

Internet Engineering Task Force
Internet Draft
Expires: January 2000

Mukul Goyal
Arian Duresi
Raj Jain
Chunlei Liu

Raj Jain is now at Washington University in Saint Louis, jain@cse.wustl.edu <http://www.cse.wustl.edu/~jain/>

Effect of Number of Drop Precedences in Assured Forwarding

Status of Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet- Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet- Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This document is not to be taken as a finished product. Some of the sections are rough and are included in order to obtain comments from the community that will benefit future iterations of this document. This is simply a step in the ongoing conversation about this document.

Abstract

This informational draft explores the issue of fair allocation of excess network bandwidth between congestion sensitive and insensitive flows in an Assured Forwarding traffic class. In the absence of any mechanism to distinguish between out-of-profile traffic of congestion sensitive and insensitive flows, congestion insensitive flows will get most of the excess network bandwidth. However, if out-of-profile packets of congestion sensitive and insensitive flows are 'colored' differently, network can be configured so as to give better treatment to excess packets of congestion sensitive flows and achieve fair allocation of excess network bandwidth. With a view to clearly distinguish between out-of-profile packets of congestion sensitive and insensitive flows, three levels of drop precedence are required. However, if the network operates close to its capacity, three levels of drop precedence are redundant as there is not much excess bandwidth to be

shared.

This is a revised and comprehensive version of our previous study presented at the March 1999 IETF Meeting at Minneapolis. The ps and pdf versions of this document with all the figures are available at: <http://www.cis.ohio-state.edu/jain/ietf/dpstd2.htm>

1 Introduction

Differentiated Services (DS) aims to provide *scalable service differentiation in the Internet* that can be used to permit *differentiated pricing of Internet service* [1]. This differentiation may either be quantitative or relative. DS is scalable as traffic classification and conditioning is performed only at network boundary nodes. The service to be received by a traffic is marked as a code point in the DS field in the IPv4 or IPv6 header. The DS code point in the header of an IP packet is used to determine the Per-Hop Behavior (PHB), i.e. the forwarding treatment it will receive at a network node. Currently, formal specification is available for two PHBs - Assured Forwarding [2] and Expedited Forwarding [3]. In Expedited Forwarding, a transit node uses policing and shaping mechanisms to ensure that the maximum arrival rate of a traffic aggregate is less than its minimum departure rate. At each transit node, the minimum departure rate of a traffic aggregate should be configurable and independent of other traffic at the node. Such a per-hop behavior results in minimum delay and jitter and can be used to provide an end-to-end 'Virtual Leased Line' type of service.

In Assured Forwarding (AF), IP packets are classified as belonging to one of four traffic classes. IP packets assigned to different traffic classes are forwarded independent of each other. Each traffic class is assigned a minimum configurable amount of resources (link bandwidth and buffer space). Resources not being currently used by another PHB or an AF traffic class can optionally be used by remaining classes. Within a traffic class, a packet is assigned one of three levels of drop precedence (green,yellow,red). In case of congestion, an AF-compliant DS node drops low precedence (red) packets in preference to higher precedence (green, yellow) packets.

Multiple levels of drop precedence can be used to mitigate the effect of round-trip time on TCP flows [4] and to achieve fair allocation of excess network bandwidth among congestion sensitive TCP and insensitive UDP flows. In this study, we perform wide ranging simulations with two and three levels of drop precedence (or colors) in order to understand the factors influencing fair allocation of excess network resources among congestion sensitive and insensitive flows.

