Design of a Medium Access Control Protocol for a WDMA/TDMA Photonic Ring Network

J. Fransson, M. Johansson, M. Roughan
Software Engineering Research Centre, RMIT University,
Level 2, 723 Swanston St, Carlton, Vic 3053, Australia

L. Andrew[1]
Department of Electrical and Electronic Engineering,
The University of Melbourne, Parkville, Victoria 3052, Australia.

M.A. Summerfield
Australian Photonics Cooperative Research Centre, Department of Electrical and Electronic Engineering,
The University of Melbourne, Parkville, Victoria 3052, Australia.

Abstract— **We present a new medium access control protocol for a photonic ring network based on a combination of wavelength- and time-division multiple access. The protocol can achieve a maximum throughput in excess of 90%, though simulations of this protocol show that it has a practical limit of around 70% throughput. By comparison, the alternative proposed protocol, slotted ALOHA, has a maximum throughput of around 36% – almost half that of our protocol.**

I. Introduction

Wavelength Division Multiplexing (WDM) is rapidly becoming the technology of choice for high-capacity optical networks because it provides a graceful upgrade path for accessing the enormous fibre bandwidth, as well as advantages in terms of scalability and network management [1]. A great deal of attention has been focused on demonstrations of long-distance WDM transmission and wide-area transport networks in which the present high costs of WDM components are justified. Less attention has been devoted to smaller-scale local, campus and metropolitan area networks (LANs, CANs and MANs) in which the requirements are somewhat different. In smaller networks, key features are flexibility (of bandwidth use and configuration), compatibility with existing networks and protocols, and low cost.

MAWSON (the Metropolitan Area Wavelength Switched Optical Network) is a WDM data network demonstrator currently under development within the Australian Photonics Cooperative Research Centre [2]. The architecture and design of the MAWSON network focus upon the requirements of LANs: low cost, flexibility, robustness, and ease of use and management. However, since single-mode fibre and single-frequency transmitters are used, the geographic area which could be serviced by a single MAWSON network is potentially very large (even without optical amplifiers). Thus the principles demonstrated by the prototype network may be applicable to a range of networks from local to metropolitan scale.

The architecture of MAWSON as described below prevents most standard Medium Access Control (MAC) protocols from being used. The only standard protocol which has been proposed as an alternative is slotted ALOHA which has poor performance. This paper describes the design and performance of a MAC protocol specifically designed for MAWSON.

This paper is structured as follows. Section II describes the MAWSON network architecture. Section III describes the proposed MAC protocol. Section IV discusses the need for synchronization between wavelength channels in the network, and the limitations this imposes. Section V describes the performance of the protocol, both the bounds imposed by the protocols overheads, and the simulated performance. Section VI then concludes the paper.

[1]Formerly of CATT, RMIT University, 124 Latrobe St, Melbourne, Victoria 3000, Australia.

II. Network Architecture

A four-node, four-wavelength MAWSON ring is illustrated in *Figure 1*. The network infrastructure is a passive optical network (PON) consisting of a fibre-optic ring with WDM add-drop multiplexers (ADM) located at each access point. Each ADM drops a single, fixed wavelength, but allows any number of wavelengths to be added simultaneously. In the demonstration network the ADMs are simple passive optical components which are illustrated in *Figure 2*.

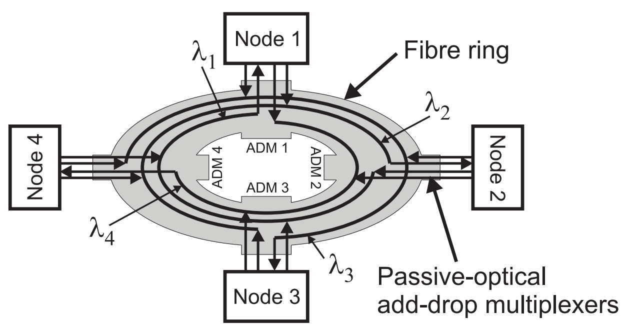

Fig. 1. Four-node MAWSON ring.

