Computer Networks 83 (2015) 235–248

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Short-term fairness in slotted WDM rings

Andrea Bianco [a,*], Davide Cuda [b], Jorge M. Finochietto [c]

[a] *Dip. di Elettronica e Telecomunicazioni, Politecnico di Torino, Italy*
[b] *Cisco Systems, Lausanne, Switzerland*
[c] *Universidad Nacional de Córdoba – CONICET, Argentina*

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form 1 February 2015
Accepted 21 March 2015
Available online 15 April 2015

Keywords:
WDM
Fairness
Metro architecture
Optical packet switching

ABSTRACT

Single-hop WDM ring networks are promising architectures for future broadband access and metro networks. However, ring networks exhibit significant fairness issues, which must be handled by a fairness enforcing protocol. Fairness is usually ensured over a time window of several network propagation delays. Thus, data flows might experience large access delays which might be not compatible to support time-sensitive applications. We solve this issue proposing the Multi-Fasnet protocol, which is able to enforce fairness in a relative short time scale, in the order of few propagation delays, without trading off the aggregated throughput network performance. We discuss Multi-Fasnet limitations and propose several novel strategies that achieve high and fair network throughput as well as low, bounded and fair access delays.

1. Introduction

Most of today's metropolitan area networks are based on circuit-switched Synchronous Optical NETwork/Synchronous Digital Hierarchy (SONET/SDH) architectures. Initially designed to transport constant-rate voice traffic, SDH networks are unable to cope with today's constantly increasing bursty data traffic demands. Ethernet-based metropolitan area networks are arguably more flexible, scalable and cost-effective than legacy SDH architectures [1–3]. Despite these advantages, Ethernet-based MANs still represent an *opaque* solution. Indeed, optical technology is used exclusively to support point-to-point connections between nodes and each node must perform optical-to-electrical (O/E/O) conversion, and electronically process the entire traffic for routing/switching. Since O/E/O conversion represents the largest cost when operating optical fiber networks [4], reducing or eliminating O/E/O conversion of in-transit traffic is a key design objective for future MANs.

Optical packet switching (OPS) architectures have been proposed as candidates to meet the requirements of more dynamic and demanding future networks. However, truly header-based packet switching in the optical domain is not mature yet, and, likely, too complex to design. Instead, single-hop Wavelength Division Multiplexing (WDM) architectures are a viable approach [5–15] to provide all-to-all connectivity among nodes distributed over a fiber ring. These networks usually operate in a time slotted fashion: slots propagate on the ring and each node can add/drop data to/from in-transit slots by means of one optical transmitter and receiver pair. Tunability is required at least on one end to enable single-hop, all-optical, connections between nodes.

Since nodes can add traffic only exploiting empty slots, data collision is avoided but upstream nodes can reduce (or even block) the transmission opportunities of downstream nodes. Thus, a fairness control scheme must be adopted to provide equal access opportunities to all nodes. Classical fairness definitions mainly refer to throughput fairness,

---

* Corresponding author.
E-mail addresses: andrea.bianco@polito.it (A. Bianco), davide.cuda@gmail.com (D. Cuda), jfinochietto@efn.uncor.edu (J.M. Finochietto).

http://dx.doi.org/10.1016/j.comnet.2015.03.014

which is usually obtained over fairly large time scales to avoid compromising the overall network performance. Our major novel contribution is to propose new solutions that provide short-term throughput fairness, i.e, over a relatively short time scale, in the order of few network propagation delays, without deteriorating the overall network performance. As such, nodes can access network resources with bounded delays even under (transient) network overloaded conditions, without penalizing time-sensitive, interactive, "mice" (i.e., low-bandwidth) flows as it would happen with traditional, long-term, fairness control schemes.

The paper is organized as follows. Section 2 introduces the related work and motivates the novelties of our approach. We describe in Section 3 the considered network architecture and in Section 4 the fairness issues arising in this network. In Section 5 we recall the original Fasnet protocol and we discuss its adaptation to the considered network. In Section 7 we present simulation results under different traffic scenarios. Finally, we derive conclusions in Section 8. The main notations used in the paper are summarized in Table 2.

The novel contributions include a complete paper restructuring and rewriting, a formal fairness definition, more extensive results under several traffic scenarios, discussions on receiver allocation strategies and on how to cope with nodes with different traffic requirements and a sensitivity analysis of protocol parameters.

The novel contributions include a complete paper restructuring and rewriting, a formal fairness definition, more extensive results under several traffic scenarios, discussions on receiver allocation strategies and on how to cope with nodes with different traffic requirements and a sensitivity analysis of protocol parameters.

2. Related work and motivations

Fairness is a well studied subject both in electronic [16,17] and optical networks [18,19].

It is well known that ring and folded bus topologies introduce unfairness in node access opportunities. In this paper, we focus on fully distributed solutions capable of solving the unfairness of ring and bus topologies, disregarding any centralized access control scheme. Furthermore, to avoid imposing any constraint on input traffic, reservation based approaches are also not considered.

Fairness protocols were proposed in the past for both ring based [20,21] and bus based electronic networks [22]. The design of fairness protocols in a WDM multichannel network imposes new challenges with respect to traditional electronic single channel networks. Indeed, since nodes are typically equipped with one transmitter only, coordination among access to different WDM channels is required to ensure good overall network performance. Furthermore, the specific network architecture considered in this paper presents some additional constraints. Indeed, the techniques proposed in WDM rings [21] as extensions of fairness protocols devised for electronic networks (e.g., MetaRing, ATMring) cannot be directly used in the context of folded bus topologies, where a node is in the same position when accessing network resources on all the available channels. Thus, traditional solutions must be adapted to the studied scenario.

The MetaRing protocol, originally proposed to address fairness in ring networks where a single channel is available, provides good throughput performance and has already been adapted to the WDM scenario [21]. Later, the Multi-MetaRing protocol was introduced as a further extension in the context of folded bus based networks [23].

The vast majority of the previously proposed distributed fairness schemes [20,21,23] ensure fairness on a time scale in the order of several propagation delays. Consider that the network end-to-end propagation delay is typically fairly larger than the average packet transmission time. For example, in a metropolitan area network, a 50 km span leads to a 250 μs propagation delay, while a 1000-bit packet at 1 Gbit/s lasts 1 μs. Thus, hundreds packets can be transmitted in a time corresponding to a network propagation delay, and throughput fairness is obtained only after a large number of packet transmissions. Among the long-term fairness schemes, we refer in this paper for performance comparison purposes to Multi-MetaRing, which was shown to be suited to the WONDER network and able to provide high throughput efficiency in [23].

Multi-MetaRing is representative of access schemes offering almost ideal throughput efficiency and fairness, but operating over long time scales, hence being less suited to react to fast traffic dynamics. Instead, the approach proposed in this paper acts on a *shorter* time scale of few round-trip propagation delays without compromising the overall network throughput. Providing fairness on a shorter time scale permits to keep bounded access delays for low-bandwidth time-critical applications, even in highly loaded conditions, as shown in the performance results section.

3. Network architecture

We consider a WDM optical packet network named WONDER [24]. Fig. 1 presents the architecture of the WONDER network. WONDER comprises N nodes connected through two counter-rotating WDM fiber rings. Each ring conveys W wavelengths, with $N \geqslant W$. Differently from traditional bidirectional dual ring networks, one ring is used for transmission only, while the second ring is used for reception only. Transmission wavelengths are switched to the reception ring, thanks to a loop-back fiber, in a folding point, as shown in Fig. 1. Note that, this folding point can be created on a dynamically selected node, exploiting an Optical Switch (OSW) [25].

The network operates in a synchronous, time slotted fashion, and slots propagate on the bus. To avoid data collision, network nodes access the WDM slotted ring by inserting fixed size packets in empty slots, whose duration T_{slot} is determined by technological constraints, Transmitted fixed-size packets travel on the transmission ring up to the folding point, where they are switched to the reception ring. The proposed architecture logically behaves as a folded bus network. Although this prevents the exploitation of space reuse of ring networks, it permits

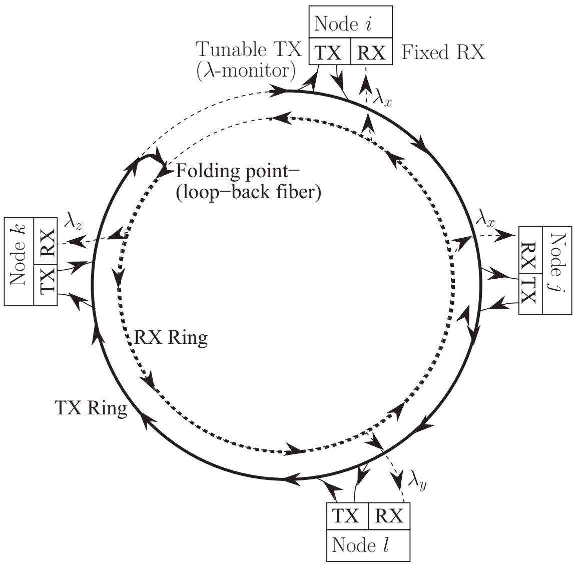


Fig. 1. Network architecture.

to reduce transmission impairments due to noise recirculation typical of all-optical ring networks [26], simplifying the node architecture. This advantage largely overcomes the slot reuse capability of ring networks.

