# Fixed Size and Variable Size Packet Models in an Optical Ring Network: Complexity and Simulations

Dominique Barth[1], Johanne Cohen[2], Lynda Gastal[3],
Thierry Mautor[1], and Stéphane Rousseau[1]

[1] PRiSM, UVSQ, 45 Av. des Etats-Unis, F-78035 Versailles, France
{Dominique.Barth,Thierry.Mautor,Stephane.Rousseau}@prism.uvsq.fr
[2] LORIA, Campus Scientifique BP2039, F-54506 Vandoeuvre les Nancy, France
Johanne.Cohen@loria.fr
[3] LRI, Bt 490 Université Paris Sud, 91405 Orsay Cedex, France
gastal@lri.fr

**Abstract.** In this paper, we compare the use of two packet models in slotted optical ring networks: a model where each packet has to be routed in consecutive slots, and a model where the slots that form a packet can be routed independently. We first focus on the algorithmic complexity of the related problems. Then, we give the results we obtain with an OMNET simulator in terms of packets' overdelay and jitter.

## 1 Introduction

All optical packet networks represent a challenging and attractive technology to provide a large bandwidth for future networks. Convergence between Internet and optical networks is also a major key point in the conception of the future telecommunication networks [7]. As mentioned by the ITU [5], the architectural choices for the interaction between IP and optical network layers, particularly the routing and signaling aspects are keys to the successful deployment of next generation networks (NGN). This is also a major topic of some working groups of the IETF [6]. This study concerns the DAVID project[4] we participate in. The DAVID optical network is a WAN interconnecting optical MANs, each one consisting of disjoint optical slotted rings [1]. Metropolitan optical rings have also been studied in the IST-1999-10402 European project METEOR. One of the question addressed in the DAVID project and in some other studies was the choice of the packet format and size transported by the slotted ring [2,10,9,8]. In this context, we compare here the communication efficiencies of a simple optical ring network under two packet models: a fixed size and a variable size packet models. A packet to be emitted by a node on the ring consists here of a sequence of slots. At each communication step, each node accesses the slot of the ring located on him. If the data contained by this slot is intended for him, the node

[4] Data and Voice Integration over DWDM, European project of the 5th PCRD.

C. Aykanat et al. (Eds.): ISCIS 2004, LNCS 3280, pp. 238–246, 2004.

reads it and removes the data from the slot (it becomes empty). If the slot is empty, the node can use it to send a new slot of a packet to be sent to a given destination. We consider two models to be compared here. In the *fixed size packet* model, all the slots composed of the same packet can be sent independently on the ring (i.e., each slot is considered as an independent packet of fixed size 1). In the *variable size packet* model, the sequence of slots corresponding to a packet have to be sent contiguously on the ring. The advantage of this second model is that there is a jitter equal to $sz(P) - 1$ and thus no need of memory in the destination node to reconstitute each packet. But, the sending protocol in each node needs to be much more complex.

The purpose of this study is to compare variable size packet and fixed size packet models from three points of view:

- Determining the complexity of obtaining an optimal scheduling in these two models;
- Evaluating performances of each model in terms of packet overdelay and jitter;
- Giving a distributed management algorithm for the variable size model.

Regarding the first two points, as we want the difficulty to be linked only to the models behaviors and not to the network control, we assume slotted rings with only one wavelength. We focus here on a network consisting of one only ring. We also give some results about two rings connected by a switch.

## 2 Models and Problems

Let us consider a ring network $R$ connecting $N$ nodes with only one slotted wavelength (with $K$ slots on each link). Each node on this ring could have some packets to send to another nodes. Each such packet $P$ is characterized by the origin node: $or(P)$, the destination node: $dest(P)$, the size in number of slots: $sz(P)$, the time at which it becomes available in $or(P)$: $dispo(P)$ and the distance (number of slots = number of steps) from $or(P)$ to $dest(P)$: $dist(P) = K * ((dest(P) - or(P)) \bmod N)$. We also consider the time $First(P)$ (respectively, $Last(P)$) of packet $P$ at which its first slot (respectively, last slot) is sent on the ring. Different measures can be defined and applied on each packet $P$. The delay, defined by $Delay(P) = (dist(P)+Last(P)-dispo(P))$, represents the time between the arrival of the packet on the origin node and the end of its transmission, i.e., the end of its reading on the destination node. The jitter, defined by $Jitter(P) = (Last(P) - First(P))$, represents the time between the beginning and the end of the emission of a packet. The overdelay, defined by $OvDel(P) = Last(P) - Dispo(P) - sz(P) + 1$, represents the difference between the delay of $P$ and its minimal possible delay (equal to $dist(P) + sz(P) - 1$). This measure, inspired from some works on scheduling, is interesting to compare the delays of packets of different sizes.