The passive ADM is based upon an optical circulator, a fibre grating and a fibre coupler. This structure has previously been employed in a laboratory demonstration of a four-channel spectrum-sliced WDM network [3]. As shown in *Figure 2*, all input wavelengths initially pass from port 1 to port 2 of the circulator. The fibre grating reflects all light within a fixed bandwidth around wavelength λ_i, so that a signal at that wavelength is reflected back to port 2 of the circulator and emerges at port 3. Port 3 is thus the ADM drop port. Signals at any wavelength(s) may be added to the ring via the coupler, which would typically have an asymmetric coupling ratio in order to present minimal loss to the "through" channels, at the expense of additional insertion loss for the added channel(s).

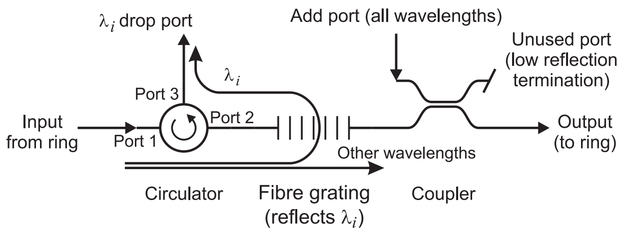

Fig. 2. The passive add-drop multiplexer.

A node will only receive information which was transmitted on the dropped wavelength. It is assumed that each node is equipped with a WDM laser-array transmitter (see, for example, [4]) capable of transmitting at all wavelengths used in the network. Such a transmitter is not only capable of switching very rapidly between wavelengths, but also of broadcasting common data on a number of wavelengths simultaneously. It is not considered necessary or desirable for a node to be able to send different data on different wavelengths at the same time.

Since each node receives on only one wavelength, multiple source nodes wishing to transmit to a single destination node must do so at different times to avoid data collisions, *i.e.* time-division multiple access (TDMA) is required in addition to WDMA. The structure of the TDMA is linked to the MAC protocol which is described in Section III below.

A. Features and Advantages of MAWSON

The most important requirement for any practical LAN technology is that it be low-cost. WDM technologies in general do not meet this requirement at this time, however in designing MAWSON we have attempted to select components with good potential for cost reduction over a five-year time-frame. Our design does not depend upon any continuously-tunable transmitters or receivers. Each node requires only one modulator, and associated high-speed electronics, independent of the number of wavelengths in the network. Since the ADM is part of the network infrastructure, no optical filters are required in the nodes. Thus, given a standardized set of wavelengths and a suitable laser array device, the hardware in all nodes is identical, supporting mass-production and simplified inventory handling. We have elected not to provide a channel-sensing capability at each node, such as is assumed in [5], to reduce costs. We have chosen instead to develop access protocols which do not depend upon such a facility.

MAWSON has a number of other features desirable in LAN applications. It is extremely flexible and robust: unlike switching technologies (e.g. switched Ethernet, ATM) it is not critically dependent upon correct operation of any one element; and unlike current ring-based networks, a single node failure should not disrupt the entire network. Nodes in the MAWSON network only process information at the rate of a single channel; no electronics operating at the aggregate rate are required. We have designed a configuration protocol (not described in this paper) which allows MAWSON nodes to automatically determine the network topology, including node/wavelength mappings and propagation delays. We plan to extend this protocol to support automatic fault isolation. Thus MAWSON will be capable of self-configuration, and be simple to manage and maintain.

III. The Request/Allocation Protocol

Since a transmitter cannot sense the transmission channel, it cannot *a priori* avoid transmitting when another node in the ring is transmitting, leading to a receiver collision. The retransmission cost when a collision occurs leads to a reduction in performance which can be seen in systems which allow such collisions such as ALOHA, or slotted ALOHA [6], with maximum throughputs of 18% and 36%, respectively.

The aim of MAC protocols is to provide fair and efficient access to the medium (regardless of scale), without introducing excessive loss or delay.

The specific properties of MAWSON prevent the use of protocols such as Carrier Sense Multiple Access with Collision Detection (CSMA/CD) as used in Ethernet. The carrier cannot be sensed to prevent collisions, nor can collisions be detected as they occur, but must be detected by higher layer protocols. This property is more characteristic of long delay networks (such as satellite networks) where the intrinsic delay makes carrier sensing impractical.

The protocol proposed here is a collision free protocol which uses allocation by the receiver. Each node which wishes to transmit to a particular receiver must *request* bandwidth, and wait for an *allocation* before it can transmit. The protocol will thus be referred to as the *Request/Allocation Protocol* (RAP).

