This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2010 proceedings

Fair Rate Degradation in Flow-Aware Networks

Robert Wójcik[1], Jerzy Domżał[2], Andrzej Jajszczyk[3]
[1]*Student Member, IEEE*, [2]*Member, IEEE*, [3]*Fellow, IEEE*

***Abstract*—In this paper, it is shown that the admission control routine in Flow-Aware Networks (FAN) may lead to severe fair rate degradation, which negatively impacts the performance of streaming applications. In order to prevent this negative behavior, the limitation mechanism is proposed. The aim of the mechanism is to limit the maximum number of new flows that may be admitted on a link between any two consecutive network's auto-measurements. The solution is efficient, viable and dramatically reduces the fair rate degradation.**

***Index Terms*—Flow-Aware Networks, FAN, service differentiation, QoS.**

I. Introduction

Flow-Aware Networking (FAN) is a QoS architecture proposal for the IP networks. To assure ceratin level of guaranteed bandwidth some admission control procedures must be applied. In FAN, admission control is measurement-based. Moreover, as FAN does not use any kind if signaling, network routers are not aware of the incoming flow characteristics. This fact makes the admission decisions more challenging than in case of, e.g., IntServ supported IP or ATM, where transmission parameters are more or less known *a priori*.

FAN intends to provide a minimum level of resources for each active flow. It does that by blocking new flows when congestion indicators exceed their fixed thresholds. It is assumed that those thresholds define the minimum level of service on each FAN link. However, as shown in the paper, this assumption cannot be made, as when many new flows arrive at the same instant, the thresholds are significantly exceeded. To eliminate the problem, we propose using a limitation mechanism which not only improves QoS assurance capabilities, but also enhances the scalability of the FAN architecture.

The remainder of the paper is organized as follows. Section II introduces the reader to the concept of Flow-Aware Networking, explaining in particular, how measurement-based admission control works. Section III exposes a fair rate degradation problem of FAN networks, i.e., inability to ascertain assumed Quality of Service when the number of incoming connections is large. Afterwards, in Section IV a solution to resolve the problem is proposed, while Section V concludes the paper.

II. Flow-Aware Networking

The idea of Flow-Aware Networking as an approach to assure Quality of Service in packet networks was initially introduced in [1] and, then, presented as a complete system in 2004 [2] as a result of inconveniences regarding the current QoS architectures [3]. The main goal of FAN is to enhance the current IP based networks by improving their performance under congestion. To achieve that, the following traffic management mechanisms are introduced: measurement-based admission control [4] and fair queuing with priorities [2], [5]. The former keeps flow rates sufficiently high by applying blocking criteria to new flows when congestion is noticed. The latter ensures that the bandwidth is fairly shared between all active flows, while ensuring negligible packet delays for flows emitting at lower rates. Compared to other QoS architectures, FAN requires no signaling and is scalable [6], [7]. FAN is also a net neutral friendly QoS architecture as explained in [8]. Finally, FAN can cooperate with the underlying optical layers to provide fast and efficient recovery in case of network element failures [9].

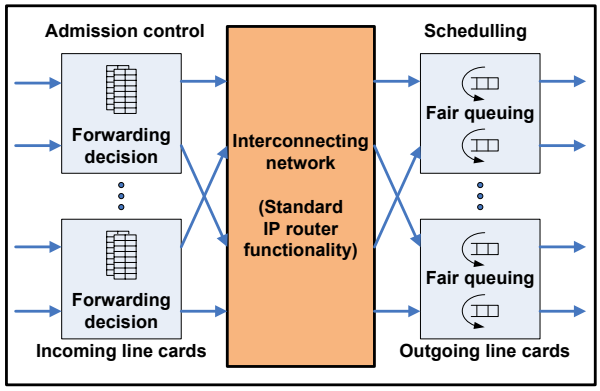
The goal of FAN is to enhance the perceivability of the current IP network. In order to do that, FAN introduces a unique router, named: cross- protect router (XP router). Figure 1a) shows a concept diagram of an XP router. FAN adds only two blocks to the standard IP router, namely the admission control and scheduling blocks. The former is placed in the incoming line cards of the router, whereas the latter is situated in the outgoing line cards. This device alone is responsible for providing admission control and fair queuing.

