Internet Traffic Shaping for IP over WDM Links with Source Output Buffering or Multiple Parallel Wavelengths

J. Aracil, M. Izal and D. Morató
Dept. Automática y Computación
Public University of Navarra, Spain
javier.aracil,mikel.izal,daniel.morato@unavarra.es

September 13, 2000

Abstract

Since the number of wavelengths per fiber is growing in an exponential fashion the overflow traffic can be routed through overflow lightpaths, thus providing an ideal network with near-infinite capacity and almost no-buffering. Such unprecedented bandwidth growth in the network backbone is only limited by the processing speed of the electronic elements. Even though multiple parallel high-speed channels (lightpaths) are provided between IP routers the switching speed of the latter is an order of magnitude below the lightpath transmission speed. As a result, minimizing transfer delay is not only a matter of forwarding traffic as fast as possible but to shape traffic so that the input queues of the destination routers do not overflow. Even though it is desirable to exploit the WDM capabilities to forward traffic in parallel channels in order to nearly eliminate the router output buffering, it turns out that the extreme burstiness of Internet traffic is even increased by routing part of the traffic through a backup channel. Instead, the use of source output buffering for traffic shaping purposes proves more beneficial. In this paper, we examine the typical scenario of a static WDM network with several wavelengths between IP routers. In a simple configuration of a primary and overflow lightpath the results show that if 3% of the traffic is routed through the overflow lightpath then the packet forwarding speed in the destination router should be increased in 20% in order to obtain the same transfer delay as with the single lightpath configuration with source output buffering.

1 Introduction

Even though Internet traffic presents high variability the most part of current IP backbones are based on constant bandwidth links between routers. Indeed the most part of high-speed Internet backbones, like the VBNS [1] or the European TEN-155 [2] are formed by constant bandwidth ATM links. The ATM is being used for Internet traffic transport with IP over ATM Permanent Virtual Circuits (PVCs). Dynamic bandwidth allocation service classes, like the Available Bit Rate (ABR)

[1] http://www.vbns.net
[2] http://www.dante.net/ten-155.html

are currently not being used since neither the standards nor the available equipment are mature enough. Such static link bandwidth configuration, although simple and straightforward, is not sufficient to absorb the traffic bursts that are due to the self-similar nature of Internet traffic. As a result the user applications face the packet loss and delay commonly experienced in the current Internet.

Optical technologies open new ways for the provision of dynamic bandwidth allocation at the link layer. Even though current WDM networks provide static allocation of resources there is a trend toward provisioning circuit-switched solutions (on-demand lightpaths) in a second generation of optical networks and, ultimately, optical packet switching. However, commercially available WDM networks today provide multiple parallel channels between backbone IP routers thus offering primary and alternate paths for network protection. A WDM network architecture with multiple static parallel channels is depicted in figure 1.