In order to implement RAP, time is divided into uniform slots. The slot structure is shown in *Figure 3.* The slot is divided into a header and a data section; the header section starts with a minislot used for clock synchronization, and then contains $N-1$ request/allocation (R/A) minislots for an N node network, while the data section is divided into M Data MiniSlots (DMSs). There is one R/A minislot preallocated to each node that may transmit on the channel — they are preallocated to avoid collisions. R/A minislot $i \neq j$ on wavelength j is used to request DMSs for use by node i on wavelength j, and to allocate DMSs to node j on wavelength i.

Figure 3 shows the exact setup of the slot for one channel of an N node network. The slot structure on each channel is identical, and synchronized, although

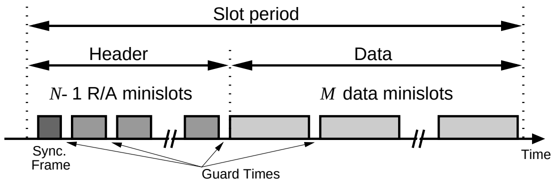

Fig. 3. The structure of the slots. The slot begins with a header section which contains a synchronization minislot and $N-1$ R/A minislots. The second part of the slot is a data section divided into M data minislots.

the R/A minislots on each channel are allocated to different nodes via a cyclic permutation so that no node need transmit on more than one channel at once during header transmission (illustrated in *Figure 5*).

Each R/A minislot consists of several fields: synchronization, Beginning Of Frame (BOF), request/allocation, and End Of Frame (EOF) fields, as shown in *Figure 4*. The minislots are separated by guard time to take into account timing errors between nodes. In a system with M DMSs a node may request from 0 to M DMSs and therefore $\lceil \log_2(M+1) \rceil$ bits are required in the request field, where $\lceil x \rceil$ is the smallest integer not less than x. Furthermore, the ith allocation bit indicates to a node that it has been allocated the ith of the DMSs and so M bits are required to allow separate allocation of all of the DMSs. Thus each R/A minislot requires

$$N_{R/A} = N_{\text{gu}} + N_{\text{syn}} + N_{\text{bof}} + \lceil \log_2(M+1) \rceil + M + N_{\text{eof}}$$

bits. This is the overhead of the RAP protocol.

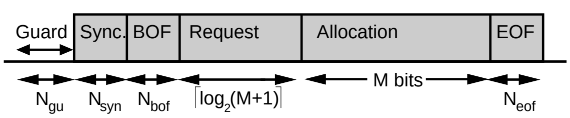

Fig. 4. The structure of a minislot. The slot begins with a preamble for synchronization purposes, followed by a BOF indicator. Next the frame contains $\lceil \log_2(M+1) \rceil$ bits for requests, then M bits for allocations. The frame ends with an EOF indicator which may include a cyclic redundancy check if required.

Each node maintains separate buffers for each possible destination node, and sends a request when the buffer contains packets to be transmitted. The node requests enough DMSs (up to the maximum M) to transmit the packets currently stored in the buffer.

When a node receives requests it performs allocations in a round robin, *i.e.* it cyclically allocates the DMSs to each node *in sequence* until either all of the requests or all of the DMSs are allocated. The last node to receive an allocation is recorded and the allocations for the next slot begin at the next node in the sequence, ensuring fair allocation of the resources, even under uneven load or overload conditions.

The requests must be processed, and so the allocation occurs in the slot following the one in which the request was received, and transmission begins in the slot following allocation. Thus the Request/Allocation/Transmission (R/A/T) procedure is a three stage procedure taking place over three slots.

IV. Synchronization

The protocol requires synchronization of the slots on all channels to avoid transmission collisions. Synchronization imposes a significant limitation on the protocol. Either the slot period is no longer than the total propagation time of light on the ring, or substantial guard times must be included around headers and between DMSs.

The most natural and efficient solution is when the slot period equals the propagation time, then the slot structure *fits* into the ring in the sense that the entire slot may be notionally present within the ring, and the slots may be placed back-to-back. The slots on each channel may be considered to be something like a ribbon rotating around the ring, onto which each node may transmit in allocated minislots — for example see *Figure 5*.