WONDER can provide Terabit bandwidths. However, to reduce electronic complexity and ensure scalability, nodes should not process the full network bandwidth. Although solutions proposed in [5,7,27] show that several transceivers per node may increase the network throughput, more transceivers also imply increased electronic complexity and power consumption in network nodes, making this solution too costly for increasing data rates. Hence, WONDER exploits WDM to provide several transmission channels over the same fiber, but the complexity of WONDER nodes is limited to equipping each node with a single, slot by slot tunable, transmitter (TT), and a single, fixed, receiver (FR), both running at the speed of a single channel. This node architecture is usually referred to as TT–FR. The tunable transmitter is needed to achieve full connectivity among nodes. Fast tunability, if needed, can be accomplished by means of an array of fixed lasers with a fast selector, as described in [26]. This classical solution permits to assume that transceiver tunability can be obtained on a slot by slot basis with negligible performance losses due to the short tuning time with respect to the 1μ slot duration.

As shown in Fig. 1, each node receiver is tuned to a specific, fixed, wavelength. Since $N \geqslant W$, several nodes may receive data on the same WDM channel. In a time slot, at most one packet can be transmitted by a node in one of the W available slots (one slot for each channel). Nodes exploit WDM to partition the traffic directed to disjoint subsets of destination nodes, each subset comprising nodes receiving on the same wavelength. To avoid performance losses due to the Head-of-the-Line (HoL) blocking problem, each node is exploiting a Virtual Output Queue (VOQ) memory architecture [29]. Nodes tune their transmitters to the receiver's destination wavelength, establishing a single-hop all-optical connection lasting one time slot. Therefore, channel resource sharing is achieved according to a dynamic Time Division Multiple Access (TDMA) scheme.

Access decisions are based on a channel inspection capability (similar to the carrier sense functionality in Ethernet), called *λ-monitor*: At the beginning of each time slot, each node senses which wavelengths have not already been used by upstream nodes, and, if any, it transmits on one of the available wavelengths in the current time slot. Thus, to avoid data collision, a *multi-channel empty-slot* protocol is used, giving priority to in-transit traffic. As such, nodes at the end of the bus may experience difficulties in accessing network resources if a fairness protocol is not adopted [22].

4. The fairness issue

A well-known problem of ring and bus topologies is the different access priority of network nodes due to their position along the ring/bus. Referring to Fig. 1, nodes at the head of the bus can "flood" a given wavelength, reducing (or even blocking) the transmission opportunities of downstream nodes competing for access to that channel, thus leading to significant fairness problems. Whereas in a ring network (with $N = W$) each node has different access opportunities on each channel, in the bus case, the node at the head of the bus is in the most favorable position on all channels. Note that providing fairness in access opportunities makes also less critical the selection of the node sitting at the head of the bus.

Several fairness definitions are available, including, among others, max–min fairness [16] and proportional fairness [17]. We refer to the max–min fairness paradigm in this paper.

The WONDER architecture is characterized by a WDM multi-channel environment with the single transceiver per node constraint. We introduce the concept of *channel* fairness (i.e., fair access for all nodes competing on the same channel) and *network* fairness (i.e., fair access for a node on all channels where it must transmit information). These two fairness definitions may lead to conflicting requirements (an example will be given in Section 7.2.4) and we assess them through the channel fairness index $\mathrm{FI}_{\mathrm{ch}}$ and the network fairness index $\mathrm{FI}_{\mathrm{net}}$, respectively.

Let $\mathcal{N}$ be the set of the network nodes and $\mathcal{W}$ the set of wavelengths. $N = |\mathcal{N}|$ and $W = |\mathcal{W}|$. Let $T = [t_{ij}]$ be the envisioned traffic matrix with t_{ij} being the amount of data generated by node i addressed to node j, where $i,\ j,\ \in\ N$. Let u_{ij} be the amount of traffic transmitted by node i to node j, normalized to t_{ij} when $t_{ij} \neq 0$. We set $u_{ij} = 1$ if $t_{ij} = 0$. More precisely, $0 \leqslant u_{ij} \leqslant 1$ measures the node efficiency in transmitting traffic toward a specific destination. Let x_{jw} be a binary variable equal to 1 if node j receives on channel w, with $w \in W$. As such, the local efficiency of each node i on channel w can be evaluated as $U_{iw} = \frac{\sum_{j=1}^{N} u_{ij} x_{jw}}{\sum_{j=1}^{N} x_{jw}}$, while node i global efficiency can be computed as $U_i = \frac{\sum_{w=1}^{W} U_{iw}}{W}$.

We evaluate the channel fairness index as follows

$$\mathrm{FI_{ch}} = \min_{w}\left\{\mathrm{FI_{ch}^{(w)}} : \mathrm{FI_{ch}^{(w)}} = \frac{\min_i U_{iw}}{\max_i U_{iw}},\ w \in W\right\} \tag{1}$$

where $\mathrm{FI_{ch}^{(w)}}$ measures the local fairness among nodes on a specific channel w and $0 \leqslant \mathrm{FI_{ch}} \leqslant 1$. The closer $\mathrm{FI_{ch}}$ to 1, the more fair are the nodes' access opportunities to network channels.

On the other hand, we assess how fairly nodes can access the overall network resources measuring the total traffic transmitted by each node. Thus, we compute the network fairness index $\mathrm{FI_{net}}$ as:

$$\mathrm{FI_{net}} = \frac{\min_i U_i}{\max_i U_i} \tag{2}$$

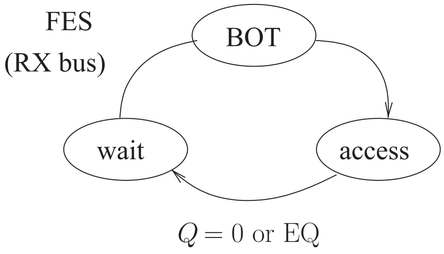
with $0 \leqslant \mathrm{FI_{net}} \leqslant 1$. The closer $\mathrm{FI_{net}}$ to 1, the more fair are the node opportunities to access network resources.

Both $\mathrm{FI_{ch}}$ and $\mathrm{FI_{net}}$ depend on the receiver-to-channel allocation and on the access protocol. The optimal allocation policy depends on many factors, including the traffic matrix T and the adopted optimality criterion (a load balancing among channels is typically sought for). As an example, a uniform receiver allocation is optimal for uniform traffic. If the traffic matrix is not uniform, several "good" allocations may exist. We informally define as *optimal* an allocation that, first, balances the load on all channels, and, second, balances the transmitter load on each channel, subject to the constraints imposed by the traffic matrix. In this paper, unless otherwise stated, we assume an optimal allocation of receivers, as discussed in [8], and focus on the fairness of the access protocol.

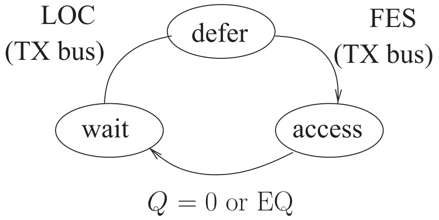
5. The Fasnet protocol in single-channel networks

Fasnet [28] is an empty-slot implicit-token passing protocol originally designed to guarantee fairness on a single-channel electronic network characterized by a slotted behavior and a folded bus topology. To provide fairness, Fasnet operates cyclically: Each cycle is associated with a chained transmission of data, called *train*, where each node can transmit at most a quota Q of packets. Note that this guarantees that all nodes have equal opportunities to access a channel. To this end, the node at the head of the bus, the *master* node, periodically transmits a first packet dubbed *locomotive* (LOC). Downstream nodes willing to access the transmission (TX) bus, monitor channel availability to detect the end of train (EOT) condition and append their transmissions to the train. The master node, instead of monitoring the TX bus, senses the reception (RX) one to detect the EOT condition triggering the transmission of a new LOC on the TX bus. LOCs can be either transmitted in-band or out-of-band. However, in the case of out-of-band transmission, an additional channel and transceiver per node are needed. For these reasons, we assume in-band LOC transmission. As a result, the EOT condition can be detected by sensing a First Empty Slot (FES) after at least one or more busy slots.

The master node manages the Fasnet dynamic by regulating the LOCs transmission and thus, the cycle (train) dynamics. Fig. 2(a) shows the abstract behavior of the master node. The master node is normally in the *wait* state. When it detects the FES condition on the RX bus, it moves to the Begin Of Train (BOT) state, transmitting a LOC on the TX bus. Then, it renews its quota Q and moves immediately to the *access* state in which it transmits its data traffic decreasing the quota by one for each packet transmission until either it exhausts its quota ($Q = 0$) or it empties its queue (EQ). Finally, it moves back to the *wait* state.



(a) Master node state diagram.



(b) Downstream node state diagram.

Fig. 2. Fasnet node state diagram.

The behavior of downstream nodes is shown in Fig. 2(b). Downstream nodes remain in the *wait* state until they sense the transit of a LOC on the TX bus. After a LOC transit, a node moves to the *defer* state, where it waits for the FES after a LOC on the TX bus. This corresponds to the EOT and makes the channel available for transmission, triggering transition from the defer state to the *access* state. When this transition happens, the quota Q is renewed and the node can transmit data without being interrupted by upstream traffic until it *releases* the channel because either it empties its queue (EQ) or $Q = 0$. Then, it moves back to the *wait* state.