Naming FAN devices as cross-protect routers is a result of mutual cooperation and protection, which exists between both additional blocks. The admission control block limits the number of active flows in a router, which improves the queuing algorithm functionality, and reduces its performance requirements. It is vital that queuing mechanisms operate quickly, as for extremely high speed links the available processing time is limited. On the other hand, the scheduling block provides admission control with the information on congestion status on the outgoing interfaces. The information is obtained from, for example, current queues occupancy. The mutual protection contributes to the smaller protected flow list and active flow list sizes, which significantly improves FAN's scalability.

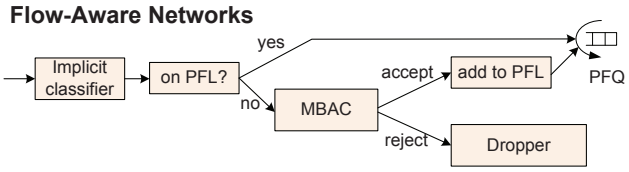
Figure 1b) presents the operation of FAN. All incoming packets are, firstly, classified into flows. The flow identification process is implicit and its goal is not to divide flows into different classes, but only to create an instance on which the service differentiation will be performed. Then, all the flows that are currently in progress, i.e., are present on the Protected Flow List (PFL) are forwarded unconditionally, whereas all new flows are subject to admission control. The admission control in FAN is measurement based (MBAC) which implies

R. Wójcik, J. Domżał and A. Jajszczyk are with the Department of Telecommunications, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Kraków, Poland (e-mail: {robert.wojcik, jdomzal, jajszczyk}@kt.agh.edu.pl).

978-1-4244-6404-3/10/$26.00 ©2010 IEEE



(a)



(b)

Fig. 1. Cross-Protect Router: architecture (a) and operation (b)

that the accept/reject decisions are based only on the current link congestion status. If a new flow is accepted, it is put onto the PFL list and then all forthcoming packets of this flow are forwarded without checking the status of the outgoing link by MBAC.

There are two per-flow fair queuing algorithms proposed for FAN: Priority Fair Queuing (PFQ) and Priority Deficit Round Robin (PDRR). Both algorithms have, logically, one priority queue and a secondary queuing system. They are intended to realize fair sharing of link bandwidth to elastic flows and priority service to streaming flows. The latter (PDRR) was primarily suggested to speed up commercial adoption since it improves the algorithm complexity from $O(log(N))$ to $O(1)$; where N is the number of currently active flows. However, it has been shown that both scheduling algorithms have similar performance [10].

To provide admission control block with the proper link status, two indicators are measured periodically: priority load (PL) and fair rate (FR). The two parameters used for flow admission criteria are estimated in the scheduling block of the XP router. Both indicators are measured periodically.

An estimation of the priority load is derived from Equation 1. Variables $pb(t)$ represent the values of a counter, incremented on the arrival of each priority packet by its length in bytes, at time t. (t_1, t_2) is a measured time interval (in seconds), and C is the link bit rate. Priority load, therefore, represents the sum of the lengths of priority packets transmitted in a certain time interval, divided by the duration of that interval, and normalized with respect to the link capacity.

$$PL = \frac{(pb(t_2) - pb(t_1)) \times 8}{C(t_2 - t_1)} \quad (1)$$

Equation 2 is used to calculate the fair rate, which is an estimation of the rate currently realized by backlogged flows. In other words, it is the rate available to each flow at the moment. To estimate the fair rate, a fictitious flow emitting single byte packets is considered. In an idle period, that flow could transmit at the link rate. Otherwise, the number of bytes that could have been transmitted is given directly by the evolution of virtual time. In Equation 2, $vt(t)$ is the value of the virtual time at time t, (t_1, t_2) is the measurement interval, S is the total idle time during the interval and C is the link capacity.

$$FR = \frac{max\{S \times C, (vt(t_2) - vt(t_1)) \times 8\}}{(t_2 - t_1)} \quad (2)$$

The more detailed explanation of the formulas (1) and (2) and all required definitions may be found in [2].

