AWG-based Optoelectronic Router with QoS Support

Annette Böhm, Magnus Jonsson and Kristina Kunert

Böhm A, Jonsson M, Kunert K. AWG-based Optoelectronic Router with QoS Support. In: 2005 International Conference on Wireless and Optical Communications Networks: March 6 - 8, 2005, Dubai, United Arab Emirates. Piscataway, NJ: IEEE; 2005. p. 23-28.

DOI: http://dx.doi.org/10.1109/WOCN.2005.1435982
Copyright: IEEE

AWG-based Optoelectronic Router with QoS Support

Annette Böhm, Magnus Jonsson, and Kristina Kunert
School of Information Science, Computer and Electrical Engineering, Halmstad University
Box 823, S-301 18 Halmstad, Sweden.

Abstract – **We present a router architecture with electronic queuing and a passive optical multi-channel network, which is based upon a single-hop star topology with an Arrayed Waveguide Grating (AWG) at its center. The AWG's property of spatial wavelength reuse and both fixed-tuned and tunable transceivers enable simultaneous control and data transmission. Our proposed medium access control protocol supports traffic differentiation and utilizes EDF (Earliest Deadline First) to schedule the traffic from input ports to output ports on the router. Our simulations show that the router treats QoS (Quality of Service) traffic well.**

Keywords – **Arrayed Waveguide Grating, Quality of Service, Optoelectronic router, IP Router, MAC protocol, WDM**

I. INTRODUCTION

The performance of data- and telecommunication equipment must keep abreast of the increasing network speed. At the same time, it is necessary to deal with the internal interconnection complexity, which typically grows by N^2 or $N\log N$, where N is the number of ports. This requires new interconnection technologies to be used internally in the equipment. Optical interconnection technology is a promising alternative that enables both high throughputs and presents solutions to the problem of complex interconnection networks (due to the fact that light beams can cross each other without interference). On the other hand, electronic solutions are still needed to cope with the increasing demands on QoS (Quality of Service) support. An overview of optical interconnections in communication equipment is found in [1].

In this paper, we present a router architecture with electronic queuing and QoS handling, while the traffic in the router is transported from input ports to output ports through a passive optical network. The network is a single-hop Wavelength Division Multiplexing (WDM) star network with an Arrayed Waveguide Grating (AWG) at its center. Our proposed medium access control (MAC) protocol relies on traffic scheduling centralized in one node, which communicates through a physical control channel. The AWG's property of spatial wavelength reuse and the combination of fixed-tuned and tunable transceivers at the input ports and output ports of the router, enable simultaneous control and data transmission. The MAC protocol's main function is to handle three different traffic classes: high-priority real-time (HRT), standard-priority real-time (SRT) and non-real-time (NRT) traffic. HRT can, for instance, be used for remote surgery over the network, while SRT is the main class of real-time traffic. Prioritization between different packets within a real-time traffic class is then done using EDF (Earliest Deadline First) scheduling.

Extensive research has been done in the field of single-hop WDM star networks, where broadcast-and-select networks based on the Passive Star Coupler (PSC) were the center of attention [2]. Opposed to the PSC, the AWG is a wavelength-sensitive device that makes it possible to reach selected parts of the network and increase its concurrency and throughput. Optical Cross Connect (OXC) architectures and Optical IP routers including AWG components have been proposed (see, e.g., [3]) but research on more pure AWG-based architectures and protocols for packet switching is limited. A comparison of the PSC and the AWG is presented in [4], while the network proposed in [5] combines the two components in a hybrid solution. AWG-based single-hop networks are presented in [6] and [7], where the latter one focuses on the development of a MAC protocol. A common feature of the related works mentioned, is the lack of real-time support. In our proposed system, each port of the router has its individual control channel. We make use of the AWG's property of spatial wavelength reuse to maximize the concurrency of both control and data traffic, and time-slotted access to enable collision-free transmissions. The high level of concurrency can be compared to control channel-based WDM star networks, where the transmission of control packets is serialized instead [8] [9]. Together with the centralized scheduling algorithm, our architecture enables support of real-time traffic with the possibility of both delay and throughput guarantees.

The rest of the paper is organized as follows: The AWG-component and the proposed router architecture are described in Section II, while the MAC-protocol and its scheduling algorithm are explained in Section III. Our solution is evaluated in a simulation, which is discussed in Section IV and conclusions are drawn in Section V.