5.1. Fasnet maximum throughput

Fasnet does not achieve a maximum throughput $TH_{max} = 100\%$, even for uniform traffic, due to the idle time T_{idle} between two consecutive trains transmission, where T_{idle} (normalized to T_{slot}) depends on the network physical size. Although all slots in each train are used back-to-back in overload, the master node detects the end of the current train only when the FES is sensed on the RX channel. Indeed, only one train at the time travels along the network and a new LOC is sent only when no packets are traveling in the network. Hence, each train is followed by a number of empty slots equal to the propagation delay from the master transmitter to the master receiver. Thus, the maximum achievable throughput under uniform traffic in overload is bounded by the ratio between the maximum train length, equal to $N \times Q$ slots, and the cycle duration, equal to $N \times Q + T_{idle}$ slots:

$$TH_{max} = \frac{N \times Q}{N \times Q + T_{idle}} \tag{3}$$

Given that T_{idle} is protocol independent, being only a function of the physical size of the network measured in slots, the higher the value of Q, the higher the maximum achievable throughput.

5.2. Fasnet delays

Fasnet guarantees an upper bound to the access delay, i.e., the time since a packet reaches the head of the transmission queue until it starts being transmitted. The worst-case access delay under uniform traffic conditions can be easily estimated by considering a packet arriving at the head of a node queue immediately after the channel has been released, i.e., the train has just left the node. In overload conditions, the packet will gain access in the next train, after a number of time slots equal to

$$D_{overload} = (N - 1) \times Q + T_{idle} \tag{4}$$

This is also the time scale over which the protocol is able to ensure throughput fairness among nodes.

Instead, if we assume that the network is lightly loaded, each node empties its queue before exhausting its quota. Therefore, the average access delay for a packet that just missed a train, named average worst-case delay at low loads D_{WC}, measured in time slots, is:

$$D_{WC} \approx \sum_{i=1}^{N} Q_i^* + T_{idle} \tag{5}$$

where Q_i^* is the effective average quota used by node *i*. As long as the network is stable, i.e. not in overload, each node throughput is equal to its load:

$$TH = \frac{Q_i^*}{\sum_{i=1}^{N} Q^* + T_{idle}} = \rho_i \tag{6}$$

Thus, when the network is lightly loaded, the real quota Q_i^* used by each node during each train cycle can be derived solving a system of N equations, one for each node. Assuming uniform traffic, $Q_i^* = Q^* \forall i$, then $TH = \frac{N \times Q^*}{N \times Q^* + T_{idle}} = \rho$. Thus,

$$Q^* = \frac{\rho}{1 - \rho} \times \frac{T_{idle}}{N} \tag{7}$$

At low loads, being $\sum_{i=1}^{N-1} Q_i^* \ll T_{idle}, D_{WC} \approx T_{idle}$ and Q^* does not depend on Q but only on the load. The worst-case access delay at low load is not a function of Q because nodes mostly end transmissions due to EQ. At high network loads, a "mouse" node has to wait for approximately $D_{overload}$ time slots before being able to transmit, being strongly penalized in term of access delays.

5.3. The impact of the quota

Fasnet permits to control fairness by tuning the value of Q granted to each node. When the same quota Q is given to all N network nodes, Fasnet enforces max–min throughput fairness, i.e., when the total traffic exceeds the network capacity, the same bandwidth ("fair share") is allocated to the nodes that offer a load in excess of the fair share. If all nodes are in deep overload, the available bandwidth is evenly divided among nodes. Note that fairness is ensured if the node queue size is at least as large as the node quota. Otherwise, node performance may be limited by the fact that not more than a number of packets equal to the queue size can be transmitted every cycle. We will assume in our simulation results that the node queue size is larger than the quota assigned to each node.

If different quotas are given to different nodes ("skewed quotas"), nodes with a larger quota will receive a larger bandwidth share in overload. However, assigning skewed quotas implies a centralized network control, whereas we seek for a fully distributed, traffic un-aware, solution. Thus, we only consider skewed quotas for comparison purposes in the remainder of the paper.

In summary, the value of Q should be chosen as a compromise between throughput efficiency (Section 5.1), small access delays and time scale of fairness control (Section 5.2). We will propose several strategies which are able to ensure high throughput even if using low Q, so as to achieve fairness in a relatively short time scale, thus ensuring bounded delays for time-sensitive mice flows.

6. Multi-Fasnet protocol: Fasnet in a WDM network

In a multi-channel network, the Fasnet behavior must be replicated over the different wavelengths, to obtain fairness on all channels. In other words, to control each of the W wavelengths, we use W trains, one for each channel. We name *Multi-Fasnet* this multi-channel access protocol.

Let Q_{iw} be the quota of node *i* on channel *w*. $Q_{iw} = Q$ if not differently stated. Furthermore, let RTT be the ring propagation delay from the master node to the last node on the transmission bus, measured in time slots. Thus, in a folded-bus network:

$$T_{idle} = 2 \times RTT + 2 \tag{8}$$

because (i) both the transmission and the reception ring must be traversed, (ii) one time slot is used by the locomotive, and (iii) folding the transmission to the reception bus implies a small propagation delay rounded to one full time slot.

The maximum throughput for the multi-channel WONDER network can be estimated computing the maximum throughput (Eq. (3)) on each channel *w* and averaging over W. Thus,

$$TH_{max} = \frac{1}{W} \times \sum_{w=1}^{W} \left(\frac{\sum_{i=1}^{N} b_{iw} \times Q_{iw}}{\sum_{i=1}^{N} b_{iw} \times Q_{iw} + T_{idle}} \right) \tag{9}$$

where b_{iw} is a binary variable equal to 1 if node *i* transmits on channel *w* and depends on (i) the node to channel allocation and (ii) the traffic matrix.

We present now the new strategies that permit to cope with the WDM scenario and, at the same time, to ensure high throughput and low delays. All these strategies assume that nodes may transmit a packet only if they sense a free slot.

6.1. Quota accumulation strategy

The first strategy is named quota accumulation strategy. In a multi-channel network more channels can become available for transmission at the same time when a node detects a FES condition in more than one channel. However, being equipped with a single tunable transmitter, nodes can access at most one wavelength per time slot. Thus, we label as train conflict the situation in which more than one channel is available for transmission at a given node in the same time slot, i.e., a node detects a FES condition in more than one channel. We solve train conflicts selecting for transmission the channel associated with the longest queue.

When train conflicts occur, nodes may release a channel although they still have both quota available and packets to transmit. Thus, train conflict might lead to unfairness in the channel access, because nodes are un-aware of the global network traffic pattern and cannot make any optimal choice on which channel to use. By allowing nodes to accumulate quota, fairness can still be achieved, but in more than one cycle. Nodes are allowed to transmit on the next cycle at most Q_{iw} packets plus the remaining (or residual) quota of the previous cycle(s). To avoid excessive quota accumulation the maximum quota that a node can accumulate on a channel is bounded by the node current queue length on the corresponding channel.

6.2. Free-access strategy

The Multi-Fasnet protocol forces nodes to access channels in a cyclic order, as the original Fasnet protocol. Thus, a node has to wait the next cycle to be able to transmit its packets, and access delays are mainly bounded by T_{idle}, at low loads. We describe here the free-access strategy which allows to significantly reduce access delays when the network is lightly loaded.

To reduce the average access delay, a proper fairness access scheme should be inactive (i.e., work similarly to an *empty-slot* mechanism) when the network is lightly loaded, while it should regulate access to network resources when the network is overloaded, ensuring equal access opportunities to all nodes. The free-access strategy allows a node to transmits packets even though the EOT has already passed, provided that this node has some quota left and free slots are detected on the bus to avoid data collision. This approach is similar to the Simple access scheme proposal [30]. As in the Fasnet protocol, on a given channel, a node renews its quota and start transmitting when the FES after the LOC is detected. However, differently from Fasnet, if a node i releases channel w because of EQ or a train conflict, it may later transmit new packets in the same cycle, provided that $Q_{iw} > 0$.

Note that besides reducing the access delay at low loads, the free access strategy also reduces the probability of accumulating quota, because nodes can access the channels even when they have lost their turn on a channel within the current cycle. At high load, the possibility of accessing the channel even if the EOT has passed is rarely exploited because few free slots are available.

6.3. Train retransmission strategies

Although the free-access strategy reduces delays, the maximum throughput of Multi-Fasnet with free access strategy (MFf) is still limited, in overload, by the channel idle time T_{idle}.

To improve throughput performance without increasing Q we need to reduce the idle time T_{idle}. To this end, we assume that on each channel w, the master node has the possibility to retransmit new trains without waiting for the EOT on the RX bus. Hence, we assume that the master node schedules a new train on channel w every $C_w(n)$ time slots, where $C_w(n)$ is a decrement counter initialized to a specific value when the n-th locomotive is transmitted on channel w. Thus, the $(n+1)$-th train on channel w is scheduled if either $C_w(n) = 0$, or if the master node senses an EOT (as in the original Multi-Fasnet protocol).