If the current FR is lower than the pre-defined minimum FR (min_FR), or the current PL is greater than the maximum pre-defined PL (max_PL), the congestion state is noticed. Under congestion, only those flows that were already active can be forwarded. Each router maintains the protected flow list (PFL) and all flows that are on the list are forwarded regardless of the congestion status. If a flow is not signed on the PFL, and the link is congested, it is blocked and needs to wait for the resources to be available again.

It has been documented in [11] that FAN provides superior transmission quality for VoIP based streaming flows, and that the emerging problem of waiting for the resources can be overcome by introducing differentiated blocking. Moreover, the flushing mechanism [12] can also decrease the waiting times for streaming flows in congested networks [12], [13]. However, streaming flows are not only affected by long waiting times. They also suffer from low FR and, unfortunately, FAN is susceptible to FR degradation under heavy congestion.

The purpose of this work is to expose and document that negative aspect of FAN, i.e., FR degradation, and to present a viable solution to enhance the architecture. It is shown, through simulations, that by simple means we can improve performance of the FAN networks, greatly reducing the FR degradation.

III. Fair Rate Degradation

The cross-protect mechanism, in FAN, is supposed to provide at least a minimum fair transmission rate to all the active flows. To achieve that, each time the measured FR drops below the min_FR threshold, the admission control starts blocking all new connections. Therefore, this procedure does not guarantee to maintain the min_FR value under congestion since proper actions are undertaken only after the min_FR boundary is crossed.

In theory, the fair rate should be allowed to drop below the threshold only slightly before the admission control block starts functioning. Unfortunately, in practice, the FR drops might be significant.

To present the problem, we performed numerous simulations using the ns-2 simulator [14]. We analyzed a 100 Mbit/s

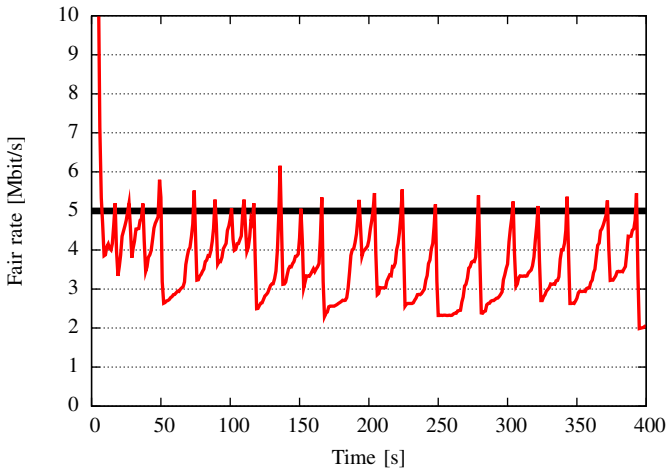


Fig. 2. Measured FR values over time on a congested FAN link

FAN link. The min_FR parameter was set to 5% of the link capacity (5 Mbit/s) and was measured every second. We provided the traffic pattern with the Pareto distribution for calculating the volume of traffic to be sent by each flow (15 MB on average). We used the exponential distribution for generating the time intervals between the beginnings of the transmissions of the flows. The duration of each simulation run was set to 800 s. The warm-up period was considered until the FR value reached the min_FR threshold for the second time. 95% confidence intervals were calculated by using the Student's t-distribution.