II. ROUTER ARCHITECTURE

A. The Arrayed Waveguide Grating

The AWG, also known as Wavelength Grating Router (WGR) or Phased Array (PHASAR) [10], consists of an

0-7803-9019-9/05/$20.00 ©2005 IEEE.

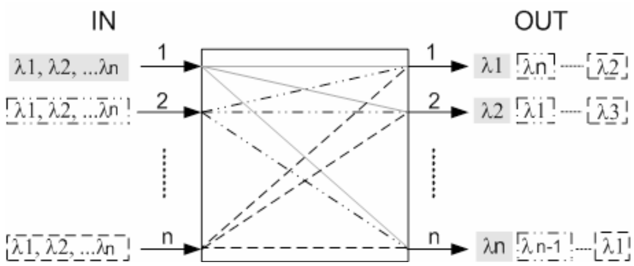

Fig. 1. Wavelength routing pattern of an $N \times N$ AWG.

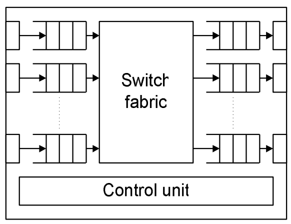

Fig. 2. Example of a router architecture

arrangement of optical waveguides on a substrate plate, which gives it special spatial and spectral properties for highly effective future photonic communication networks. AWGs are often used as a passive wavelength routing devices in optical communication networks. An $N \times N$ AWG has N input ports (input waveguides) and N output ports (output waveguides). Offering full connectivity, it can simultaneously accept N wavelengths at each input port and route each one of them to a specific output port without collision (Fig. 1) [10]. To be able to send and receive on all possible wavelengths, the nodes attached to the AWG need tunable transmitters and tunable receivers at the input and output ports, respectively. To make multiple communication channels possible, an AWG uses spatial wavelength reuse (Fig. 1), i.e. every wavelength can be used on all output ports at the same time without collision. This kind of routing makes AWGs efficient and improves the throughput and latency behavior of optical networks [4].

B. The architecture

A general router can be viewed as a number of input and output ports and their corresponding queuing systems (see Fig. 2). The switch fabric is the heart of the router and is responsible for carrying packets from input ports to desired output ports. In low-performance routers, the switch fabric is often just a simple bus, while pure high-performance routers often have some kind of complex multi-stage interconnection network to implement an experienced all-to-all connectivity.

In our router architecture, the AWG network acts as switch fabric. The proposed network architecture consists of one $N \times N$ AWG, integrated in a hub with a processing unit used for the centralized scheduling (further on denoted as "Protocol Processor", PP) and $N - 1$ pairs of input and output ports (Fig. 3). The AWG component is the center of the physical WDM star network, and the PP and each pair of input and output ports are connected to the AWG by two fibers, one for transmission and one for reception.

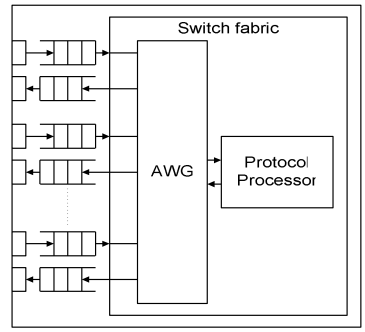

Fig. 3. The proposed router architecture with the AWG network acting as a switch fabric

The network is based upon the principle of CC-TT-TR-FT-FR (Control Channel - Tunable Transmitter - Tunable Receiver - Fixed Transmitter - Fixed Receiver) [2]. The fixed receiver and transmitter are used to establish physical control channels between the router's input and output ports and the PP. Each of the $N - 1$ fixed transmitter-receiver pairs will send on an individual wavelength. On the control channels, the PP receives all information needed to run its scheduling algorithm, and sends control information back to the ports. In addition to the control wavelength, one of the remaining $N - 1$ wavelengths can, in each time slot, be used for transmission of a data packet.

The PP transmits and receives data via wavelength array components. The array component on the transmitting side of the PP is attached to a $(N - 1) \times 1$ combiner, which demultiplexes the $N - 1$ different wavelengths onto one fiber. Correspondingly, a $1 \times (N - 1)$ splitter separates incoming wavelengths before they enter the $N - 1$ array of receivers.