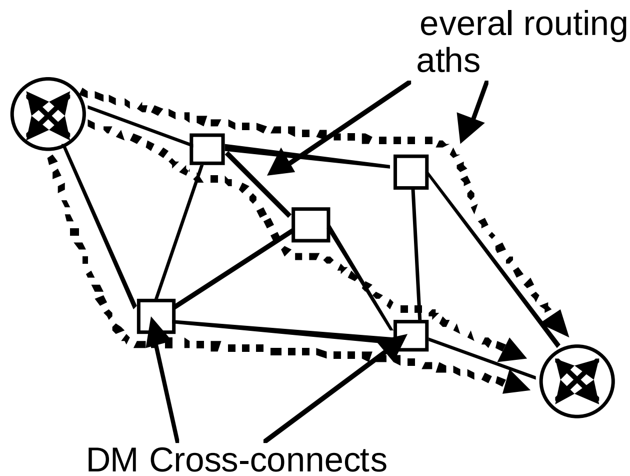

Figure 1: WDM network

The WDM network architecture of figure 1 is most common for a first generation optical networks scenario, a foreseeable evolution of current IP over ATM/Frame Relay networks, in which IP routers use the WDM layer as a link layer with multiple parallel channels, several of those being used for protection or overflow traffic, which leads to a network design with little buffering at the routers and a number of alternate paths to absorb traffic peaks. The same scenario is normally assumed in deflection routing networks, which are based on the principle of providing nearly no buffering at the network interconnection elements but several alternate paths between source and destination, so that the high-speed network becomes a distributed buffering system. The advantage is that buffer requirements at the routers are relaxed, thus simplifying the electronic design. In the WDM case, we note that the backup channels can be used to provide an alternate path for the overflow traffic as proposed in [1]. Rather than handling the traffic burstiness via buffering, leading to delay and packet loss in the electronic bottleneck, the backup channels can be used to absorb overflow traffic bursts. Interestingly, the availability of multiple parallel channels between source and destination routers resembles the telephone network scenario, in which a number of alternate paths (tandem switching) are available between end offices. The reason for providing multiple paths is not only for protection in case of failure of the direct link but also availability of additional bandwidth for the peak hours.

While the use of multiple parallel channels has clear advantages for bandwidth provisioning during congestion epochs we note that out of sequence packets may be delivered at the destination.

As a consequence, a sorting algorithm is needed so that packets are delivered in order to the end application. At Gbps speed such sorting procedures may become the throughput bottleneck. Alternatively, the number of out-of-sequence packets can be minimized by performing deflection at the flow or circuit level instead of doing so on a packet-by-packet basis [2]. Such proposals are consistent with the foreseeable evolution of gigabit IP routers, which provide flow-switching solutions for IP traffic such as Multi-Protocol Label Switching (MPLS). Indeed, MPLS allows per-flow deflection since the packet route is determined by the MPLS tag and not by the destination IP address. Thus, the entire flow can be deflected to the overflow lightpath and no packet misordering occurs since individual packets belonging to the same flow will not be routed through different paths.

Even though the effect of out-of-sequence packets poses significant challenges for the use of multiple parallel channels we choose not to address the topic and focus on the statistical features of overflow traffic, which are also of fundamental importance to evaluate network performance. Precisely, in the early stages of deployment of telephone networks A. K. Erlang found that the overflow traffic can no longer be regarded as Poissonian. On the contrary, the overflow traffic burstiness is higher than in the Poissonian model. The method of treating the overflow traffic is called Equivalent Random Theory [3]. Equivalent Erlagian models for blocking probability calculations can be established, that incorporate the overflow load effect. The overflow traffic can be characterized by a Poissonian input with higher intensity, in order to account for the burstiness of the latter.

Regarding IP over WDM, we note that multiple parallel channels between IP routers (for example SONET VCs) are also common in the Internet, and the availability of such multiple parallel channels between routers is likely to grow with the use of WDM technology. However, the network impact of the overflow traffic in the Internet case, that presents long-range dependence features, still remains an open issue. Thus, our aim is to analyze the case for Internet traffic in the presence of multiple parallel WDM channels. Since such WDM channels are high speed we explicitly focus on the traffic multiplex of a very large population of users and not on the packet level structure of the multiplexed traffic. Regardless of the use of flow or circuit level deflection policies, which serve to reduce the number of out-of-secuence packets, we identify a number of traffic engineering issues that significantly affect network performance. Indeed, our findings show that even though the correlational structure of the overflow traffic (long range dependence) remains nearly the same the traffic burstiness in the overflow lightpath increases. Thus, the destination router receives a very bursty traffic through one of its input channels. As a result, it is not always beneficial to route the overflow traffic through the overflow lightpaths, but on the contrary use some buffering (shaping) at the source router output. In order to have a better understanding of the above let us first review and analyze the Internet traffic features.

2 Internet traffic characteristics and WDM link model

Internet traffic presents long-range dependence, as opposed to independent-increments models such the Poisson packet arrival process. Let us present an intuitive, rather than mathematically complete, description of the statistical features of Internet traffic and the implications in network engineering. For a more analytical treatment of Internet traffic the reader is referred to [4] and references therein.