Figure 2 demonstrates the problem as it shows the measured fair rate values over time on a severely congested FAN link, with 2500 flows arriving with the intensity of, on average, 3.5 flows per second. As we can see, FR drops well below the min_FR threshold (5% of the link capacity, marked with a solid flat line). Such situations occur because between two consecutive FR measurements many new flows arrive and are admitted before the router realizes that the admission control should be in the blocking state.

These FR drops are obviously unwanted since they are dangerous to streaming applications which require a certain available bandwidth. The whole concept of FAN is that this bandwidth (min_FR) can be provided for such flows. However, Figure 2 shows that FAN fails in providing this key quality.

There are two approaches to mitigate the problem. One is to reduce the interval between two consecutive measurements of the fair rate. If the FR is estimated more often, statistically fewer flows are admitted between the measurements and the system reacts more quickly. The problem with this method is that frequent estimations require more computational power from the router's CPU. This issue becomes even more significant in core networks, as those devices deal with numerous flows and must react almost instantly. The second, straightforward, but viable approach is to limit the number of flows that can be admitted between two measurements. We discuss this solution in the next section.

IV. Limitation mechanism

In the literature, numerous admission control mechanisms have been proposed over the years, mainly for Integrated Services, Differentiated Services or call admission control procedures in ATM. The PhD dissertation of A. W. Moore [15] contains the detailed comparison of them. However, most of the proposals rely on the fact that at least a limited information about the incoming flow is available through signaling. As FAN does not use any kind of signaling, those methods are not applicable.

Some of the admission control mechanisms also notice the problem of over admitting. In [16], it is stated that the system needs to wait for a period of time after any change of the number of connections in progress happens, before the link congestion status can be re-estimated. The author proposes a timescale solution, i.e., to regard a time interval as a function of the number of active flows. As the number grows considerably, the interval is decreased to reduce the probability that a situation in a link changes significantly within that interval. Unfortunately, as already mentioned, increasing the frequency of measurements imposes more strict demands on the router CPUs. Therefore, we argue and prove in this section, that by providing even the simplest limitation mechanism, we can resolve the over-admitting problem while not increasing the routers' computational power requirements.

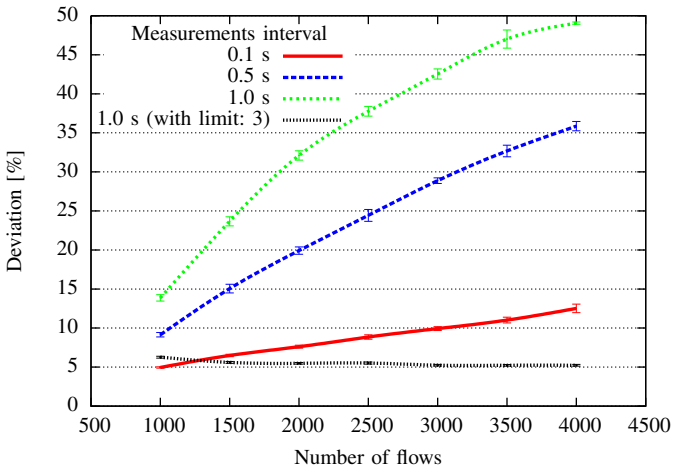
The limitation mechanism in FAN enhances the functionality of the admission control block. The idea is that between any two consecutive measurements, only a limited, fixed number of new flows may be admitted. This approach protects the admission control block from over-admitting, i.e., from allowing too many new flows to acquire access to the link, which, consequently, degrades the FR.

To provide limitations, we need to introduce only a simple counter, incremented on arrival of each new flow, and reset on each measurement. When it reaches a certain number, all new flows are rejected. This way, the extra CPU power required is hardly noticeable, while the benefits are significant.

