Blocking Probability in Pipeline Forwarding Networks

Mario Baldi, and Andrea Vesco

Original scientific paper

***Abstract*—As multimedia communications continue to grow steadily on the Internet, pipeline forwarding (PF) of packets provides a scalable solution for engineering delay-sensitive traffic while guaranteeing deterministic Quality of Service (QoS) with high resource utilization. In PF networks resource reservation, while ensuring deterministic QoS on a per-flow basis, can result in a not null blocking probability. A reservation request may fails due to enough resources being available but not during the proper time frames. This work analyses blocking probability of reservation requests since it affects the capability of utilizing network resources to carry traffic with deterministic QoS. The blocking probability and, consequently, the achievable network utilization are analytically derived on general topology PF networks as function of the traffic intensity given the traffic matrix and the network routing. The correctness of the blocking models is also assessed by simulation in different scenarios. This work represent a valuable contribution over previous analytical models of the blocking probability as their application to real size scenarios is impractical due to their computation complexity.**

***Index Terms*— Pipeline forwarding, blocking probability, quality of service support.**

I. Introduction

Various multimedia applications, such as voice-over-IP, videoconferencing and 3D video communication, are becoming widely available. Traffic generated by these applications is referred to as delay-sensitive as timely packet delivery is important for them to work properly.

Packet networks, originally designed for data applications, are not engineered to tightly control the delay packets experience in routers where they might contend for resources, e.g., transmission capacity, consequently be queued for a variable time, and possibly be dropped. Sophisticated packet scheduling mechanisms can be introduced with various trade-off points between complexity and efficiency in the deployment of network resources [6]. The differentiated service model (DiffServ) is widely deployed in today's networks due to its limited complexity, although it imposes hard limitations on the network utilization by traffic with Quality of Service (QoS) requirements [1]. This is acceptable as long as the amount of such traffic is small compared to best effort one, i.e., it utilizes a small fraction of the network capacity [3]. Several studies found delay sensitive traffic to be constantly on the rise [2].

Ultimately, DiffServ based solutions become critical if such growth is faster than technology enables proportionally more powerful network infrastructures at the same cost per capacity unit.

The pipeline forwarding (PF) of packets [4] is a possible scalable solution for engineering delay-sensitive traffic on the Internet [5] while providing deterministic QoS [6][7] without penalizing network utilization.

In PF networks the resources are reserved on a per-flows basis to guarantee deterministic QoS. The availability of resources depends on the solution of a distributed scheduling problem. Thus, the reservation requests may experience a not null blocking probability also when resources are still available. As a result a blocking probability affects the capability of utilizing network resources to carry traffic with deterministic QoS.

The PF blocking issues have been analytically addressed in [4], more recently in [8], and extensively assessed by simulation in [9]. However, the computation complexity of previous analytical models makes their application impractical to real size scenarios. This work overcomes such limitation by adopting a novel approach in the blocking analysis. Blocking is expressed as a function of the model of the reservation request arrival process at the network ingress and the model of the reservation holding process, given the traffic matrix, the network topology and routing.

Firstly, it is shown how to relate the traffic intensity with the steady state distribution of the number of active reservations on network links. Then the general form of the blocking probability is derived as function of this steady state distribution on single and multi-hop routes. Finally it is shown how to estimate network-wide blocking in a given scenario as function of the mean arrival rate of reservation requests on all the available network routes. Notably, it is claimed that under the proper assumptions the blocking models, here derived, provide an upper bound to the blocking probability. Simulation results are also provided to verify the correctness of the models.

The paper is organized as follow. Section VI discusses PF by presenting its operating principles and the blocking problem in details. The blocking probability analysis and the assumptions it is related to, are presented in Section III while the reservation-level simulator and extensive results are presented in Section IV. In Section V the related work and

Manuscript received November 4, 2013; revised April 3, 2014.

Mario Baldi is with the Department of Computer and Control Engineering, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy.

Andrea Vesco is with the Istituto Superiore Mario Boella (ISMB), Via P.C. Boggio, 61, 10129 Torino, Italy.

and the differences with the previous approaches are discussed. Finally, conclusions are drawn in Section VI.

II. Pipeline forwarding

A. Operating principles

The pipeline forwarding is a well-known optimal method widely adopted in computing and manufacturing. In its networking implementation [4] all nodes are synchronized with a common time reference (CTR), while utilizing a basic time period called time frame (TF). The TF duration T_f may be derived, for example, as a fraction of the UTC second received from a time-distribution system such as the global positioning system (GPS) and Galileo. The CTR structure depicted in Fig. 1 is composed of TFs grouped into time cycles which are further grouped into super cycles; each super cycle lasts for one UTC second.

In PF networks TFs are partially or totally reserved on a per-flow basis during a resource reservation phase. It follows that packets are timely moved along their path and served at well-defined instants at each node. Nodes therefore operate as they were part of a *pipeline*, from which the technology's name is derived.

The basic pipeline forwarding operation is regulated by two simple rules:

1. all packets that must be sent in TF t by a node must be in its output ports' buffers at the end of TF t-1;
2. a packet p transmitted in TF t by a node n must be transmitted in TF $t+d$ by node n+1, where d is a predefined integer called *forwarding delay*, and TF t and TF $t+d$ are referred to as the *forwarding TF* of packet p at node n and node n+1, respectively; the value of the forwarding delay is determined at resource-reservation time when the TF $t+d$ is scheduled. d is large enough to satisfy previous rule.

Moreover two options of the basic pipeline forwarding operation are possible. When node n deploys *immediate forwarding*, the forwarding delay — measured in TFs — is equal to the propagation delay $\tau_{n,n+1}$ between node n and $n+1$ plus one TF for all packets forwarded by node n. When node n deploys *non-immediate forwarding* it may use different forwarding delays for packets belonging to different flows. Thus the forwarding delay ranges from $\tau_{n,n+1} + \text{TF}$ to $\tau_{n,n+1} + k \cdot \text{TF}$.

In PF networks, a *schedule* is the set of TFs along a path of subsequent nodes that satisfy the forwarding rules. A *synchronous virtual pipe* (SVP) is created by allocating an *available schedule*, *i.e.*, a set of TFs not reserved to other flows, along a route of subsequent nodes, reserved for forwarding a pre-reserved amount of bits – for this reason called *forwarding TFs*. As a result a SVP is an independent time-invariant channel with deterministic upper bounded delay, assured bandwidth and loss/late probability equal to zero, under the assumption that incoming traffic is policed and shaped not to exceed the total capacity of the reservation. Thus, traffic crossing the network into a SVP experiences (*i*) bounded end-to-end delay, (*ii*) low delay jitter, upper bounded by one TF, and (*iii*) neither congestion nor resulting loss. It is worth noting that the deterministic QoS, in term of bandwidth and end-to-end delay, provided by a SVP is independent on the network utilization. That is, if an SVP can be created from end to end, the QoS provided does not depend on the actual network utilization, and it does not negatively affect the QoS provided by the already created SVPs.