A packet arrival process can be characterized by the number of bytes or packets that arrive in fixed duration intervals. Now consider a sample interval of duration δ milliseconds and the number of bytes per interval coming from a generic traffic stream, which is the random variable X_i. The average traffic rate in larger time intervals can be obtained by averaging the random variable X_i so we obtain the aggregated packet arrival process $S_i = (X_{i+1} + ... + X_{i+m})/m$ for all positive i and m. The limit behavior of an averaged sum of large number of random variables is dictated by the Central Limit Theorem (CLT) that states that the variance of the aggregated process decays with the inverse of the number of variables in the sum. In other words, as we observe the network traffic in larger time intervals we notice that traffic burstiness decays rapidly. This is clearly observed in figure 2, where a synthetic traffic trace obtained from a Poisson process is compared to a real traffic trace. Several time scales (100 ms, 1 s and 4 s.) are displayed, showing that burstiness for Internet traffic occurs at any time scale.

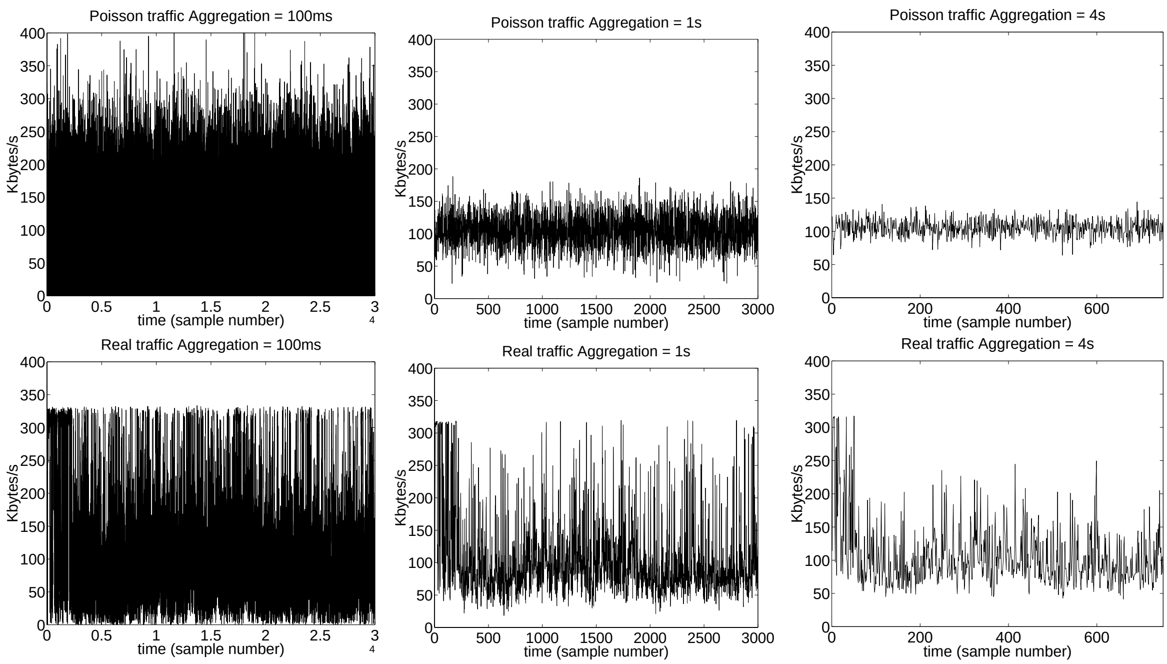

Figure 2: Poisson and FBT traffic in several time scales

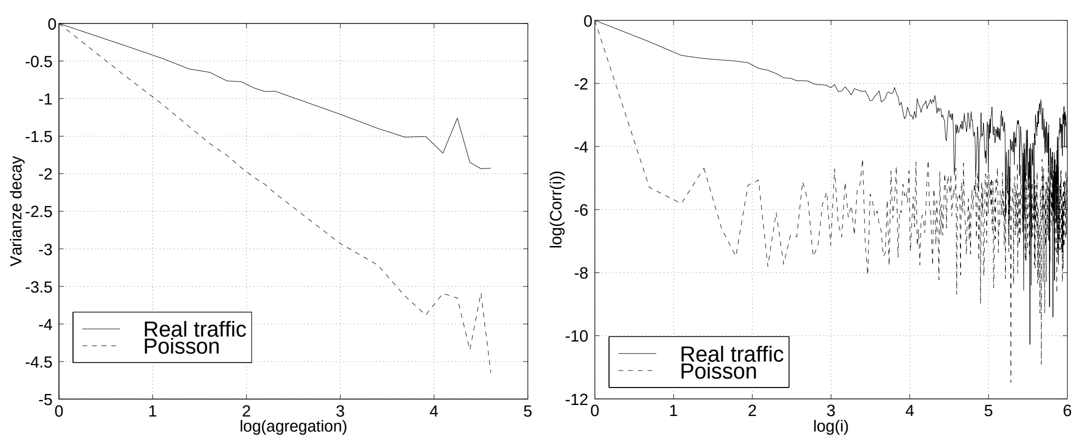

Figure 3: Variance versus aggregation level and correlation

Figure 3 shows the log-log plot of variance versus aggregation level (time scale) for the random variable number of bytes per interval and the correlation for Poisson and real trace cases. Such variance can be interpreted as a measure of the traffic variability. We observe a slow decay of the variance for the real trace and, on the other hand, we note that the real traffic trace is strongly correlated, with a hyperbolic decay of the correlation function. While the Poisson model is independent, thus meeting the conditions of the CLT, the real traffic trace presents long range dependence features. The slope of decay of the correlation(β) gives the Hurst parameter ($H = 1 - \beta/2$). A Hurst parameter value of 0.5 is characteristic of an independent increments process.

A traffic model that presents long-range dependence is the Fractional Brownian Traffic (FBT), which has been widely used to model Internet traffic [5]. The FBT is an scaled version of the Fractional Brownian Motion (FBM), which is a Gaussian stochastic process that shows the long range dependence features observed in Internet traffic. The number of bytes arriving to the link in the interval (0,t) can be modeled by the stochastic process $B(0,t) = mt + \sqrt{ma}Z_t$, where Z_t is a normalized FBM. The parameter m is just the average number of bytes per interval of duration t while the parameter a represents the variance divided by the mean, thus providing a characterization of the traffic variability. The third parameter for the FBT is the Hurst parameter H.

