Halmstad University Post-Print

Fibre-ribbon pipeline ring network with distributed global deadline scheduling and deterministic user services

Carl Bergenhem, Magnus Jonsson and Jörgen Olsson

N.B.: When citing this work, cite the original article.

©2001 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Bergenhem C, Jonsson M, Olsson J. Fibre-ribbon pipeline ring network with distributed global deadline scheduling and deterministic user services. In: International Conference on Parallel Processing Workshops, 2001. Los Alamitos: IEEE Computer Society; 2001. p. 311-318.

DOI: <http://dx.doi.org/10.1109/ICPPW.2001.951967>

Copyright: IEEE

Post-Print available at: Halmstad University DiVA
<http://urn.kb.se/resolve?urn=urn:nbn:se:hh:diva-2735>

Fibre-Ribbon Pipeline Ring Network with Distributed Global Deadline Scheduling and Deterministic User Services

Carl Bergenhem
Computers and
Communications lab,
Halmstad University,
Halmstad, Sweden.
carl.bergenhem@ide.hh.se

Magnus Jonsson
Computers and
Communications lab,
Halmstad University,
Halmstad, Sweden.
magnus.jonsson@ide.hh.se

Jörgen Olsson
Volvo Technological
Development Corporation,
Göteborg, Sweden.
jool@vtd.volvo.se

Abstract

This paper introduces a novel, fair medium access protocol for a pipelined optical ring network. The protocol provides global deadline scheduling of packets. Requests for sending packets are sent by the nodes in the network to a master node. The master uses the deadline information in the requests to determine which packet is most urgent. Arbitration is done in two steps, the collection and distribution phases. The protocol is therefore called two-cycle medium access (TCMA). The network is best suited for LANs and SANs (system area networks) such as a high speed network in a cluster of computers or in an embedded parallel computer. Offered services in this network include best effort messages, guarantee seeking messages, real-time virtual channels, functions used in parallel processing. These are possible without additional higher level protocols. A simulation analysis of the network with the protocol is presented. Further analysis shows minimum slot length and fairness of the protocol.

1. Introduction

The contribution put forward by this paper is a novel medium access protocol that uses the deadline information of individual packets, queued for sending in each node, to make decisions, in a master node, about who gets to send. The new protocol may be used with a previously presented network topology; the control channel based fibre ribbon pipeline ring (CC-FPR) network [1]. Simulations of the protocol prove its validity.

The proposed medium access protocol provides the user with a service for sending best effort messages,

which are globally deadline scheduled. The global deadline scheduling is a mechanism that is built into the medium access protocol. No further software in upper layers is required for this service. Other networks may have upper layer protocols added to them to give them better characteristics for real-time traffic, but it is hard to achieve fine deadline granularity by using upper layer protocols.

Real-time services in the form of best effort messages, as mentioned above, guarantee seeking messages, and real-time virtual channels (RTVC) are supported for single destination, multicast and broadcast transmission by the network. There is also a service for non real time messages. The network also provides services for parallel and distributed computer systems such as short messages, barrier synchronisation, and global reduction. Support for reliable transmission service (flow control and packet acknowledgement) is also provided as an intrinsic part of the network [3].

The network with the proposed protocol is best suited for LANs and SANs (system area networks) where the number of nodes and network length is relatively small. This is important since the propagation delay adversely affects the medium access protocol. Examples of suitable applications are embedded systems (e.g., for use as an interconnection network in a radar signal processing system) and cluster computing.

Motorola OPTOBUS™ bi-directional links with ten fibres per direction are used but the links are arranged in a unidirectional ring architecture where only $\lceil N/2 \rceil$ bi-directional links are needed to close a ring of N nodes. Fibre-ribbon links offering an aggregated bit rate of several Gbits/s have reached the market [4]. The increasingly good price/performance ratio for fibre-ribbon links indicates a great success potential for the proposed type of networks.