2 Simulation Configuration and Parameters

The simulations performed in this study use the network configuration shown in Figure 1. Here, customers 1 through 10 send data over the link between Routers 1 and 2 using the same AF traffic class. Traffic is one-

dimensional with only ACKs coming back from the other side. Customers 1 through 9 carry an aggregated traffic coming from 5 Reno TCP sources. Customer 10 gets its traffic from a single UDP source sending data at a rate of 1.28 Mbps. Common configuration parameters are detailed in Table 1. All TCP and UDP packets are marked green at the source before being 'recolored' by a traffic conditioner at the customer site. The traffic conditioner consists of two 'leaky' buckets (green and yellow) that mark packets according to their token generation rates (called reserved/green and yellow rate). In two color simulations, yellow rate of all customers is set to zero. Thus, in two color simulations, both UDP and TCP packets will be colored either green or red. In three color simulations, customer 10 (the UDP customer) always has a yellow rate of 0. Thus, in three color simulations, TCP packets coming from customers 1 through 9 can be colored green, yellow or red and UDP packets coming from customer 10 will be colored green or red. All the traffic coming to Router 1 passes through a Random Early Drop (RED) queue. The RED policy implemented at Router 1 can be classified as *Single Average Multiple Threshold RED* as explained in next section.

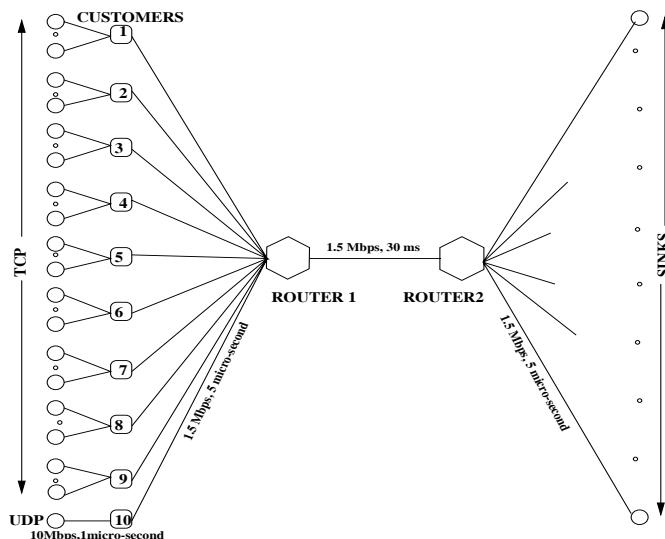


Figure 1: Simulation Configuration

We have used NS simulator version 2.1b4a [5] for these simulations. The code has been modified to implement the traffic conditioner and multi-color RED (RED.n).

3 Calculating Random Early Drop (RED) Average Queue Length

In RED, the drop probability of a packet depends on the average queue length which is an exponential average of instantaneous queue length at the time of the packet's arrival [6]. The drop probability increases linearly from 0 to max_p as average queue length increases from min_th to max_th . With packets of multiple colors, one can calculate average queue length in many ways and have multiple sets of drop thresholds for packets of different colors. In general, with multiple colors, RED policy can be implemented as a variant of

Table 1: Simulation Configuration Parameters

| | | | |
|--------------------------------------|---------------|--------------------------------------|----------------|
| Simulation Time | 100 s | Link between Customers and Router 1: | |
| TCP Window | 64 packets | Link Bandwidth | 1.5 Mbps |
| IP Packet Size | 576 bytes | One Way Delay | 5 microseconds |
| UDP Rate | 1.28 Mbps | Drop Policy | DropTail |
| Queue Size (for all queues) | 60 packets | Link between Router 1 and Router 2: | |
| Link between Customers and UDP/TCPs: | | Link Bandwidth | 1.5 Mbps |
| Link Bandwidth | 10 Mbps | One Way Delay | 30 miliseconds |
| One Way Delay | 1 microsecond | From Router 1 To Router 1 | |
| Drop Policy | DropTail | Drop Policy | RED_n DropTail |
| Link between Router 2 and Sinks: | | Link Bandwidth | 1.5 Mbps |
| | | One Way Delay | 5 microseconds |
| | | Drop Policy | DropTail |

one of four general categories:

- *Single Average Single Threshold (SAST)*
- *Single Average Multiple Thresholds (SAMT)*
- *Multiple Average Single Threshold (MAST)*
- *Multiple Average Multiple Thresholds (MAMT)*

Single Average Single Threshold RED has a single average queue length and same *min th* and *max th* thresholds for packets of all colors. Such a policy does not distinguish between packets of different colors and can also be called *color blind* RED.