The fixed tuned and tunable transmitters on the $N - 1$ input ports of the router are each connected to a 2×1 combiner, that demultiplexes the two wavelengths (one fixed wavelength to the PP and a tunable one for data traffic to the nodes) onto one fiber. At the output port, a 1×2 splitter is connected to one fiber with incoming traffic.

III. MEDIUM ACCESS CONTROL PROTOCOL

The main task of the MAC protocol is to delegate the available wavelengths between the router's input ports based on control information, and to make sure that the hard real-time messages with the shortest deadlines are sent first. The fibers interconnecting the ports and the protocol processor are assumed to have a maximum length of a few tens of meters, i.e., in the case of a distributed router. In

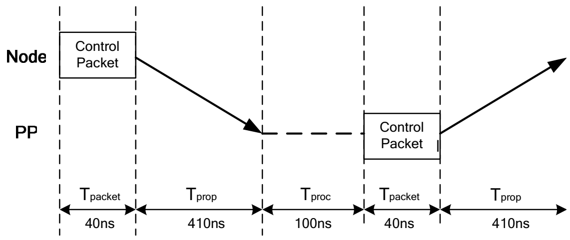

Fig. 4. Control packet propagation.

other words, the propagation delay is assumed to be a fraction of a time slot.

A. The control traffic

The PP needs to receive information about the packets to be scheduled for the following time slot. In each time slot, each input port sends a control packet to the PP over the dedicated control channel. The PP has a memory to store information about packets that did not get permission to be sent. Due to the memory, the control information does not need to be sent twice.

Each of the router's input ports has four different message queues per possible output port. One queue for HRT messages whose control information is already stored at the PP and one for the remaining HRT packets (further on denoted as "control queue" and "non-control queue" respectively). Corresponding queues exist for standard and non-real-time traffic; non-real-time is defined as SRT traffic with infinite deadlines. In each time slot an input port can receive one or more new packets that are put in the queue that corresponds to the desired traffic class. Although the incoming bit rate is the same as the bit rate in our internal router network, the packet generation at the input might be rather bursty due to, e.g., division into smaller cells. It is therefore important that control information about more than one packet at a time can be sent. The queues are sorted by deadline, and a control message with information about the first four packets at a time is sent to the PP, requesting permission to send data. The HRT queues are always checked first. In case of an empty HRT queue, the SRT queue is considered.

The control message contains information about traffic class, destination port and deadline for each of the four packets in question. As a control packet is simple and contains little information, one slot can be used to send control traffic to the PP, to run the scheduling algorithm and to send control messages back to the ports. Considering a 64 × 64 AWG, the length of a control packet (T_{pack1}) to the PP is less than 150 bits (1 bit for the traffic class, 6 bits to identify the destination and 20 bits to state the deadline for each of the four packets, plus some bits of overhead), i.e. about 60 ns, assuming a bit rate of 2.5 Gbit/s. A control message from the PP back to the ports (T_{pack2}) is even shorter, about 50 bits (6 bits to state the destination for the accepted packet and, if an output port can expect incoming traffic, 6 bits to identify the sending port, plus overhead). This corresponds to about 20 ns. Accessing the PP's memory to fetch and store information about unscheduled packets can be done during the transmission time of the control packets and does not add additional time. At the end of the control slot, the pairs of input/output ports tune their respective transmitters and receivers according to the control information, which takes about 100 ns, (T_{tune}).

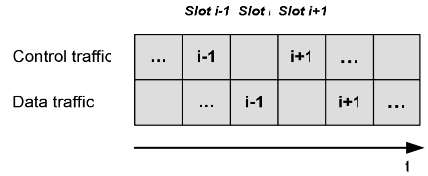

Figure 5. Control packet propagation.

Assuming a bit rate of 2.5 Gbit/s, a 400 ns processing time (T_{proc}) at the PP and defining one slot (T_{slot}) as 1 µs, i.e. 2500 bits, leads to the following calculation of the propagation time, T_{prop}:

$$T_{prop} = 0.5 \cdot (T_{slot} - T_{pack1} - T_{pack2} - T_{proc} - T_{tune}) = 210 \text{ ns}. \qquad (1)$$

In other words, a slot of 1 µs leaves enough communication time to support a fiber length of 40 m between the PP and a router port (see Fig. 4).

