Preferential Treatment of Acknowledgment Packets in a Differentiated Services Network

K. Papagiannaki⋆, P. Thiran†,‡, J. Crowcroft⋆, C. Diot†
⋆: UCL, †: Sprint ATL, ‡: EPFL

Abstract. In the context of networks offering Differentiated Services (DiffServ), we investigate the effect of acknowledgment treatment on the throughput of TCP connections. We carry out experiments on a testbed offering three classes of service (Premium, Assured and Best-Effort), and different levels of congestion on the data and acknowledgment path. We apply a full factorial statistical design and deduce that treatment of TCP data packets is not sufficient and that acknowledgment treatment on the reverse path is a necessary condition to reach the targeted performance in DiffServ efficiently. We find that the optimal marking strategy depends on the level of congestion on the reverse path. In the practical case where Internet Service Providers cannot obtain such information in order to mark acknowledgment packets, we show that the strategy leading to optimal overall performance is to copy the mark from the respective data packet into returned acknowledgement packets, provided that the affected service class is appropriately provisioned.

1 Introduction - Motivation

There have been several proposals for implementing scalable service differentiation in the Internet. Such architectures achieve scalability by avoiding per-flow state in the core and by moving complex control functionality to the edges of the network. A specific field in the IP header (the DS field) is used to convey the class of service requested. Edge devices perform sophisticated classification, marking, policing (and shaping operations, if required). Core devices forward packets according to the requirements of the traffic aggregate they belong to [3].

Within this framework, several schemes have been proposed, such as the "User Share Differentiation (USD)" [2], the "Two-Bit Differentiated Services" architecture [14], and "Random Early Drop with In and Out packets" (RIO) [5]. Preferential treatment can be provided to flow aggregates according to policies defined per administration domain. All those schemes (along with the work of the Differentiated Services Working Group of the IETF [3]) refer to unidirectional flows. One of the reasons behind this is the fact that service providers have control only over the forward path of their traffic, and not over the reverse. Therefore, every effort to provision reverse paths of bi-directional flows has to be initiated by the other-end provider or the destination hosts.

The Transmission Control Protocol (TCP) is the dominant data transport protocol in the Internet [4], [13]. TCP is a bi-directional transport protocol, which uses 40-bytes acknowledgment packets (ACKs) for reliable data transmission and to control its sending rate. Data packet losses are detected through duplicate or lack of acknowledgments on the reverse direction. Such an event is followed by a reduction in transmission rate, as a reaction to possible congestion on the forward path.

Modeling TCP shows that TCP flows suffer when the path serving their ACKs is congested [12]. Therefore, marking data packets belonging to connections facing congestion on their reverse path may not prove adequate for a connection to reach its performance target, due to ACK losses on the congested reverse path. We are interested in identifying ways in which data and acknowledgment packets have to be marked so that connections with different levels of congestion on forward and reverse paths can still achieve their performance goals.

In this paper, we evaluate, through a thoroughly devised experimental plan, TCP performance in a Differentiated Services network featuring congestion on forward and/or reverse paths. We build a testbed that offers three classes of service, namely Premium, Assured and Best-Effort [14]. We examine cases when forward and/or reverse paths are congested, and vary the marks carried by data and acknowledgment packets. We quantify the effect that each one of those factors has on TCP throughput. Our goals are (i) to assess whether such an effect exists, and (ii) to identify the optimal marking strategy for the acknowledgments of both Premium and Assured flows.

To the best of our knowledge, this problem has been addressed only by simulation in [11]. This latter study simulated a dumb-bell topology to investigate the effect of marking acknowledgment packets, and analyzed the behavior of a single flow, for which multiple marking schemes were applied. Our study diverts from [11] in that: (i) we use a testbed instead of a simulator, (ii) we use a more complex network topology with one level of aggregation, which features combinations of congested/uncongested forward/reverse paths for each class of service, (iii) we try out all possible combinations of data - acknowledgment packet markings for all classes of service, (iv) we identify which factors influence TCP throughput in our experiments and quantify their effect, and (v) we propose optimal acknowledgment marking strategies which lead to better Premium and Assured throughput regardless of the level of congestion on the network.

The organization for the rest of the paper is as follows. Section 2 explains the experimental plan and the associated statistical model that we have adopted in order to quantify the effect of the congestion levels and acknowledgment marking strategies on the throughput of the Premium and Assured flows. Section 3 describes the actual testbed on which the measurements were carried out, according to the plan elaborated in Section 2. Section 4 provides the analysis of the collected data for Premium and Assured flows and discusses which factors (or interaction thereof) influences throughput the most. Section 5 identifies the optimal acknowledgment marking strategies in networks where congestion can be predicted. In practice, this may not be possible, so Section 6 proposes a sub-optimal strategy which is independent of the network congestion and achieves throughput values close to the optimal ones found in Section 4. We conclude the paper with a summary of our main results, and we discuss whether marking ACKs would be practical in a Differentiated Services network.

2 Experimental Design and Methodology

In the statistical design of experiments, the outcome is called the *response variable*, the variables or parameters that affect the response variable are called *factors*; the values that a factor can take are called *levels*; the repetitions of experiments are called *replications* [9] [10].

