Design for an Efficient Architecture for a Reflective Memory System and its Implementation

Il-Joo Baek *, Soo-Young Shin, Jae-Young Choi, Tae-Rim Park and Wook-Hyun Kwon

* Department of Electrical and Computer Engineering, Seoul National University, Seoul, Korea
(Tel : +82-2-873-2279; E-mail: ij100@cisl.snu.ac.kr)

Abstract: This paper proposes an efficient network architecture for reflective memory system (RMS). Using this architecture, the time for broadcasting a shared-data to all nodes can be significantly shortened. The device named topology conversion switch (TCS) is implemented to realize the network architecture. The implemented TCS is applied to the ethernet based real time control network (ERCnet) to evaluate the performance.

Keywords: Reflective Memory System, Token Ring, Topology Conversion Switch, Real time control network

1. INTRODUCTION

Reflective memory system (RMS) is one of the well known solutions to solve the various problems occurring when data are shared in distributed nodes under the multi process environment [1]. Main idea of the RMS is that the shared data are copied and updated automatically to each memory of the remote nodes. As soon as a shared data are updated to the memory of a node, the data are transmitted to all other nodes within certain amount of time. One of the important factors to decide the performance of the RMS is the shared data transmission time (SDTT) during which the shared data are automatically transmitted to the all nodes.

There have been various existing results about the RMS such as network shared memory (NSM), shared common RAM network (SCRAMNET+) and VME microsystems international corporation (VMIC)'s RMS [2][3][4]. Especially, all these results have adopted the ring topology. When data are passed through a node in the traditional ring network, the data are copied into a buffer and retransmitted to the next node. The procedure for copying the shared data and transmitting them to other nodes has significant influence on the SDTT. Therefore the SDTT increases as the number of nodes and size of data increases in the traditional ring network. In contrast, the NSM proposed unidirectional slotted ring architecture to reduce the SDTT. However, this architecture makes a limit to the amount of data. The SCRAMNET+ and VMIC RMS have adopted the same traditional data transmission method. Therefore they have the same delay problem as the traditional ring network has.

In this paper, a new efficient ring network architecture for the RMS is proposed. In the network architecture, received data from a previous node are directly transmitted to the next node with neither buffering nor retransmitting. Therefore the SDTT is significantly reduced. Also, there is no limit to the amount of data that can be passed through a node because it has no fixed size buffer between receiver and transmitter of the node. To realize this idea, the topology conversion switch (TCS) is designed and implemented. The implemented TCS is applied to the ethernet based real time control network (ERCnet) to evaluate the performance of the network architecture. The ERCnet is the industrial control network using the Reflective Memory (RM) for the nuclear power plant.

This paper is organized as follows. Section 2 describes the conceptual overview of the RMS and discusses the existing results. The network architecture is described in detail and the TCS is introduced in Section 3. Section 4 provides the brief introduction of the ERCnet and the discussion on the performance evaluation. Section 5 presents concluding remarks.

2. EXISTING RESULTS

2.1 Overview of the RMS

Distributed shared memory (DSM) falls into two classifications: software based DSM and hardware based DSM [6]. In the software based DSM, each node has its own local memory and shared data consistency is managed by massage transmission and virtual addressing concept [7]. Therefore it has advantages of both shared memory system and message passing system. In the contrast, the hardware based DSM makes physically distributed shared memory space have the same physical address [8]. By this reason, it can provide a transparency to a programmer and a user.

The RMS is a branch of hardware based DSM and therefore it has merits of both the shared memory system and the massage passing system. Also it is easy for user and programmer to develop the system because the all distributed shared memories are treated like one physically independent memory. The basic idea of the RMS is based on automatic update of remote shared-memory copies [5]. In other word, a shared data updated to a node are automatically copied and transmitted to the same address in each memory of the all connected nodes. This is called RM update. So, the nodes can

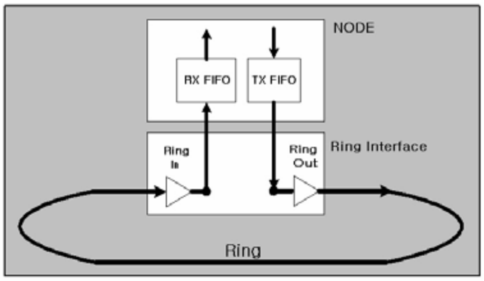

Fig. 1 Traditional ring network

use a shared data immediately without sending any request message. The RMS is composed of a dual ported memory physically distributed and logically mapped into a global, shared address space. The RM updates can occur over different types of interconnection network and the size of the shared data also can be decided various ways [1].