Non-pipelined packets, *i.e.*, packets that are not part of a SVP, can be transmitted during any unused portion of a TF, whether it is not reserved or it is reserved but currently unused, for example, because flows with reserved resources generate fewer packets than expected. A large part of Internet traffic today is generated by TCP-based elastic applications (*e.g.*, file transfer, e-mail, WWW) that do not require a guaranteed service in term of end-to-end delay and jitter. Such traffic can be dealt with as non-pipelined, *i.e.*, best effort traffic, and can benefit from statistical multiplexing. Each PF node performs statistical multiplexing of best-effort traffic forwarding these packets during any unused TF portions. In principle any service discipline can be applied to this traffic, for example, the DiffServ model.

As a result, SVPs are not at all TDM-like circuits. They are virtual channels providing guaranteed service in terms of bandwidth, delay, and delay jitter, but fractions of the link capacity not used by the traffic crossing the network inside a SVP can be utilized by non-pipelined packets.

B. The blocking problem

In PF networks, resources must be reserved in the form of

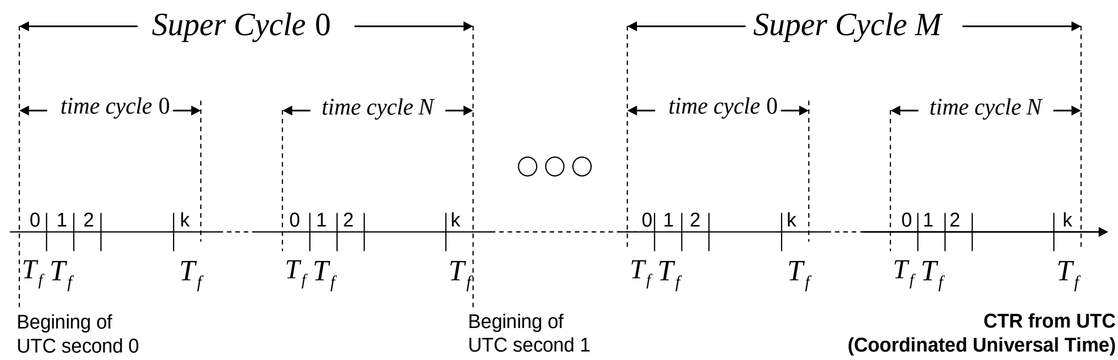

Fig. 1 Common Time Reference structure.

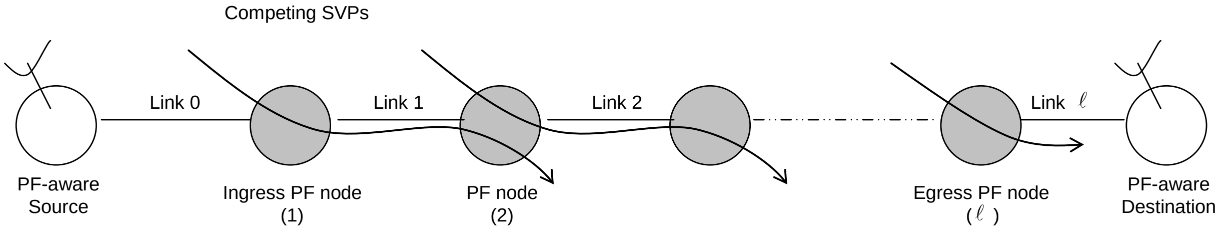

Fig. 2 ℓ -hop route w_r with PF-aware end systems.

transmission capacity during specific TFs to provide QoS guarantees on a per-flow basis. Given the route a flow takes through the network, the identity of the TFs reserved on a link is bound to the identity of TFs reserved on the previous link because of PF operating principles. For example, if *immediate forwarding* is deployed, once the identity of a TF (*i.e.*, the index, or position, of the TF inside the time cycle) to be reserved on a link is fixed, the identity of the corresponding TFs on all the subsequent links of the route is uniquely determined. For example, if the forwarding delay is determined to be *d* between each pair of nodes, once TF *j* is chosen on the first link, TF *j*+*d* must be reserved on the second one, *j*+2•*d* on the third one, and so on. Therefore, TFs reserved on a link impose constraints on reservations to be performed on adjacent links.

Hence, reserving resources for a flow requires solving a *scheduling problem* to find an available schedule on the route from source to destination. It might happen that, even there are TFs with available resources on the various links, they identity do not match the timing (i.e., the constraints on their position within the time cycle) resulting from the chosen forwarding delay (i.e., ultimately from PF operating principles). In this case the reservation request is said to be *unschedulable*. In other words, the reservation request can be *blocked* (and therefore rejected) even if enough resources are available on all the links on that flow's route.

This paper studies the probability for a resource reservation to be blocked. The *blocking probability* with respect to the average network utilization is an index of the efficiency in the utilization of resources as it provides an indication of how much traffic can be handled by the network before new flows have a non-negligible probability of being rejected.

Unschedulability does not exist in asynchronous packet networks because resources (*e.g.*, transmission capacity, buffer space, processing capacity) are reserved based solely on their not being already booked, independently of the specific time at which they will be used (which is, in fact, unknown). This does not imply that network resources can be fully reserved to traffic flows and reservation requests cannot be blocked. In fact, resource reservation is based on various heuristic procedures, commonly called admission control, that maintain network utilization (very) low in order to control the QoS with practical solutions, such as DiffServ.

III. Blocking Analysis

Table I summarizes the notation adopted in the following analysis.

TABLE I
SYMBOLS

TC	Number of TFs in the time cycle.
w_r	A generic route on the network.
ℓ	Number of PF nodes in route w_r .
λ_{w_r}	Mean arrival rate of reservation requests on route w_r .
λ_j	Mean arrival rate of reservation requests on link *j*.
μ	Mean holding time of reservations, *i.e.*, SVPs.
ρ_j	Traffic intensity on link *j*.
$\overline{x}_i^{(j)}$ and $x_i^{(j)}$	State of an available, *i.e.*, not reserved, and not available, *i.e.*, reserved, respectively TF *i* on link *j*.
$\pi_i^{(j)}$	Probability of having *i* active reservations, *i.e.*, SVPs, on link *j*.
$\Pi_j = \left\{\pi_0^{(j)}, \pi_1^{(j)}, ..., \pi_{TC}^{(j)}\right\}$	Steady state distribution of the number of active reservations, *i.e.*, SVPs, on link *j*.
S	Traffic matrix.
R	Routing matrix.