In our case, we are interested in two *response variables*, namely the throughput of the Premium flows, henceforth denoted by y, and the throughput of the Assured flows, denoted by y'.

We will study the influence of four *factors*, which are:

- the *marking of the ACKs for the Premium flows*, denoted by P,
- the *marking of the ACKs for the Assured flows*, denoted by A,
- the *marking of the ACKs for the Best-Effort flows*, denoted by B,
- the *existence or absence of congestion* on the forward and/or reverse path, denoted by C.

Each of these four factors can take three *levels*, which are as follows:

- for each factor P, A and B, the levels will be p, a and b, based on whether the acknowledgment packet of the corresponding flow is marked as Premium, Assured or Best-Effort respectively,
- for C, we will distinguish the three following levels: f (the forward path is congested, but not the reverse path), r (the reverse path is congested, but not the forward path) and t (both forward and reverse paths are congested). We chose not to consider the rather trivial case where both forward and reverse paths are not congested, since in this case none of the other factors (marking strategies) will affect the response variables (throughputs of Assured and Premium flows).

The experiments consist in measuring throughputs of Premium and Assured flows under all possible combinations of the four factors P, A, B and C. Since each factor has three levels, there are $3^4 = 81$ possible combinations, and the design is called a full 3^4 design [9]. Each experiment will be *replicated* three times, so that a total of 243 experiments will be carried out. The throughputs achieved by flows during each experiment are not independent; the throughput of a Premium flow depends on the number of flows sharing the allocated Premium capacity, while the throughput of an Assured flow depends on the number of Assured and Best-Effort flows sharing the same path. Each one of those experiments is uniquely defined by a combination of five letters i, j, k, l, m ($i, j, k \in \{p, a, b\},\ l \in \{f, r, t\},\ m \in \{1, 2, 3\}$), where the first four are the corresponding level of factor $P = i$, $A = j$, $B = k$ and $C = l$, and where $R = m$ refers to one of the three replications. For example, y_{aapt3} will denote the throughput of the Premium flow measured on the third experiment ($R = 3$) with Premium and Assured acknowledgments marked as Assured ($P = A = a$), with Best-Effort acknowledgments marked as Premium ($B = p$), and with both forward and reverse paths congested ($C = t$).

A 3^4 experimental design with 3 replications assumes the following model for the response variable of the Premium flows, for all 243 possible combinations of the indices i, j, k, l, m [10]:

$$y_{ijklm} = x^0 + x_i^P + x_j^A + x_k^B + x_l^C + x_{ij}^{PA} + x_{ik}^{PB} + \ldots + x_{kl}^{BC} + x_{ijk}^{PAB} + \ldots + x_{jkl}^{ABC} + x_{ijkl}^{PABC} + \epsilon_{ijklm} \quad (1)$$

where the terms of the right hand side are computed as follows [9] [10]:

- $x^0 = \sum_{i,j,k,l,m} y_{ijklm}/243$ is the average response (throughput) over the 243 experiments,
- $x_i^P = \sum_{j,k,l,m} y_{ijklm}/81 - x^0$ is the difference between the throughput averaged over the 81 experiments taken when factor P takes level i, with $i \in \{p, a, b\}$, and the average throughput x^0. It is called the *main effect* due to factor P at level i. A similar expression holds for x_j^A, x_k^B and x_l^C,
- $x_{ij}^{PA} = \sum_{k,l,m} y_{ijklm}/27 - x_i^P - x_j^A + x^0$ is called the *effect of (2-factor) interaction* of factors P at level i and A at level j. A similar expression holds for the five other pairs of factors,
- $x_{ijk}^{PAB}, \ldots, x_{jkl}^{ABC}$, and x_{ijkl}^{PABC} are the *effects of (3- and 4-factors) interaction*, and are computed using similar expressions,
- ϵ_{ijklm} represents the experimental error in the mth experiment (residual), $1 \leq m \leq 3$, which is the difference between the actual value of y_{ijklm} and its estimate computed as the sum of all the above terms.

The model for the throughput of the Assured flows is identical, with all terms in (1) denoted with a prime to distinguish them from the variables linked to Premium flows.

The importance of each factor, and of each combination of factors, is measured by the proportion of total variation in the response that is explained by the considered factor, or by the considered combination of factors [9]. These percentages of variation, which are explicited in the Appendix of the full version of this paper [15], are used to assess the importance of the corresponding effects and to trim the model so as to include the most significant terms.

Model (1) depends upon the following assumptions: (i) The effects of various factors are additive, (ii) errors are additive, (iii) errors are independent of the factor levels, (iv) errors are normally distributed, and (v) errors have the same variance for all factor levels [9]. The model can be validated with two simple "visual tests": (i) the normal quantile-quantile plot (Q-Q plot) of the residuals ϵ_{ijkml}, and (ii) the plot of the residuals ϵ_{ijklm} against the predicted responses y_{ijklm}, y'_{ijklm}. If the first plot is approximately linear and the second plot does not show any apparent pattern, the model is considered accurate [9]. If the relative magnitude of errors is smaller than the response by an order of magnitude or more, trends may be ignored.