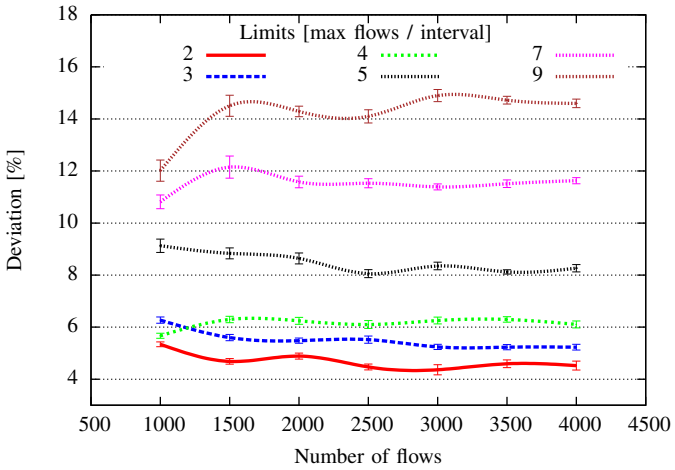
To present the problem numerically and to show how limitations can mitigate FR degradations, we define a mean deviation from the min_FR threshold, as follows:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{|min\_FR - FR_i|}{min\_FR}\cdot 100\% \qquad (3)$$

where FR_i are the measured FR values over time. This parameter shows how much measured FR values differ from the min_FR during the total measurement time (simulation time). As, in all cases, we simulate only the overloaded links, the ideal FR values should oscillate around the threshold and the deviation should be near zero. Figure 3 shows the mean deviation of the measured FR with respect to the number of active flows. All the simulation scenario parameters were presented in the previous section. The number of active flows is bond with the intensity of their arrival by the following formula:



(a)



(b)

Fig. 3. Mean deviation of the measured FR from the min_FR threshold with respect to: a) the measurement interval length, b) the maximum number of flows accepted in one interval.

$$\frac{N}{\lambda} = T = const \tag{4}$$

where: N is the number of active flows, λ the intensity of their arrival and T is the simulation time. The intensity was set in the simulator by changing the mean interval between the beginnings of the transmissions of new flows, according to the formula:

$$\lambda = \frac{1}{t_e} \tag{5}$$

where t_e is the interval to the beginning of the next flow's transmission obtained from the exponential distribution. 10 simulations were performed for each case, to calculate the 95% confidence intervals using the Student's t-distribution.

As shown in Figure 3a), when limiting is not applied (three rising curves), the deviation rises along with the number of active flows, and is greater when larger measurement intervals are set. Both dependencies are natural. The number of active flows, associated with the statistical intensity of their arrival, impacts the number of flows which request the resources every second, while the measurement interval impacts the duration of that arrival. Both factors contribute to the fact that more or fewer flows may be over-admitted. However, when the limitation is used (the flat line, 3 flows per measurement in this case) we observe almost constant deviation, and much smaller than that when limits are not applied. The fact that the deviation does not increase with the number of active flows helps to administer the network, as the operators can keep the links in a proper condition regardless of the current network overload.

The limit of 3 flows per measurement was chosen experimentally. Figure 3b) shows how various limits impact the FR deviation under the same conditions. We have noticed, that choosing a limit too strict (low) results in under-feeding the link, as the link serves flows faster than they can be admitted. Such a case happened when we applied the limit of 1 flow per measurement. On the other hand, choosing a high limit does not solve the problem well, as the deviations start to rise. In Figure 3b) we can observe that limiting the admission of new flows to 2–4 per measurement is sufficient under these network conditions. The deviation is around 5-6% and is totally acceptable.

Right now, it is only a matter of experiments to pick the right limit. In the presented simulation scenario, picking the limit of 2-4 flows per measurement is adequate, however, in other situations, especially on links with a different capacity, it will differ. That is why it is necessary to investigate the limit selection process in future. Moreover, we plan to develop a mechanism to automatically choose the proper limit and embed it into the FAN routers. In such a way the system will become more robust and invulnerable to faulty set-ups. In Section V, some additional information regarding the future plans are presented.