The variable size packet (VSP) model has a jitter constraint: for each packet $P$, we want $Jitter(P) = sz(P) - 1$. This implies that all the slots of a packet have to be sent contiguously on the ring. Let us remark that, in this model, the overdelay of a packet $P$ is equal to: $OvDel(P) = First(P) - Dispo(P)$. In the fixed size packet (FSP) model there is no such constraint and all the slots of a same packet can be sent independently of the others. Thus, the jitter is in this case an evaluation parameter of the behavior of the network.

As previously indicated, several objective functions can be considered for this problem (*Overdelay*, *Jitter*, *Cmax*). In this section, we have decided to focus on two of these objectives. The first criterion (*Cmax*) is the classical *MS makespan* criterion that consists of minimizing the number of steps needed to send all packets to their destination, and thus the time of reception of the last packet. The second criterion called *MD OvDel* consists of minimizing the maximal overdelay of the packets. In order to formalize these two problems, we consider static instances: each node contains a finite and given set of packets. Moreover, we consider a centralized model where each node has a global vision of the network. In terms of complexity, we have thus to consider the following problems:

*Problem 1.* MS-VSP-Scheduling. Let there be a ring $R$, a set $S$ of packets, an integer $B$. Does there exist a scheduling of $S$ on $R$ such that all the packets have reached their destinations after at most $B$ steps, with the constraint $Jitter\ (P) = sz(P) - 1$ for each packet $P$?

*Problem 2.* MD-VSP-Scheduling. Let there be a ring $R$, a set $S$ of packets, an integer $B$. Does there exist a scheduling of $S$ on $R$ such that the maximal OverDelay $OvDel(P)$ over all packets $P$ is less than or equal to $B$, with the constraint $Jitter(P) = sz(P) - 1$ for each packet $P$?

Of course, we have the same problems for the FSP model called MS-FSP-Scheduling and MD-FSP-Scheduling (with the constraint on the jitter is removed).

### 2.1 Problem Complexity

Now, we focus on the complexity of MS-VSP-Scheduling problem.

**Theorem 1.** *The MS-VSP-Scheduling problem is NP-complete.*

*Proof.* It is easy to see that MS-VSP-Scheduling problem belongs to NP since a nondeterministic algorithm needs only to guess a scheduling $S$ on $R$ and check in polynomial time that the delay of all packets in $S$ is less than or equal to $B$.

We reduce 3-PARTITION to MS-VSP-Scheduling problem. 3-PARTITION is defined as follows [3].

*Problem 3.* 3-PARTITION Problem. Let there be a finite set $C$ of $3m$ elements $\{a_1, \ldots, a_{3m}\}$, a bound $\beta \in Z^+$ and a size $s(a) \in Z^+$ such that $\sum_{a \in C} s(a) = m\beta$. Can $C$ be partitioned into $m$ disjoint sets $C_1, C_2, \ldots, C_m$ such that, for $1 \leq i \leq m$, $\sum_{a \in C_i} s(a) = \beta$ ? Recall that 3-PARTITION is NP-complete in the strong sense. Let there be a finite set $C$ of $3m$ elements, a bound $\beta \in Z^+$

and a size $s(a) \in Z^+$ such that $\sum_{a \in C} s(a) = m\beta$ be an arbitrary instance of 3-PARTITION denoted by $I$. We transform instance $I$ into an instance $I'$ of MS-VSP-Scheduling problem. Instance $I'$ is composed of a ring $R$ with 4 nodes numbered $1, 2, 3, 4$ with $K$ slots on each link, where $K = m(\beta+1)-1$. The set $S$ of packets has $3m$ elements $p_1, \ldots, p_{3m}$ such that for $1 \leq i \leq 3m$, $sz(p_i) = s(a_i)$, $dispo(p_i) = 0$, $or(p_i) = 2$, $dest(p_i) = 3$. At time 0, ring $R$ contains some packets that are already transmitted.

- The link connecting node 4 and node 1 contains $K$ slots called $slot^{4\to1}(1), \ldots, slot^{4\to1}(K)$ and there exists a packet $z_0$ of size 1 such that packet $z_0$ is contained by slot $slot^{4\to1}(K)$.
- The link connecting node 1 and node 2 such that it contains $K$ slots called $slot^{1\to2}(1), \ldots, slot^{1\to2}(K)$ and there are $m$ packets $z_1, \ldots, z_m$ of size 1 such that for $1 \leq i \leq m$, packet $z_i$ is contained by slot $slot^{1\to2}(K - i \times (\beta+1))$.