In Single Average Multiple Thresholds RED, average queue length is based on total number of packets in the queue irrespective of their color. However, packets of different colors have different drop thresholds. For example, if maximum queue size is 60 packets, the drop thresholds for green, yellow and red packets can be {40/60, 20/40, 0/10}. In these simulations, we use Single Average Multiple Thresholds RED.

In Multiple Average Single/Multiple Threshold RED, average queue length for packets of different colors is calculated differently. For example, average queue length for a color can be calculated using number of packets in the queue with same or better color [4]. In such a scheme, average queue length for green, yellow and red packets will be calculated using number of green, yellow + green, red + yellow + green packets in the queue respectively. Another possible scheme is where average queue length for a color is

Table 2: Two Color Simulation Parameters

| Simulation ID | Green Rate | Max Drop Probability {Green,Red} | Drop Thresholds {Green,Red} | Green Bucket Size (in Packets) |
|---------------|------------|-------------------------------------|--------------------------------|-----------------------------------|
| 1-144 | 12.8 kbps | {0.1,0.1} | {40/60,0/10} | 1 |
| 201-344 | 25.6 kbps | {0.1,0.5} | {40/60,0/20} | 16 |
| 401-544 | 38.4 kbps | {0.1,1} | {40/60,0/5} | 2 |
| 601-744 | 76.8 kbps | {0.5,0.5} | {40/60,20/40} | 32 |
| 801-944 | 102.4 kbps | {0.5,1} | | 4 |
| 1001-1144 | 128 kbps | {1,1} | | 8 |
| 1201-1344 | 153.6 kbps | | | |
| 1401-1544 | 179.2 kbps | | | |

calculated using number of packets of that color in the queue [7]. In such a case, average queue length for green, yellow and red packets will be calculated using number of green, yellow and red packets in the queue respectively. Multiple Average Single Threshold RED will have same drop thresholds for packets of all colors whereas Multiple Average Multiple Threshold RED will have different drop thresholds for packets of different colors.

4 Experimental Design

In this study, we perform full factorial simulations involving many factors:

- **Green Traffic Rates:** Green traffic rate is the token generation rate of green bucket in the traffic conditioner. We have experimented with green rates of 12.8, 25.6, 38.4 and 76.8 kbps per customer. These rates correspond to a total of 8.5%, 17.1%, 25.6% and 51.2% of network capacity (1.5 Mbps). In order to understand the effect of green traffic rate, we also conduct simulations with green rates of 102.4, 128, 153.6 and 179.2 kbps for two color cases. These rates correspond to 68.3%, 85.3%, 102.4% and 119.5% of network capacity respectively. Note that in last two cases, we have oversubscribed the available network bandwidth.
- **Green Bucket Size:** 1, 2, 4, 8, 16 and 32 packets of 576 bytes each.
- **Yellow Traffic Rate (only for three color simulations):** Yellow traffic rate is the token generation rate of yellow bucket in the traffic conditioner. We have experimented with yellow rates of 12.8 and 128 kbps per customer. These rates correspond to 7.7% and 77% of total capacity (1.5 Mbps) respectively. We used a high yellow rate of 128 kbps so that all excess (out of green rate) TCP packets are colored yellow and thus can be distinguished from excess UDP packets that are colored red.

Table 3: Three Color Simulation Parameters

| Simulation ID | Green Rate | Max Drop Probability {Green,Yellow,Red} | Drop Thresholds {Green,Yellow,Red} | Yellow Rate | Bucket Size (in Packets) | |
|---------------|------------|--|---------------------------------------|-------------|-----------------------------|--------|
| | | | | | Green | Yellow |
| 1-720 | 12.8 kbps | {0.1,0.5,1} | {40/60,20/40,0/10} | 128 kbps | 16 | 1 |
| 1001-1720 | 25.6 kbps | {0.1,1,1} | {40/60,20/40,0/20} | 12.8 kbps | 1 | 16 |
| 2001-2720 | 38.4 kbps | {0.5,0.5,1} | | | 2 | 2 |
| 3001-3720 | 76.8 kbps | {0.5,1,1} | | | 32 | 32 |
| | | {1,1,1} | | | 4 | 4 |
| | | | | | 8 | 8 |

