Diffserv's Assured Forwarding PHB: What Assurance does the Customer Have?

B. Nandy, N. Seddigh, P. Pieda
Computing Technology Lab, Nortel Networks
Ottawa, Canada
Email:{bnandy, nseddigh, ppieda}@nortelnetworks.com

ABSTRACT: This paper examines issues related to assuring the throughput performance parameters of a Service Level Agreement in an Assured Forwarding based Diffserv-capable IP network. In an effort to understand the kinds of quantifiable parameters that can be included in a contract, the analysis of detailed studies on throughput are presented. Seven different factors are discussed which can bias bandwidth assurance for equal paying customers. Design options for various Traffic Conditioning schemes at the edge of the network to mitigate the effect of these factors are also discussed. The scope of network for which such schemes are viable is assessed. There are open issues that need to be addressed before such schemes can achieve their desired effect on the general Internet.

1.0 Introduction

Traditional IP networks offer users best-effort service. In this model, all packets compete equally for network resources. However, this best-effort service cannot provide any predictability and reliability in end-to-end packet delivery, making it unsuitable for real-time and business deemed mission-critical applications. This requires the ability to provide Quality of Service (QoS) i.e., to offer service differentiation based on the requirements of users and applications.

The Differentiated Services (Diffserv) architecture [1] has recently become the preferred method to address QoS issues in IP networks. This packet marking based approach to IP-QoS is attractive due to its simplicity and ability to scale. An end-to-end differentiated service is obtained by concatenation of per-domain services and Service Level Agreements (SLAs) between adjoining domains along the path that the traffic crosses in going from source to destination. Per domain services are realized by traffic conditioning at the edge and simple differentiated forwarding mechanisms at the core of the network. Two of the more popular proposed forwarding mechanisms are Expedited Forwarding (EF)[3] and Assured Forwarding (AF)[4] Per Hop Behavior (PHB). Traffic conditioning includes classification, metering, policing and shaping.

The basic concept of AF-based services is appealing as it proposes simple mark and drop mechanisms to realize IP QoS. The AF approach will provide better than best-effort service by controlling the drop preference of packets at the time of congestion. AF provides an interesting alternative that may enable service offerings at lesser cost for audio, video, web and other applications.

The AF PHB draft proposes four classes and three-drop preferences per class. AF is an extension of the RIO scheme [5], which uses single FIFO queue and two-drop preferences. Most of the current studies of differentiated drop mechanisms are based on the RIO approach. The same issues are applicable to IETF's AF proposal. The discussion of this paper assumes a RIO-like framework.

The basis of the RIO mechanism is RED-based (Random Early Detect) differentiated dropping of packets during congestion at the router. In RIO, traffic profiles for end-users are maintained at the edge of the network. When user traffic exceeds the contracted target rate, their packets are marked out-of-profile. Otherwise, packets are marked in-profile. The RIO scheme utilizes a single queue. All user packets are directed to and serviced from the same queue.

Two sets of RED thresholds are maintained, one each for in-profile and out-of-profile. Two separate average buffer occupancy calculations are tracked, one for in-profile packets and one for in-profile plus out-of-profile packets. The possibility of dropping in-profile packets depends only on the buffer occupancy of in-profile packets while the possibility of dropping out-of-profile packets depends on the buffer occupancy of in-profile plus out-of-profile packets. This scheme gives the appearance of two coupled virtual queues within a physical queue.

Although the IETF Diffserv Working Group is not exploring service related issues, we argue that it is necessary to examine, evaluate and understand the kinds of end-to-end services that could be created for an end user using AF-like PHB. Questions include: Can the SLA contracts have quantitative assurances of any form? If so, what are the SLA parameters that can be assured? This paper examines the throughput assurance issues and assesses the kind of quantitative assurance that can be given in a SLA contract. The issues with two other typical performance parameters: delay and latency are also discussed briefly.

2.0 Throughput Assurance Issues

Three common performance parameters included as part of a SLA include: throughput, drop probability, and latency. In Section 3, the possibility of providing quantitative assurances for each of these parameters is discussed. To better understand difficulties with quantitative assurances in an AF-like framework, this section focuses on issues related to one of these parameters – throughput. Seven factors are identified that affect a provider's ability to contract soft bandwidth guarantees. Results of a de-

tailed implementation-based study on five of these factors are reported in a recently submitted paper [7]. The seven factors are:

Impact of Round Trip Time (RTT): Since TCP utilizes a self-clocked sliding window based mechanism, any bandwidth guarantee is a function of RTT. Aggregate flows with different RTTs, despite having identical target rates, will get different shares of the bandwidth. For over-provisioned networks [7][10], the flow aggregates will achieve their target rate irrespective of their RTTs. This is because in-profile traffic is protected and out-profile traffic will be dropped before any in-profile packets are dropped. However, there will be an unfair sharing of the excess bandwidth in favour of those target aggregates with lower RTTs. In the under-provisioned case, neither of the aggregated flows will achieve their target. However, the high RTT flows will be further away from the target than the flows with low RTT.