The construction of our instance of MS-VSP-Scheduling problem is completed by setting $K = m(\beta+1)-1$ and $B = m(\beta+1)-1$. It is easy to see how this construction can be accomplished in polynomial time. All that remains to be shown is that the desired partition exists if and only if there exists a scheduling having all the stated properties.

First, assume that there exists a partition $C' = \{C_1, C_2, \ldots, C_m\}$ such that for $1 \leq i \leq m$, $\sum_{a \in C_i} sz(a) = \beta$. Let $i$ be an integer such that $1 \leq i \leq m$. We consider the partition $C_i$. Without losing generality, we set $C_i = \{a_{i^1}, \ldots, a_{i^t}\}$, where $t$ depends on the $C_i$, we will construct a scheduling such that

for each $j \in [1, \ldots, t]$, $First(p_{i^j}) = (\beta+1) \times (i-1) + \sum_{\alpha=1}^{j-1} sz(p_{i^\alpha})$.

Note that the overdelay of packet $p_{i^j}$ is $Last(p_{i^j}) = First(p_{i^j}) + sz(p_{i^j})$. Since $Last(p_{i^j}) < B$, we construct a scheduling having all the stated properties. Figure 3 is an example considering only nodes 1 and 2 and the link between them. In this example, $m = 3$, $\beta = 4$, $B = 14$.

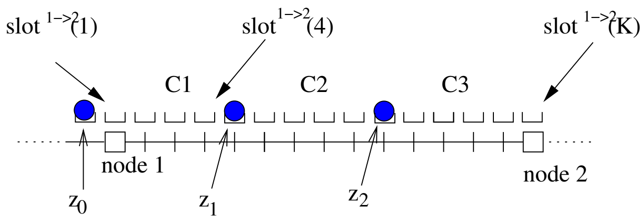


**Fig. 1.** Proof of transformation in Theorem 1 for $m = 3$ and $\beta = 4$

Conversely, assume that there exists a scheduling $S$ on $R$ such that the overdelay over all packets of $S$ is less or equal to $B$. Now, we will construct a partition of $m$ elements $C_1, C_2, \ldots, C_m$ such that for $1 \leq i \leq m$, $\sum_{a \in C_i} s(a) = \beta$. First, we focus on the property of such a scheduling. Each packet $p_i$, $1 \leq i \leq 3m$

satisfies the following property: $OvDel(p_i) \leq B$. It implies that all packets $p_i$, $1 \leq i \leq 3m$, are inserted before the slot containing packet $z_0$ arrives at node $n_2$.

Now, we construct set $C_i$, $1 \leq i \leq m$: $a_j$ belongs to $C_i$ if $(\beta+1) \times (i-1) < First(p_j) < i \times (\beta+1)$. First assume that $i > 2$. By construction, packets $z_{i-1}$ and $z_i$ arrive at time $(i-1) \times (\beta+1)$ and $i \times (\beta+1)$ respectively. Since, for all packets $P$ have $sz(P) - 1$ as jitter, all slots containing a part of $p_j$ should be contiguous: they must be between slots containing packets $z_{i-1}$ and $z_i$. Thus, we get the following property $\sum_{j:a_j \in C_i} sz(p_j) \leq \beta$ and $\sum_{a_j \in C_i} s(a_j) \leq \beta$. We apply the same argument for the case where $i = 1$. So set $\{C_1, C_2, \ldots, C_m\}$ is partition of $C$ such that for $1 \leq i \leq m$, $\sum_{a \in C_i} s(a) = \beta$.

So, we have shown that the desired partition exists if and only if there exists a scheduling having all the stated properties. This concludes the proof. □

## 3 One Ring Distributed Model

Let us consider the ring network presented in Section 2. In an off-line model, with a centralized management, finding a scheduling that minimizes either the delay or the jitter cannot be obtained in a polynomial time. From now on, we study a distributed management protocol in an on-line model. Quite obviously, in such a model, minimizing the criteria previously defined, cannot be obtained in polynomial time. Considering the two models (FSP and VSP) we present a management protocol in order to send packets contained in nodes to their destinations. But first, we present an access control strategy in each node to decide which packet can be sent and which one first.

In a distributed model, each node sees only the slot upon it at each step. If this slot indicates the number of slots in the free zone, the node knows if it can send a packet of a given size or not. Thus, the node knows the maximal packet length that can be put into this free zone. The access control protocol has to choose which packet or which subset of packets to put into it, if it is possible. When a slot contains data (or packets) it is considered as taken. In this case, when a node meets such a slot, it takes the packet if this packet is destined for it. The slot becomes free again and the node can use it. Thus, when a node is in front of a free zone, it tries to put one of its packets into it. However, this free zone could not be large enough to send any packet. In such a case, the node can reserve this zone in order to merge it with another contiguous one. Thanks to this reservation, the node is sure to find it free again in the next turn. For the other nodes, the slot seems to be taken but it does not contain any packet. In such a packet, the origin node is also the destination.