The physical ring network is divided into three rings or channels (see Figure 1). For each fibre ribbon link, eight fibres carry data, one fibre is used to clock the data, byte by byte, and one is used for the control channel. Access is divided into slots like in an ordinary TDMA (Time Division Multiple Access) network. The control channel ring is dedicated for bit-serial transmission of control packets, which are used for the arbitration of data transmission in each slot. The clock signal on the dedicated clock fibre, which is used to clock data, also clocks each bit in the control packets. Separate and dedicated clock- and control fibres simplify the transceiver hardware implementation in that no clock recovery circuitry is needed [5]. The control channel is also used for the implementation of low-level support for barrier-synchronisation, global reduction, and reliable transmission.

The ring can dynamically (for each slot) be partitioned into segments to obtain a pipeline optical ring network [2]. Several transmissions can be performed simultaneously through spatial bandwidth reuse, thus achieving an aggregated throughput higher than the single-link bit rate (see Figure 2 for an example). Even simultaneous multicast transmissions are possible as long as multicast segments do not overlap. Although simultaneous transmissions are possible in the network because of spatial reuse, each node can only transmit one packet at a time.

A disadvantage with the CC-FPR protocol presented in [1] is that a node only considers the time constraints of packets that are queued in it, and not in downstream nodes. As an example (see Figure 2), Node 1 decides that it will send and books Links 1 and 2, regardless of what Node 2 may have to send. This means that packets with very tight deadlines may miss their deadlines. The novel network presented here does not suffer from this problem.

The rest of the paper is organised as follows. Section 2 presents the novel medium access protocol. Section 3

Figure 1: A Control Channel based Fiber Ribbon Pipeline Ring network.

presents the user services available in the network, while Section 4 examines some aspects of implementation such as minimum slot length. Section 5 describes the simulations carried out on best effort traffic. Finally conclusions are presented in Section 6.

2. Two-cycle medium access protocol

The two phases of medium access for the proposed protocol are collection phase and distribution phase (see Figure 3). Therefore it is referred to as the two-cycle medium access protocol (TCMA protocol). As can be seen, the protocol is time division multiplexed to share access between nodes. The basic time unit is called a slot and the minimum slot size is analysed in Section 4. Slots are organised into cycles with a predefined number of slots. The number of slots is chosen so that each node is master at least once per cycle (in this section assumed to be one per cycle).

The role of being network master is cycled around the ring. Thus all nodes are identical. The role as master is passed on to the next downstream node at the end of the slot. Every node detects when the clock signal is interrupted at the end of the slot and increments a counter which determines the next master.

There are two types of TCMA control packets, which are used in each of the two phases (see Figure 4). A complete collection phase packet will contain a start bit and a total of $N - 1$ requests that are added one by one by each node. The master receives its own request internally. Each request consists of three fields. The "prio"-field contains the priority level of the request which is further described below. Nodes use the link reservation and destination fields to indicate destination node(s) and which links must be traversed to reach the

Figure 2: Example where Node 1 sends a single-destination packet to Node 3, and Node 4 sends a multicast packet to Node 5 and Node 1.

destination node. For the link reservation field, each bit corresponds to one link and tells whether the link is reserved (1) or not (0). The destination field has one bit for each node in a corresponding way. Since a node may write several destination nodes into the destination field, multicast or broadcast is possible.

In the distribution phase packet, the "result of requests"-field contains the outcome of each node's request. This is the only field, in this phase, which contains network arbitration information. The other fields are used for services such as reliable transmission ("ACK/NACK"- and "flow control" fields) and global reduction (the "Extra information"-field). These are described later.

In the collection phase, the current master initiates by generating an empty packet, with a start bit only, and transmits it on the control channel. Each node appends its own request to this packet as it passes and then passes the packet on to the next node. The master receives the complete request packet (see Figure 4) and may then process this to determine which requests will be fulfilled.

The time until deadline (referred to as laxity) of a packet is mapped, with a certain function, to be expressed within the four-bit limitation of the current version of TCMA's priority field. A shorter laxity of the packet implies a higher priority of the request. The result of the mapping is written to the priority field. One priority level is reserved (15 in the proposed implementation of the protocol) and used by a node to indicate that it does not have a request. If so, the node signals this to the master by using the reserved priority level and writes zeros in the other fields of the request packet.