The FBT is the most interesting model for network dimensioning and control. First, the FBT has been shown to model accurately the Internet traffic coming from the multiplex of a large number of users [5]. Secondly, a variety of methods exist to generate synthetic traces of FBTs, that are useful for simulation purposes [6][7]. Third, the single-server single-queue model has been extensively analyzed and there is a closed analytical expression for the queue distribution tail, which is given by:

$$P(Q_{length} > q) = e^{-\frac{(c-m)^{2H} q^{2-2H}}{2K(H)^2 am}} \tag{1}$$

Where c is the server capacity and and $K(H) = H^H(1-H)^{1-H}$

The above formula shows that the queue length for a given utilization factor depends upon two parameters: the Hurst parameter H and the variance divided by the mean a. In other words, the queue length depends on the correlational structure of the input traffic and the traffic variability. The latter is not surprising, since the queue length for the well-known M/G/1 models [8] depends on the coefficient of variation (variance divided by squared mean). Intuitively, the more variability in the input traffic the higher the chance to find large traffic bursts introducing large delays in the queueing system buffer. The queue length dependence with both parameters can be observed in figure 4 that shows the queue tail distribution versus the a coefficient and Hurst parameter H for a value of the utilization factor of 0.7.

The results show that queue length is not only dependent on the parameter H but also on the traffic marginal distribution variability.

2.1 WDM overflow link model

The WDM link model under analysis is depicted in figure 5, showing the multiple parallel channels configuration presented in figure 1. The destination router receives traffic from such channels through separate input queues. Such input queues are necessary for the destination router since packet forwarding is done in the electronic domain, so that buffers are mandatory for speed adaptation purposes with the optical network. In fact, current Internet routers have separate queues per port in order to absorb traffic burstiness with a feasible electronic design. The input traffic could be