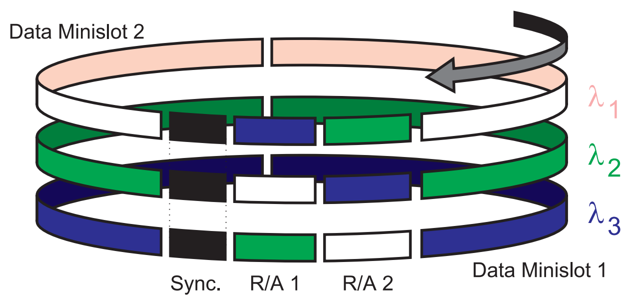

Fig. 5. The slots on the data channels may be thought of as ribbons with the locations of minislots marked on them. The slots on each channel are synchronized. The shade of each R/A minislot shows which node this minislot is allocated to – note that these cyclically permute so that a node only needs to place requests and allocations onto one channel at a time.

Figure 6 (a) shows a time-space diagram of the slot-structure for a three node network with the slot period equal to the propagation time, *i.e.* it shows how the slot moves through time and space. Note that the diagram is really an unwrapped cylinder – the top edge of the diagram joins to the bottom – the header forming a continuous spiral around the cylinder.

In contrast, if the slot period is longer than the propagation time then the slot cannot fit into the ring. Using the analogy of *Figure 5* the ribbon is now longer than the ring on which it sits. In some sense only part of the ribbon may be on the ring at any one time. We might try to have a spool of ribbon leaving the ring at some point, for instance at a node in the network, but this is contrary to the principle of the MAWSON ring which is intended to be a passive network. A node would have to have significant active equipment to remove all of the channels from the network and delay these signals for some period of time.

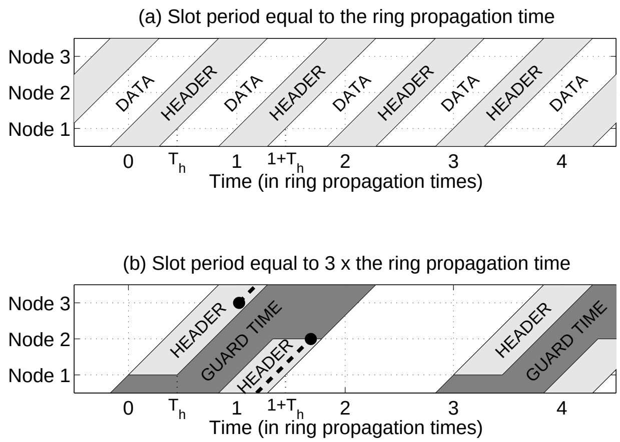

Fig. 6. A time-space diagram of the slot structure. In (a) a slot period equals the ring propagation time. Note that the diagram is really an unwrapped cylinder, so that the headers actually form an unbroken spiral around the cylinder. T_h is the time to transmit the header. In (b) the slot period is longer than the ring propagation time. In this case the headers can no longer form an unbroken spiral. The dashed line illustrates how Node 3 would transmit a request or allocation to Node 2.

A feasible alternative would be to have a separate notional spool for each channel, located at the node which receives that channel. This requires no additional hardware since each ADM already removes the signal on its own receive channel. However, this complicates synchronization of the channels substantially. In particular, a wide guard time must be inserted around the header and between DMSs to flush unwanted data from the ring.

Figure 6 (b) illustrates the guard time. The figure shows the location (in time and space) of the header, but it should be noted that this represents the header on all three channels, not all of which exist over the extent of the header in the diagram. For instance, light traveling to Node 3 along Channel 3 is removed from the network by the ADM at Node 3, and therefore the header on Channel 3 is removed at Node 3. However in order for Node 3 to transmit requests to Node 2, the header on Channel 2 continues from Node 3 until it reaches Node 2. The dashed line in the figure illustrates the transmission of a request or allocation from Node 3 to Node 2 along Channel 2.

The wide guard time around the header is required to prevent receiver collisions. For instance, if Node 1 began transmitting data to Node 2 on Channel 2 at time $t \in (T_h, 1)$, then Node 2 would receive the data before receiving its header. Worse yet, if Node 1 continued transmission of data past time 1.0, there would be a receiver collision between the data and the header on Channel 2. Introducing a guard time of the propagation delay plus the header transmission time in the ring isolates the header from the data section. The header transmission time must still be less than the ring propagation time because the individual minislots in the header are not similarly isolated.

The DMSs are also separated by a guard time to isolate them from the other DMSs. However, the transmitting node does not change during one of the DMSs, so that the minislot may be as long as desired.