Finding the correct value of $C_w(n)$ can be challenging. Indeed, on the one hand LOCs retransmitted too frequently result in shorter trains that can overlap while traveling along the bus, preventing last nodes from transmitting their data. On the other hand, LOCs retransmitted with a low frequency result in longer trains which might be left partially empty leading to throughput losses. We present two different strategies to update $C_w(n)$.

6.3.1. Fixed-Length Train (FLT) strategy

When the FLT retransmission strategy is used, $C_w(n)$ is always initialized to $C_w(n) = N \times Q \forall n,\ w$. The FLT strategy ensures that nodes access channels cyclically, according to a statistical Time Division Multiple Access (TDMA) scheme which degenerates to a fixed TDMA scheme in overloaded uniform traffic. Thus, no slots are left empty, unless when train conflict occurs. However, problems might arise if the value of the node quota does not match the node traffic needs (e.g., uniform quota values under non-uniform traffic patterns). In this case, trains may be left partially empty and throughput losses are experienced.

6.3.2. Dynamic-Length Train (DLT) strategy

Since short-term traffic needs (i.e., in the order of ms-s) are normally highly-variable and unknown, it may be impossible to properly tune the train length to match traffic statistics in the FLT strategy. To overcome this difficulty, the DLT strategy adapts the train length to match traffic needs, by deriving the next train length from the current train utilization. This results in a much more flexible solution with respect to FLT, where a constant $C_w(n)$ value is used, preventing any adaptation to traffic variations. However, being a dynamic strategy, DLT is expected to perform not better that FLT with a train length perfectly matched to the traffic pattern, assumed static and known. Let $0 \leqslant \mu_w(n) \leqslant 1$ be the ratio among the number of used slots and the total length of the n-th train on channel w.

If $\mu_w(n) < 1$, the current train is left partially empty because nodes do not have enough traffic to fill the whole train, hence the length of the next train can be decreased. Instead, if the current train is full, the length of the next train should likely be increased to account for possibly larger node transmission needs. Then, we dynamically update to $C_w(n)$ depending on $C_w(n-1)$ and on $\mu_w(n)$. As such, at the n-th locomotive retransmission, $C_w(n)$ is initialized to:

$$C_w(n) = \begin{cases} C_w(n-1) \times (1+I) & \mu_w(n-1) = 1 \\ C_w(n-1) \times (1-D) & \mu_w(n-1) < 1 \end{cases} \qquad (10)$$

where I and D are two protocol parameters that denote the increase and decrease steps of the train length. The impact of I and D is studied in Section 7.2.5, where it is shown that small performance differences can be highlighted by varying these parameters. A minimum train length equal to N is enforced to cope with under-loaded situations.

7. Performance evaluation

7.1. Simulation scenario

We present performance results obtained by simulation. Simulation runs exploit a proprietary simulation environment developed in the C language. Simulator correctness has been verified against known results in the literature and in situations were the results can be easily predicted. Statistical significance of the results is assessed by running experiments with an accuracy of 3% under a confidence interval of 95%.

We consider a network with $W = 4$ wavelengths and a total of $N = 16$ nodes, uniformly allocated to the different channels, if not differently specified. Nodes are evenly spaced on the ring, with a node-to-node distance equal to 1.6 km, for a total ring length of about 25.6 km. Slots last 1 μs, corresponding to a fixed packet size of about 1250 bytes at 10 Gbit/s, and to a node-to-node propagation delay equal to 8 time slots. Thus, the ring *RTT* (distance from the first to the last node) is equal to 120 slots. Each node maintains W separate FIFO queues, one for each channel. Each FIFO can store a maximum of about $L = 120{,}000$ fixed-size packets, corresponding to 150 MB to ensure that the available queue is always larger than the node quota in all the studied scenarios, including the case in which quota is cumulated, when the network is not overloaded. Three different values of quota ($Q = 10, Q = 100, Q = 1000$) are usually considered in simulation results. For the DLT strategy, if not differently stated, I and D are set respectively to 0.3 and 0.1. These values are motivated by the sensitivity analysis performed in Section 7.2.5.

Let the element t_{ij} of the traffic matrix T, $1 \leqslant i, j \leqslant N$, represents the average packet generation rate from node i to node j, normalized to the channel rate. Let $\rho \in [0, 1]$ represents the total network load normalized to the available capacity. We consider the uniform traffic case and some unbalanced traffic scenarios: The *1-server*, the *1-mouse* and the *triangular traffic* scenarios.

7.1.1. Uniform traffic

For the uniform traffic pattern, the traffic matrix is:

$$t_{ij} = \rho \times \begin{cases} 0 & \text{iff } i = j \\ \frac{W}{N}\frac{1}{N-1} & \text{otherwise} \end{cases} \qquad (11)$$

7.1.2. 1-Server traffic

In the 1-server traffic pattern, the node at the head of the bus acts as a server, and all other nodes as clients: An access network with a video distribution server located at the central office can be a suitable example of this scenario. The traffic between clients represents background traffic due to other running applications, as, for instance, peer to peer. More precisely, the server transmits at a high rate, equal to the capacity of one wavelength, with equal probability to the other $N-1$ client nodes. The remaining network capacity is shared by client nodes. One wavelength is devoted to transmissions from all clients to the server. Therefore, the server receives and transmits traffic saturating one full wavelength, and each client receives from the server and transmits to the server the same amount of information. Being $\mathcal{C}$ ($\mathcal{S}$) the set of client (server) nodes, the traffic matrix can be expressed as:

$$t_{ij} = \rho \times \begin{cases} 0 & \text{iff } i = j \\ \frac{1}{N-1} & \text{iff } i \in \mathcal{S} \wedge j \in \mathcal{C} \\ \frac{1}{3}\frac{W-1}{N-1} & \text{iff } i \in \mathcal{C} \wedge j \in \mathcal{S} \\ \frac{2}{3}\frac{W-1}{(N-1)(N-2)} & \text{iff } i \in \mathcal{C} \wedge j \in \mathcal{C} \wedge i \neq j \end{cases} \qquad (12)$$

7.1.3. 1-Mouse traffic

In the 1-mouse traffic scenario, all nodes are equally loaded, except the last node on the bus, which generates a small amount of traffic. The traffic matrix element is:

$$t_{ij} = \rho \times \begin{cases} 0 & \text{iff } i = j \\ \frac{W-t_m}{(N-1)^2} & \text{iff } i \in \{1, \ldots, N-1\} \\ \frac{t_m}{N-1} & \text{iff } i = N \end{cases} \qquad (13)$$

where t_m is the traffic generated by the mouse node N. In the simulation results, we set $t_m = 0.1$. This traffic scenario highlights the access difficulties of a node willing to transmit low bandwidth delay sensitive data when the network load is high because of the traffic exchanged by other nodes.

7.1.4. Triangular traffic

For the triangular traffic case, to simplify the traffic description, we consider a network composed of $N = 8$ nodes and $W = 2$ channels. In this scenario the traffic load is balanced between the different channels in reception but not in transmission.

Each node injects a different amount of traffic destined to the first half and the second half of network nodes, depending on its position along the bus. For instance, the first (last) node sends 90% (10%) of its traffic to nodes $\{1, 2, 3, 4\}$, and the remaining 10% (90%) to nodes $\{5, 6, 7, 8\}$. The triangular traffic matrix is described by the following equations, where nodes are partitioned in two sub-sets:

$$t_{ij} = \rho \frac{W}{N} \begin{cases} 0 & \text{iff } i = j \\ \frac{0.9-(i-1)\times\delta}{N/W-u_{ij}} & \text{iff } \forall i,\ j \leqslant \frac{N}{W} \\ \frac{0.1+(i-1)\times\delta}{N/W-u_{ij}} & \text{iff } \forall i,\ \frac{N}{W}+1 \leqslant j \leqslant N \end{cases} \tag{14}$$

where $u_{ij} = 1$ if i and j belong to the same sub-set and $u_{ij} = 0$ otherwise, and $\delta = \frac{0.9-0.1}{N-1}$. Inside each node subset, destinations are chosen with the same probability. If nodes $\{1,2,3,4\}$ receive on the first channel, and nodes $\{5,6,7,8\}$ receive on the second channel, traffic load is balanced among the two channels. With this node allocation, on channel $w-1$ $(w = 2)$, the traffic transmitted on each channel is triangularly skewed as we move from nodes at the head (end) of the bus to nodes located at end (head) of the bus.

7.2. Simulation results

We present aggregated network performance averaging them over the W channels. Network performance are reported in terms of network throughput, normalized to the network capacity, and average access delay, normalized to the time slot duration. Performance per node are evaluated as node throughput, normalized to the node transmitter capacity, average access delays and fairness. We evaluate fairness computing the FI_{ch} and the FI_{net} at $\rho = 1$ and $Q = 100$. Table 1 reports these values for the different scenarios.