Request priority is a central mechanism of the TCMA protocol. A wider priority field, in each request (see Figure 4), is possible and would provide higher resolution of priority. This would probably have a slight advantageous affect on performance. Further evaluation

Figure 3: The two phases, collection and distribution, of the TCMA protocol. Notice that the network arbitration information, for data in slot $N+1$, is sent in the previous slot, slot N . Observe that the lengths of the phases, and placement in time, in the diagram are not to scale.

is out of the scope of this paper. Two mappings between deadline and priority, logarithmic and linear (see Figure 5), have been simulated. Results show a negligible difference in performance of throughput, packet-loss, and latency. Further evaluation of how the performance is affected by different mappings is therefore put beyond the scope of this paper. For the simulations presented in Section 5, logarithmic mapping is used.*This mapping gives higher resolution of laxity, the closer to its deadline a packet gets.

Packets queued locally in nodes are sorted by laxity and distance and each node selects its most urgent packet as the request. In the case that there are several packets that are equally urgent, the packet that is destined furthest and possible to transmit in the next slot is selected. Nodes will not request transmission of a packet that will pass the master since the clock signal is interrupted there and data cannot pass. A node will only make a request that may be possible to fulfil regarding RTVCs (see sections 3.5) in the own or other nodes that would use links in the path of the packet that the node would want to send. This implies that a request will only be rejected if requests from other nodes are more urgent. Slots belonging to RTVCs do not need to be "requested" since all nodes know which slots are already reserved.

When the completed collection phase packet arrives back at the master, the requests are processed. There can only be N requests in the master, as each node gets to send one request per slot. The list of requests is sorted in the same way as the local queues. The master traverses the list, starting with the request with highest priority (closest to deadline) and then tries to fulfil as many of the N requests as possible. In case of priority ties, the request with the largest distance to its destination is chosen. If there still is a tie, then requests from upstream nodes (closer to the master) have priority over other request.

When the master has scheduled the requests, it distributes the result to all nodes in the distribution phase. In this phase, the master node, and only the master

Figure 4: Contents of the TCM control packets. Notice the possibility in the distribution phase packet to send information other than for medium access arbitration.

node, has the possibility to use the other fields in the distribution phase packet, such as sending acknowledges for packets sent during the previous slots. For further explanation of this, see Section 3.4. When all nodes have received the results of the request, each node is ready for the beginning of the next slot where data may be transmitted. A request was granted if the corresponding bit in the “request result field” of the distribution phase packet contains a “1”.

The advantage of the TCMA protocol is that the deadline requirements for packets from all nodes are taken into account and, for one packet from each node, thus considered at a global level. Since the global queue that is used for deciding which packets will be sent, consists of one request per node (in the current implementation), there will be situations where lower priority packets are sent even though there are higher priority packets queued at another node. This situation is referred to as a priority inversion and is due to the fact that the global queue of requests does not have the “complete picture” of the individual queues in the nodes. Since the priority for a message is dynamically increased as the laxity decreases, the TCMA protocol implements an approximation of the optimal “earliest deadline first” algorithm. The limitation is, as stated above, that only one message from each node is considered in each slot. However, for each node, it is always the most urgent message that is considered.

3. User services

The user services described below are: best effort messages (Section 3.1), non real time messages (Section 3.2), guarantee seeking messages (Section 3.3), reliable transmission (Section 3.4), real time virtual channels (Section 3.5), barrier synchronisation (Section 3.6) and global reduction (Section 3.7).

Figure 5: Two different deadline-to-priority mappings that were tested. For the linear transformation, deadlines longer than 14 slots are all mapped to priority level 14.