- Yellow Bucket Size (only for three color simulations): 1, 2, 4, 8, 16, 32 packets of 576 bytes each.
- Maximum Drop Probability: Maximum drop probability values used in the simulations are listed in Tables 2 and 3.
- Drop Thresholds for red colored packets: The network resources allocated to red colored packets and hence the fairness results depend on the drop thresholds for red packets. We experiment with different values of drop thresholds for red colored packets so as to achieve close to best fairness possible. Drop thresholds for green packets have been fixed at {40,60} for both two and three color simulations. For three color simulations, yellow packet drop thresholds are {20,40}.

In these simulations, size of all queues is 60 packets of 576 bytes each. The queue weight used to calculate RED average queue length is 0.002. For easy reference, we have given an identification number to each simulation (Tables 2 and 3). The simulation results are analyzed using ANOVA techniques [8] briefly described in Section 7.

5 Performance Metrics

Simulation results have been evaluated based on utilization of reserved rates by the customers and the fairness achieved in allocation of excess bandwidth among different customers.

Utilization of reserved rate by a customer is measured as the ratio of *green* throughput of the customer and the reserved rate. Green throughput of a customer is determined by the number of green colored packets received at the traffic destination(s). Since in these simulations, the drop thresholds for green packets are kept very high in the RED queue at Router 1, chances of a green packet getting dropped are minimal and ideally green throughput of a customer should equal its reserved rate.

The fairness in allocation of excess bandwidth among n customers sharing a link can be computed using the following formula [8]:

$$Fairness\ Index = \frac{(\sum x_i)^2}{n \times \sum (x_i^2)}$$

Where x_i is the *excess* throughput of the i th customer. Excess throughput of a customer is determined by the number of yellow and red packets received at the traffic destination(s).

6 Simulation Results

Simulation results of two and three color simulations are shown in Figure 2. In this figure, a simulation is identified by its *Simulation ID* listed in Tables 2 and 3. Figures 2a and 2c show the fairness achieved in allocation of excess bandwidth among ten customers for each of the two and three color simulations. Figures 2b and 2d show the utilization of reserved rate by each of ten customers for each simulation.

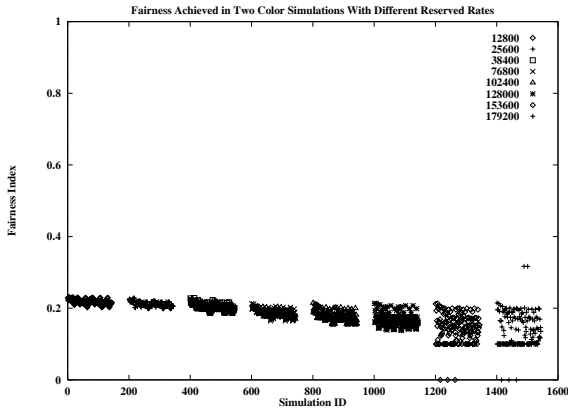
It is clear from figure 2a that fairness is not good in two color simulations. With three colors, there is a wide variation in fairness results with best results being close to 1. Note that fairness is zero in some of the two color simulations. In these simulations, total reserved traffic uses all the bandwidth and there is no excess bandwidth available to share.

As shown in Figures 2b and 2d, there is a wide variation in reserved rate utilization by customers in two and three color simulations. Figure 3 shows the reserved rate utilization by TCP and UDP customers. For TCP customers, we have plotted the average reserved rate utilization in each simulation. Note that in some cases, reserved rate utilization is slightly more than one. This is because token buckets are initially full which results in all packets getting green color in the beginning. Figures 3b and 3d show that UDP customers have good reserved rate utilization in almost all cases. In contrast, TCP customers show a wide variation in reserved rate utilization.