The traffic scenarios reported in the next subsections are selected to highlight the following aspects of the various strategies. The uniform traffic scenario permits to introduce the different strategies, to compare them against Multi-MetaRing, and to discuss the main features of the FLT and DLT strategies, such as the sensitivity to the quota value, the importance of quota accumulation, and delay performance. The 1-server traffic permits to show the limitations of the FLT strategy when the train length is not matched to the traffic pattern. Instead, the 1-mouse traffic highlights the Multi-MetaRing limitations due to long-term (in the order of several RTTs) fairness control. When Multi-MetaRing is adopted, a mouse flow exhibits more than one order of magnitude larger access delays, due to traffic generated by overloading upstream elephant mice flows. Finally, the triangular traffic permits to analyze the effect on network performance of the receiver to channel allocation policies. We use the most flexible DLT strategy to derive performance results in this case.

Table 1
Fairness indices for the different traffic scenarios.

Traffic scenario	MFf		FLT		DLT	
	FI_{ch}	FI_{net}	FI_{ch}	FI_{net}	FI_{ch}	FI_{net}
Uniform (Accumulated Quota)	1	1	1	1	1	1
Uniform (Fixed Quota)	1	1	1	1	0.87	0.87
1-server	0.78	0.61	0.77	0.61	0.85	0.78
1-mouse	0.87	0.87	0.98	0.98	0.97	0.97
Triangular (Optimal RX)	0.95	0.85	0.96	0.85	0.96	0.85
Triangular (Optimal RX/TX)	1	1	1	1	1	1
Triangular (Optimal quota)	1	1	1	1	1	1

Table 2
Main notations used in the paper.

Symbol	Meaning	Default value
N	Number of nodes	16
W	Number of wavelengths	4
T	Traffic matrix	Uniform
t_{ij}	Data generated at node i directed to node j	1/N
T_{slot}	Slot duration	1 μs
T_{idle}	Time between two train transmissions in time slots	242
Q	Quota	100
Q_{iw}	Quota of node i on channel w	
RTT	Ring propagation delay in time slots	120
TH	Network throughput	[0, 1]
ρ	Network offered load	[0, 1]
$C_w(n)$	Train length on channel w at time n	
I	Train length increase step	0.3
D	Train length decrease step	0.1

7.2.1. Uniform traffic scenario

Fig. 3 shows the delay vs throughput for Multi-MetaRing, the Multi-Fasnet protocol (MF) and the Multi-Fasnet protocol adopting the free access strategy (MFf). We choose Multi-MetaRing because it provides throughput fairness, albeit on a relatively longer time scale than MFf. Multi-MetaRing also operates under the control of a transmission quota, which was chosen to be equal to 1936, the minimum value needed to guarantee 100% throughput in all traffic conditions (see [21] for details).

MFf, Multi-Fasnet with the free access strategy, is able to ensure an average access delay two orders of magnitude lower than the Multi-Fasnet protocol, and comparable to Multi-MetaRing. Indeed, when the network is lightly-loaded, if a new packet arrives within the current cycle, nodes transmit exploiting their residual quota without waiting for the following train.

As discussed in Section 6, the Multi-Fasnet maximum throughput is dramatically affected by the value of the

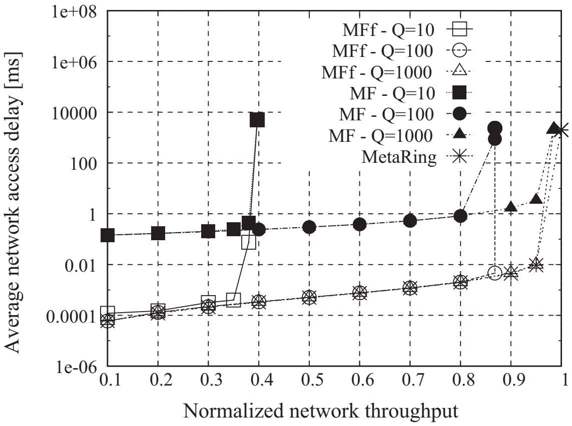


Fig. 3. Average access delay of Multi-Fasnet (MF), Multi-Fasnet with free-access (MFf) and Multi-MetaRing, under uniform traffic for $RTT = 120$. Fasnet quota $Q \in \{10, 100, 1000\}$.

quota. The maximum achievable throughput evaluated using Eq. (9) is, respectively, $TH_{max} = 0.4$ for $Q = 10$, $TH_{max} = 0.87$ for $Q = 100$, and $TH_{max} = 0.98$ for $Q = 1000$, correctly matched by simulation. According to Eq. (7), when the network is lightly loaded, the average transmission delay is independent of the quota value. Indeed, the train length depends on the input traffic and the network size only. At high loads, the average delay depends on the access delay plus the time needed to traverse the queue length L. Under uniform traffic, in overload, all nodes access channel w every $D_{WC_w} = N \times Q + T_{idle}$ and transmit Q packets. Thus, the average delay is equal to $L/Q \times D_{WC_w} + Q/2$. Therefore, the average delay in overloaded conditions is approximately equal to 4830 ms for $Q = 10$, 2260 ms for $Q = 100$ and 2450 ms for $Q = 1000$, for a slot duration of 1 µs. Similar trends with different absolute values can be observed in smaller size networks, where the same throughput is obtained with smaller quotas. When the free-access strategy is adopted, access delays at low load become independent of T_{idle} and are reduced by an order of magnitude. Thus, in the remainder of the paper, we only consider MFf, the Multi-Fasnet protocol adopting the free access strategy.

In Fig. 4, Multi-Fasnet, FTL and DLT (all adopting the free access strategy) are analyzed when adopting a value of quota $Q = 100$ and $Q = 10$. Under uniform traffic, both FLT and DLT strategies improve the network utilization, because they are able to cope with the RTT-induced idle time. Furthermore, being perfectly matched to the uniform traffic scenario, the FLT strategy achieves the highest maximum throughput and its performance are almost independent of the value of Q. The minor throughput losses are still induced by the train conflict effect. The DLT strategy presents some throughput losses with respect to the FLT strategy; the trains are left partially empty because their average length is larger than the optimal one (equal to $N \times Q$ under uniform traffic pattern). Note that the DLT is significantly more robust to the chosen quota value, since performance is less affected than those of MF and FLT by the chosen quota value, although increasing Q still provides some benefits. Table 1 shows how Multi-Fasnet and FLT ensure ideal fairness in the case of uniform traffic.

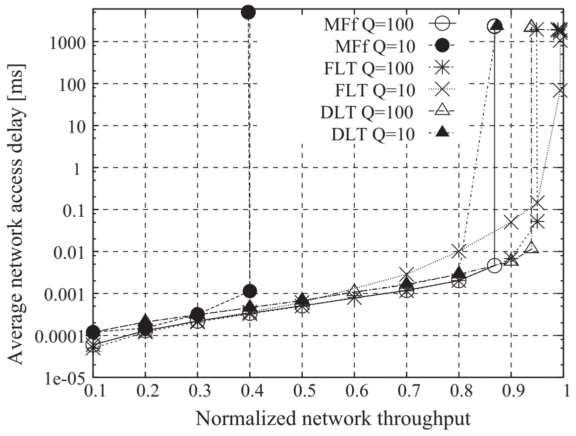


Fig. 4. Multi-Fasnet, FLT and DLT performance under uniform traffic for $Q = 10$ and $Q = 100$.

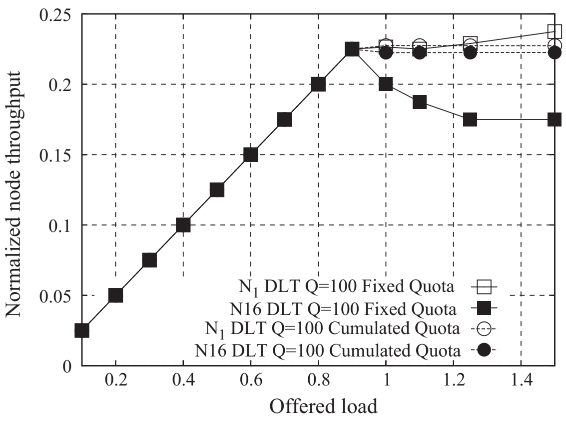


Fig. 5. Comparison between the first and the last node throughput under uniform traffic for the DLT protocol with and without quota accumulation.

Fig. 5 shows the first and last node throughput when the quota is renewed at each train reception (no quota accumulation), and when the quota is accumulated, for the DLT strategy. When quota is not accumulated, the last node suffers throughput degradation in overload and $FI_{ch} = FI_{net} = 0.87$. Losses are independent of the channel, i.e., the last node experiences the same losses on all the channels due to the overlap of too frequently retransmitted trains. Indeed, when trains are short, the last nodes might not be able to access the network for several transmission cycles, and have to wait until the master node decreases the train transmission frequency (increasing the train length), to recover for the lost cycles. If the quota is not accumulated, transmission opportunities are lost, and cannot be recovered later, creating throughput unfairness. When quota is accumulated, nodes can recover the lost transmission opportunities in the next cycles seizing their quota in the longer train.