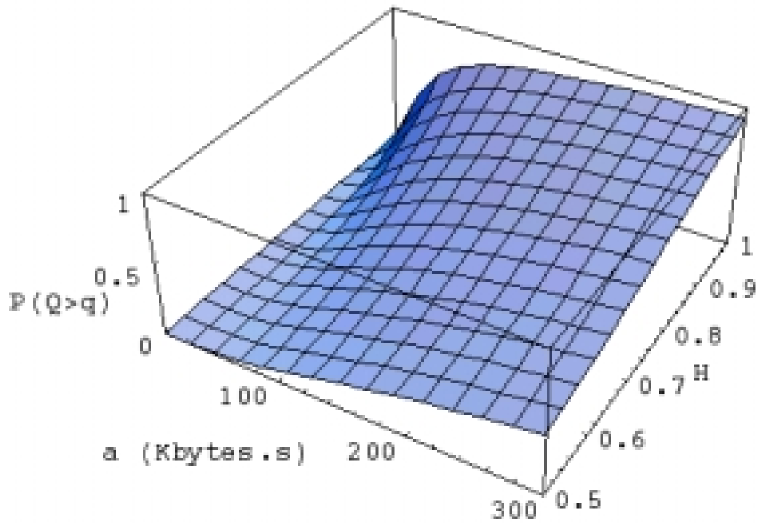

Figure 4: Sensitivity of queueing delay versus H and a

also multiplexed in a single input queue but then the multiplexer should operate at a rate which is equal to the sum of the input ports peak rate. Such design is clearly infeasible for Gigabit routers.

The destination router can be modeled as a server that performs the packet forwarding operations. Indeed routers are normally characterized by the number of packets per second that they can forward. Such packet forwarding involves the usual operations of destination address recognition and routing table lookup for next hop resolution. As a first approximation, we consider that the input buffers will be served by a service discipline that is proportional to their mean offered load to the destination router input buffers. Thus, the utilization factor remains the same for the different queues and the router capacity is fairly divided. On the other hand, the model also serves to the purpose of fair comparison of both primary and overflow traffic impact on a single server single queue system with the same utilization. We compare two different network configurations, both possible with static WDM layer. The first alternative is to route overflow traffic through the overflow lightpaths, so that no output buffering at the source router is necessary. The second alternative is to provide a buffer at the server (namely shaping traffic at origin), but not to de-multiplex the input traffic into two different lightpaths.

We note that the queuing delay in the router input buffers is the most important contribution to the end-to-end delay, even in the single channel case, since the optical channel capacity is assumed to be much higher than the router processing speed. Since the traffic load for both channels presents different statistical features such queuing delays also differ, and will be subject of our analysis.

3 Results and discussion

We evaluate the average message delay versus the utilization factor for the single lightpath and dual lightpath cases, with different traffic overflow thresholds. The delay is defined as the time interval between the moment the packet enters the lightpath for transmission at point X_{in} and the

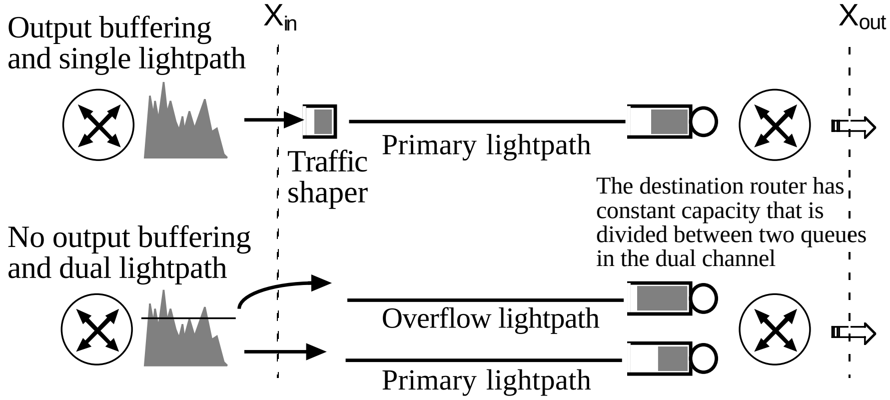

Figure 5: WDM link model with overflow channels

moment the packet leaves the destination router for next hop transmission at point X_{out} (see figure 5). The input traffic is an FBT with the a parameter fitted to the empirical values reported in [5] from traces obtained at Bellcore. Furthermore, an experimental measurement of the IP over ATM access link of Public University of Navarra [9] shows agreement with the Bellcore data. The Hurst parameter takes on the values 0.5, 0.7, 0.9, showing a range of un-correlated to highly correlated traffic. We note that the typical value of the H parameter for Internet traffic is in the vicinity of 0.7 [5]. FBT traces are obtained generating an FGN with the method described in [6] and proposed by Hosking in [10].