Large Number of Active Flows: The total number of active flows in the core of the network and the buffer allocation plays an important role [6] in determining the TCP throughput for individual flows as well as flow aggregates. With an increased number of active micro-flows in the core of the network, TCP throughput of a single flow will fluctuate. The effectiveness of RED parameters is partially dependent on the number of active flows. Large number of active TCP flows will cause the queue length to cross the RED *maxth* value and drop multiple packets causing timeout. This will also lead to unfair sharing of TCP bandwidth. Engineering of RED parameters are key to this problem. For a given set of RED parameters, the end-to-end TCP flow behavior will change as the number of active flows will be changing with time.

Impact of Non-responsive Flows There are two cases [7][9]: (a) responsive (TCP) and non-responsive (UDP) flows share the same class with identical drop precedence and (b) flows share the same class with different drop precedence. In the first case, non-responsive flows can starve the responsive flows in an under-provisioned network; but the responsive flows will reach the target rate for over-provisioned network. In the second case, the responsive flows can be protected from the non-responsive flows. It should be also noted that if UDP and TCP share the same physical queue, UDP traffic is susceptible to delay variation.

Size of Target Rate: In an over-provisioned network, the excess bandwidth will get distributed equally irrespective of the target rate[7]. This may not be an acceptable solution, as the customer with high target rate will expect a higher share of the excess bandwidth. It will be interesting to observe the degradation in bandwidth share among flows aggregated with different target rates in under-provisioned network.

Number of Flows in an Aggregate: This is of interest since the service agreements will be on aggregated traffic and various business houses will contract a target rate with a service provider. It is possible that some organization will have thousands of flows in a target aggregate while others will have hundreds of flows. In an over-provisioned network, the aggregate with large number of flows will get more share of the excess bandwidth [7].

Variation in TCP Stack: Different TCP stacks like Reno, SACK and new Reno have different ways of handling packet drops [8]. This causes different levels of aggressiveness to maintain throughput in the face of a packet drop. Thus, two users with same packet drop probability but different TCP stacks can obtain different throughput.

Variation in Packet Size: Flows with same RTT but different packet sizes will achieve target rate but the excess bandwidth will be split unfairly in an over-provisioned network [7].

3.0 SLA Performance Parameters

In Diffserv-capable IP networks, a typical Service Level Agreement [2] will be specified in two parts: (a) Traffic Conditioning related components i.e., service performance parameters, scope of the service etc. and (b) general service characteristics like availability, pricing and billing etc. Some of the expected service performance parameters by the customers are throughput, drop probability and latency. In this section, the possibilities of providing quantitative throughput assurances are discussed in detail. The possibilities of drop and latency assurances are briefly discussed.

Throughput:
As discussed in the previous section, for over-provisioned networks, the target rates of aggregated flows are achievable irrespective of most of the issues. However, the degree to which excess bandwidth is fairly distributed depends a great deal on various factors. In fact, it is questionable if the SLA should attempt to specify any quantitative distribution guideline for the excess bandwidth. Discussion in Section 2.0 also indicates that as the network approaches an under-provisioned state, various factors play an important role in determining the extent to which aggregates of TCP flows achieve or don't achieve their target rates. The tendency to maximize profit and minimize over-provisioning will cause hotspots in the network. While tools are being developed to assist with network management and provisioning issues in a Diffserv-capable network, given the one-to-anywhere nature of Internet services, it would be prudent for network providers to assume that periodic, sustained under-provisioning will occur. The approaches taken by various researchers to mitigate the impacts of some factors on throughput are outlined in next paragraph.

Various packet marking schemes have been proposed [5][10][11][12] for Diffserv but the suitability of these schemes are yet to be studied. The marker based on an Average Rate Estimator [5] at the traffic conditioner decides which packet is to be marked for higher drop preference depending on if the target rate has been exceeded or not. This marking scheme is suitable for an average target rate. Yeom and Reddy [10] has suggested an algorithm that improves fairness among the individual flows with different RTTs within an aggregation. The marking scheme based on two rates and three colour [11] is suitable for a service where peak rate is enforced separately from the committed information rate. This packet-marking scheme is suitable to handle the inherent burstiness of TCP sources. Kim [12] proposes a means by which individual flows within an aggregate can fairly share the target rate for that aggregate.

We suggest that some of the factors can be mitigated by proper traffic conditioning at the edge of the network. For example: (a) To alleviate the impact of RTT on the flow bandwidth, the flow RTTs can be tracked and differential drop technique can be used to compensate for the high RTT flows. In fact, if the RTTs for flows in the same customer aggregate are different then per-flow RTT measurement is required. (b) To alleviate a large throughput fluctuation due to large number of active TCP flows, an admission control mechanism based on active flow count is necessary. All these solutions will require per flow (or per policy) queuing and state tracking at the edge of the network.

The effect of unfair sharing caused by factors such as the packet size and size of the target rate can be mitigated through intelligent traffic conditioners. However, the techniques used would imply that each edge router would have knowledge about these different factors at other edge routers. Thus, some form of communication is required between the Traffic Conditioners. However, the solutions can be complex and raise scalability concerns.