Delay fairness is shown in Fig. 6 by reporting the access delay for the first and last node on the bus under uniform

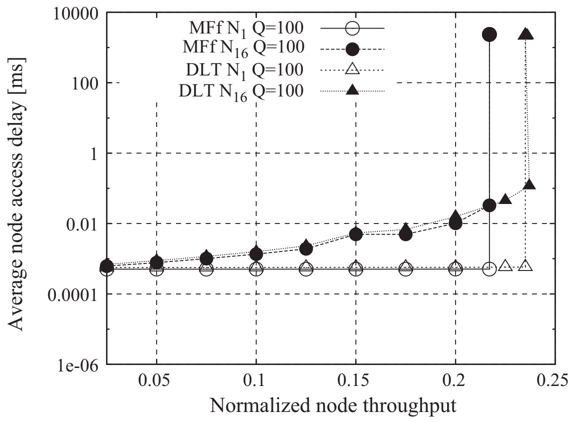


Fig. 6. Comparison between the first and the last node performance for $Q = 100$ for the Multi-Fasnet with free access and the DLT strategy under uniform traffic.

traffic. Remarkably, even if using the dynamic DLT strategy, which implies that a larger number of locomotives has to be transmitted in-band, access delays are not affected. Both Multi-Fasnet and DLT slightly penalize the last node because of the free access strategy. However, the absolute value of the average delay is fairly small and it does not increase significantly for the last node, even in high load.

In summary, FLT and DLT strategies provide better performance than Multi-Fasnet in uniform traffic. Delays are well controlled by all protocols. Quota accumulation is needed to avoid penalizing nodes at the end of the bus.

7.2.2. 1-Server traffic scenario

Let us now compare the Multi-Fasnet protocol and the FLT and DLT strategies under the 1-server traffic scenario to highlight DLT flexibility.

The server throughput is plotted against the offered load in Fig. 7. As expected, although the FLT strategy is very efficient under uniform traffic, it shows limited performance under unbalanced traffic conditions, while Multi-Fasnet suffers because of the idle time between two following cycles. The DLT strategy is instead able to match the train length to the input traffic pattern, achieving a larger throughput, clearly outperforming FLT and Multi-Fasnet. Specifically, the server is able to achieve larger values of throughput (about 0.9) before being starved.

When the network is in deep overload, the throughput of all nodes converges to the same value, according to the max–min fairness paradigm, unless a different quota is assigned to the server.

The fairness indices FI_{net} and FI_{ch} are quite far from 1 because the server experiences more losses than clients. As performance increases employing the DLT strategy, both FI_{net} and FI_{ch} get closer to 1.

7.2.3. 1-Mouse traffic scenario

The aim of the 1-mouse traffic scenario is to show that the DLT strategy is able to ensure low and bounded access delay for low-bandwidth delay-sensitive traffic thanks to the provided short-term fairness control scheme. Since all the nodes but the last one are highly loaded, the 1-mouse traffic is a worst-case scenario for this mouse traffic. Fig. 8 compares the first and the last node (the mouse) access delay for the DLT strategy, Multi-Fasnet and the Multi-MetaRing protocol [23].

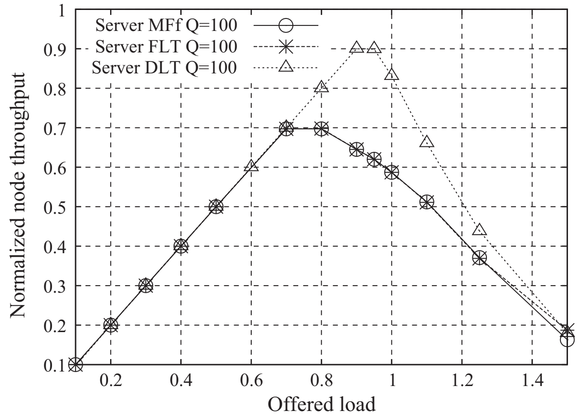


Fig. 7. Server throughput for the Multi-Fasnet, the FLT strategy and the DLT strategy in the 1-server scenario (Q = 100).

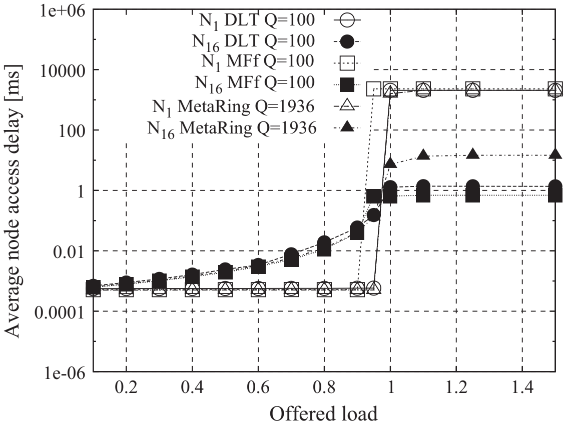


Fig. 8. First and last node average access delay for the Multi-MetaRing, the Multi-Fasnet, and the DLT strategy in the 1-mouse scenario.

When the network is overloaded, as shown in the right part of the plot, the DLT strategy, employing lower values of quota Q, ensures an access delay of roughly 1 ms, which is an order of magnitude less than the access delay of the Multi-MetaRing protocol, because the DLT strategy achieves fairness on a much shorter time interval. Multi-Fasnet performs similarly. This scenario proves the importance of providing throughput fairness on a relatively short time scale of few RTTs to guarantee access delays in the order of few ms to lightly loaded nodes when the network is highly loaded.

Real networks do not operate permanently in overload regime. However, the overload traffic case is significant because is representative of network behavior under transient overload situations due to a sudden traffic increase or to network reconfiguration. In these conditions, time sensitive traffic would suffer of QoS degradation if not adopting short-term fairness control.

7.2.4. Triangular traffic scenario

Being strongly unbalanced, the triangular traffic offers different solutions to the receivers' channel allocation problem, while keeping the load balanced among channels. We consider three different node to channel allocations, in a network with $N = 8,\ W = 2$.

In the first scenario, nodes are allocated on the different channels only to balance the receiver load. The first four nodes receive on channel 1, while the last four nodes receives on channel 2. We dub this solution *RX optimal* allocation.

In the second scenario, we balance both the traffic load among the different channels and the amount of traffic injected by each node on the different channels. Among all the available solutions which balance the traffic load among the different channels, we select the one able to equalize the node transmitter load on the different channels. We dub this solution *RX/TX optimal* allocation. As an example, odd nodes receive on channel 1 while even nodes receive on channel 2.

In the latter scenario, the same node to channel allocation of the first scenario is adopted, but we exploit skewed quotas: The quota values of each node are set on each channel proportionally to the amount of traffic transmitted by the node on the corresponding channel. Let Q_{iw} be the value of quota of node i on channel w. We assume that each node has a total (over the W channels) quota equal to $Q_{tot} = W \times Q$. Thus, $Q_{iw} = \frac{\rho_{iw}}{\rho_i} \times Q_{tot}$, where ρ_{iw} is the load generated by node i on channel w and $\rho_i = \sum_{w=1}^{W} \rho_{iw}$ is the aggregate load generated by node i. In this case, each time node i renews its quota on channel w, the quota value on that channel is set to Q_{iw}. This solution is named *optimal quota*.

Fig. 9 shows the overall normalized throughput for the three above described RX allocation when the DLT strategy is adopted, being better suited to unbalanced traffic scenarios. Node allocation strategies do not affect the overall network performance if the channel load is balanced. However, when the RX optimal allocation is adopted, nodes at the beginning and at the end of the bus (nodes 1 and 8) suffer some throughout losses, as shown in Fig. 10(a). Due to the triangular traffic matrix, nodes at bus edges experience a strong unbalance between the traffic and end up using their transmitter inefficiently. Indeed, on the channel where these nodes are highly loaded, transmissions always end because of quota exhaustion, while, on the channel where they are lightly loaded, the channel is released because of empty queue. As a consequence, as load increases, on one of the two channels bus-edge nodes do not have enough quota to accommodate all the traffic, while on the other channel they have a quota larger than needed. Both $\mathrm{FI}_{\mathrm{ch}}$ and $\mathrm{FI}_{\mathrm{net}}$ reflect this unfairness. Indeed, $\mathrm{FI}_{\mathrm{ch}} \simeq 0.95$, because the fairness within a channel is quite good independently of the adopted Fasnet version. However, the large difference in the amount of traffic transmitted by edge node with respect the middle nodes leads to a low $\mathrm{FI}_{\mathrm{net}} \simeq 0.85$.