Data transmissions take place in the following time slot. Due to the fixed tuned and individual wavelength of the control channel, it is possible to handle data and control traffic in parallel. The AWG allows control traffic from all the input ports to be sent to the PP simultaneously. Therefore, the medium access control does not result in more than one single slot of delay and does not have any impact on the data traffic flow (se Fig.5).

As each pair of input and output ports has its own control channel, clock synchronization on the packet level is not required. Control messages from the PP are sent out simultaneously, and as interconnection distances are assumed to be short, the propagation delay due to dispersion is negligible. Hence, all ports are synchronized on incoming control packets and further synchronization is not needed. (A discussion of timing and dispersion in WDM star networks can be found in [11].)

B. The scheduling algorithm

After having sorted its four queues according to the EDF (Earliest Deadline First) algorithm, each of the router's input ports sends information about the first four packets in the "non-control" HRT queue to the PP and moves the packets to the corresponding HRT "control queues". As long as there are packets queued in any "non-control" HRT queue, the SRT traffic has to wait. This guarantees that the HRT traffic does not get disturbed or compromised in any situation.

The PP is responsible for accepting or denying requests. To avoid receiver collisions, there must be no more than one transmission to a single destination in any given time slot and an approach to select the right set of transmissions is proposed below. When the PP has received all the control messages, it sorts them (together with the control information stored in its memory) by deadline, checks them one by one and determines which of the requested transmissions that can be accepted without causing a receiver conflict. As soon as a transmission to a certain output port is accepted, any other request for communication with this destination is denied. When the algorithm has found the right set of packets, the PP sends out individual messages to the input ports and stores the remaining requests in the memory.

Each input port receives information about which output port it is permitted to send traffic to. It picks the packet out of the "control queue", sends it and discards it from the queue. The control message also contains information about which other source the node can expect traffic from in the following slot.

C. Feasibility analysis

In order to determine the performance characteristics for high-priority real-time traffic, a deterministic analysis of throughput guarantee, delay and latency is given below. Even when all the input ports want to send data to the same output port in any given time slot, there is always at least one packet (the one with the shortest deadline) that will be guaranteed access to the medium, i.e. a capacity of 1 can be guaranteed for high-priority real-time traffic.

One can look upon the channel as a periodic task, while the network would constitute a CPU or processing system (from a scheduling point of view). The capacity, C_i, would be the worst-case-execution-time (WCET) for the task (the logical channel) with index *i.* Before we describe the feasibility test, we make the following definitions:

According to basic EDF theory [12] the ***Utilization*** of periodic real-time traffic is defined as

$$U = \sum C_i / P_i \text{ , where } P_i \text{ is the period of task } i. \quad (2)$$

The ***Hyperperiod*** for a set of periodic tasks is defined as the length of time from when all tasks' periods start at the same time, until they start at the same time again.

A ***BusyPeriod*** is any interval of time in which a link is not idle.

The ***Workload function h(t)*** is the sum of all the capacities of the tasks with absolute deadline *d* less than or equal to *t*, where *t* is the number of timeslots elapsed from the start of the Hyperperiod. It is calculated as follows:

$$h(t) = \sum (1 + \left\lfloor \frac{t - d_i}{P_i} \right\rfloor) C_i \text{ , where } d_i \le t. \quad (3)$$

A ***feasible system state*** is a system state with all logical channels in the system being feasible.

Following the discussion from above, and the new definitions, testing if a new logical channel can be added is therefore equivalent to testing if the new state is still feasible, given that the new logical channel has been added. The feasibility test is done in two steps, each step being a test of its own.

First Constraint: The utilization of the link has to be less than or equal to one (100%), following the discussion above about guaranteed access to the medium.

Second Constraint: For all values of *t*, the workload function *h(t)* has to be less or equal to *t*.

The second constraint, in the form given above, does not lend itself out particularly well to computation. It is shown in [13] how to reduce the time and memory complexity of the second constraint check. If $h(t) \le t$ in the first busy period of the Hyperperiod in the supposed schedule to come, then $h(t) \le t$ for all *t*. The following upperbound of the interval to be checked is therefore an improvement of the algorithm above:

$$t \text{, such that } 1 \le t \le BusyPeriod, \quad (4)$$

where *BusyPeriod* is the first BusyPeriod in the schedule at the start of the Hyperperiod. Furthermore, one needs not check every integer from the first timeslot, but only the integers *t* where

$$t \in \bigcup_{i=1} \{mP_i + d_i : m = 0,1,2...\} \quad (5)$$

IV. Simulation Analysis

Simulations were carried out to analyze delay, throughput and deadline-miss-ratio of the two real-time traffic classes by varying the HRT and the SRT traffic intensity respectively.