Being the overflow traffic defined as the input traffic that exceeds the primary lightpath bandwidth, we can view such overflow traffic as the peaks exceeding a cut-off value, which is the primary lightpath speed. We consider a range of cut-off values from one and a half times to two times the input traffic mean. Since such cut-off value equals the primary lightpath peak rate, we note that cut-off values below the input traffic mean would make the utilization factor of the primary lightpath grow beyond 100% in the single wavelength case. The utilization factor of the traffic shaper can be calculated as the peak rate divided by the average input traffic. Thus, the primary lightpath utilization factor is in the range from 66% (one and a half times the input traffic mean) to 50% (two times the mean). Even though the utilization factor is not large there are still peaks to be routed through the overflow lightpath. Figure 6 shows the original input traffic process and the overflow traffic for several cut-off values.

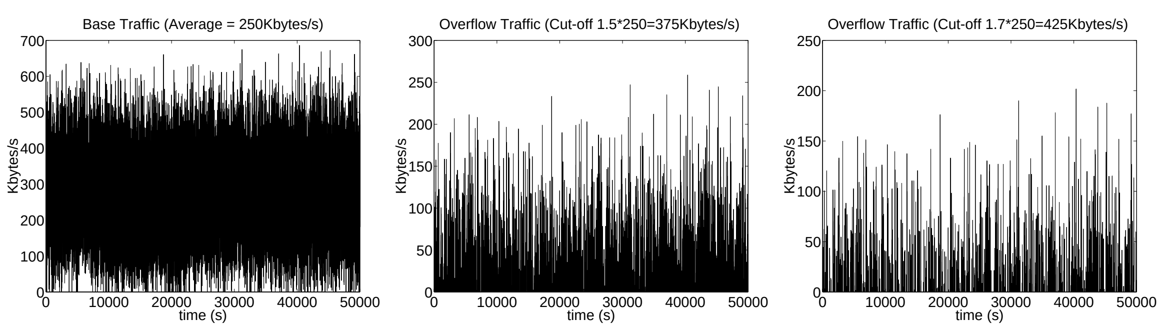

Figure 6: Original input and overflow traffic

Figure 7 shows the average delay for both network scenarios, dual lightpath case and no buffering

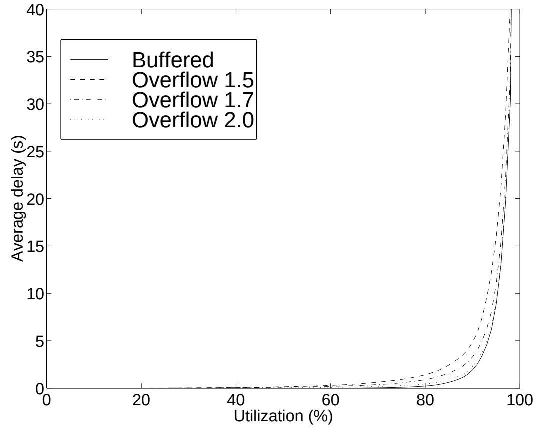

Figure 7: Average transfer delay in the overflow lightpath and input buffer system

and single lightpath and buffering (shaping) at the source router output. We note that the curves are similar, showing that both solutions are nearly the same as far as transfer delay is concerned. This is simply because a small percentage of the input traffic is routed through the overflow lightpath. As the cut-off threshold increases less traffic is routed through the overflow lightpath, as can be observed in figure 7. Indeed, the extreme burstiness of Internet traffic produces traffic bursts which are far away from the mean, even though the mass of the distribution is centered around the mean. Table 1 presents the traffic percentage routed through the overflow and primary lightpath, showing a small percentage of traffic in the overflow lightpath. Recall that values of H in the vicinity of 0.7 are the most usual in the Internet.