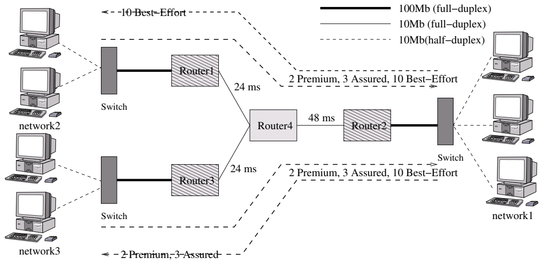

Fig. 1. Experimental Setup. The network consists of 3 ISP networks interconnecting at a Network Access Point. Each network generates a predefined traffic mix and directs its flows to selected networks so that different forward/reverse path congestion levels are achieved.

3 Network setup

In this section, we describe the way we design the network testbed, which allows us to realize the experimental plan described in Section 2.

3.1 Designing the Testbed Topology

The four factors we identified in the previous section are: factors P, A, and B, denoting the marks of the packets acknowledging Premium, Assured and Best-Effort packets respectively, and factor C, denoting whether forward and/or reverse TCP paths are congested. The different levels of factors P, A, and B can be easily realized, since they correspond to differential packet marking. Factor C is more problematic. We want to be able to test all levels in a single network, with a limited number of experiments.

By the term "congestion", we describe the condition of a link where contention for resources leads to queue build-ups and possible packet losses. Clearly, the more flows a link serves, the more likely it is that this link will reach a state of congestion. Under this assumption, we implement congestion on specific links using different numbers of flows on different links.

Our goal is our network to offer all three levels of congestion, we wish to investigate, for both Premium and Assured flows. A dumb-bell topology cannot implement all three levels f, r and t of factor C for both Premium and Assured flows, even if we assume that the two types of flows are isolated. For instance, we can test levels f and r by having the connecting link congested in one direction and not the other, but we cannot simultaneously test level t, which would require the connecting link to be congested in both directions.

Therefore our network must feature a minimum of four routers. A "Y"-topology, such as the one depicted in Figure 1, is capable of offering all three levels of factor C for both Premium and Assured flows in a single network.

If Network 2 and Network 3 initiate two Premium, three Assured, and ten Best-Effort flows towards Network 1, then the forward path for both Premium and Assured flows initiated in those two networks will be congested, because of the bottleneck link they have to share in order to reach Network 1. On the other hand, if the same number of flows is issued by Network 1, then the forward path of both Premium and Assured flows from Network 1 will not be congested, due to the limited number of flows following that path. Therefore, Premium and Assured flows initiated within the bounds of Network 1 will have no congestion on their forward path. They will face congestion on their reverse path, because of the Premium and Assured flows from the other two networks (level r of factor C).

Flows generated within the bounds of Networks 2 and 3 face congestion on their forward path. Existence and absence of congestion on the reverse TCP paths will give us the other two levels for factor C. More specifically, if we congest the path leading to Network 3, and leave the path to Network 2 uncongested, then we are capable of realizing levels f and t for the flows initiating in Network 2, and Network 3 respectively.

We achieve that by directing the Premium, and Assured flows of Network 1 towards Network 3, and the Best-Effort flows of Network 1 towards Network 2. In that way, the path from Network 1 to Network 2 accommodates the Best-Effort traffic of Network 1 and the acknowledgment traffic of Network 2. Thus, Premium and Assured flows from Network 2 face congestion on their forward and no congestion on their reverse path (level f of factor C). On the other hand, the path from Network 1 to Network 3 is loaded with both the Premium, and Assured traffic of Network 1, as well as the acknowledgment traffic of all 15 flows of Network 3. In other words, flows generated by Network 3 will face congestion on both forward and reverse paths (level t of factor C).

Origin → Destination Network / factor C	**Premium flows**	**Assured flows**
Network2 → Network1	on forward path only (f)	on forward path only (f)
Network3 → Network1	on both paths (t)	on both paths (t)
Network1 → Network3	on reverse path only (r)	on reverse path only (r)
Network1 → Network2	*This path does not carry any Premium or Assured flows.*	

Table 1. Routes realizing the different levels of factor C for Premium and Assured flows.

One may wonder why the reverse path of flows from Network 3 to Network 1 is congested, while the forward path from Network 1 to Network 3 is not, since they actually follow the same route. This can be explained by the difference in size between packets flowing on both directions: packets on the reverse path of flows from Network 3 to Network 1 are acknowledgment packets (ACKs), which are much smaller in size than the data packets flowing on the forward path from Network 1 to Network 3. The queuing time of ACKs can therefore be quite large compared to their processing time when data packets have to be served ahead of them, making the reverse path for flows from Network 3 to Network 1 congested, even if few data packets are present in the buffer. On the other hand, the delay of a data packet is mostly due to transmission and not to queuing (since not many data packets occupy that buffer), so that the forward path from Network 1 to Network 3 appears not congested to these flows. The resulting topology is presented in Fig. 1.