For the simulations we chose a 64 × 64 AWG, which leads to a router with 63 input ports and 63 output ports (saving one pair of input and output ports for the PP). Further we assumed periodic traffic for both traffic classes with a period and deadline of 100 time slots. The curves are attained by running the simulator for 1000 time slots with a new set of parameters for each data point in the curve. The statistical results are computed from slot 400 and forward. Periodic traffic channels in the simulator are treated as follows. In order to simulate incoming traffic over the whole period, a random offset is computed, that decides in which slot (in each period throughout the simulation) a certain channel will be requested for transmission. The distribution of source addresses and destination addresses is randomized with an even distribution. Certain further assumptions were made:

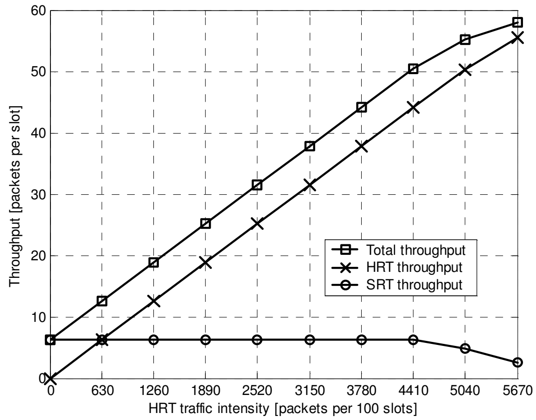

Fig. 6. Throughput of HRT and SRT traffic and total throughput against HRT traffic intensity.

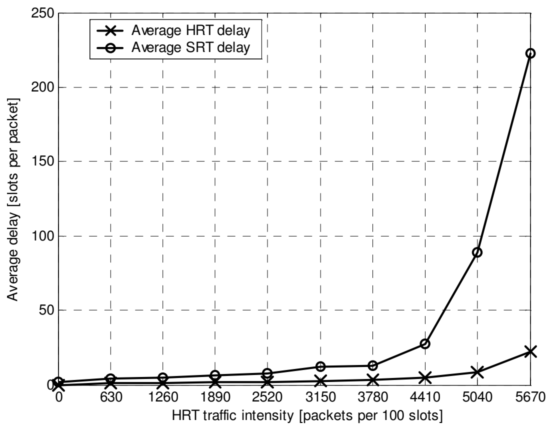

Fig. 7. Average delay of HRT and SRT traffic against HRT traffic intensity.

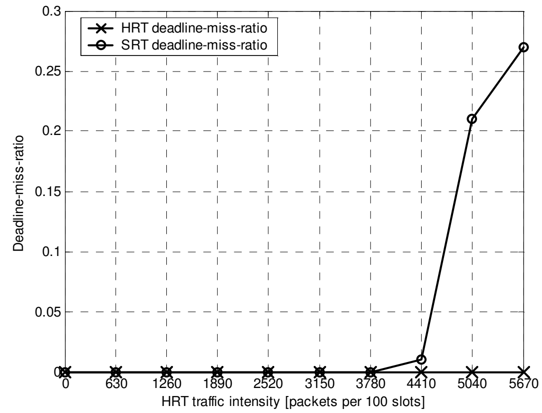

Fig. 8. Deadline-miss-ratio of HRT and SRT traffic against HRT traffic intensity.

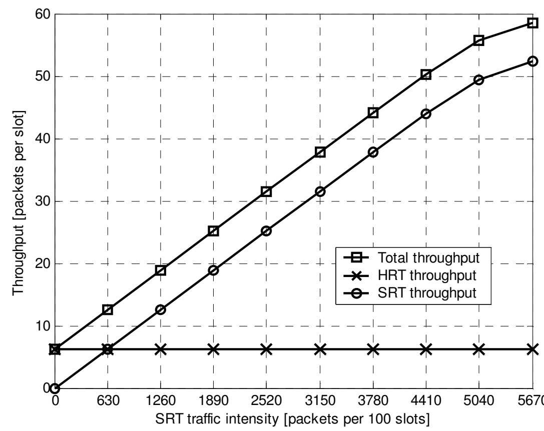

Fig. 9. Throughput of HRT and SRT traffic and total throughput against SRT traffic intensity.