The blocking analysis is organized in three incremental steps (*i*) the model is defined and assumptions identified, then (*ii*) the arrival of resource reservation requests on a link in isolation is modelled and later used as the basis for the calculation of the (*iii*) blocking probability on a multi-hop route involving multiple links, which proves

Theorem. *The blocking probability* $P_{w_r}(B)$ *experienced by reservation requests issued by PF-aware sources on a* ℓ *-hop route* w_r *, such as the one depicted in Fig. 2, is given by*

$$P_{w_r}(B) = \sum_{k=0}^{TC} \pi_k^{(0)} \cdot \left(1 - \prod_{j=1}^{\ell} \sum_{z=0}^{TC} \left(1 - \frac{z}{TC}\right) \pi_z^{(j)}\right)^{TC-k} \quad (1)$$

The probability $\pi_i^{(j)}$ *of finding i active reservations, with* $1 \le i \le TC$ *, on link j is given by*

$$\pi_i^{(j)} = \frac{\pi_0^{(j)}}{i!} \rho_j^{\ i} \quad (2)$$

where $\rho_j = \lambda_j / \mu$ *is the traffic intensity on link j,* λ_j *is the mean arrival rate of reservation requests on link j and the*

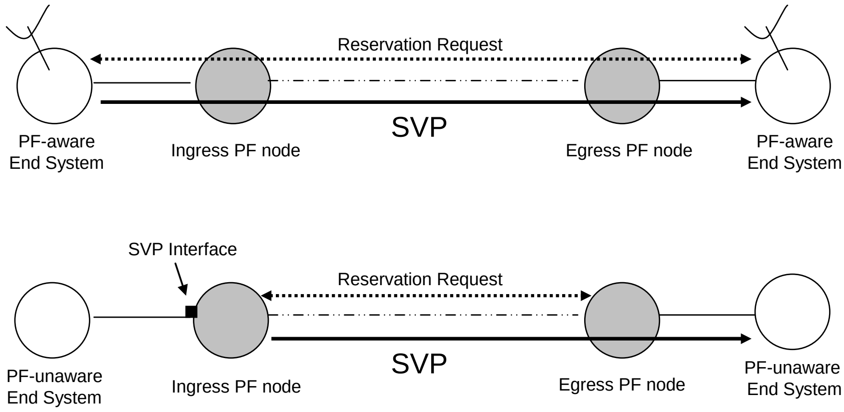

Fig. 3 PF-aware (top) and PF-unaware (bottom) end systems scenarios.

probability for all the TFs on link j to be available is given by

$$\pi_0^{(j)} = \frac{1}{\sum_{z=0}^{TC} \frac{\rho_j^{\,z}}{z!}} \tag{3}$$

under the assumption of statistical independence of TF status.

The network-wide blocking probability $P_{net}(B)$ can be then expressed as function of the blocking probability $P_{w_r}(B)$ on the υ routes $w_1, w_2,..,w_\upsilon$ existing in the network as

$$P_{net}(B) = \sum_{r=1}^{\upsilon} \lambda_{w_r} \cdot P_{w_r}(B) \Big/ \sum_{r=1}^{\upsilon} \lambda_{w_r} \tag{4}$$

where λ_{w_r} is the mean arrival rate of reservation requests on route w_r .

A. Model Assumptions and Notation

The blocking analysis relies on the following assumptions:

- The PF nodes operate in accordance to the immediate forwarding option.
- υ possible routes $w_1, w_{2,}.., w_\upsilon$ exist on the network.
- The propagation delay between PF nodes is zero. Thus, the forwarding TFs of a packet are *i* and *i*+1 at two subsequent nodes on the route to the destination.
- Resources are reserved, when setting up an SVP, on a per-flow basis.
- PF-aware synchronous end systems take part in resource reservation and are end points of SVPs, as shown in Fig. 3 (top).
- Since it is not realistic that in the near future all end systems will be PF-aware, a scenario in which end-system are PF-unaware must be considered. Traditional end-systems are not involved in resource reservation that is performed on their behalf by a so called *SVP interface* at the ingress of the PF network. SVP interfaces and PF nodes at the egress of the PF network act as SVP end points, as shown in Fig. 3 (bottom). SVP interfaces are also responsible for forwarding packets sent by PF-unaware sources during reserved TFs. In this scenario the blocking probability can be computed based on the section of the path on which PF is being deployed.
- An end system is connected to each network edge link. An edge link can be either ingress or egress link. An ingress link connects a PF-aware end system with the ingress PF node whilst an egress link connects and egress PF node with a PF-aware end system.
- The SVPs are unidirectional, as depicted in Fig. 3.
- The end systems perform reservation requests according to a Poisson process with mean arrival rate λ . The reservation request arrival model rate does not take into account any reaction to rejection — such as immediate request repetition.
- The reservation holding times are statistically independent and exponentially distributed. All reservations have the same mean reservation holding time $1/\mu$.
- Each reservation request requires the allocation of one TF in a time cycle on each link of a ℓ -hop route. Hence, the maximum number of active flows on a link is TC. This assumption does not affect the generality of the analysis as it is valid also in case TF fractions are being reserved. This can be done by defining a TF to be available when there is an unreserved TF portion sufficient to transmit the required amount of bits. Moreover, the analysis can be easily extended to the case in which the sources request for the reservation of multiple TFs in a time cycle.
- An end system performs reservation requests towards different destinations in accordance to the traffic matrix *S*. The traffic matrix has dimensions s *x* s, where s is the number of network edge links. The matrix is square because network links are assumed to be full-duplex. Each matrix element $S(m,n)$ represents the fraction of requests for reserving one time frame for an SVP from ingress link *m* to the egress link *n*. Each row of the matrix sum to one.
- A routing protocol routes reservation requests towards their destination. The solution of the routing problem is stored in a three-dimensional routing matrix *R* having

dimension s x l x s, where l is twice[1] the number of links in the network, including the edge ones. Elements $R(n,j,m)$ are set to 1, if link j belongs to the route from m to n, and 0 otherwise. Each possible matrix $R(:,:,m)$ having dimensions s x l provides the routing matrix from the ingress link m toward all possible egress links.

- Each link has one channel. The analysis can be easily extended to the case of multiple channels per link, such as in Wavelength Division Multiplexing (WDM) networks.
- The switching architecture of the PF nodes is crossbar-based hence strictly non-blocking.
- The PF network processes reservation requests one by one and the time for processing requests is assumed negligible.
- A reservation request is blocked if an available schedule cannot be found from end to end.
- If multiple schedules are available on a route, one is chosen randomly, such that the TF occupancy is distributed across the time cycle.
- The TF status, *i.e.*, reserved vs. available, is statistically independent.

B. Distribution of the number of active reservations on a link in isolation

The mean arrival rate of reservation requests on all the available routes of a general topology network is here calculated under the above assumptions. Then it is used to derive the mean arrival rate of reservation requests on a single link and to estimate the steady state distribution of the number of active reservations on a single link.