Because the constraint on the jitter does not occur in the FSP scheduling model, there is no need to have a free zone as large as a packet to send it. Indeed, a packet can be sent using severals disjoint free zones. Thus, in the distributed management protocol, a packet is put in the first possible free slot. However, in the VSP scheduling model, we assume that a node receiving a packet of $S$ slots creates a free zone of $S$ slots. Similarly, a node putting a packet of $L$ slots into a free zone of $S > L$ slots creates after it a free zone of $S - L$ free slots. The consequence of such a basic distributed control is that the ring can become a cyclic sequence of free zone steps, each one being too short to be used by any node. Consequently, we have to enhance this distributed control in order to be able to merge consecutive free zones.

We now present our distributed management protocol in the VSP model.

### 3.1 Distributed Management Protocol in the VSP Model

In the VSP model, packets have to be sent so that the constraint on the jitter is respected. Before putting a packet, the node has to be sure to access a free zone large enough to put the packet in it. A free zone is made of a set of free slots. Each free slot of a free zone indicates the number of consecutive free slots left. Then when a node receives the first slot of a free zone, it is informed of the length of the free zone. However, as we said before, the ring can become a cyclic sequence of small free zones in which none of the packets can be put without being split.

**First Part of the Distributed Management Protocol in the VSP Model.**

1. When a node receives a taken slot:
   (a) If the packet inside the zone is destined for this node, then it takes it and considers the slot as free.
   (b) If the packet inside the zone is not destined for itself, then it does nothing.
2. When a node receives a reserved slot:
   (a) If the slot has been reserved by this node, then it makes it free again and considers this slot as such. It updates the length of the whole free zone, in the case where several reserved zones are adjacent.
   (b) If the packet has not been reserved by this node, then it considers this slot as not free.
3. When a node receives a free slot:
   (a) If the node has packets to send and the free zone indicated by the slot is large enough to put in one of its packets, then it sends a packet according to the $FIFO$ strategy. It updates the length of the free zone remaining.
   (b) If the node has packets to send but the free zone indicated by the slot is not large enough to put in one of its packets, then:
      i. If this free zone is just after with a free zone reserved by this node then it reserves it.
      ii. If the node has not reserved a free zone yet then it reserves this one.
      iii. In other cases, the node lets this free zone go as it is.
   (c) If the node does not have packets to send, then it lets the slot free.

When a node reserves a free zone, it is able to know the number of adjacent free zones and so the total length of the reserved zone. Then when this reserved zone reaches again the node, each slot has the same number of free slots as before.

One problem remains using this protocol. Some packets cannot be sent because they are too long. It is possible to have packets remaining in nodes and to be not able to send them to their destination. For example, consider a two node ring network with two free zones of length one. Assume that each node has to send one packet of length two. Then they have to reserve at least two contiguous slots and merge them in one free zone of length two. But they reserve one slot each other and this prevents them from reserving two contiguous slots. This is a deadlock problem.

**Second Part of the Protocol.** At each time slot, a single node has the highest priority. This priority allows it to reserve a zone that belongs to another node if and only if this zone follows a free zone. The priority changes periodically among nodes in order to ensure some fairness. In an off-line VSP distributed model, the distributed management protocol, using the priority, ensures each packet to reach its destination and, by using the priority, it ensures to merge at least two contiguous free zones in two turns, or at least three contiguous zones in $3 + [\#nodes - 1]/(\#nodes)$ turns. Proofs are available in [4].

### 3.2 Simulation Results

In this section, we present results from simulations. We have considered a ring of 10 nodes and 9 slots between each pair of nodes. Thus, the nodes can use 90 slots to send their packets. We focus on three kinds of traffic. The first one, called *traffic C*, is characterized by a long period of small packets and a short period of large packets. The second one, called *traffic G*, is characterized by a short period of small packets and a long period of large packets. The third one, called *traffic I*, is characterized by a short period of small packets and a short period of large packets.

**Results Analysis.** Even if many measures can be investigated with such a simulation model, we have focused on two parameters. The first one is the average overdelay of the packets in order to study the influence of the length of packets on the average overdelay. The second one is the average jitter. In Figure 2, we compare the average overdelay average in the FSP and the VSP models. We study the average jitter value according to traffic type.