- All data messages have a length of one time slot.
- The queues in all ports and in the PP s memory are of infinite length.
- Packets are generated at the beginning of a time slot.
- Delay and deadline are expressed in number of time slots, from the point in time when a packet is generated until its transmission starts.

Fig. 6-8 show the router s behavior when keeping the SRT traffic intensity at a constant rate of 10 %, i.e. 630 packets per period of 100 time slots. The HRT traffic intensity is increased from 0 % to 90 % (5670 packets per period of 100 time slots). At the highest HRT traffic rate, the total traffic load in the system is 100 %.

Fig. 6 shows that an increasing HRT traffic load leads to steady increase in HRT throughput, while the SRT traffic remains on a constant level before experiencing a slight decrease at HRT rates greater than 4410 packets per period. Due to its lower priority, SRT traffic is filling the capacity not needed by HRT traffic and at high HRT traffic intensities, SRT packets get starved. The total throughput reaches up to a value of about 58 packets per slot.

The behavior of the average delay in Fig. 7 can be explained accordingly. With an increased HRT traffic intensity, the prioritized HRT traffic is affected the least, while the SRT traffic type experiences an increase in delay.

As can be seen in Fig. 8, the slight increase in average delay at high HRT traffic intensities does not affect the HRT deadline-miss-ratio. This shows that the HRT capacity of the network is not challenged by the imposed traffic intensities. The sudden increase in average delay for SRT packets leads to a deadline-miss-ratio of about 0.27 at the maximum HRT traffic intensity of 5670 packets period.

The second set of figures (Fig. 9-11) are obtained by having a constant HRT traffic rate of 10 %, i.e. 630 packets per period, while the SRT traffic intensity is varied between 0 % and 90 % (i.e. between 0 and 5670 packets per period).

The throughput of the two traffic classes, together with the total throughput, is shown in Fig. 9. As the intensity of the SRT traffic increases, its throughput reaches a maximum of about 52 packets per input port and slot at its highest traffic intensity (5670 packets per period). Due to

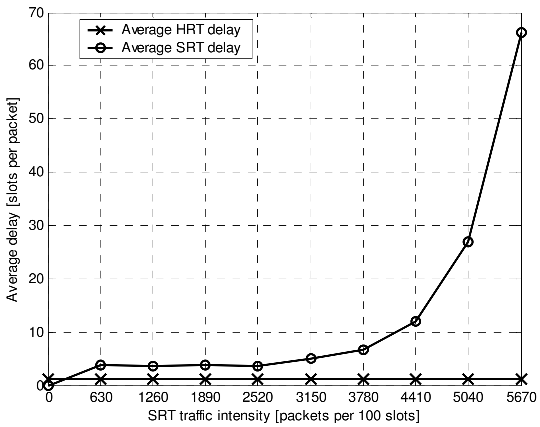

Fig. 10. Average delay of HRT and SRT traffic against SRT traffic intensity.

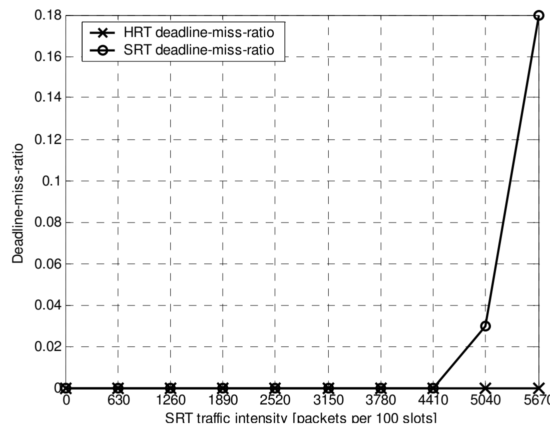

Fig. 11. Deadline-miss-ratio of HRT and SRT traffic against SRT traffic intensity.

its high priority, the throughput of the HRT traffic is not affected and remains constant at 6.3 packets per input port and slot. Comparing the total throughput in this figure to the total throughput in Fig. 6 shows that a 100 % utilization of the system, in both cases, leads to a total throughput of about 58 packets per node and slot.