Table 1 displays the level of congestion for each class of service at each network. For example, a Premium flow originating at Network 3 will have both its forward and reverse paths congested, whereas an Assured flow originating at Network 2 will face congestion on its forward path only, because the path from Network 1 to Network 2 is lightly loaded.

3.2 Testbed Configuration

The testbed derived from the topology described in the previous section (Fig. 1) consists of four routers, Linux PCs with kernel version 2.2.10 that supports differentiated services [1]. Delay elements have been added on the links interconnecting the four routers, so that the bandwidth-delay product is large enough for the TCP flows to be able to open up their windows and reach their steady state (maximum window size is set to 32KB). Each network features two or three HP-UX workstations. Long-lived TCP Reno flows are generated using the "netperf" tool [8]. They last 10 minutes and send 512 bytes packets. Netperf reports the achieved throughput by each flow for the whole duration of the experiment (there is no warm-up period). The choice of long-lived flows was made so as to avoid transient conditions, and to focus on TCP flows in their steady state, where acknowledgment packet marking will have the greatest impact.

End hosts in our testbed did not have the capability of marking packets, and therefore edge routers perform policing, marking, classification and appropriate forwarding. The packets are policed based on their source and destination addresses using a token bucket. The Assured service aggregate is profiled with 2.4 Mbps and any excess traffic is demoted to Best-Effort. The Premium flows are profiled with 1 Mbps each, and any excess traffic is dropped at the ingress (shaping should be performed by the customer). There is no special provision for Best-Effort traffic, except from the fact that we have configured the routers in such a way so that it does not starve.

Each outgoing interface is configured with a Class Based Queue (CBQ) [7] consisting of a FIFO queue for Premium packets and a RIO queue for Assured and Best Effort packets. The RIO parameters are 35/50/0.1 (min_{th}, max_{th}, max_p) for the OUT packets and 55/65/0.05 for the IN packets. We chose those values so that: (i) the minimum RED threshold is equal to 40% of the total queue length, as recommended in [6], (ii) the maximum threshold for OUT packets is much lower than the one for IN packets to achieve higher degree of differentiation between Assured and Best-Effort packets, as suggested in [5], and (iii) the achieved rate for each one of the three classes of service is close to the profiled one.

Lastly, in order to enable cost-efficient analysis of all the possible scenarios, we assume that all networks implement the same acknowledgment marking strategy. We believe that this assumption has little influence on the results we observe.

4 Analysis of Experimental Results

We now apply the methodology outlined in Section 2 on the collected experimental results and identify the most important factors for Premium and Assured flows. Due to lack of space, we refrain from presenting the raw experimental results in this paper, and encourage the reader to refer to the full version of this paper [15].

4.1 Analysis of Variance (ANOVA)

The effects of various factors and their interactions can be obtained utilizing the methodology described in Section 2. It is really important to understand the complexity behind the relations among the four factors. Acknowledging the packets of a certain class of service with specific marks does not only influence the throughput of the flows belonging to that service class, but it also affects the flows belonging to the class, which is going to be utilized by the acknowledgment packets.

Component	Percentage of Variation
Main effects (total)	48.69%
Congestion (SSC/SST)	**31.23%**
Premium ACKs (SSP/SST)	**10.43%**
Assured ACKs (SSA/SST)	**4.42%**
BE ACKs (SSB/SST)	**2.6%**
First-order interactions (total)	37.17%
Premium ACKs - Congestion (SSPC/SST)	**32.26%**
Assured ACKs - Congestion (SSAC/SST)	0.59%
BE ACKs - Congestion (SSBC/SST)	1.11%
Premium ACKs - Assured ACKs (SSPA/SST)	**2.04%**
Premium ACKs - BE ACKs (SSPB/SST)	0.86%
Assured ACKs - BE ACKs (SSAB/SST)	0.19%

Table 2. Analysis of Variance table for Premium flows. Effects that account for more than 1.5% of total variation are represented in bold. Notations between parentheses are detailed in [15].

Component	Percentage of Variation
Main effects (total)	89.51%
Congestion (SSC/SST)	**86.38%**
Premium ACKs (SSP/SST)	0.54%
Assured ACKs (SSA/SST)	0.8%
BE ACKs (SSB/SST)	**1.77%**
First-order interactions (total)	8.89%
Premium ACKs - Congestion (SSPC/SST)	0.21%
Assured ACKs - Congestion (SSAC/SST)	**7.86%**
BE ACKs - Congestion (SSBC/SST)	0.16%
Premium ACKs - Assured ACKs (SSPA/SST)	0.5%
Premium ACKs - BE ACKs (SSPB/SST)	0.09%
Assured ACKs - BE ACKs (SSAB/SST)	0.04%

Table 3. Analysis of Variance table for Assured flows. Effects that account for more than 1.5% of total variation are represented in bold. Notations between parentheses are detailed in [15].

Table 2 details the percentages of variation apportioned to the main effects, and to first-order interactions for Premium flows, which amount to 85% of the total variation. Second and third-order interactions turn out to be negligible. Respective results specific to the Assured flows are presented in Table 3. From Table 3, we can see that the role of congestion is much more important for Assured flows than for Premium flows. Furthermore, the main effects and the first-order interactions are adequate to justify 98.4% of the existing variation.