Table I shows the mean percentage of time in which FR drops below 90% (a) and 80% (b) of the min_FR. This characteristic is very important for streaming applications which require a certain amount of bandwidth to be available. Being aware of the fact that FR boundaries can be and are constantly crossed, the network administrator might want to set the threshold a bit higher, e.g., to provide a guaranteed level of 5% of the link's bandwidth, a min_FR value could be set to, say 7%. However, as observed in Figure 2 this approach may be deceiving, as FR degradations are uncontrollable and unforeseeable.

As can be seen in Table I, by introducing limitations (marked rows), we can drastically reduce the FR degradation. When limitations are present, the FR drops below 90% of its minimum threshold 5 to 10 times less than in the comparable situation (measurements once every second). The outcome is even more convincing in the second case, as the FR value hardly ever drops below the 80% of the min_FR threshold, which is a firm result. Similarly, as in case of deviation, this characteristic is almost independent of the number of flows when limitations are applied.

TABLE I
THE PERCENTAGE OF TIME IN WHICH FR DROPS BELOW 90% (A) AND 80% (B) OF THE MIN_FR THRESHOLD

Measurement interval	Number of flows						
	1000	1500	2000	2500	3000	3500	4000
90% (a)							
0.1	10.96 ± 0.92	25.18 ± 1.39	34.01 ± 1.26	41.60 ± 1.02	47.32 ± 1.33	52.59 ± 1.55	58.06 ± 1.85
0.5	34.82 ± 1.26	60.28 ± 1.81	73.49 ± 0.85	80.12 ± 1.08	83.22 ± 0.44	86.73 ± 0.55	87.26 ± 0.32
1.0	53.61 ± 1.74	76.69 ± 0.65	85.00 ± 0.48	87.49 ± 0.33	89.01 ± 0.51	91.69 ± 0.96	92.29 ± 0.19
1.0 (limit: 3)	**13.90 ± 1.11**	**10.49 ± 0.85**	**8.89 ± 0.97**	**9.16 ± 0.84**	**7.97 ± 0.75**	**8.21 ± 0.76**	**7.58 ± 0.94**
80% (b)							
0.1	0.03 ± 0.06	1.18 ± 0.28	3.50 ± 0.71	7.21 ± 0.71	11.56 ± 0.99	15.64 ± 1.98	21.62 ± 2.59
0.5	9.07 ± 1.52	30.32 ± 2.12	47.53 ± 1.78	59.04 ± 2.09	68.40 ± 0.82	74.00 ± 1.50	76.11 ± 0.56
1.0	24.62 ± 1.30	56.46 ± 2.11	72.73 ± 1.19	77.79 ± 0.54	82.42 ± 0.57	85.83 ± 0.93	84.84 ± 0.32
1.0 (limit: 3)	**0.10 ± 0.06**	**0.01 ± 0.03**	**0.01 ± 0.03**	**0.00 ± 0**	**0.00 ± 0**	**0.00 ± 0**	**0.00 ± 0**

For the purpose of comparison, Table I also shows the times when we increase the frequency of measurements. They show that reducing the inter-measurement time even 10 times does not provide better performance than introducing a simplest limitation. This, essentially, proves that increasing the frequency of measurements is a much worse option to mitigate the FR degradation problem than the limitations. Finally, even if it would be possible to provide a proper frequency of measurements, still the dependency on the number of active flows remain and cannot be neglected.

V. CONCLUSION

Flow-Aware Networking is a simple and efficient architecture to provide Quality of Service differentiation in the IP networks. However, this proposition is relatively new and still needs some improvements or additional mechanisms.

We have shown that frequent degradations of the FR may occur on FAN links when there are too many flows attempting to acquire the access to the link's bandwidth. To prevent those degradations, either FR needs to be measured more often, or we need to introduce some sort of limitations. The first option, as explained, consumes much more router's CPU power which is undesirable. Limitations, on the other hand, are viable, easy to implement and the benefits from introducing them are remarkable.