Let

$$W = \lambda \cdot S \tag{5}$$

be a matrix having dimensions s x s whose element $W(m,n)$ represent the mean arrival rate λ_{w_r} of reservation requests on route w_r from ingress link m to egress link n, where $r = [(m-1)\cdot s + n]$.

Let also define

$$O(m,j) = \sum_{k=1}^{n} S_{m,k} \cdot R_{k,j,m} \tag{6}$$

a matrix having dimensions s x l whose element $O(m,j)$ provides the fraction of requests entering the network through edge link m that traverse link j. Note that the total fraction of requests routed through link j from every ingress to every egress link can be calculated as the sum of elements in column j of matrix O.

Since the sum of two Poisson processes with rate parameters λ_1 and λ_2, respectively, is still a Poisson process with rate parameter $\lambda_1 + \lambda_2$ and given the assumptions the analysis relies on, the arrival process of reservation requests at link j in isolation is still Poisson with mean arrival rate $Q(j)$, where

$$Q = \lambda \cdot \left[O^T \cdot I \right]^T \tag{7}$$

I being a column vector of length l whose elements are all set to one. The traffic intensity vector L, whose element $L(j)$ represent the traffic intensity on link j can be obtained by multiplying each element of Q by the mean reservation holding $1/\mu$.

Since

1. the number of reserved TFs on link j is related to the arrival process of reservation requests,
2. the number of reserved TFs decreases when a reservation terminates, *i.e.*, an SVP is torn down and the corresponding resources released,
3. the maximum number of TFs simultaneously reserved on link j is TC and the number of reserved TFs coincides with the number of active reservations, *i.e.*, SVPs, on link j,

the number of active reservations on link j in isolation can be represented by a stochastic process described by the Markov[2] chain *M/M/TC/TC*. In accordance to the notation adopted in [10] the fourth value describing the Markov chain is equal to TC because it coincides with the maximum number of active reservations on the link, *i.e.*, clients in service. The state-transition-rate diagram of the resulting TC-*server loss system* is depicted in Fig. 4.

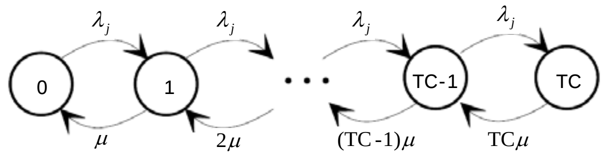

Fig. 4 State-transition-rate diagram for the TC-*server loss system* M/M/TC/TC.

The steady state distribution of the number of active reservations on link j is solved in [10] by

$$\pi_i^{(j)} = \frac{\pi_0^{(j)}}{i!}\left(\frac{\lambda_j}{\mu}\right)^i = \frac{\pi_0}{i!}\rho_j{}^i \quad \forall i = 1,2....,TC \tag{8}$$

where

$$\pi_0^{(j)} = \frac{1}{\sum_{z=0}^{TC}\frac{1}{z!}\left(\frac{\lambda_j}{\mu}\right)^z} = \frac{1}{\sum_{z=0}^{TC}\frac{\rho_j{}^z}{z!}} \tag{9}$$

As discussed in [10] the ergodicity of the Markov chain is assured for $0 < \mu < \infty$ and $0 < \lambda < \infty$.

Let N be a random variable representing the number of active reservations on link j. Thus, the average number of active reservations $E[N]$ can be derived through the *Little* theorem as

$$E[N] = \frac{\lambda_j}{\mu}\left(1 - \pi_{TC}^{(j)}\right) \tag{10}$$

In the reminder of this paper $E[N]$ is referred to as the

[1] A full-duplex link connecting node A and node B is modeled by two links, one from A to B and another one from B to A, into the routing matrix.

[2] The blocking probability analysis is based on the steady state distribution of the number of active reservations on the network links. In this work a Markovian approach is adopted to devise such distribution, but any other approach could be used as long as it is compatible with the assumptions set forth.

average utilization of a link j.

It is worth noting that with the Markovian approach the steady state distribution of active reservations on a link is derived under the assumption that a reservation request is blocked, hence rejected, only if TC TFs are simultaneously reserved on that link. However, as explained in Section II.B, due to unschedulability, a reservation request crossing an ℓ-hop route potentially experiences blocking even if resources are still available. Consequently, the blocking model presented here estimates correctly the steady state distribution of the number of active reservations on each link only as long as the end-to-end blocking probability is negligible. As traffic increases to the levels where unschedulability results in non-null blocking probability, the model overestimates the number of active reservations resulting from the corresponding traffic intensity.

However the ultimate objective of this analysis is to provide a model to estimate the minimum value of average utilization of the links — related to the traffic intensity at the network ingress — at which the end-to-end blocking probability is not negligible. This value represents the maximum network utilization achievable when providing deterministic QoS with acceptable (negligible) probability of rejecting reservation requests, as further discussed in Section IV. As a result, since in the region of interest unschedulability does not take place, or it is at least negligible, the presented model based on the Markovian approach can be used to provide meaningful and reliable results.

It is also worthwhile highlighting that the following analysis relies on the statistical independence of TF status, *i.e.*, TF reservations, on both the same link and consecutive links, as in [4]. Although TF reservations on a given link are not statistically independent in general, the assumption is reasonable under the conditions considered in this work. In fact, as discussed above, this paper focuses on network load levels above which the blocking probability becomes non negligible. Because in such operating conditions reservation requests are (mostly) accepted, independence of TF reservations on each link stems from the independence of reservation requests – featured by our end system model. The statistical independence of TF reservations on consecutive links will be discussed in the next section.

C. Blocking Probability on an ℓ-Hop Route

The blocking probability experienced by reservation requests issued by a source on a ℓ-hop route depends on the existence of an end-to-end available schedule. When immediate forwarding is deployed, given a forwarding TF at the source, the schedule, *i.e.*, the forwarding TFs on all the links traversed, is univocally determined.

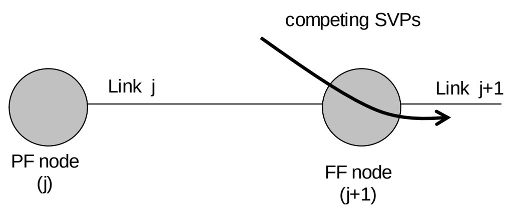

Fig. 5 A 2-hop sub-route.

Considering a 2-hop sub-route such as the one depicted in Fig. 5, packets forwarded in TF i by PF node j are forwarded in TF i+1 by the subsequent PF node j+1. All possible schedules can be expressed as $\forall i \in [0,TC), \left\{\bar{x}_i^{(j)}, \bar{x}_{(i+1)\bmod TC}^{(j+1)}\right\}$.