As it can be seen in Figure 2, the overdelay is better in the FSP model than in the VSP model. Also, the FSP model accepts a larger data flow than the VSP model before the network is saturated. The network is saturated when the delay increases indefinitely. Then the average overdelay is not bounded. In the VSP model, when the data flow is over 0.13 amount of packets per time unit, the network is saturated, whereas in the FSP model the network can accept at least 0.2 amount of packets per time unit before being saturated. The amount of packets per time unit is the number of packet of size 1 generated by time unit. Referring to Figure 2, the average jitter increases as a function of the data traffic. In fact, as expected, the average jitter is better in traffic $C$ where there are smaller packets and deteriorates as the number of large packets increases.

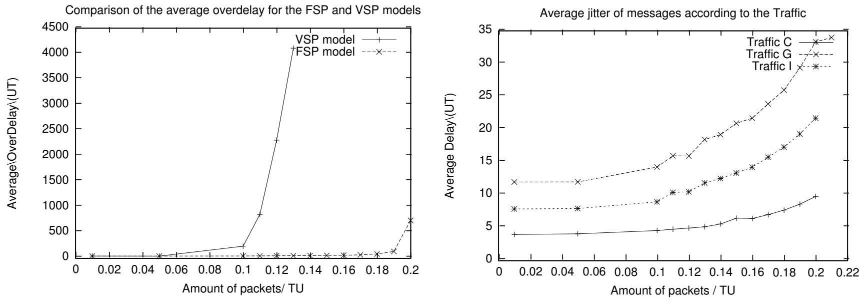


**Fig. 2.** Comparison of average overdelay in FSP and VSP models and average jitter under different types of traffic

## 4 Conclusions and Perspectives

We have shown that finding optimal packets scheduling is difficult in both FSP and VSP models. For this last one, in an online and distributed context, there is also the problem of merging consecutive free zones to avoid live-locks. The cost of a protocol realizing this concatenation can be evaluated in terms of resource use, and the one we give here ensures that at each step, about 80% of the slots are used to carry data, and this is not to the detriment of the overdelay of large packets.

## References


1. Barth, D., Cohen, J., Fragopoulou, P., Hébuterne, G.: Wavelengths Assignment on a Ring All-Optical Metropolitan Area Network. In: Proceedings of the 3rd Workshop on Approximation and Randomization Algorithms in Communication Networks (ARACNE 2002), Rome Italy. (2002)
2. Bianco, A., Bonsignori, M., Leonardi, E., Neri, F.: Variable-Size Packets in Slotted WDM Ring Networks. Optical Network Design and Modelling (ONDM 2002), Torino, Italy. (2002)
3. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Series of Books in the Mathematical Sciences. W.H. Freeman and Company (1979)
4. Gastal, L., Mautor, T., Rousseau, S., Barth, D., Cohen, J.: Comparison of Fixed Size and Variable Size Packet Models in an Optical Ring Network. Technical report, PRiSM, UMR 8144, University Versailles-Saint Quentin, 45 avenue des Etats Unis, 78035 Versailles Cedex, France. (2004)
5. International Telecommunications Union. Convergence between Internet and Optical Networks. ITU News Magazine **2** (2002) `http://www.itu.int/itunews/issue/2002/06/convergence.html`
6. Jain, R., et al.: IP over Optical Networks: A Summary of Issues. IPO and MPLS Working Groups Internet Draft for IETF. (2001) `http://www.cse.ohio-state.edu/~jain/ietf/ftp/draft-osu-ipo-mpls-issues-01.pdf`

7. Rajagopalan, B., Pendarakis, D., Saha, D., Ramamoorthy, R.S., Bala, K.: IP over Optical Network: Architectural Aspects. IEEE Communications Magazine **38** (2000) 94–102
8. Yao, S., Yoo, S.J.B., Mukerjee, B.: A Comparison Study between Slotted and Unslotted All-Optical Packet-Switched Network with Priority-Based Routing. In: Proceedings of OFC 2001, Anaheim, California (2001)
9. Yu, X., Cheng, Y., Qiao, C.: A Study of Traffic Statistics of Assembled Burst Traffic in Optical Burst Switched Networks. In: Proceedings of the IEEE Optical Networking and Communications Conference (OPTICOMM), Boston, Massachusetts (2002) 149-159
10. Zapata, A., Düser, M., Spencer, J., de Miguel, I., Bayvel, P., Breuer, D., Hanik, N., Gladisch, A.: Investigation of Future Optical Metro Ring Networks Based on 100-Gigabit Ethernet (100GbME). ITG Fachbericht, Vol. 175. VDE, Berlin, Germany (2003) 89-98