It is important to understand the scope (i.e., topological extent) of the service. Various traffic conditioning may be feasible for all traffic between an ingress point and an egress point or a set of egress points. Moreover, some of the solutions are feasible at a local node but a global view required for end-to-end service. This will require communication among Traffic Conditioners and extensive measurements and state tracking. Further study is needed to address the scalability issues with this approach. It appears as though SLA contracts should try to avoid specifying sharing of excess bandwidth in an over-provisioned network as well as stating how bandwidth would be distributed in an under-provisioned network.

Drop Probability:
Since the Diffserv framework is intended to operate on bilateral agreements between two neighboring domains, an owner of a domain can obtain a service level agreement with its neighbor on drop probability. For example, the agreement may assure three different packet drop probabilities depending on committed rate, excess rate . If the bandwidth usage by the aggregated flow is less than committed rate, 2% packets will be dropped; if it is between committed and excess rate, 5% packets will be dropped and beyond that 10% packets will be dropped. Such drop assurance may be possible in a single domain, but it is not clear how the end-to-end drop probabilities (as specified in SLAs) of aggregate flows going anywhere in the Internet across multiple domain can be ensured for a customer.

Latency:
Guarantees on end-to-end delay bounds for VPNs (between two fixed pre-determined points) crossing multiple domains remain a possibility. However, it implies each domain must have a clear idea of delay patterns for AF traffic within the network. The filter for that VPN and its profile must be available at the ingress border router for each domain on the router from source to destination. A protocol is needed for inter-domain communication of such information. A one-to-anywhere service cannot by definition supply a delay bound since the total delay is a factor of the delays incurred in variable number of intermediate domains.
Within a particular router, specification in relative packet forwarding urgency among the classes can be controlled by appropriate scheduling mechanism among the queues.

4.0 Conclusions

This paper focuses on technical challenges with providing quantitative assurances for performance parameters in Assured Forwarding based Diffserv-capable IP networks. Three performance parameters are examined. One of the key performance parameters – throughput – is examined in detail and results presented to show how seven different factors can affect the throughput guarantee. The discussion shows that in over-provisioned networks, while target rates are achievable irrespective of the seven factors, there is unfair sharing of excess bandwidth for equal paying customers. Further, in under-provisioned networks, there is unfair degradation for equal paying customers. If clauses for such scenarios need to be included in SLAs, further study is required to develop solutions to mitigate the effects of such factors. The issues with excess bandwidth sharing can be addressed by intelligent Traffic Conditioning at the edge. However, the solutions may be complex, not scalable and not applicable to Internet (one-to-any network). The paper finally points to further work required to address issues with end-to-end quantitative guarantees for the other two performance parameters – packet drop and latency.

5.0 Acknowledgements

We would like to thank Alan Chapman for his comments on the initial version of this draft and Jamal Hadi Salim

for various discussions during diffserv prototype development.

6.0 References

[1] Blake, S. Et al, "An Architecture for Differentiated Services", RFC 2475, December 1998

[2] Bernet, Y. Et al, “A Framework for Differentiated Services”, Internet Draft, <draft-ietf-diffserv-framework-02.txt>, February 1999.

[3] Jacobson, V. Et al, “An Expedited Forwarding PHB”, Internet Draft <draft-ietf-diffserv-phb-ef-01.txt>, November 1998

[4] Heinanen J., Baker F., Weiss W., and Wroclawski J., “Assured Forwarding PHB Group”, Internet Draft, <draft-ietf-diffserv-af-05.txt>, February 1999.

[5] Clark D. and Fang W., “Explicit Allocation of Best Effort Packet Delivery Service”, http://diffserv.lcs.mit.edu/exp-alloc-ddc-wf.ps, 1998

[6] Morris, R., “TCP Behavior with Many Flows”, IEEE International Conference on Network Protocols, October 1997, Atlanta, Georgia

[7] Seddigh, N., Nandy, B., Pieda, P, “Bandwidth Assurance Issues for TCP flows in a Differentiated Services Network”, submitted to Globecom’99.

[8] Fall, K. and Floyd, S., “Simulation Based comparisons of Tahoe, Reno and SACK TCP”, Computer Communication Review, 26(3), July 1996.

[9] Ibanez J, Nichols K., “Preliminary Simulation Evaluation of an Assured Service”, Internet Draft, draft-ibanez-diffserv-assured-eval-00.txt>, August 1998

[10] Yeom, I and Reddy N, "Impact of marking strategy on aggregated flows in a diff-serv network," submitted to IWQoS’99, http://dropzone.tamu.edu/~ikjun/papers.html

[11] Heinanen, J and Guerin, R, “A Two Rate Three Color marker”, <draft-heinanen-diffserv-trtcm-00.txt>, March 1999.

[12] Kim H, “A Fair Marker”, <draft-kim-fairmarker-diffserv-00.txt>, Internet draft, April 1999