3.1. Best effort messages

The TCMA protocol supports best effort messages [6] as it is. With this service, messages at nodes are accepted for transmission without any guarantees to the user except that the network will try to meet all deadlines. Messages are queued in the node according to deadline and transmitted according to deadline order in a global queue generated by the TCMA protocol. The message may fail, e.g., due to congestion in the network. When the deadline of a message expires the user may be notified. Usually the user will choose to have the message removed from the queue but also has an option to keep the message queued despite its expired deadline. This type of communication, best effort messages, is simulated in Section 5.

3.2. Non real time messages

The non real time message service is suitable for users that do not require real time constraints at all. Such a user could be a bulk file transfer. The messages do not contain any timing requirements and are queued according to destination and transmitted in first come first served order. Non real time messages are similar to best effort messages but have lower priority and are only sent when there is no best effort message to be sent from the node.

3.3. Guarantee seeking messages

Guarantee seeking messages normally have hard timing constraints. If the communication system cannot guarantee the timing constraints of a guarantee-seeking message, the owner of the message should be aware of it immediately. In the proposed network, a guarantee is only given if enough deterministic bandwidth (slots) owned by the node, through an RTVC (See Section 3.5), is free before the deadline of the message. In other words, an RTVC is first set up for the segments of the ring (possibly the whole ring). All slots booked for guarantee seeking messages are then tracked.

3.4. Low-level support for reliable transmission

The proposed network has low level support for reliable transmission [3]. Network control information, acknowledge/negative acknowledge and flow control, is sent in the control channel instead of in ordinary data packets. This results in less or no overhead in the data channel, i.e., better bandwidth utilisation. The field in the control packet named ACK/NACK contains bits that correspond to the N packets that may have been received by the current master during the previous N slots. The

ACK/NACK information is also sent when a node is master. If a packet was correctly received (correct checksum) a "1" is written in the position of the ACK/NACK field that corresponds to the slot that the packet was received. If a faulty packet or no packet was received, a "0" is written. All nodes must keep track of their transmissions and can therefore resolve the meaning of the bits in the ACK/NACK field. In this way the nodes can be notified if their packet was correctly received or has to be retransmitted. The latency for a node to send and receive ACK/NACK is bounded and deterministic which is desirable.

The $4(N - 1)$ bits in the flow control field relate to independent logical connections. Put simply, each node can have up to four logical connections in average, with low-level support for flow control, depending on the implementation. The master sets the bit corresponding to a logical connection to "1" if it is to be halted temporarily, else the bit is set to "0". This service may be combined as required with any of message sending services.

3.5. Real-time virtual channels

Logical connections with guaranteed bit rate and bounded latency can be realised in the network by using slot reserving. Such connections are referred to as RTVCs [6]. Either the whole ring is reserved for a specific node in a slot, or one or more segments of the ring are dedicated to some specific node(s), i.e., a single slot can be allocated for several RTVCs at the same time. Slots are organized into cycles with a predefined number of slots. Nodes keep track of the current slot index in the cycles. A slot that has been reserved for an RTVC, guarantees that transmission is possible every cycle, thus guaranteed bit rate. Several slots may be reserved for an RTVC in order to increase the guaranteed bit rate. Initially, when the use of RTVCs is enabled, each node has J non-reserved slots where it is master, giving a cycle length of $N \cdot J$ slots, where N is the number of nodes. In order to always have bandwidth for best effort traffic controlled by the TCMA protocol, only $J - 1$ slots are reserveable.

When a node wants to reserve a slot for an RTVC, it searches for slots where the required links are free, so allocation of a new segment can be done. First, the node's own slots are searched. If not enough slots could be allocated for the reservation, the search is continued in other nodes. In this case, the node broadcasts a packet containing a request to all other nodes to allocate the desired segment in their slots. The packet contains information about the links required and the amount of slots needed. Each node then checks if any of its own slots have the required free links. All nodes send a packet back to the requesting node to notify which slots,

if any, that have been allocated. When the requesting node has received the answers, it decides if it is satisfied with the number of allocated slots. If not, it sends a release packet. Otherwise, it can start using the reserved slots immediately. If so, the node notifies other nodes by broadcasts the details of the RTVC. It should also send a release packet if more slots than needed were allocated. A node wishing to set up an RTVC can thus "borrow" slots from other nodes. Internally in the nodes, each RTVC has its own queue for packets. All nodes have information about which slots are reserved, for RTVCs between pairs of nodes.