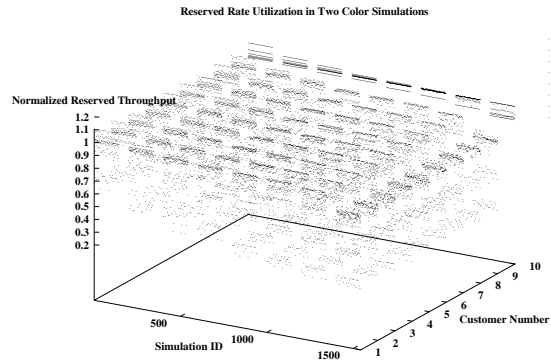
In order to determine the influence of different simulation factors on the reserved rate utilization and fairness achieved in excess bandwidth distribution, we analyze simulation results statistically using Analysis of Variation (ANOVA) technique. The next section gives a brief introduction to ANOVA technique used in the analysis. In later sections, we present the results of statistical analysis of two and three color simulations.

7 Analysis Of Variation (ANOVA) Technique

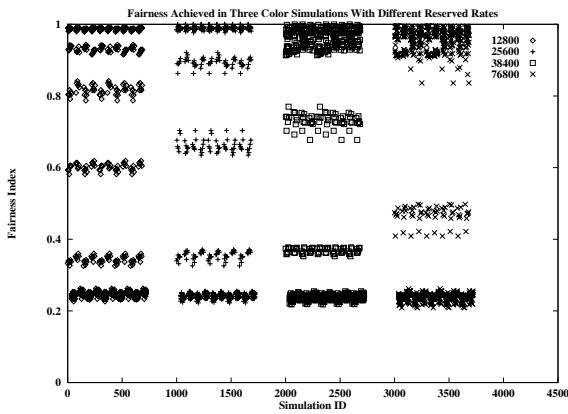
The results of a simulation are affected by the values (or *levels*) of simulation factors (e.g. green rate) and the *interactions* between levels of different factors (e.g. green rate and green bucket size). The simulation



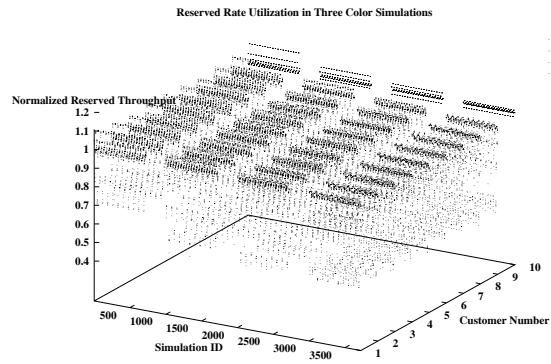
(a)



(b)



(c)



(d)