Dropping all interactions that explain less than 1.5% variation as negligible, and using model (1), Premium and Assured throughput can be described with the following formulae:

$$y_{ijklm} = x^0 + x_i^P + x_j^A + x_k^B + x_l^C + x_{ij}^{PA} + x_{il}^{PC} + \epsilon_{ijklm}, \tag{2}$$

$$y'_{ijklm} = x'^0 + x'^B_k + x'^C_l + x'^{AC}_{jl} + \epsilon'_{ijklm}, \tag{3}$$

for $i, j, k \in \{p, a, b\},\ l \in \{f, r, t\},\ and\ 1 \leq m \leq 3$.

4.2 Visual tests

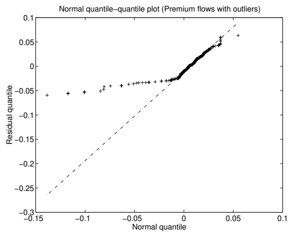

Figure 2. Normal Q-Q plot for residuals.

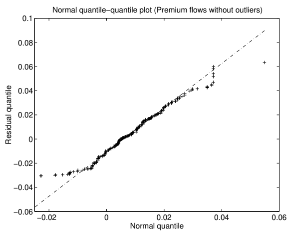

Figure 3. Normal Q-Q plot for residuals (no outliers).

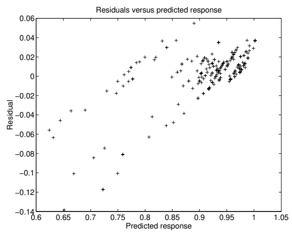

Figure 4. Residuals versus predicted response.

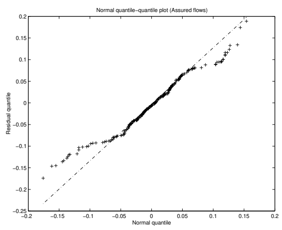

Figure 5. Normal Q-Q plot for residuals

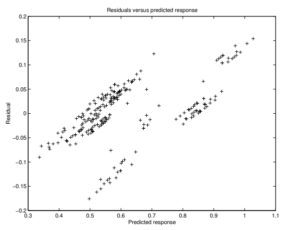

Figure 6. Residuals vs. predicted response

To test whether (2), and (3) are indeed valid models, we perform the tests described at the end of Section 2. Figure 2 presents the Q-Q plot between the residuals and a normal distribution. Since the plot is not linear, we cannot claim that the errors in our analysis are normally distributed. However, if we take out the 20 outliers (out of 243 values), then the new Q-Q plot displays a linear relation between the two distributions (Figure 3). Investigating the data, we see that those 20 outlying values do not occur for the same experiment, and therefore the error introduced by ignoring them does not challenge the validity of our analysis. We suspect that those outlying values are due to SYN packets getting lost at the beginning of the experiments, when all the flows start simultaneously. A flow will normally need 6 seconds (TCP Reno timeout) to recover from such a loss, delaying its transmission while other flows increase their sending rates.

The second visual test is the plot for the residuals ϵ_{ijklm} versus the predicted response y_{ijklm}, and is presented in Figure 4. No apparent trends appear in the data. This validates our model, and allows us to conclude that the throughput of a Premium flow in a Differentiated Services network is mostly affected by 1) congestion on forward/reverse paths, 2) the interaction between congestion and the Premium acknowledgment marks, and 3) the Premium acknowledgment marks themselves.

The visual tests for the Assured flows are presented in Figures 5, and 6, and also confirm the accuracy of model (3) for the Assured throughput. Therefore, an Assured flow is mainly affected by 1) congestion on its forward and reverse path, 2) the interaction between the congestion level and the Assured acknowledgment packet marks, and 3) the BE acknowledgment packet marks.

5 Optimal Marking Strategies

In the previous section we modeled the Premium and Assured flow throughput utilizing a small number of parameters and proved that TCP throughput is sensitive to acknowledgment marking. In this section we identify the *optimal acknowledgment marking strategy* for each class of service, which we define as the marking algorithm which results in the highest throughput values for the class of service in consideration.