3.6. Barrier synchronisation

Barrier Synchronisation (BS) is an operation to control the flow of processes in a distributed processing system. A logical point in the control flow of an algorithm is defined, at which all processes in a processing group must arrive at before any of the processes in the group are allowed to proceed further. When, during execution, a BS point is encountered in the application program, the node broadcasts the encountered BS_id in the control packet when the node is master. In this way all nodes are notified that the node has reached the BS point. Nodes belonging to a different BS group can ignore the broadcast, but nodes belonging to the same group, i.e., has the same id, will make a note in an internal table. The control packet contains a field in which BS_id can be sent (see Figure 4). The id field contains 8 bits, which permits ids ranging from 1 to 255. When the field is zero no BS command is sent.

When a node participating in the BS group has received the correct BS_id from all the participants it knows that all the other nodes are at the same executing point and may proceed. The worst case latency, for a node that reaches the BS point until it can broadcast this to the other nodes, is one cycle. Clearly the implications of sending BS information in the control channel is both bounded latency and better bandwidth utilisation of the data channel. The whole BS mechanism is handled by the communication interface, transparent to the calling user processes.

In the description above, static allocation of barrier synchronization BS_ids is assumed. The programmer (or the compiler) allocates the required parameters for BS and GR off-line, before runtime. With minor adjustments, dynamic allocation is also possible but is not investigated further in this paper.

3.7. Global reduction

Global Reduction (GR) is similar to barrier synchronisation where data is collected from distributed processes when they signal their arrival at the

synchronisation point. A global operation, e.g., sum or product, is performed on the collected data so that only a single value is returned. At the end of the GR all participating nodes have access to the same data. As in the case of BS we assume that the programmer (or the compiler) statically allocates the necessary parameters, off-line, before runtime.

The GR command requires the following parameters: operation, length, data type, GR_id, and the ids of the participating nodes. The last two parameters are similar to the parameters for the BS command. The operation parameter tells the nodes what operation (sum, product, max, min, etc.) should be performed on the received data. The data type parameter indicates how the data field in the control packet (see Figure 4) is to be interpreted and the length parameter tells the length of the data field. In the studied case, the data field is 320 bits long and may facilitate the transfer of, e.g., up to five double precision floating point numbers (IEEE-754). Having multiple data items in the data field gives the opportunity to have, e.g., vector operations. Other data types may also be distributed. Except for the additional fields and the global function, the nodes treat GR commands in the same way as BS commands. Further on the same reasoning of performance advantages also holds for GR.

The type bit tells whether the control packet contains a BS or a GR command, and hence whether data is contained in the data field or not (see Figure 4). The data field is currently only used for data reduction, but may be used for, e.g., sending short messages.

Figure 6: The diagram shows the relation between total network length and minimum slot length, for three different numbers of nodes.

4. Implementation aspects

T_{tcma} is the time required to complete network arbitration according the TCMA protocol. This also sets the minimum possible slot length for the network. T_{tcma} scales for increasing network length and number of nodes as follows:

$$T_{tcma} = T_{collection} + T_p + T_{selection} + T_{distribution} \quad (1)$$

and is explained below. The master requires a finite time for processing the collected requests to select which requests may be sent. Part of this processing is sorting the incoming requests. This can be done as they arrive by checking them bit by bit and thus the sorting time is incorporated in the time for the collection phase, $T_{collection}$. When requests have been sorted, the list of requests has to be traversed to select which packet that may be sent. The time for this is denoted as $T_{selection}$ and is assumed to be $N \cdot 30$ ns. Propagation delay, T_p , is part of required arbitration time only once since the master depends on feedback, i.e., the requests from the other nodes only in the collection phase. L is the total length of the ring, while V is the propagation speed through the optical fibre and is assumed to be $2 \cdot 10^8$ m/s. The delay through each node (approximately 1 bit time per node) is neglected in the propagation delay. The total propagation delay around the ring is:

$$T_p = \frac{L}{V} \quad (2)$$

From Figure 4 one can see the size of the fields in the current implementation of the control packets which lead to:

$$T_{collection} = \frac{1 + (2N + 4)(N - 1)}{C} \quad (3)$$

and

Figure 7: Best effort packet latency vs. throughput for a varying number of nodes.