Figure 2: Simulation Results

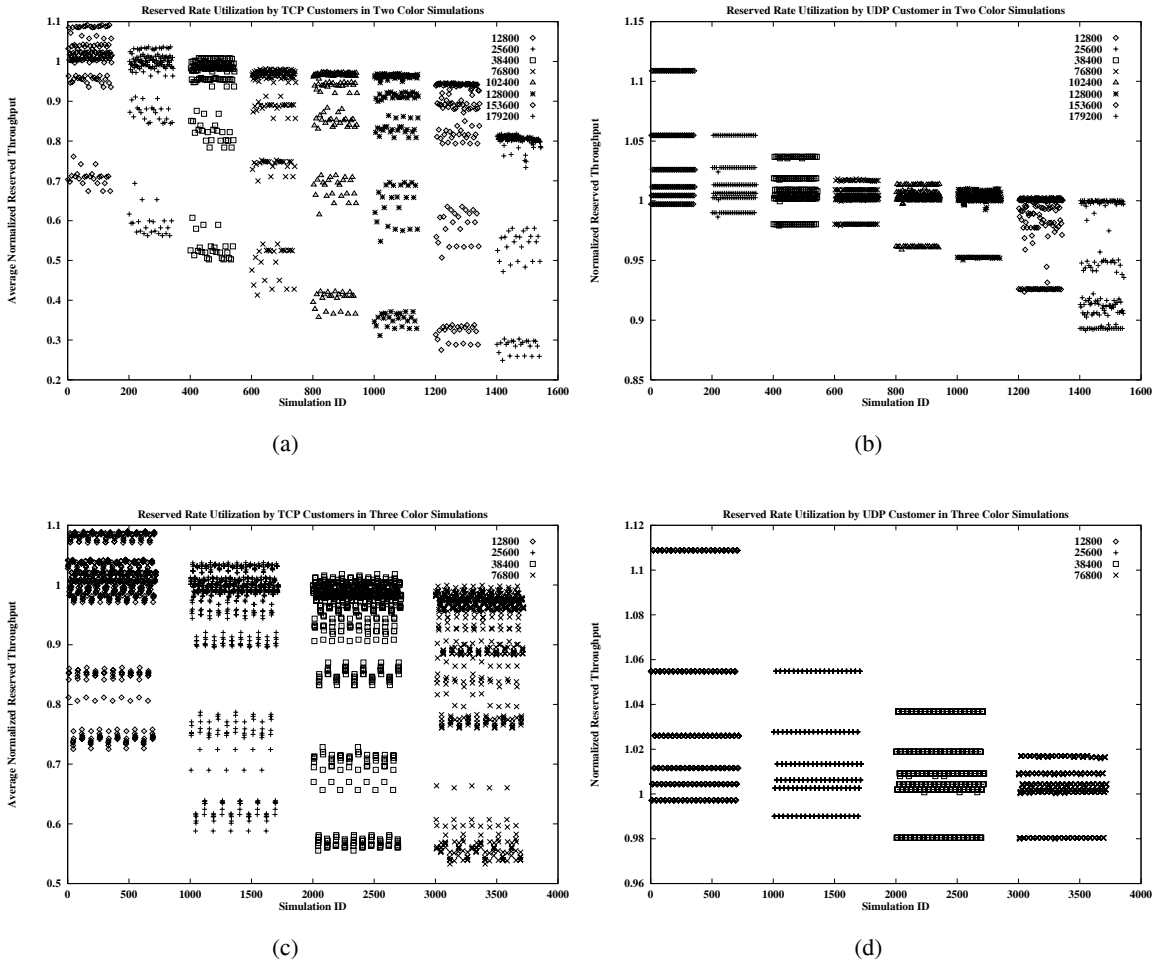


Figure 3: Reserved Rate Utilization by TCP and UDP Customers

factors and their levels used in this simulation study are listed in Tables 2 and 3. *Analysis of Variation* of simulation results is a statistical technique used to quantify these effects. In this section, we present a brief account of *Analysis of Variation* technique. More details can be found in [8].

Analysis of Variation involves calculating the *Total Variation* in simulation results around the *Overall Mean* and doing *Allocation of Variation* to contributing factors and their interactions. Following steps describe the calculations:

- Calculate the *Overall Mean* of all the values.
- Calculate the individual effect of each level a of factor A , called the *Main Effect of a* :

$$\text{Main Effect}_a = \text{Mean}_a - \text{Overall Mean}$$

where, Main Effect_a is the main effect of level a of factor A , Mean_a is the mean of all results with a as the value for factor A and *Overall Mean* is the overall mean of simulation result values.

The main effects are calculated for each level of each factor.

- Calculate the *First Order Interaction* between levels a and b of two factors A and B respectively for all such pairs:

$$\text{Interaction}_{a,b} = \text{Mean}_{a,b} - (\text{Overall Mean} + \text{Main Effect}_a + \text{Main Effect}_b)$$

where, $\text{Interaction}_{a,b}$ is the interaction between levels a and b of factors A and B respectively, $\text{Mean}_{a,b}$ is mean of all results with a and b as values for factors A and B , Main Effect_a and Main Effect_b are main effects of levels a and b respectively.