Introducing such wide guard times is inefficient, and so we only consider the case where the slot period is the same as the propagation delay in the ring. The slot period may be less than one ring propagation time, but this simply wastes capacity because the R/A/T cycle requires three ring propagation times. Fitting two slots into the ring length leads to extra overhead without reducing the minimum delay characteristics of the system. It is preferable to have multiple DMSs within a single slot, to allow more than one node to transmit during each R/A/T cycle.

The limitation that the slot be the same length as the ring means that a long ring is needed to attain an efficient protocol. For instance a 1 km ring at 100 Mbits/s has a propagation delay corresponding to a ~500 bit slot — a slot barely sufficient to transmit all of the header information, and therefore with very little room to transmit data. There will also be an absolute minimum length for the ring determined by the length of the header, which is roughly proportional to the number of nodes in the system.

We will discuss, in the following section, the minimum practical lengths of the ring for the protocol to be efficient. Extra ring length can be easily and cheaply added in the form of fibre on a spool.

It should also be noted that the limitation described above is reduced for higher speed networks. At 1 Gbit/s the slot period is reduced by an order of magnitude, and the minimum ring length is similarly reduced, making the limitation almost negligible.

V. Performance

This section describes the performance of the protocol. We assume here that the slot period is equal to the propagation delay in the ring. The performance described below assumes 100 Mbit/s transmission, but there is one transmission channel per node, and therefore the transmission capacity of the network scales directly with the number of nodes in the network, *i.e.* the network capacity is 100 Mbits per second per node, full duplex.

The allocation procedure is inherently fair, and so the performance measures we will concentrate on are scalability, throughput, delay and loss. The scalability and maximum throughput of the protocol are described immediately below. Delay and loss measured using simulation are presented in Section V-B, where the protocol is demonstrated to have good performance under loads of up to 70% of the maximum.

A. Performance Bounds

This section describes the overhead introduced by RAP, and therefore the upper bounds on the throughput of MAWSON. The header requires

$$N_{\text{overhead}} = S + (N-1) \times N_{R/A}$$

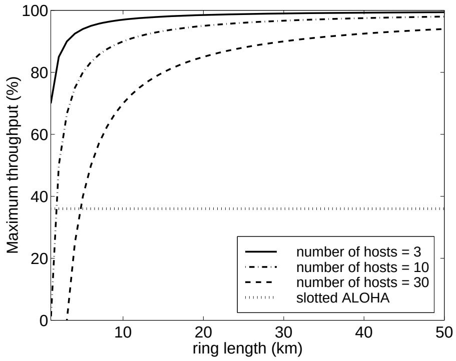

Fig. 7. The maximum attainable throughput for different ring lengths for 3, 10 and 30 node networks, and for a slotted ALOHA system.

bits where there are N nodes, M DMSs, and S is the number of bits used in slot synchronization. We make the following conservative assumptions about the MAWSON network: $N_{\text{gu}} + N_{\text{syn}} + N_{\text{bof}} + N_{\text{eof}} = 40$, so that

$$N_{R/A} \quad = \quad 40 + \lceil \log_2(M+1) \rceil + M.$$

and $S = 50$ bits.

In the following we assume $M = 7$ as this is a reasonable choice for the number of DMSs. A low number such as 1 or 2 leads to delays because a node may have to wait several cycles before it receives an allocation, whilst a high number leads to a large R/A overhead. Choosing $M = 2^n - 1$ is logical, because the number of request bits is $\lceil \log_2(M+1) \rceil$. Note that in a real network the number of DMSs would ideally vary with the number of nodes in the network, but that we have kept this fixed here to illustrate the variation in overhead with different numbers of nodes.

Figure 7 shows the maximum attainable throughput for the above network for 3, 10 and 30 nodes for different ring lengths. As expected the maximum throughput increases with increasing ring length as the R/A overhead decreases in proportion to the total slot length. It is noteworthy that with approximately 1km of fibre per node the maximum throughput is approximately 90%. In contrast the maximum throughput for slotted ALOHA is approximately 36%, regardless of ring length.

B. *Simulated Performance*

The ring network was simulated under a variety of scenarios, and ring designs. The simulated traffic was made up of Poisson packet arrivals with packet sizes drawn from a uniform distribution with minimum packet size 160 bits and maximum packet size 9000 bits. Other simulations confirmed similar results with other packet size distributions (for example one drawn from the sizes of Ethernet packets on our own Ethernet). The packets' destinations were chosen randomly. The simulations used finite buffers of size 256 kbits per node, per channel.