$$T_{distribution} = \frac{6N + 325}{C} \quad (4)$$

where N is the number of nodes and C is the bit rate of the links. For the Motorola Optobus, $C = 800$ Mb/s. The minimum slot length (T_{icma}) is plotted in Figure 6. As we can see from the figure, the minimum slot length increases with increased network length and number of nodes. This is just as can be expected from a network that requires feedback in its medium access protocol.

5. Simulation analysis

The simulation analysis is done by discrete time computer simulation. Networks of 8, 16, 32 and 64 nodes were simulated. Each best effort packet is given a relative deadline at generation which is decremented each time slot. The packet is queued until it is either sent successfully or, when deadline reaches zero, is deemed lost and removed from the queue. On request from the application a packet could be kept in the queue even after the deadline is passed but this is not simulated.

Some further assumptions for the simulations:

- Messages are one packet long and take one time slot to send. The term packet and message is therefore used synonymously.
- Uniform traffic is assumed, i.e., all nodes have equal probability of message generation and uniformly distributed destination addresses. This implies that, on average for an ideal ring network, it is theoretically possible to transmit two packets each slot, since the packet on average is destined "half way" around the ring, which is $N/2$ hops. However, this disregards protocol effects, which lowers the average utilisation, which we will see later. An example of protocol effect is that a node may not send past the master since the clock is interrupted

Figure 8: Total throughput of best effort packets vs. packet generation intensity.

there. The pipelined ring topology of the network suggests that it is very effectively utilised when traffic is mostly destined one hop to the nearest neighbour such as in some types of radar signal processing [7] [8]. If this special mode of traffic were simulated, which it is not, the effect would simply be higher throughput because of aggregation, i.e., more packets would be sent during one slot.

- Messages were generated according to a poisson process and all messages were of single destination type.
- The deadline of all best effort packets is set at generation to 800 slot times ahead, which would equate to a deadline of 4 ms with a slot time of 5 μ s (4kBytes per packet if Motorola Optobus is used).
- Physical effects, such as the propagation delay through the optical fibre, and the time required to detect the end of a slot, were assumed to be less than one slot time and are therefore neglected.
- Message latency is defined as the time elapsed from the moment a message is generated until the entire message is received in the receiver.
- Infinite size of message queues is assumed.
- The simulator is run for a total of 100 000 slot times and starts to log statistics at 20 000 slot times.
- The "total packet generation intensity" is the generation intensity for the network regarded as a whole, not of the individual nodes. That is, the sum of all individual packet generation intensities at nodes is the total packet generation intensity.

Figure 7 shows the best effort packet latency for varying levels of network throughput. At a useful level of packet latency, the network has an average throughput of approximately 1.6 packets per slot. This is 60 % better than the theoretical limit, of one packet per slot, for

Figure 9. Packet throughput vs. packet generation intensity for 16 nodes. The other curves represent latency for the 15 different distances between source and destination.

networks without spatial reuse. Similar results for varying number of nodes are obtained.