When trying to reserve a forwarding TF i at node j for a new SVP, TF i+1 may be unavailable at node j+1 because it had been previously reserved for a competing SVP set up on a different path (see Fig. 5). It should be noted that an existing SVP sharing the same path until node j will not represent a possible source of contention on TF i+1 at node j+1: immediate forwarding operation ensures that since such SVP is not using TF i at node j, it is not using TF i+1 at node j+1. Hence, it can be concluded that:

1. a reservation request may compete for scheduling at the output link of the subsequent PF node only with SVPs entering from different input ports of that PF-node.
2. reservation requests traversing the same subset of network links may compete for scheduling only at the first they share on their path.

Therefore, the probability for TF i+1 to be available on an output link of a PF node should be calculated considering the subset of competing SVPs given the average utilization of that link. However, having assumed statistical independence of TF reservations on different links, our model considers all active SVPs as having potentially reserved TF i+1. TF reservations on different links cannot be assumed to be statistically independent in the general case given that the forwarding TFs of an SVP on the subsequent links of its path are interdependent. However, the statistical independence assumption is reasonable – if reservation requests are statistically independent, which is the case in many practical cases – as long as the number of routes across the network is large, as it is in real-world global networks. As a result, the presented model is overestimating the probability for TF i+1 to be unavailable on an output link of PF node j+1, *i.e.*, it provides an upper-bound on the blocking probability.

Under the above assumptions, the probability for a schedule to be available given that TF i is the forwarding TF at node j, *i.e.*, the probability of finding the TF i+1 available on the output link of PF node j+1, can be expressed as follows

$$\begin{aligned} P(A^{(j+1)} \mid \bar{x}_i^{(j)}, \Pi_{j+1}) &= \pi_0^{(j+1)} + P(\bar{x}_{i+1}^{(j+1)} \mid 1)\cdot \pi_1^{(j+1)} + \\ & P(\bar{x}_{i+1}^{(j+1)} \mid 2)\cdot \pi_2^{(j+1)} + \ldots + P(\bar{x}_{i+1}^{(j+1)} \mid TC)\cdot \pi_{TC}^{(j+1)} \end{aligned} \tag{11}$$

where $P(\bar{x}_{(i+1)\bmod TC}^{(j+1)} \mid k)$ is the probability for TF i+1 to be available given that k TFs are reserved. Thus

$$P(A^{(j+1)} \mid \bar{x}_i^{(j)}, \Pi_{j+1}) = \sum_{z=0}^{TC} P(\bar{x}_{(i+1)\bmod TC}^{(j+1)} \mid z)\cdot \pi_z^{(j+1)} \tag{12}$$

where

$$P(\bar{x}_{(i+1)\bmod TC}^{(j+1)} \mid z) = \frac{\binom{TC-1}{z}}{\binom{TC}{z}} = 1 - \frac{z}{TC} \tag{13}$$

as it can be derived through combinatorial enumeration.

Therefore

$$P(A^{(j+1)} \mid \bar{x}_i^{(j)}, \Pi_{j+1}) = \sum_{z=0}^{TC} \left(1 - \frac{z}{TC}\right) \pi_z^{(j+1)} \qquad (14)$$

Note that (14) applies for each TF chosen as a forwarding TF at node *j*, *i.e.*, the probability of schedule availability is invariant on the forwarding TF at an upstream node under the assumptions made. Whenever a reservation request is received and being processed, it is blocked if for each potential forwarding TF at the upstream node there is no available schedule. Moreover, the number of schedules over which an available schedule is searched depends on the number of available TFs, *i.e.*, potential forwarding TFs, at the upstream node. Therefore, the probability for a reservation request to be blocked can be expressed as follows:

$$P(B) = \sum_{k=0}^{TC} \pi_k^{(j)} \cdot \left[1 - P(A^{(j+1)} \mid \bar{x}_i^{(j)}, \Pi_{j+1})\right]^{TC-k} \qquad (15)$$

Where the steady state distribution of the number of active reservations Π_j on the output link of node *j* is obtained from (8) and (9) given the mean rate of reservation requests arriving on node *j* and their mean holding time. The upper-bound blocking probability (resulting from (14) being the worst-case probability for a schedule to be available) on a 2-hop sub-route can be derived substituting (14) in (15) as follow:

$$P_U(B) = \sum_{k=0}^{TC} \pi_k^{(j)} \cdot \left(1 - \sum_{z=0}^{TC} \left(1 - \frac{z}{TC}\right) \pi_z^{(j+1)}\right)^{TC-k} \qquad (16)$$

Considering now a whole ℓ-hop route w_r, all possible end-to-end schedules can be expressed as $\forall i \in [0, TC), \left\{\bar{x}_i^{(0)}, \bar{x}_{(i+1)\bmod TC}^{(1)}, \ldots, \bar{x}_{(i+\ell)\bmod TC}^{(\ell)}\right\}$. The worst-case probability for a schedule to be available on w_r for a given forwarding TF at the PF-aware source is the compound probability of schedule availability on all the two-hop sub-routes from source to destination. Under the independence assumption made, such compound probability is obtained by multiplying the probability of schedule availability on all the 2-hop sub-routes on route w_r as follow:

$$\begin{aligned} & P(A^{(\omega_r)} \mid \bar{x}_i^{(0)}, \Pi_{w_r}) = \\ & \prod_{j=1}^{\ell} P(A^{(j)} \mid \bar{x}_{(i+j-1)\bmod TC}^{(j-1)}, \Pi_j) = \\ & \prod_{j=1}^{\ell} \sum_{z=0}^{TC} \left(1 - \frac{z}{TC}\right) \pi_z^{(j)} \end{aligned} \qquad (17)$$

Note that (17), that is derived recursively starting from (14), applies to any TF chosen as a forwarding TF at the source, *i.e.*, under the assumptions made the end-to-end probability of schedule availability is invariant on the forwarding TF at the source.

Whenever a reservation request is generated and processed by the network, it is blocked if for each potential forwarding TF at the source a schedule is not available on the whole route w_r. The number of schedules over which an available schedule is searched from end to end depends on the number of potential forwarding TFs, at the source. The blocking probability for a reservation request on a ℓ-hop route w_r can thus be expressed as

$$P_{w_r}(B) = \sum_{k=0}^{TC} \pi_k^{(0)} \cdot \left[1 - P(A^{(\omega_r)} \mid \bar{x}_i^{(0)}, \Pi_{w_r})\right]^{TC-k} \qquad (18)$$

By substituting (17), that provides the worst-case probability for a schedule to be available end-to-end, in (18) we obtain an upper-bound to the blocking probability as expressed in (1), which proves the **Theorem** at the beginning of this Section. Eq. (8) and (9) can be used in (1) to estimate the steady state distribution of the number of active reservations $\Pi_j = \left\{\pi_0^{(j)}, \pi_1^{(j)}, \ldots, \pi_{TC}^{(j)}\right\}$ on each link of route.