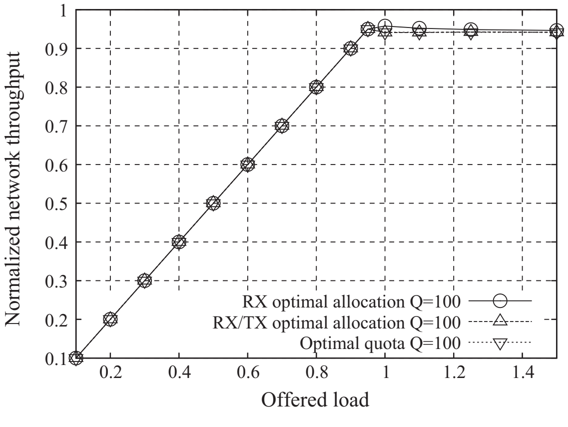
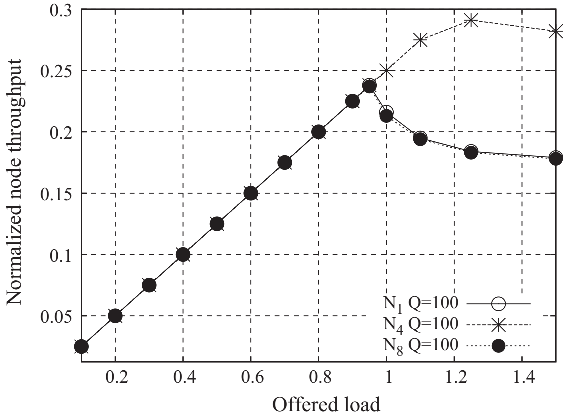
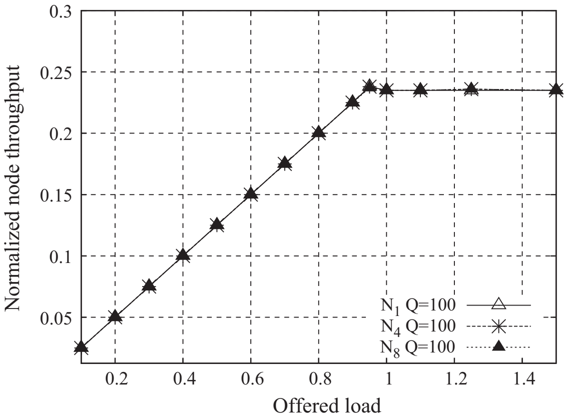


Fig. 9. Normalized throughput for the DLT strategy with $Q = 100$ under triangular traffic, with an optimal allocation of receivers only, of both receivers and transmitters, and of receivers only but selecting the optimal value of quota on each channel.



(a) RX optimal allocation



(b) RX/TX optimal allocation

Fig. 10. Nodes 1, 4 and 8 throughput for the DLT strategy, under triangular traffic with $Q = 100$, when adopting the RX optimal allocation (top), the RX/TX optimal allocation (bottom), and the RX optimal allocation with skewed quotas (bottom).

Fig. 10(b) shows that the unfairness disappears when nodes are allocated considering both the receivers' and the transmitters' load. The same result is achieved when the quota is adapted to each node traffic demand. Thus, properly tuning the quota value on the different channels permits to achieve fairness even if the TX/RX allocation is not optimal. Operating on the quota means operating at the protocol layer rather than at the physical layer, avoiding modifications to the node architecture. Indeed, to provide a re-allocation capability, nodes must be equipped with slowly (i.e., not on a timeslot basis) tunable receivers and a queue per destination node instead of a queue per channel is needed. Furthermore, a rather complex re-allocation protocol must be defined [8], increasing network complexity. Operating on the quota values could be an interesting alternative, but it requires a centralized network control as well as a traffic knowledge/estimation.

Indeed, properly tuning the quota may permit to cope with scenarios in which the node to channel allocation problem cannot be optimally solved. For simplicity, let consider a network composed of $N = 4$ nodes and $W = 2$ channels, with the following traffic matrix:

$$T = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.3 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.1 & 0.2 \\ 0.4 & 0.1 & 0.2 & 0.3 \end{pmatrix}$$

The possible allocations which balance the traffic on the different channels from the receivers' point of view are:

$$\left\{[N_1, N_2]_{\lambda_1}, [N_3, N_4]_{\lambda_2}\right\}\left\{[N_1, N_3]_{\lambda_1}, [N_2, N_4]_{\lambda_2}\right\} \times \left\{[N_1, N_4]_{\lambda_1}, [N_2, N_3]_{\lambda_2}\right\}$$

None of the above allocation is able to equalize the transmitter load on the different channels. For example, with the first allocation node N_1 (N_2) generates 0.3 (0.5) on λ_1, whereas node N_3 (N_4) generates 0.7 (0.5) on λ_2. In general, the traffic matrix can prevent to evenly balance the load on the different channels from the receivers' perspective. On the other hand, the unfairness could be reduced by selecting a proper value of quota on each channel, as it is has been shown in the case of triangular traffic.

7.2.5. Sensitivity analysis

We conclude our analysis focussing on the DLT strategy to show how the choice of D and I affects protocol performance. The upper graph of Fig. 11 shows the throughput versus delay performance when $D = 0.1$ and for I ranging from 0.1 to 0.5, under uniform traffic. The lower graph shows the throughput versus delay performance for $I = 0.3$ and D ranging from 0.1 to 0.5, always under the uniform traffic scenario. First, the performance of the DLT strategy is almost independent of the values of I and D. When the network is lightly loaded, trains are usually short and close to their minimum value (remember that a minimum train length equal to N has been imposed to leave at least one available slot to each node). In overload conditions, small values of I ensure slightly larger throughput, because they reduce the probability of transmitting long trains, which could be left partially empty. On the other hand, too short trains have a negative effect on delays: The last nodes need to accumulate more quota to recover for the train conflicts. Thus, fairness will be achieved on a larger time scale. A similar but less evident behavior is exhibited for large values of D: decreasing the train length when an empty slot is detected on the reception bus implies that the probability of leaving additional empty slots on the following trains decreases, but also that the last nodes suffer more from the train conflict effect.

This is better shown in Fig. 12, where the difference between the access delay of the first and the last node is plotted for a load equal to 0.9. I (D) ranges from 0.1 to 0.5 with $D = 0.1$ ($I = 0.3$). A large value for I and a small value for D reduce the delay unfairness, as they favor longer trains, which reduce train conflicts. Therefore, delays increase, especially for the last node, although the absolute delay values are very similar. In summary, I and D values should be chosen to slightly trade delays and throughput. However, the protocol is shown to be robust to variations of these parameters, making less critical to set up protocol parameters for network designers and managers.

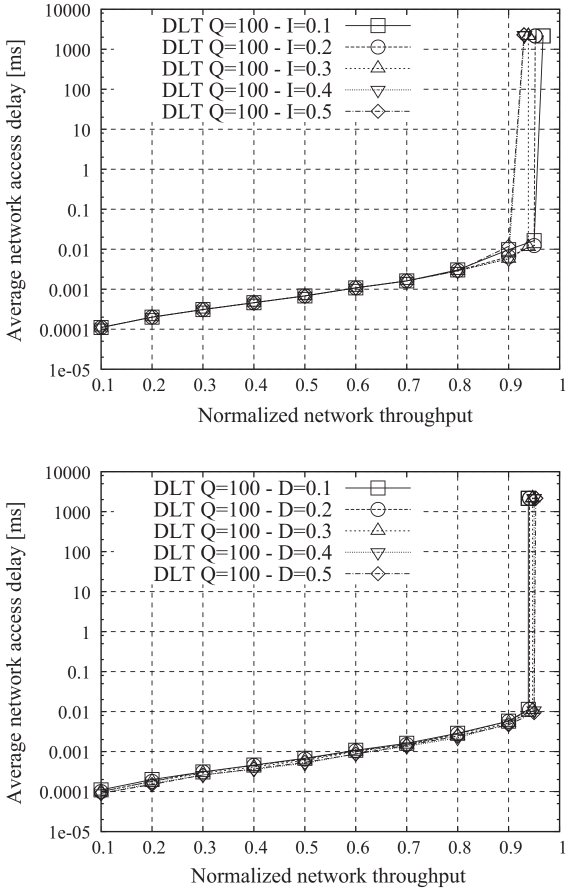


Fig. 11. DLT performance under uniform traffic for $D = 0.1$ and I ranging from 0.1 to 0.5 (top), and for $I = 0.3$ and D ranging from 0.1 to 0.5 (bottom).

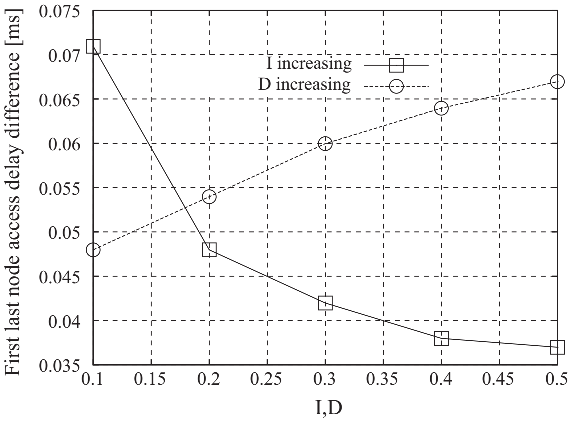


Fig. 12. Difference in the average access delay between the first and the last node for the DLT strategy under uniform traffic for $\rho = 0.9$, with I and D ranging from 0.1 to 0.5.

8. Conclusions

This paper was motivated by the idea that single-hop WDM packet networks can be devised as a cost-effective and feasible solution to fulfill future traffic demands of broadband access systems and metropolitan networks. Among the challenges that these networks introduce, the definition of distributed access protocols that can guarantee both high throughput, good fairness and bounded delays is crucial to demonstrate the efficiency of these networks in handling data at the packet level.

We introduced the fairness problem and discussed Multi-Fasnet, the adaptation of the Fasnet protocol to the WDM scenario. We highlighted the limitations of Multi-Fasnet in terms of throughput and delays. Furthermore, to overcome Fasnet limitations, we proposed several new strategies, namely the quota accumulation, the free access and the DLT strategies, which, coupled together, shows robustness to parameter setting, high throughput, low delays and good fairness properties on a short time-scale, an important feature for time-critical mice flows. Furthermore, DLT permits to operate at the protocol level to dynamically balance the load among channels following traffic fluctuations.