Figure 8 shows the packet throughput against packet generation intensity. Throughput has a linear relation to packet intensity up to the point of saturation, which can be seen in the figure as the point on the plot where throughput starts to decrease with increasing packet intensity. As packet intensity passes the point of saturation it becomes increasingly difficult for the packet transmission scheduler to effectively utilise the bit rate of the network. This is because it is always increasingly difficult to schedule packets the further their destinations are [2]. When the network is saturated each node's queues will tend to always contain far destined packets with short deadlines. These cannot always be scheduled together with the shorter destined packets because of conflicting links. When two packets have the same priority level the packets with the further destination has priority. Thus slot utilisation will decrease which is expressed in the decline of the throughput after the "summit" in the figure. The simulation shows that the utilisation of slots does not perform as well as may be theoretically expected (two packets transmitted per slot, because of pipelining). This is attributed to the policy of using deadline to selecting packets for transmission. The policy is not as bandwidth conserving as, e.g., the pure FWS policy [2] which can achieve higher average throughput. This is the cost of having good support for real-time traffic and fairness as will be seen shortly.

Figure 9 shows total packet throughput and average packet latency for each destination distance plotted against packet generation intensity. The simulation is for 16 nodes. Concluded from this simulation is that TCMA treats packets fairly regardless of the distance to the destination even when the network approaches and is saturated. Observe that there are 15 (for $N - 1$ distances to destination) plots for latency but that these overlap and appear as one.

6. Conclusions

This paper presents a medium access protocol with global deadline scheduling for an optical ring network. The user services include best effort messages, non real time messages, guarantee seeking messages, real time virtual channels (RTVCs), group communication such as barrier synchronisation and global reduction, and services for reliable transmission. RTVCs and guarantee seeking messages are realised through slot reservation. Simulation results of the medium access protocol with best effort traffic have been presented. The protocol is shown to be fair even when the network is saturated and it supports spatial reuse to achieve throughputs higher than one. The network with the presented protocol is

suitable for applications such as in embedded systems, e.g., for use as interconnection network in a radar signal processing system, or as a high performance network for use in a LAN environment. Also worth mentioning is that the network can be built today using fibre-optic off-the-shelf components

Acknowledgement

This work is part of M-NET, a project financed by SSF (Swedish Foundation for Strategic Research) through ARTES (A Real-Time network of graduate Education in Sweden).

References

- [1] M. Jonsson, "Two fibre-ribbon ring networks for parallel and distributed computing systems," *Optical Engineering*, vol. 37, no. 12, pp. 3196-3204, Dec. 1998.
- [2] P. C. Wong and T.-S. P. Yum, "Design and analysis of a pipeline ring protocol," *IEEE Transactions on communications*, vol. 42, no. 2/3/4, pp. 1153-1161, Feb./Mar./Apr. 1994.
- [3] C. Bergenheim and J. Olsson, "Protocol suite and demonstrator for a high performance real-time network," *Master thesis, Centre for Computer Architecture (CCA), Halmstad University, Sweden*, Jan. 1999.
- [4] D. Bursky, "Parallel optical links move data at 3 Gbits/s," *Electronic Design*, vol. 42, no. 24, pp. 79-82, Nov. 21, 1994.
- [5] C. Bergenheim, "A demonstrator for a CC-FPR network," *Technical report 0010, School of Information Science, Computer and Electrical Engineering (IDE), Halmstad University, Sweden*, Feb. 2000. In Swedish.
- [6] K. Arvind, K. Ramamritham, and J. A. Stankovic, "A local area network architecture for communication in distributed real-time systems," *Journal of Real-Time Systems*, vol. 3, no. 2, pp. 115-147, May 1991.
- [7] M. Jonsson, B. Svensson, M. Taveniku, and A. Åhlander, "Fiber-ribbon pipeline ring network for high-performance distributed computing systems," *Proc. International Symposium on Parallel Architectures, Algorithms and Networks (I-SPAN'97)*, Taipei, Taiwan, Dec. 18-20, 1997, pp. 138-143.
- [8] M. Taveniku, A. Åhlander, M. Jonsson, and B. Svensson, "The VEGA moderately parallel MIMD, moderately parallel SIMD, architecture for high performance array signal processing," *Proc. 12th International Parallel Processing Symposium & 9th Symposium on Parallel and Distributed Processing (IPPS/SPDP98)*, Orlando, FL, USA, Mar. 30 - Apr. 3, 1998, pp. 226-232.