An issue of interest is how many DMSs should be used for a particular number of nodes. *Figure 8* shows average waiting times (the time between a packet arriving at a node, and being transmitted onto the network), and loss probabilities as a function of the number of DMSs used, for 3 node rings of various ring lengths. Note that for a 10 km ring, 2 data minislots (1 per possible transmitter) is a quite satisfactory solution, both in terms of delay and loss. However, for a 10 km ring, as the overhead increases with more DMSs the loss and delay increase.

Figure 9 shows the waiting time distributions for a 20km ring with 3 DMSs under varying loads. The finite buffer imposes an upper limit on buffering delay which has little effect for loads of 70 Mbps/node or less, but which has a significant effect for loads of 75 Mbps/node or greater. Thus the maximum practical throughput of the network is around 70% rather than the >90% bound, but it is normal in networks not to be able to achieve the theoretical upper bound.

Finally we show that the protocol is scalable. *Figure 10* shows the scaling behavior of the protocol for a fixed ring length of 20 km. For larger numbers of nodes the knee occurs at a lower throughput, indicating a gradual reduction in capacity which would typically be overcome by an increase in ring length.

As stated above the protocol is inherently fair because of the allocation procedure, and simulation confirmed this conclusion, though space limitations prevent us from presenting the simulation results.

C. *Discussion of Results*

Longer rings improve throughput, but the tradeoff is that longer rings introduce longer delays, an effect enhanced because the request allocation cycle introduces delays in addition to buffering and transmission delays. This additional delay is twice the time for a signal to traverse the ring (once for the

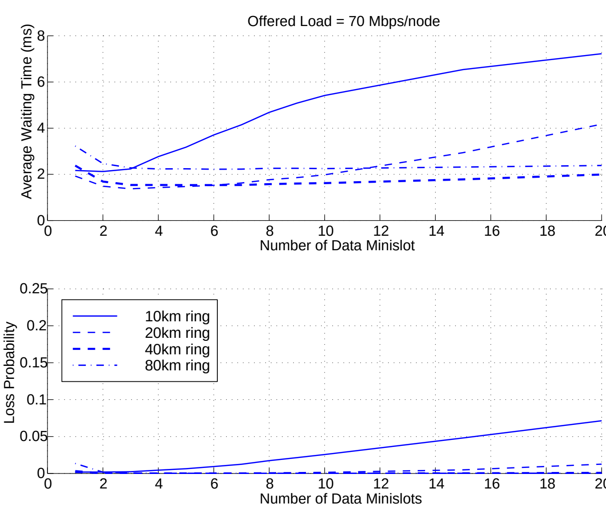

Fig. 8. The waiting time and probability of loss vs. number of DMSs for a three node network. Note that 95% confidence intervals around the results are $\leq \pm 0.25$ ms for the waiting times, and $\leq \pm 0.0025$ for the loss probabilities.

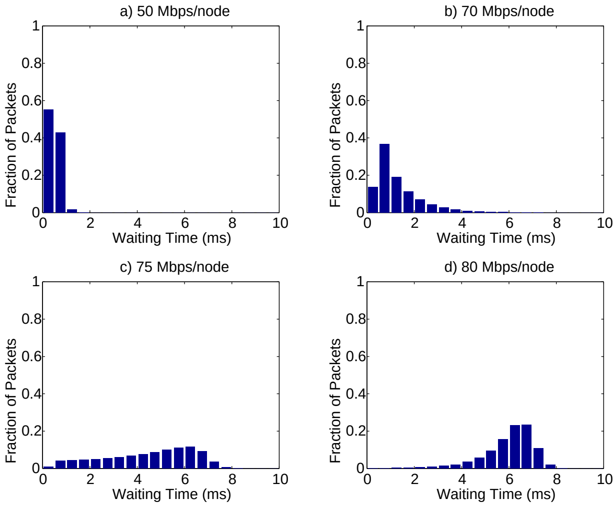

Fig. 9. Delay distribution, fibre length 20km, 3 DMSs.