- Calculate the *Total Variation* as shown below:

$$\text{Total Variation} = \Sigma(\text{result}^2) - (\text{Num_Sims} \times (\text{Overall Mean})^2)$$

where, $\Sigma(\text{result}^2)$ is the sum of squares of all individual results and *Num_Sims* is total number of simulations.

- The next step is the *Allocation of Variation* to individual main effects and first order interactions. To calculate the variation caused by a factor A , we take the sum of squares of the main effects of all levels of A and multiply this sum with the number of experiments conducted with each level of A . To calculate the variation caused by first order interaction between two factors A and B , we take the sum of squares of all the first-order interactions between levels of A and B and multiply this sum with the number of experiments conducted with each combination of levels of A and B . We calculate the allocation of variation for each factor and first order interaction between every pair of factors.

Table 4: Main Factors Influencing Reserved Rate Utilization Results

| Factor/Interaction | Allocation of Variation (in %age) | | | |
|-----------------------------------|-----------------------------------|--------------|---------------|--------------|
| | 2 Colors | | 3 Colors | |
| | TCP Customers | UDP Cusotmer | TCP Customers | UDP Customer |
| Green Rate | 18.46% | 47.86% | 10.36% | 19.39% |
| Green Bucket Size | 77.14% | 22.58% | 81.88% | 63.22% |
| Green Rate - Green Bucket Size | 3.65% | 28.76% | 3.34% | 17.39% |

8 ANOVA Analysis for Reserved Rate Utilization Results

Table 4 shows the Allocation of Variation to contributing factors for reserved rate utilization results. As shown in figure 3, reserved rate utilization of UDP customers is almost always good for both two and three color simulations. However, in spite of very low probability of a green packet getting dropped in the network, TCP customers are not able to fully utilize their reserved rate in all cases. The little variation in reserved rate utilization results for UDP customers is explained largely by bucket size. Large bucket size means that more packets will get green color in the beginning of the simulation when green bucket is full. Green rate and interaction between green rate and bucket size explain a substantial part of the variation. This is because the definition of rate utilization metric has reserved rate in denominator. Thus, the part of the utilization coming from initially full bucket gets more weight for low reserved rate than for high reserved rates. Also, in two color simulations for reserved rates 153.6 kbps and 179.2 kbps, the network is oversubscribed and hence in some cases UDP customer has a reserved rate utilization lower than one. For TCP customers, green bucket size is the main factor in determining reserved rate utilization. TCP traffic because of its bursty nature is not able to fully utilize its reserved rate unless bucket size is sufficiently high. In our simulations, UDP customer sends data at a uniform rate of 1.28 Mbps and hence is able to fully utilize its reserved rate even when bucket size is low. However, TCP customers can have very poor utilization of reserved rate if bucket size is not sufficient. The minimum size of the leaky bucket required to fully utilize the token generation rate depends on the burstiness of the traffic.

9 ANOVA Analysis for Fairness Results

Fairness results shown in figure 2a indicate that fairness in allocation of excess network bandwidth is very poor in two color simulations. With two colors, excess traffic of TCP as well as UDP customers is marked red and hence is given same treatment in the network. Congestion sensitive TCP flows reduce their data rate in response to congestion created by UDP flow. However, UDP flow keeps on sending data at the same rate as before. Thus, UDP flow gets most of the excess bandwidth and the fairness is poor. In three

Table 5: Main Factors Influencing Fairness Results in Three Color Simulations

| Factor/Interaction | Allocation of Variation |
|---|-------------------------|
| Yellow Rate | 77.15% |
| Yellow Bucket Size | 10.78% |
| Interaction between Yellow Rate and Yellow Bucket Size | 9.85% |