It is worth remarking that Multi-Fasnet does not strictly need to be operated with a slotted access. Indeed, variable-size packets can be accommodated in an asynchronous access, with no time slots nor time reference to be shared among nodes. However, the free-access strategy cannot be supported in this case. This interesting feature of Multi-Fasnet is left for future investigation.

Acknowledgments

We wish to acknowledge Fabio Neri contribution to this paper. Fabio suddenly passed away in April 2011 and we wish to dedicate this paper to his memory.

References

[1] Metro Ethernet Network, Comparison to Legacy SONET/SDH MANs for Metro Data Service Providers. <http://metroethernetforum.org/PDF_Documents/WP_SPBusinessCase_Final071403.pdf>.

[2] Ethernet in the MAN: New Technology Creates New Revenue Opportunities, Sorrento Networks, 2003. <http://www.cascaderange.org/presentations/Ethernet_in_the_MAN.pdf>.

[3] H. Schmidtke, A. Gibbemeyer, Five Reasons to Adopt Layer 2 Ethernet Switching over DWDM Networks Now, 2006. <http://networks.usa.siemens.com/Whitepapers/DWDM_Whitepaper.pdf>.

[4] M. Maier, Optical Switching Networks, Cambridge University Press, 2008.

[5] N.L. Sauze, A. Dupas, E. Dotaro, L. Ciavaglia, M.H. Nizam, A. Ge, L. Dembeck, A Novel, Low cost optical packet metropolitan ring architecture, in: European Conference on Optical Communication (ECOC 2001), Amsterdam, The Netherlands, 2001.

[6] I.M. White, M.S. Rogge, K. Shrikhande, L.G. Kazovsky, A summary of the HORNET project: a next-generation metropolitan area network, IEEE J. Sel. Areas Commun. 21 (9) (2003) 1478–1494.

[7] C. Jelger, J. Elmirghani, A slotted MAC protocol for efficient bandwidth utilization in WDM metropolitan access ring networks, IEEE J. Sel. Areas Commun. 21 (8) (2003) 1295–1305.

[8] A. Bianco, J.M. Finochietto, G. Giarratana, F. Neri, C. Piglione, Measurement based reconfiguration in optical ring metro networks, IEEE/OSA J. Lightwave Technol. 23 (10) (2005) 3156–3166.

[9] M.C. Yuang, I. Fen Chao, B.C. Lo, HOPSMAN: an experimental optical packet-switched metro WDM ring network with high-performance medium access control, IEEE/OSA J. Opt. Commun. Network. 2 (2) (2010) 91–101.

[10] http://www.intunenetworks.com/home/shape-up/core_innovation/.

[11] http://www.ist-mains.eu/.

[12] D. Chiaroni, et al., Demonstration of the interconnection of two optical packet rings with a hybrid optoelectronic packet router, in: European Conference on Optical Communication (ECOC 2010), Torino, Italy, September 2010.

[13] T. Bonald, R.M. Indre, S. Oueslati, C. Rolland, Throughput delay trade-offs in slotted WDM ring networks, in: BROADNETS 2010, Athens, Greece, October 2010.

[14] B. Uscumlic, I. Cerutti, A. Gravey, P. Gravey, D. Barth, M. Morvan, P. Castoldi, Optimal dimensioning of the WDM unidirectional ECOFRAME optical packet ring, Photon. Network Commun. 22 (3) (2011) 254–265.

[15] A. Bianco, T. Bonald, D. Cuda, R. Indre, Cost, power consumption and performance evaluation of metro networks, IEEE/OSA J. Opt. Commun. Network. 5 (1) (2013) 81–91.

[16] D. Bertsekas, R. Gallager, Data Networks, Prentice-Hall, 1992.

[17] L. Massoulie, J. Roberts, Bandwidth sharing: objectives and algorithms, IEEE/ACM Trans. Network. 10 (3) (2002) 320–328.

[18] K. Kyeong Soo, H. Okagawa, K. Shrikhande, L.G. Kazovsky, Unslotted optical CSMA/CA MAC protocol with fairness control in metro WDM ring networks, in: IEEE GLOBECOM 2002, vol. 3, 2002, pp. 2370–2374.

[19] M. Ajmone Marsan, A. Bianco, E. Leonardi, M. Meo, F. Neri, MAC protocols and fairness control in WDM multi-rings with tunable transmitters and fixed receivers, IEEE/OSA J. Lightwave Technol. 14 (6) (1996) 1230–1244.

[20] S. Breuer, T. Meuser, Enhanced throughput in slotted rings employing spatial slot reuse, in: IEEE INFOCOM'94, Toronto, Canada, June 1994.

[21] M. Ajmone Marsan, A. Bianco, E. Leonardi, F. Neri, S. Toniolo, MetaRing fairness control schemes in all-optical WDM rings, in: IEEE INFOCOM'97, Kobe, Japan, April 1997.

[22] N. Bouabdallah, A. Beylot, E. Dotaro, G. Pujolle, Resolving the fairness issue in bus-based optical access networks, IEEE J. Sel. Areas Commun. 23 (8) (2005) 1444–1457.

[23] A. Bianco, D. Cuda, J.M. Finochietto, Multi-MetaRing fairness control in a WDM folded-bus architecture, Comput. Networks 62 (2014) 152–161.

[24] A. Antonino, A. Bianco, A. Bianciotto, V. De Feo, J.M. Finochietto, R. Gaudino, F. Neri, WONDER: a resilient WDM packet network for metro applications, Opt. Switch. Network. 5 (1) (2008) 9–28.

[25] A. Bianco, D. Cuda, J.M. Finochietto, F. Neri, C. Piglione, Multi-Fasnet protocol: short-term fairness control in WDM slotted MANs, in: IEEE GLOBECOM 2006, San Francisco, CA, USA, November 2006.

[26] A. Carena, V. De Feo, J.M. Finochietto, R. Gaudino, F. Neri, C. Piglione, P. Poggiolini, RingO: an experimental WDM optical packet network for metro applications, IEEE J. Sel. Areas Commun. 22 (8) (2004) 561–1571.

[27] M. Herzog, M. Maier, M. Reisslein, Metropolitan area packet-switched WDM networks: a survey on ring systems, IEEE Commun. Surv. Tutor. 6 (2) (2004) 2–20.

[28] J.O. Limb, C. Flores, Description of Fasnet – a unidirectional local–area communication network, Bell Syst. Techn. J. 61 (7) (1982).

[29] N. McKeown, A. Mekkittikul, V. Anantharam, J. Walrand, Achieving 100% throughput in an input-queued switch, IEEE Trans. Commun. 47 (8) (1999) 1260–1267.

[30] J.O. Limb, A simple multiple access protocol for metropolitan area networks, in: ACM Symposium on Communications Architectures & Protocols, Philadelphia, PA, USA, September 1982.

Andrea Bianco is Full Professor and Vice Head at the Dipartimento di Elettronica e Telecomunicazioni of Politecnico di Torino, Italy. He has co-authored over 180 papers published in international journals and presented in leading international conferences in the area of telecommunication networks. He was technical program co-chair of HPSR (High Performance Switching and Routing) 2003 and 2008, DRCN (Design of Reliable Communication Networks) 2005, IEEE ICC 2010 ONS Symposium, Networking 2015, and IEEE Globecom 2015 Next Generation Networking Symposium. He has been TPC member of several conferences, including IEEE INFOCOM, IEEE GLOBECOM, IEEE ICC, HPSR, ONDM and Networking. His current research interests are in the fields of protocols and architectures of all-optical networks and switch architectures for high-speed networks. Andrea Bianco is a IEEE Senior Member.

Davide Cuda is Engineer at CISCO systems, Lausanne, Switzerland. He received his Ph.D. in Electrical Engineering from Politecnico di Torino in February 2009, and he graduated from Politecnico di Torino in November 2005. From August 2007 to August 2008 he visited the Networks Research Laboratory of Prof. Biswanath Mukherjee at University of California, Davis, USA. He was Post-Doc at Politecnico di Torino in 2009, at the Italian National Research Council from January 2010 to March 2011 and at Orange Labs, Paris from March 2012 until March 2013. Before moving to CISCO systems in November 2013, he held a Post-Doc position at Telecom ParisTech, Paris. He participated in several Italian and European projects related to optical networks. His main research interests includes asynchronous scheduling algorithms, power and scalability issues of all-optical switching architectures, and design of distributed control algorithms for all-optical switching fabrics.

Jorge M. Finochietto holds a MS and PhD in Electronics Engineer from Universidad Nacional de Mar del Plata, Argentina, and from Politecnico di Torino, Turin, Italy, respectively. His research interests are in the field of performance evaluation, high-speed networking and switching, and optical and wireless networks. He is Full Professor at Universidad Nacional de Córdoba, Argentina (UNC), and he is also Senior Researcher at the National Research Council (CONICET) of Argentina. From 2005 to 2007, he was a post-doc student at the Telecommunication Network Group of Politecnico di Torino. Since 2007 he is member of the Digital Communications Group at UNC. He has been involved in several national and international research projects in the fields of communication networks. He has co-authored over 50 papers published in international journals and presented in leading international conferences.