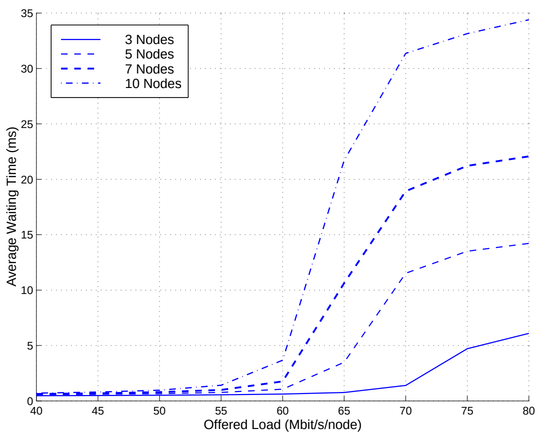

Fig. 10. Scalability of the network for a 20km ring with three DMSs.

request, once for allocation). This leads to extra delays of 1.0e-05 seconds per km for a packet arriving at an empty buffer, which is insignificant compared with the buffering delays at the nodes.

Even this minimum delay could be reduced by making cyclic allocation of any unallocated DMSs so that a new packet arriving at a node may be immediately transmitted in one of these extra minislots.

Furthermore, using prioritized queues, nodes can ensure an upper bound on delays for delay-sensitive traffic. Any protocol such as ALOHA, or CSMA/CD that allows collisions, and therefore retransmissions, cannot guarantee such an upper bound.

VI. Conclusion

A MAC protocol for the MAWSON network architecture has been presented. The protocol is efficient – it introduces around 10% overhead. In practise the protocol can provide throughputs of up to 70% of the maximum bit rate with no substantial loss, and introduces delays on the order of only a few milliseconds. Furthermore, the protocol has a fixed upper bound on delay, and is intrinsically fair.

Note that although 70% throughput of a 100 Mbit/s link does not sound impressive by modern standards, there is one channel per node which represents 70 Mbits/s full duplex per node. The assumptions used in the simulations were conservative – the actual MAWSON network will probably have less overhead, and therefore higher throughput.

The protocol has other desirable properties such as:

- **Bit rate transparency:** two nodes with higher speed transmission capabilities may transmit at this higher speed during the DMSs with no penalty.
- **Link layer transparency:** any protocol such as the Logical Link Control protocol of IEEE 802.2 may be used to transmit the packets within the DMSs.
- **Robustness to node failure:** a node failure does not effect transmission between other pairs of nodes.

There are a number of other features which could be easily incorporated into the above protocol if the specific need arises, for example:

- Semi-permanent DMS reservation could provide virtual circuits for high-speed real-time services.
- One or more DMSs could operate on an ALOHA basis, *i.e.* any node could transmit during that minislot, with the understanding that a collision would result in the loss of the packet. This might be ideal for traffic which is sensitive to delay but can tolerate some loss, such as uncompressed voice traffic.
- Acknowledgments of the number of correctly received packets could be built into the R/A minislots, as a fourth phase of the R/A/T cycle.

The features listed above may be included in a later version of MAWSON if there is a perceived need.

Acknowledgments

M.Roughan gratefully acknowledges the support of Ericsson Australia.

References

[1] *Special Issue on Multiwavelength Technology and Networks, Journal of Lightwave Technology*, vol. 14, no. 6, June 1996.

[2] M.A. Summerfield, "Mawson: A Metropolitan Area Wavelength Switched Optical Network," in *APCC'97: The Third Asia-Pacific Conference on Communications*, Sydney, Australia, November 1997, pp. 327–331.

[3] R.D.T. Lauder, J.M. Badcock, W.T.Holloway, and D.D. Sampson, "Experimental demonstration of a WDM ring network employing a hybrid EDFA as a shared multiwavelength source," *Technical Digest of Optical Amplifiers and their Applications*, pp. 102–105, Paper TuB6, July 1997, Victoria, B.C., Canada.

[4] T.-P. Lee, C.E.Zah, R.Bhat, W.C.Young, B.Pathak, F.Favire, P.S.D. Lin, N.C. Andreadakis, C.Caneau, A.W. Rahjel, M.Koza, J.K. Gamelin, L.Curtis, D.D.Mahoney, and A.Lepore, "Multiwavelength DFB laser array transmitters for ONTC reconfigurable optical network testbed," *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 967–976, June 1996.

[5] Marco Ajmone Marsan, Andrea Bianco, Emilio Leonardi, Michela Meo, and Fabio Neri, "MAC protocols and fairness control in WDM multirings with tunable transmitters and fixed receivers," *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 1230–1244, June 1996.

[6] L. Kleinrock, *Queueing Systems*, vol. I: Theory, John Wiley and Sons, 1975.