IV. Numerical Results

This section shows how the blocking problem affects the capability of utilizing network resources to carry traffic with deterministic QoS through numerical results devised from both the analytical models and simulations. The latter enable validating the correctness of the model in providing an upper-bound on the blocking probability in various network scenarios following the assumptions the analysis relies on.

A call level simulator is used to run experiments with different numbers of TFs per time cycle (which enables evaluating how the CTR structure affects blocking) and different network load levels. First, a given network load is emulated by generating random reservations for each network link; then a reservation request is issued on each route and it is determined whether it would be accepted, *i.e.*, if at least one TF is available on all the links of the route to satisfy the immediate forwarding principle. Results are presented by plotting the blocking probability – computed as the ratio of the number of reservation requests not accepted over the total number considered – versus the average utilization of the links. Lines plot the analytical model, while each marker represents a simulation result.

A. Single switch

The first set of experiments is run on a single strictly non-blocking switch with four input/output ports, as depicted in Fig. 6. The sources of reservation requests are connected to the input ports of the switch (full dots) and destinations to the output ports (empty dots). Reservation requests are assumed to be uniformly distributed among all the input/output ports.

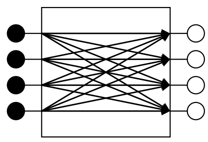

Fig. 6 4x4 strictly non-blocking PF capable switch.

Fig. 7 shows the blocking probability experienced by reservation requests issued by sources at the input of the

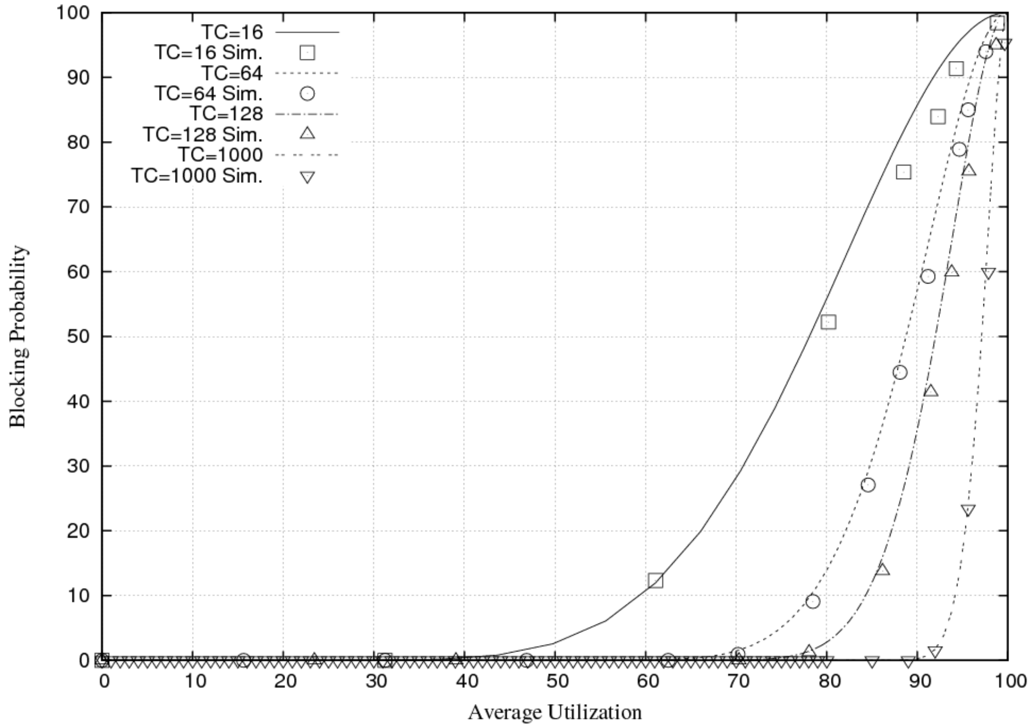

Fig. 7 Blocking probability experienced at the output of the switch by reservation requests for different numbers of TFs per time cycle (TC).

switch versus the average utilization of the output links for different number of TFs in the time cycle, from 16 to 1000. As expected, the blocking probability experienced by reservation requests generated by sources increases with increasing average link utilization and decreases with a larger number of TFs per time cycle. With a large number of TFs in a time cycle, e.g., 1000 TFs, the blocking probability is negligible as long as the average link utilization is under 90%. The blocking probability given by (4) is confirmed to be an upper-bound as the markers representing values measured during the simulations are always below the curve resulting from the analytical model.

B. *Network of switches*

The second set of experiments considers the network of Switches depicted in Fig. 8 to study blocking on a multi-hop network over which more complex scheduling is required.

In all experiments the time cycle consists of 1000 TFs and without loss of generality links are assumed unidirectional. Reservation request sources are connected to the network ingress ports (full dots) and destinations to the network egress ports (empty dots). Reservation requests generated by each source are assumed to be uniformly distributed toward the eight reachable output ports of the switches. In this traffic scenario, the links of switch G represent bottlenecks. Assuming that ρ is the traffic intensity on each switch G output link, the traffic intensity on the output links of switch

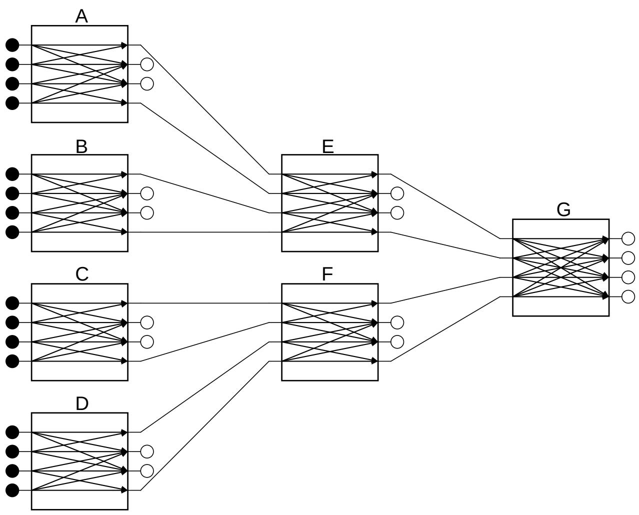

Fig. 8 Network of 4x4 strictly non blocking PF capable switches.