Congestion		**Assured ACKs**	
x_r^C	(-0.0538, -0.0442)	x_p^A	(-0.072, 0.0347)$^\alpha$
x_f^C	(0.0135, 0.0232)	x_a^A	(-0.0447, 0.0619)$^\alpha$
x_t^C	(0.0257, 0.0354)	x_b^A	(-0.0433, 0.0634)$^\alpha$
Premium ACKs		**Best-Effort ACKs**	
x_p^P	(-0.0076, 0.0020)$^\alpha$	x_p^B	(-0.0676, 0.0391)$^\alpha$
x_a^P	(0.0212, 0.0309)	x_a^B	(-0.0475, 0.0592)$^\alpha$
x_b^P	(-0.0281, -0.0184)	x_b^B	(-0.0449, 0.0618)$^\alpha$
Premium ACKs - Congestion		**Premium - Assured ACKs**	
x_{pr}^{PC}	(0.046, 0.0477)	x_{pp}^{PA}	(-0.0173, -0.0156)
x_{ar}^{PC}	(0.0221, 0.0239)	x_{pa}^{PA}	(0.0057, 0.0075)
x_{br}^{PC}	(-0.0708, -0.0690)	x_{pb}^{PA}	(0.0089, 0.0107)
x_{pf}^{PC}	(-0.0249, -0.0232)	x_{ap}^{PA}	(0.0132, 0.0149)
x_{af}^{PC}	(-0.0138, -0.012)	x_{aa}^{PA}	(-0.0069, -0.0051)
x_{bf}^{PC}	(0.0361, 0.0378)	x_{ab}^{PA}	(-0.0089, -0.0071)
x_{pt}^{PC}	(-0.0236, -0.0219)	x_{bp}^{PA}	(0.0015, 0.0032)
x_{at}^{PC}	(-0.011, -0.0092)	x_{ba}^{PA}	(-0.0014, 0.0003)$^\alpha$
x_{bt}^{PC}	(0.032, 0.0338)	x_{bb}^{PA}	(-0.0026, -0.0009)

α: indicates that the effect is not significant

Table 4. Confidence intervals for Premium flow analysis.

Congestion		**Best-Effort ACKs**	
x'^C_r	(0.1648, 0.1892)	x'^B_p	(-0.0074, 0.0169)$^\alpha$
x'^C_f	(-0.1047, -0.0804)	x'^B_a	(-0.0361, -0.0118)
x'^C_t	(-0.0966, -0.0722)	x'^B_b	(0.007, 0.0314)
Assured ACKs - Congestion			
x'^{AC}_{pr}	(0.0375, 0.0419)		
x'^{AC}_{ar}	(0.0336, 0.038)		
x'^{AC}_{br}	(-0.0777, -0.0733)		
x'^{AC}_{pf}	(-0.0224, -0.018)		
x'^{AC}_{af}	(-0.0215, -0.0171)		
x'^{AC}_{bf}	(0.0373, 0.0417)		
x'^{AC}_{pt}	(-0.0217, -0.0172)		
x'^{AC}_{at}	(-0.0187, -0.0142)		
x'^{AC}_{bt}	(0.0337, 0.0381)		

α: indicates that the effect is not significant

Table 5. Confidence intervals for Assured flow analysis.

In order to evaluate whether a factor has a positive or negative effect, we must first provide confidence intervals for the effects identified as important in the previous section. Those confidence intervals can be computed using t-values read at the number of degrees of freedom associated with the errors. Due to lack of space we refer to [9] for the techniques used to compute those intervals and the degrees of freedom associated with each one of the components in our analysis. The obtained 90% confidence intervals for Premium and Assured flows are presented in Tables 4 and 5 respectively.

If the interval contains the value 0, then the effect of the corresponding component is not statistically significant. For the rest, an interval with positive values indicates higher than average throughput, while an interval with negative values indicates lower than average throughput.

5.1 Results Specific to Premium Flows

In Section 4 we showed that the throughput of a Premium flow is mostly affected by congestion (31.23%), acknowledgments to Premium packets (10.43%), acknowledgments to Assured packets (4.42%), the interaction between acknowledgments to Premium packets and congestion (32.26%), and the interaction between ACKs directed to Premium and Assured packets (2.04%). We can further derive from Table 4 that:

1. **effect of factor C**: the average throughput of a Premium flow suffers when the reverse path is congested (the confidence interval for $C = r$ lies in the negative), and reaches high values when the forward path is also congested ($C = f$, $C = t$).
   Therefore, marking acknowledgment packets is especially important when the reverse path is congested ($C = r$). In this case, marking data packets does not affect the performance since the forward path is lightly utilized. If congestion exists on the forward path or on both forward and reverse paths ($C = f$, or $C = t$), it seems that the protection of data packets on the forward path is capable of making up for the lost (or delayed) acknowledgment packets.
2. **effect of factor PC**: when a Premium flow suffers from congestion on the reverse path, then acknowledgment packets have to be marked as Premium ($PC = pr$). Best-Effort acknowledgment packets in this case are not adequate so that Premium flows reach their performance goal ($PC = br$). Lastly, if no congestion is present on the reverse path, then even Best-Effort acknowledgment marking is capable of offering a sustained rate to Premium flows ($PC = bf$, and $PC = bt$ offer higher than the average throughput x^0).
3. **effect of factor P**: regardless of congestion, the Premium flows achieve better throughput when their acknowledgments are marked as Assured ($P = a$). In this case, acknowledgment traffic exceeding the agreed-upon profile can still get transmitted as Best-Effort,
4. **effect of factor PA**: Premium flows achieve higher throughputs when their ACKs do not utilize the same service class as ACKs directed to Assured flows (confidence intervals are negative when $PA = pp$, $PA = aa$, or $PA = bb$).

Consequently, the strategy leading to higher Premium throughput values, is to acknowledge Premium data packets with Assured packets. If the reverse TCP path is congested, though, then Premium ACKs have to be marked as Premium.[1]

[1] In our experiments we had not specifically provisioned for ACKs and therefore we would expect slightly different results for point 3, and 4 in case we had.

