwidth estimation of a path depends only on local statistics at a source (for some methods the destination also), there is no requirement of any global QoS state information exchange.

4. Admission control adjusted by effective bandwidth

The proposed framework depicted by figure 1. shows the link sharing of some source-destination pairs (e.g. S1-D1 and S2-D2). There are two types of devices: the wrx and the incoming routers. The core routers function is the same as in current networks: to forward packets to the best route. This approach is applied on stateless networks, thus, the resource reservation is not necessary. Moreover, the core router will not distribute periodically the exchange of QoS state information. However, the Incoming devices should perform the following functions:

- Computation of application's effective bandwidth.
- Estimation of the available bandwidth by local methods.
- Comparison of the available resources versus the application requirements.

1) Computation of the effective bandwidth and register the application/user requirements

The definition of the effective bandwidth associated with a general source is described in section two. It is important to mention that the effective bandwidth is associated with resources and also to sources constrains (requirements such as: delay and jitter) related to physical and logical network resources. If a delay requirement associated to a source (in this case a flow) becomes weak, then effective bandwidth is the mean load that should not exceed the capacity of the resource. Our proposal is related to the effective bandwidth without considering delay constrains.
There are different kinds of traffic sources. Therefore, the incoming device should compute the effective bandwidth of each flow based on the kind of traffic it belongs, and its specific requirements, (described on the IP header Type of Service field and the Service Level Agreement respectively).

Connection acceptance control is primarily concerned with expectations of future QoS. An approach of connection acceptance control mechanisms is to make use of prior declarations (assessed by the user or the network) and empirical averages. The idea proposed in [5] is the use of prior declarations to choose a linear function that bounds the effective bandwidth.

2) Estimation of the available bandwidth

This framework is applied to a connectionless network (IP network), with link sharing of some source-destination pairs (showed by figure 3). Therefore, the estimation of available bandwidth is mandatory.

The available bandwidth estimation on this framework is not limited to a particular method, we propose the use of very popular methods: TOPP and SLoPS (described in section three). Differences between TOPP and SLoPS as methodologies to estimate available bandwidth are related to the statistical processing of the measurements.

Most of measurement methods, and consequently most of tools related, still have open research. Some of these problems are: cross traffic, and uncertainties in host synchronization. Besides, the no consideration of some network devices such as: proxies, firewalls or switches, causes delay subestimation as well as unexpected packet drop.

3) Comparison of the available resources versus application requirements

If the path available bandwidth is greater than the application bandwidth requirement the flow will be accepted, otherwise rejected. It is important to observe that the consideration of QoS parameters such as packet loss and delay is out of the scope of this paper.

5. Conclusions and Future Work

This paper describes a framework for admission control on a packet-based Internet environment. This approach is applied on stateless networks, thus the resource reservation is not required and consequently the Quality of service is not guaranteed. However, the proposed framework improves the network QoS respect to the current Internet scheme.

The main advantage of our proposed framework is the system scalability, because there are no extra-mechanisms to be implemented at the core routers rather than the incoming ones. However, the admission mechanism avoids the QoS degradation of previously accepted applications on the network path.

It is important to remark that the use of the effective bandwidth term is applied to flows (non-aggregated applications). The use of this concept on totally positively skewed stable process is left to future work applied to groups of flows including the consideration of heavy tail behavior.

As future work, is left the implementation of this admission control proposal in a test-bed, utilizing a modification of the SLoPS measurement method, based in the available bandwidth of a single flow. Work in progress concerns to the degradation of QoS characteristics such as delay and packet loss, which is not considered on this work.

5. Acknowledgements

We thank the Automous University of Baja California as well as the PROMEP for their support in this research work.

6. References