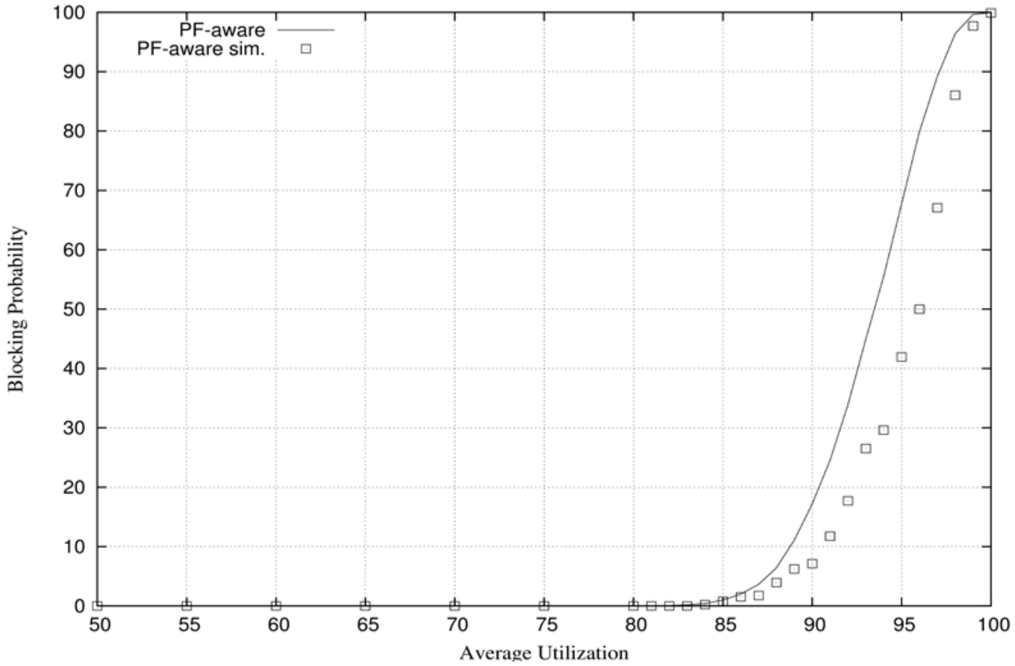

Fig. 9 Network-wide blocking probability experienced by reservation requests.

E and F that connect them to switch to G, is also ρ. The traffic intensity on the output links of switch A, B, C and D that connect them to switch E and F is ¾ ρ whereas the traffic intensity at the network ingress is ½ ρ.

Fig. 9 shows the network-wide blocking probability experienced by reservation requests versus the average utilization of the bottleneck links that are considered as representative of the overall network utilization. As with the previous graphs, the analytical upper bound is plotted as a line, while markers represent simulation results.

Since there is no interest in operating a network at a utilization level at which resource reservations have a high probability of being rejected, the most significant part of the plots is where the blocking probability becomes non-null and starts growing. This point represents the maximum fraction of link capacity that can be reserved to delay-sensitive traffic without having reservation requests experience significant blocking due to unschedulability. Note that the part of the plot for which the blocking probability is negligible coincides with the area of validity of the presented analytical model, as discussed in Section III.B.

As expected the blocking probability increases with the number of switches traversed. The lowest network utilization at which non-null blocking probability is experienced by reservation requests decreases from about 91% on a one-hop route (see Fig. 7) to 84% on a 3-hop route. It is worth highlighting that such percentage is much higher than the one commonly reserved to delay sensitive traffic, such as voice, on asynchronous networks implementing differentiated services, which is around 25% to 33% [11]. The marks corresponding to the blocking probability measured during the simulations are always below the lines devised through the blocking probability model provided by (4) in the above Theorem, thus confirming their validity as an upper-bound.

V. Related Works

Call blocking is a phenomenon limiting the performance of any network where resources are reserved to traffic. Hence, blocking probability has been widely studied for a long time and in a large number of technologies ranging from analogue phone networks, to circuit switched or time division multiplexing (TDM) networks [12], to optical switched networks [13]. However the blocking problem in PF networks is somewhat different as it is not necessarily related to the lack of resources, but to the lack of a feasible schedule of TFs through the network.

Blocking probability has been extensively studied by both simulation in various network scenarios [9] and analysis. Given that this paper focuses on analysis, only analytical blocking probability studies available in the literature are hereafter considered.

When Ofek *et al.* proposed the pipeline forwarding operation [4], they also introduced the blocking problem and provided an initial blocking probability analysis. Further work on blocking analysis has more recently been done for a single-hop scenario then extended to multi-hop scenarios [8]. These approaches and their results are presented in the following and compared to the ones of this paper.

The analysis in [4] considers the current amount of transmission resources reserved in a TF to be L, and the maximum amount of resources that can be reserved in a TF to be L_{th} and models the probability for a TF to be unavailable as $p = \mathrm{P}\left(L \ge L_{th}\right)$. Assuming that TF unavailability events are mutually independent, the probability for a reservation request to be blocked on a route of h nodes performing immediate forwarding is given by the probability that for each of the TC TFs in the time cycle, the TF is not available on at least one of the h links on the route[3], *i.e.*, $P(B) = \left[1-(1-p)^h\right]^{TC}$.

In order to devise numerical results, Ofek *et al.* assume the TF unavailability probability density function to be a truncated Gaussian distribution and calculate the blocking probability for several standard deviation and average values of the TF load, *i.e.*, mean of the Gaussian distribution.

The work in [4] also devises the blocking probability for non-immediate forwarding in case (*i*) one TF and (*ii*) up to *TC* TFs of extra forwarding delay are introduced at each hop. The calculation is impractical in the former case (*i*) above when the number of TFs per time cycle *TC* is not very small; consequently, [4] provides some results for small *TC* values, *e.g.*, *TC*=5. Not being able to use the devised analytical equations, numerical results with larger numbers of TFs per time cycle, *e.g.*, *TC*=48, are obtained in [4] through simulation.

One major shortcoming of [4] when compared to this work is relying on the assumption that the distribution of the probability of unavailable TFs is Gaussian, which is not substantiated in any way. Also, the probability p, and consequently the blocking probability $P(B)$, are not related to usable system parameters, such as the traffic matrix describing the distribution of reservation requests through the network, the reservation request generation process by end systems at the network ingress, the holding time of the reservations, and routing on the network. As a result the devised blocking model cannot be easily deployed for network dimensioning purposes.

Further work on blocking analysis that does not rely on a specific reservation probability density function has more recently been presented in [8] for a single-hop scenario first and then extended to multi-hop scenarios. A combinatorial enumeration approach is adopted to derive the blocking probability of reservation requests through a single switch as $P(B)=C_{block}/C_{total}$, where C_{total} is the total number of possible schedules through a switch and C_{block} is the total number of schedules through a switch that would lead to blocked reservation requests. In this analysis the link reservation level is modelled in terms of number of unavailable TFs per time cycle, rather than the mean of the distribution of the probability p for a TF to be unavailable, as in [4]. However, when extended across several hops, as only sketched in [8], the combinatorial analysis of the blocking probability becomes impractical. In fact, it requires the computation of a set of conditional probabilities that grows with the number of TFs in the time cycle, which makes the solution not applicable to realistic time cycle dimensions and no numerical results are provided in [8].