5.2 Results Specific to Assured Flows

We have seen in Section 4 that Assured flows' throughput is mostly affected by congestion (86.38%), the interaction between acknowledgments to Assured packets and congestion (7.86%), and acknowledgments to Best-Effort packets (1.77%). From the confidence intervals presented in Table 5, we can further derive that:

1. **effect of factor C**: Assured flows perform poorly when their forward path is congested ($C = f, C = t$).
   The Assured service is a statistical service which is contracted to face less loss than BE in case of congestion. Therefore, its throughput is affected by congestion on the forward path much more than Premium flows (a minimum throughput service).
2. **effect of factor AC**: if only the reverse path is congested ($C = r$), then flows should protect their ACKs by subscribing them to a provisioned class of service. If only the forward path is congested ($C = f$), it is better if flows receive ACKs that do not belong to a rate-limited class of service. For the third case ($C = t$), when both paths are congested, we see that it is better if Assured flows send their ACKs as BE. We believe that this result stems from the fact that the forward path contains an aggregation point which makes flows from networks 2 and 3 behave in similar ways.
3. **effect of factor B**: lastly, it is better if Best-Effort traffic does not utilize the Assured service class for its ACKs; in such a case Assured flows have to compete with Assured-marked ACKs for resources (indeed other measurements taken throughout the experiments show that such a combination leads to the largest number of retransmissions for Assured flows).

Therefore, the optimal strategy for Assured flows is to mark ACKs as Premium or Assured in networks with congested reverse paths, and as BE in networks where the reverse TCP paths are lightly utilized.

6 Practical and Efficient Marking Strategies

In the previous section, we identified the optimal acknowledgment marking strategies for Premium and Assured flows. Those strategies depend on the level of congestion on forward and reverse paths, and should therefore be specific to particular networks. It is however difficult, if not impossible, to predict the level of congestion on the reverse path (especially since this path may change in time due to routing), and marking ACKs depending on their source and destination pair imposes a rather large overhead.

In this section, we identify a sub-optimal acknowledgment marking strategy, which still leads to higher Premium and Assured throughputs, but which no longer depends on the network specifics. We then show that this marking strategy achieves throughput close to the optimal values obtained in the previous section, and outperforms the "best-effort marking" used by default.

Sections 4 and 5 showed that in cases of congestion on the reverse TCP path, ACKs have to be protected. In such cases, Premium and Assured flows benefit when their acknowledgment packets belong to a provisioned class of service. More specifically, we know from Section 5 that: (i) regardless of congestion, Premium flows perform the best when their ACKs are marked as Premium (effect of P on Premium throughput), (ii) Premium flows acknowledged with Premium packets perform better when Assured flows do not use the Premium service class for their ACKs (effect of PA on Premium throughput), and (iii) Assured flows perform better when Best-Effort flows receive BE ACKs (effect of B on Assured throughput). Therefore, marking strategies that are practical while maintaining a good performance for both Premium and Assured flows, are the ones described by $P = p$, $B = b$ and $A = a$ or $A = b$.

Throughput values achieved by Assured flows under those schemes and different levels of congestion are similar except when the reverse TCP path is congested. For the latter case, Assured acknowledgements lead to higher Assured throughput.

Therefore, a practical and yet optimal strategy seems to be the one where each flow receives acknowledgments in the same class of service. Such a strategy can be easily implemented by a network, if TCP implementations are modified so that acknowledgment packets copy the mark of the packet, they acknowledge. In this section we evaluate the performance of this particular marking strategy by comparing it with the optimal strategy identified in Section 4, and the default strategy, where all ACKs are marked as BE. Figure 7 displays the average throughput achieved by Premium and Assured flows under those three acknowledgment marking strategies.

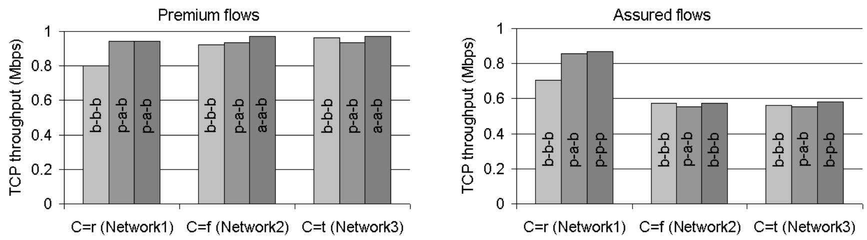

Figure 7. Premium and Assured flow throughput under 3 acknowledgment marking strategies (the default strategy, the strategy where each acknowledgment packet copies the mark of the data packet, and the optimal). Strategies are described by the combination of factors P-A-B.

From Figure 7 we see that the optimal acknowledgement marking strategy may improve performance by 20% over the throughput achieved when ACKs are marked as BE (Assured flows in Network 1 achieve 20% higher throughput when all the flows are acknowledged with Premium packets - case p-p-p - rather than with Best-Effort packets - case b-b-b). Furthermore, we see that the marking strategy, where each flow receives ACKs belonging to the same class as the data packets, performs well in most cases and independently of the level of congestion on the forward and reverse paths for both Premium and Assured flows.