In this paper we have investigated the simplest method of limitation, i.e., the static hard-coded limit, pre-set by the administrator. Despite its simplicity, the improvement is significant. In future, more complex methods of limitations are to be investigated. Moreover, we plan to create a mechanism which will automatically select the most suitable limit and embed it into the FAN routers. In such a way the system becomes more robust and invulnerable to faulty set-ups.

ACKNOWLEDGMENT

Project operated within the Foundation for Polish Science Ventures Programme co-financed by the EU European Regional Development Fund.

REFERENCES

[1] J. Roberts and S. Oueslati, "Quality of Service by Flow Aware Networking," *Philosophical Transactions of The Royal Society of London*, vol. 358, pp. 2197–2207, 2000.

[2] A. Kortebi, S. Oueslati, and J. W. Roberts, "Cross-protect: implicit service differentiation and admission control," in *Proc. High Performance Switching and Routing (HPSR)*, Phoenix, USA, 2004, pp. 56–60.

[3] J. Roberts, "Internet Traffic, QoS and Pricing," in *Proceedings of the IEEE*, vol. 92, September 2004, pp. 1389–1399.

[4] A. Kortebi, S. Oueslati, and J. Roberts, "MBAC algorithms for streaming flows in Cross-protect," in *EuroNGI Workshop*, Lund, Sweden, June 2004.

[5] ——, "Implicit Service Differentiation using Deficit Round Robin," in *Proc. International Teletraffic Congress ITC19*, Beijing, China, August/September 2005.

[6] A. Kortebi, L. Muscariello, S. Oueslati, and J. Roberts, "On the scalability of fair queueing," in *Proc. ACM HotNets-III*, San Diego, USA, November 2004.

[7] ——, "Minimizing the overhead in implementing flow-aware networking," in *Proc. of the 2005 ACM symposium on Architecture for networking and communications systems (ANCS)*. New York, USA: ACM, 2005, pp. 153–162.

[8] J. Domzal, R. Wojcik, and A. Jajszczyk, "Qos-aware net neutrality," in *Proc. First International Workshop on Neutral Access Networks (NEUTRAL)*, Cannes, France, August 2009.

[9] J. Domzal, R. Wojcik, K. Wajda, A. Jajszczyk, V. Lopez, J. A. Hernandez, J. Aracil, C. Cardenas, and M. Gagnaire, "A multi-layer recovery strategy in fan over wdm architectures," in *Design of Reliable Communication Networks (DRCN)*, Washington, USA, October 2009.

[10] A. Kortebi, L. Muscariello, S. Oueslati, and J. Roberts, "Evaluating the number of active flows in a scheduler realizing fair statistical bandwidth sharing," in *Proc. of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*. New York, USA: ACM, 2005, pp. 217–228.

[11] A. Jajszczyk and R. Wojcik, "Emergency Calls in Flow-Aware Networks," *Communications Letters, IEEE*, vol. 11, pp. 753–755, September 2007.

[12] J. Domzal and A. Jajszczyk, "The flushing mechanism for mbac in flow-aware networks," in *Proc. Next Generation Internet Networks NGI 2008*, Krakow, Poland, 2008, pp. 77–83.

[13] J. Domzal, R. Wojcik, and A. Jajszczyk, "Reliable transmission in flow-aware networks," in *IEEE Global Communications Conference (GLOBECOM)*, Honolulu, USA, 2009.

[14] The Network Simulator: ns-2, http://www.isi.edu/nsnam/ns/.

[15] A. W. Moore, "Measurement-based management of network resources," Ph.D. dissertation, University of Sambridge, April 2002.

[16] N. G. Bean, "Robust connection acceptance control for atm networks with incomplete source information," *Annals of Operations Research*, vol. 48, no. 4, August 1994.