To summarize, the analysis presented here follows a different approach for the blocking probability derivation compared to [4] and [8] and leads to a formulation that is actually usable. The blocking probability experienced by reservation requests is derived as function of the steady state distribution of the number of active reservations on each link it traverses calculated taking into account network routing, traffic matrix, and reservation request arrival process and duration, i.e., network and traffic quantities commonly taken into account when engineering a network and its traffic. This work also offers an improved multi-hop blocking analysis because it (*i*) does not assume the same mean link utilization on different links, (*ii*) provides computational feasible blocking models also for large numbers of TFs per time cycle, *e.g.*, 1000 TFs, and (*iii*) addresses the cases of both PF-unaware and PF-aware end-systems.

[3] The formulation ignores the propagation delay, which does not affect the result.

VI. Conclusions

This work presents in details the blocking problem stemming from the *pipeline forwarding* operating principles and improves upon previous blocking analysis for the case of *immediate forwarding* operation. The upper-bound blocking probability experienced by reservation requests issued by pipeline forwarding-aware and pipeline forwarding-unaware sources has been derived on general topology networks as function of the average utilization of the network links given the model of the reservation request arrival process at the network ingress, the model of reservation holding process, the traffic matrix describing the mean load of reservation requests from every ingress to every egress node and the network routing, that are in most practical cases known parameters useful for network dimensioning purpose.

Moreover the validity of the blocking models is verified by simulations in different network and traffic scenarios. Finally, the different analytical approaches of related works on the blocking problem are discussed showing the improvements offered by this analysis.

References

[1] J. C. R. Bennett, K. Benson, A. Charny, W. F. Courtney, J-Y Le Boudec, “Delay Jitter Bounds and Packet Scale Rate Guarantee for Expedited Forwarding,” *IEEE/ACM Transactions on Networking*, vol.10, no.4, pp. 529-540, August 2002.

[2] Cisco Systems, Inc., “The Zettabyte Era – Trends and Analysis,” May 29, 2013, [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni /VNI_Hyperconnectivity_WP.html

[3] M. Yuksel, K. K. Ramakrishnan, S. Kalyanaraman, J. D. Houle, R. Sadhvani, “Value of Supporting Class-of-Service in IP Backbones,” *Fifteenth IEEE International Workshop on Quality of Service*, Evanston, IL, pp. 109-112, June 2007.

[4] C-S. Li, Y. Ofek, A. Segall, K. Sohraby, “Pseudo-Isochronous Cell Switching in ATM Networks,” *Computer Networks and ISDN systems*, vol. 30, no. 24, pp. 2359-2372, December 1998.

[5] M. Baldi, G. Marchetto, Y. Ofek, "A Scalable Solution for Engineering Streaming Traffic in the Future

Internet," *Computer Networks*, vol. 51, no. 14, pp. 4092-4111, October 2007.

[6] M. Baldi, Y. Ofek, “End-to-end Delay of Videoconferencing over Packet Switched Networks,” *IEEE/ACM Transactions on Networking*, vol.8, no.4, pp. 479-492, August 2000.

[7] M. Baldi, J.C. De Martin, E. Masala, A. Vesco, "Quality-Oriented video transmission with pipeline forwarding," *IEEE Transactions on Broadcasting*, vol.54, no. 3, pp. 542-556, September 2008.

[8] V.T. Nguyen, R. Lo Cigno, Y. Ofek, M. Telek, “Time Blocking Analysis in Time-Driven Switching Networks,” *in the proceedings of INFOCOM 2008*, pp. 1804-1812, 2008.

[9] M. Baldi, Y. Ofek, “Fractional lambda switching - principles of operation and performance issues,” *SIMULATION: Transactions of the Society for Modeling and Simulation International*, vol. 80, no. 7, July 2004.

[10] L. Kleinrock, “Queueing Systems,” *A Wiley-Interscience Publication*, vol 1: Theory, pp. 105-106.

[11] A. Clauberg, “Deploying Tight-SLA services on an IP Backbone,” *Next Generation Internet Symposium*, March 2002; [Online]. Available: http://www.zdv.uni-mainz.de/ak-sys/2002/aclauber-ngi-ZKI.ppt.

[12] M. Sivakumar and S. Subramaniam, “A Performance Evaluation of Time Switching in TDM Wavelength Routing Networks,” *in the proceedings of the First International Conference on Broadband Networks (BROADNETS’04)*, 2004.

[13] K. Kuppuswamy and D. C. Lee, “Some analysis of call blocking probabilities in hierarchical WDM networks offering multiclass service," *International Journal of Communication Systems*, Volume 24, Issue 6, pp. 709-731, June 2011.

Mario Baldi received his M.S. Degree Summa Cum Laude in Electrical Engineering and his Ph.D. in Computer and System Engineering both from Politecnico di Torino. He was Assistant Professor at Politecnico di Torino from 1997 to 2002, when he became Associate Professor. He was Principal Architect at Embrane, Inc., Santa Clara, CA, from October 2010 to May 2012 and Principal Member of Technical Staff at Narus, Inc., Sunnyvale, CA in June 2012. He was Vice Dean of the PoliTong Sino-Italian Campus at Tongji University, Shanghai, China, and coordinator of the BSc programme in Information Technology Engineering from October 2006 to July 2010. He was Vice President for Protocol Architecture at Synchrodyne Networks, Inc., New York, from November 1999 till October 2001. He was head of the Computer Networks Group (NetGroup) at the Department of Computer Engineering of Politecnico di Torino from 2001 till 2007.

Mario Baldi has been Guest Professor at Tongji University, Shanghai, China, Honorary Visiting Professor at La Trobe University, Melbourne, Victoria, Australia, Adjunct Professor at University of Illinois at Chicago, Visiting Professor at Institut de Technologie du Cambodge, Phnom Penh, Cambodia, and visiting researcher at the IBM T. J. Watson Research Center, Yorktown Heights, NY, at Columbia University, New York, NY, and at the International Computer Science Institute (ICSI), Berkeley, CA.

Andrea Vesco received the M.Sc. degree in Telecommunication Engineering and the Ph.D. in Computer and System Engineering from the Politecnico di Torino in 2003 and 2009 respectively. After one year of post-doc with the Control and Computer Engineering Department as a member of the Computer Networks Group at the Politecnico di Torino, he joined the Networking Lab at the Istituto Superiore Mario Boella (ISMB) in 2010.

His research interests were on Quality-of-Service (QoS) over packet switched networks and wireless access networks. Moreover he carried out research activities on QoS over Network-on-Chip (NoC).

After ten years of experience and several publications in the networking field, he had the opportunity to focus his researches on the Smart City subject in 2013. His research effort is currently spent on the role of ICT in development and management of Smart Cities, with particular focus on sustainable mobility and energy efficiency.