Cut-off value	Overflow Traffic for $H = 0.5$	Overflow Traffic for $H = 0.7$ (typical Internet case)	Overflow Traffic for $H = 0.9$
Input traffic average	13.5%	13.6%	22.0%
1.5 times input traffic average	1.1%	1.2%	4.8%
1.7 times input traffic average	0.23%	0.31%	2.1%
2 times input traffic average	0.01%	0.03%	0.5%

Table 1: Percentage of traffic routed through the overflow lightpath

Even though the results are quite similar for the overall system it is worth evaluating the delays for the overflow traffic. Figure 8 shows the same plots depicted in figure 7 but only the transfer delay in the overflow lightpath branch is shown, in comparison to the transfer delay in the single lightpath system. The results show that a considerable delay affects the packets which are routed through the overflow lightpath. As a result, the variance in the transfer delay increases dramatically, so that late packet arrivals may even generate retransmissions due to TCP timeouts. On the other hand, we note that the delay variance effect may induce packet misordering.

The results can be explained by looking at the evolution of the H and a parameters in the overflow traffic, which are shown in figure 9. We observe that while the H parameter decays slowly in comparison with the input traffic original value the a parameter takes on values well above the original FBT values. The increased traffic variability, which can also be observed in figure 6, makes

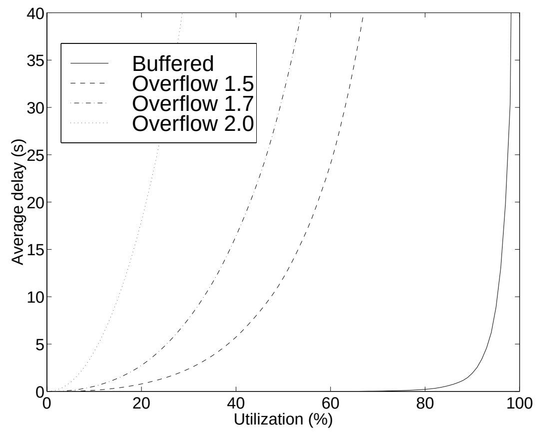

Figure 8: Average transfer delay in the overflow queue and primary queue

performance in the overflow queue decrease dramatically. Indeed, the queue length does not only depend on correlations but also on the variability of the marginal distribution (see figure 4). In this case, the extreme traffic bursts in the overflow lightpath translate into a performance drop in the overflow queue.

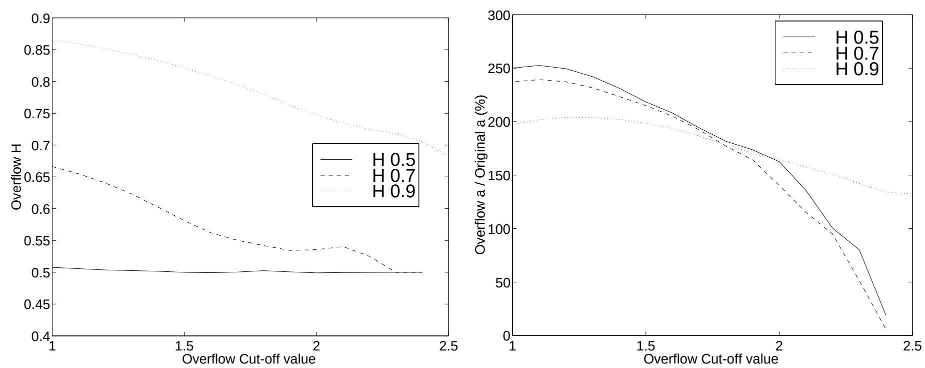

Figure 9: Values of a and H for the overflow traffic

The above results show that the overflow traffic requires a capacity increase in the destination router, which makes the utilization factor of the overflow traffic input queue decrease.Thus, our analysis leads to the following interesting result: if an overflow lightpath is in use the output buffering in the network is reduced but the destination router is forced to increase the packet forwarding capacity in order to treat packets fairly concerning transfer delay. Namely, we are posing more performance requirements in the electronic bottleneck. On the other hand, we note that since less than a 3% of the traffic is being directed to the overflow lightpath for the typical Internet case ($H = 0.7$) the traffic shaper utilization factor remains nearly the same.