Fig. 10 shows that the SRT traffic experiences an increase in delay, as the traffic load in the system increases, while the HRT traffic remains at a delay of about 1.1 slots (not including actual transmission) per packet throughout the simulation. Due to its higher priority, it is not affected by the changing level of SRT traffic.

Fig. 11 verifies that the HRT traffic meets all its deadline requirements. In spite of the increasing SRT traffic load, the deadline-miss-ratio of the HRT traffic remains zero, while the SRT deadline-miss-ratio increases considerably as its intensity approaches 90 %.

V. Conclusion

In this paper, a router architecture with electronic queuing and an AWG-based, passive optical multi-channel network have been proposed. A MAC protocol handling high-priority real-time, standard priority real-time, and non-real-time traffic, is described. A protocol processor schedules the traffic over the network in a centralized manner. Using the AWG component, enables the router s input ports to send control information simultaneously to the processor without any collision, while, at the same time, data traffic can be sent to the output ports. This level of concurrency gives short delays and makes efficient use of the bandwidth. A throughput guarantee for HRT traffic is given and proven with EDF-analysis. In a simulation, delay, throughput and deadline-miss-ratio are analyzed for varying high and standard priority real-time traffic intensities. It is shown that all incoming HRT traffic is handled without causing any deadline misses, even at high traffic intensities.

References

[1] M. Jonsson, Optical interconnection technology in switches, routers and optical cross connects, *SPIE Optical Networks Magazine*, vol. 4, no. 4, pp. 20-34, July/Aug. 2003.

[2] B. Mukherjee, WDM-based local lightwave networks, part I: single-hop systems , *IEEE Network*, vol. 6, no. 3, pp. 12-27, 1992.

[3] H. J. Chao and T.-S. Wang, An optical interconnection network for terabit IP routers, *Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2095-2112, Dec. 2000.

[4] M. Maier and A. Wolisz, Demonstrating the potential of arrayed waveguide grating based single-hop WDM networks , *Optical Networks Magazine*, vol. 2, no. 5, pp. 75-85, 2001.

[5] C. Fan, M. Maier, and M. Reisslein, "The AWG||PSC network: A performance enhanced single-hop WDM network with heterogeneous protection", In *Proceedings of INFOCOM 2003*, vol. 3, pp. 2279-2289, San Francisco, CA, USA, 2003.

[6] D. Banerjee, J. Frank, and B. Mukherjee, Passive optical network architecture based on waveguide grating routers , *IEEE Journal of Selected Areas in Communication*, vol. 16, no. 7, pp. 1040-1050, Sept. 1998.

[7] M. Maier, M. Reisslein, and A. Wolisz, A hybrid MAC protocol for a metro WDM network using multiple free spectral ranges of an arrayed waveguide grating , *Computer Networks*, vol. 41, no. 4, pp. 407-433, 2003.

[8] J. Diao and P.L. Chu, Packet rescheduling in WDM star networks with real-time service differentiation , *Journal of Lightwave Techn.*, vol. 19, no. 12 , pp. 1818-1828, Dec. 2001.

[9] J.S. Choi, Performance analysis of a bandwidth guaranteed medium access control protocol for WDM local networks , *Photonic Network Communications*, vol. 6, no. 1, pp. 69-81, July 2003.

[10] M.K. Smit and C. van Dam, PHASAR-based WDM devices: Principles, design and applications , *IEEE Journal of Selected Topics on Quantum Electronics,* vol. 2, no. 2, pp. 236-250, June 1996.

[11] G. Semaan and P. Humblet, Timing and dispersion in WDM optical star networks , In *Proceedings of INFOCOM'93*, pp. 573-577, San Francisco, CA, USA, 1993.

[12] C.L. Liu and J.W. Layland, Scheduling algorithms for multiprogramming in hard real-time traffic environments , *Journal of the Association for Computing Machinery*, vol. 20, no. 1, Jan.1973.

[13] J.A. Stankovic, M. Spuri, K. Ramamritham, G.C. Buttazzo, *Deadline Scheduling for Real-Time Systems - EDF and Related Algorithms*. Kluwer Academic Publishers, 1998.