color simulations, fairness results vary widely with fairness being good in many cases. Table 5 shows the important factors influencing fairness in three color simulations as determined by ANOVA analysis. Yellow rate is the most important factor in determining fairness in three color simulations. With three colors, excess TCP traffic can be colored yellow and thus distinguished from excess UDP traffic which is colored red. Network can protect congestion sensitive TCP traffic from congestion insensitive UDP traffic by giving better treatment to yellow packets than to red packets. Treatment given to yellow and red packets in the RED queues depends on RED parameters (drop thresholds and max drop probability values) for yellow and red packets. Fairness can be achieved by coloring excess TCP packets as yellow and setting the RED parameter values for packets of different colors correctly. In these simulations, we experiment with yellow rates of 12.8 kbps and 128 kbps. With a yellow rate of 12.8 kbps, only a fraction of excess TCP packets can be colored yellow at the traffic conditioner and thus resulting fairness in excess bandwidth distribution is not good. However with a yellow rate of 128 kbps, all excess TCP packets are colored yellow and good fairness is achieved with correct setting of RED parameters. Yellow bucket size also explains a substantial portion of variation in fairness results for three color simulations. This is because bursty TCP traffic can fully utilize its yellow rate only if yellow bucket size is sufficiently high. The interaction between yellow rate and yellow bucket size for three color fairness results is because of the fact that minimum size of the yellow bucket required for fully utilizing the yellow rate increases with yellow rate.

It is evident that three colors are required to enable TCP flows get a fairshare of excess network resources. Excess TCP and UDP packets should be colored differently and network should treat them in such a manner so as to achieve fairness. Also, size of token buckets should be sufficiently high so that bursty TCP traffic can fully utilize the token generation rates.

10 Conclusions

One of the goals of deploying multiple drop precedence levels in an Assured Forwarding traffic class is to ensure that all customers achieve their reserved rate and a fair share of excess bandwidth. It is assumed that combined reserved rate for all customers is less than the network capacity. Network should be configured in such a manner so that in-profile traffic (colored green) does not suffer any packet loss and is successfully delivered to the destination. The fair allocation of excess network bandwidth can be achieved only by

giving different treatment to out-of-profile traffic of congestion sensitive and insensitive flows. The reason is that congestion sensitive flows reduce their data rate on detecting congestion however congestion insensitive flows keep on sending data as before. Thus, in order to prevent congestion insensitive flows from taking advantage of reduced data rate of congestion sensitive flows in case of congestion, excess congestion insensitive traffic should get much harsher treatment from the network than excess congestion sensitive traffic. Hence, it is important that excess congestion sensitive and insensitive traffic is colored differently so that network can distinguish between them. Clearly, three colors or levels of drop precedence are required for this purpose. However, if the total reserved traffic is close to network capacity, three levels of drop precedence are redundant as there is not much excess capacity to be shared. Thus, utility of three levels of drop precedence in a traffic class depends on the proportion of reserved traffic to the total capacity.

References

- [1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, *An Architecture for Differentiated Services*, RFC 2475, December 1998.
- [2] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, *Assured Forwarding PHB Group*, RFC 2597, June 1999.
- [3] V. Jacobson, K. Nichols, K. Poduri, *An Expedited Forwarding PHB*, RFC 2598, June 1999.
- [4] D. Clark, W. Fang, *Explicit Allocation of Best Effort Packet Delivery Service*, IEEE/ACM Transactions on Networking, August 1998.
- [5] NS Simulator, Version 2.1 , Available from <http://www-mash.cs.berkeley.edu/ns>.
- [6] S. Floyd, V. Jacobson, *Random Early Detection Gateways for Congestion Avoidance*, IEEE/ACM Transactions on Networking, August 1993.
- [7] N. Seddigh, B. Nandy, P. Piedad, *Study of TCP and UDP Interactions for the AF PHB*, Internet Draft - Work in Progress, draft-nsbnpp-diffserv-tcpudpaf-00.pdf, June 1999.
- [8] Raj Jain, *The Art of Computer Systems Performance Analysis*, John Wiley and Sons Inc., 1991.

Authors's Address

Mukul Goyal, Arian Durrezi, Raj Jain, Chunlei Liu
Computer and Information Science Dept
The Ohio State University
2015, Neil Avenue, Room 395
Columbus, OH 43210

Phone: +1 (614) 292-3989
Fax: +1 (614) 292 2911
Email: mukul@cis.ohio-state.edu

This internet draft expires in January 2000