We evaluate the processing speed increase which is due in the router in order to achieve the same transfer delay as with the single lightpath solution. Being C_{over} the router capacity assigned to the overflow traffic and C_{prim} the capacity assigned to the primary traffic, figure 10 shows the ratio $(C_{over} + C_{prim})/C_{single}$ where C_{single} represents the capacity needed in the single lightpath system to achieve the same transfer delay, which is set to 20 ms. Such ratio is plotted against the percentage of overflow traffic that is routed through the overflow lightpath. The results show a significant increase in the router capacity (in the vicinity of 20%) with small percentages of traffic routed through the overflow lightpath.

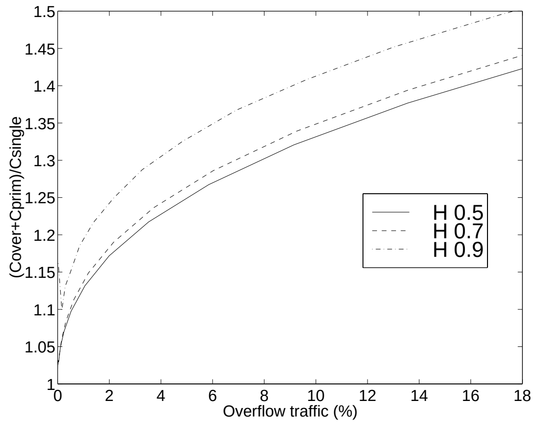

Figure 10: Percentage of capacity increased required in the router versus percentage of overflow traffic

Finally, we note that the single lightpath architecture with output shaping at the source router imposes little buffering requirements. For example, if the cut-off threshold (i.e. the primary lightpath peak rate) is one and a half times the traffic mean then the traffic shaper utilization factor is approximately 66% with the typical Internet case of $H = 0.7$, thus giving an average buffer occupancy in an infinite buffer system of 8 Kbytes.

4 Conclusions

We have analyzed a common WDM network scenario with primary and overflow lightpaths, in which the WDM spare capacity in the overflow lightpath is used to absorb traffic peaks [1], thus significantly reducing buffering in the network. The analysis leads to the following result: the use of the spare capacity in the overflow lightpath poses significant constraints in the network routers, which constitute the network bottleneck. Rather than using such spare bandwidth a single lightpath with an input shaper becomes more effective. Even though the availability of near-infinite bandwidth makes it easier to satisfy the large traffic demand of the Next Generation Internet, the traffic grooming and shaping operations are of fundamental importance to achieve the desired Quality of Service.

References

[1] B. Arnaud. Architectural and engineering issues for building an optical internet. In *Proceedings of SPIE International Symposium on Voice, Video, and Data Communications – All-Optical Networking: Architecture, Control, and Management Issues*, Boston, MA, November 1998.

[2] E. A. Varvarigos and J. P. Lang. A virtual circuit deflection protocol. *IEEE/ACM Transactions on Networking*, 7(3), June 1999.

[3] J. Bellamy. *Digital telephony.* John Wiley & Sons, New York, 1991.

[4] J. Beran. *Statistics for long-memory processes.* Chapman & Hall, New York, 1994.

[5] I. Norros. On the use of fractional brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13(6), August 1995.

[6] C. Huang, M.Devetsikiotis, I. Lambadaris, and A. R. Kaye. Fast simulation for self-similar traffic in ATM networks. In *Proceedings of the IEEE ICC95*, Seatle, June 1995.

[7] V. Paxson. Fast, approximate synthesis of fractional gaussian noise for generating self-similar network traffic. *Computer Communications Review*, 27(5), 1997.

[8] L. Kleinrock. *Queueing Systems.* John Wiley & Sons, New York, 1975.

[9] J. Aracil, D. Morató, and M. Izal. Analysis of Internet services for IP over ATM. *IEEE Communications Magazine*, December 1999.

[10] J. R. M. Hosking. Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20(12), 1984.