7 Conclusions - Discussion

Carrying out experiments in a Differentiated Services network, offering three classes of service (Premium, Assured and Best-Effort), we have investigated the effect of acknowledgment packet marking, and congested/uncongested forward/reverse paths, onto TCP throughput. We have studied and quantified the effect of the identified factors onto the throughput of the provisioned classes of service, and we have identified the levels of those factors that lead to optimal throughput values.

The analysis of the collected results shows that TCP throughput is sensitive to congestion on the reverse TCP path, thereby confirming the results obtained by modeling the TCP behavior in asymmetric networks, with slow or congested reverse TCP paths [12].

Consequently, bi-directionality of traffic must be taken into account in Service Level Specifications and special provision must be made for bi-directional flows in a Differentiated Services network. Protection of TCP data packets alone will not ensure better performance to flows, requesting a better than Best-Effort service.

Our results indicate that in cases of congested reverse TCP paths, Premium and Assured flows have to be acknowledged by preferentially treated packets. Nevertheless, there is no single best marking strategy, which leads to optimal performance for both Premium and Assured flows. Furthermore, the optimal marking strategy for each provisioned class of service requires explicit knowledge of the level of congestion on the reverse path; a piece of information which is really hard to obtain.

As a consequence, we have investigated sub-optimal marking strategies, which still lead to high throughput values for the two provisioned classes of service, and are independent of the network specifics. We have proven that the marking strategy, which fulfills those requirements and is easily implementable, is the one where acknowledgment packets carry the marks of their respective data packets. We have shown that this strategy leads to performance comparable to the optimal marking strategy, and outperforms the default strategy of Best-Effort ACKs.

How to offer better than Best-Effort services over a single network is still an open, and active area of research. No specific recommendations have been made by the appropriate bodies and investigation of the performance achieved by flows, when they request preferential treatment, is still under investigation. Furthermore, provisioning of Differentiated Services networks remains an interesting problem, still to be solved.

This paper has shown that in the case of TCP flows, provisioning of forward paths is not adequate for TCP flows to reach their target rate. Reverse paths have to be provisioned as well, and acknowledgment packets have to be appropriately marked. In other words, network providers offering Differentiated Services do not only have to draw agreements with the providers handling their egress traffic, but also have to draw agreements with the providers returning ACKs for the TCP flows initiating within their bounds. We have to notice, however, that the additional provisioning that has to be done due to the acknowledgment traffic will be rather limited, because of the limited size of the acknowledgment packets. More specifically, in our experiments our provisioned traffic classes managed to reach their target levels with differential acknowledgment marking without special provisioning of the reverse paths.

References

1. W. Almesberger. Linux Network Traffic Control - Implementation Overview. http://icawww1.epfl.ch/linux-diffserv/, April 1999.
2. A. Basu and Z. Wang. A Comparative Study of Schemes for Differentiated Services. Technical report, Bell Laboratories, Lucent Technologies, July 1998.
3. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. An Architecture for Differentiated Services. Request for Comments (Proposed Standard) 2475, Internet Engineering Task Force, October 1998.
4. K. Claffy. The Nature of the Beast: Recent Traffic Measurements from an Internet Backbone. In *INET'98*, 1998.
5. D. D. Clark and W. Fang. Explicit Allocation of Best-Effort Packet Delivery Service. *ieanep*, 6(4):362–373, August 1998.
6. S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1(4):397–413, August 1993.
7. S. Floyd and V. Jacobson. Link-sharing and resource management models for packet networks. *IEEE/ACM Transactions on Networking*, 3(4), August 1995.
8. Information Networks Division, Hewlett-Packard Company. Netperf: A Network Performance Benchmark. http://www.netperf.org, February 1995.
9. Raj Jain. *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*. John Wiley, New York, 0-471-50336-3, 1991.
10. Peter W. M. John. *Statistical Design and Analysis of Experiments*. Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104-2688, 0-89871-427-3, 1998.
11. S. Köhler and U. Schäfer. Performance Comparison of Different Class-and-Drop Treatment of Data and Acknowledgements in DiffServ IP Networks. Technical Report 237, University of Würzburg, August 1999.
12. T. V. Lakshman, U. Madhow, and B. Suter. TCP/IP Performance with Random Loss and Bidirectional Congestion. *IEEE/ACM Transactions on Networking*, 1998.
13. S. McCreary and K. Claffy. Trends in Wide Area IP Traffic Patterns - A View from Ames Internet Exchange. In *ITC'00*, Monterey, September 2000.
14. K. Nichols, V. Jacobson, and L. Zhang. A Two-bit Differentiated Services Architecture for the Internet. Internet Draft, Internet Engineering Task Force, May 1999. Work in progress.
15. K. Papagiannaki, P. Thiran, J. Crowcroft, and C. Diot. Preferential Treatment of Acknowledgment Packets in a Differentiated Services Network. Technical Report TR10-ATL-041001, Sprint Advanced Technology Laboratories, April 2001.