2.2 Traditional ring network architecture

The NSM, the SCRAMNET+ and the VMIC RMS are typical researches that use the ring network as the interconnection network for the RMS. The ring network has an advantage that makes the implementation of hardware of the RMS easier. Figure 1 is showing the inner architecture of the traditional ring network. When data are transmitted among the connected nodes, a node that received a data from a previous node storages the data into its receiving buffer and transmitting buffer respectively. Then the node retransmits the data to a next node. Therefore the evitable time delay for coping and retransmitting the data occurs whenever data pass through the nodes.

2.3 NSM

The NSM proposed a unidirectional slotted ring architecture shown in figure 2 to reduce the time delay caused in the conventional ring network. The NSM has fixed sized slots that circulate around the ring, the result are the same as if all nodes have the slots itself like shown the right picture of figure 2.
By using these slots, data copy time to a receiving buffer can be reduced.

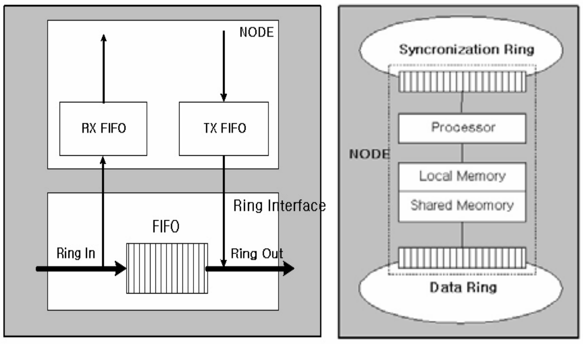

Fig. 2 Architecture of NSM

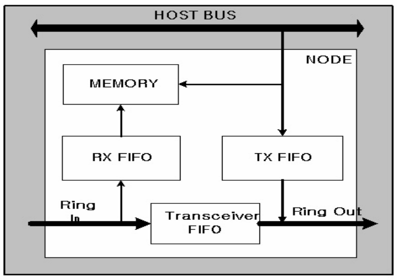

Fig. 3 Architecture of SCRAMNET+, VMIC

However this study has a structural defect that there is a limit to the amount of data the slots can handle because the size of slots are fixed.

2.4 SCRAMNET+, VMIC RMS

The SCRAMNET+ and VMIC RMS use the same data transmission used in the traditional ring network like shown in figure 1. As shown in figure 3, there is transceiver FIFO between the entrance and the exit of the ring. A received shared-data are copied into the FIFO and retransmitted by the transmitter to the next node. By this delay for coping and retransmitting, the performance of the system decreases as the number of nodes and size of data increases.

3. A NEW ARCHITECTURE

3.1 ***Broadcasting based ring network***

This paper proposes broadcasting based ring network architecture to overcome the drawbacks of the discussed related studies. There exists a switch in a ring interface like shown in figure 4. This switch makes it possible to change the ring topology into the bus topology. When a new data are updated to RM of a node, the data are stored in the node's transmitting buffer while the switch is switched into transmission mode at the same time. Then the node broadcasts the data to all nodes. The broadcasted data are bypassed through the connected nodes and come back to the original sender and then the node goes back to the reception mode by switching the switch again. This has the same effect as if all nodes are connected to bus network and therefore the data are neither buffered nor retransmitted by each node.

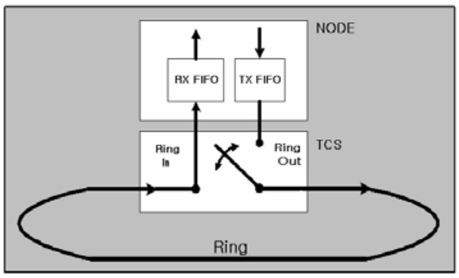

Fig. 4 Architecture of NSM

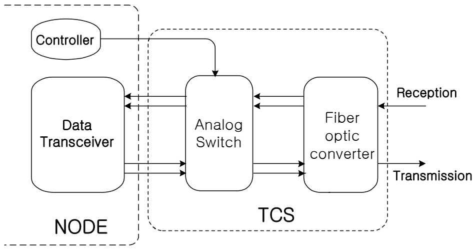

Fig. 5 Block diagram of TCS

The nodes just read the data when they are passed through themselves. By using this concept, the data processing delay per node can be reduced and therefore the RMS can share more data at high speed.

3.2 *TCS*

Topology Conversion Switch (TCS) is a kind of ring interface device. It is designed to use the ring network as broadcasting bus network. After a TCS receives a broadcasted data, it transfers them to central processing unit (CPU) of the connected node while the data are bypassed to the next nodes automatically. Block diagram of the TCS is shown in figure 5 and its prototype is shown in figure 6. It has analog switch part and fiber optic conversion part. The analog switch part are controlled by control line from the connected node. In reception mode, the data reception line from the fiber optic conversion part is diverged. It is not also connected to the data transceiver part but also to the data transmission line of the fiber optic conversion part. In transmission mode, the connection between the data reception line and the data transmission line is cut off and the data transmission line is connected to the data transceiver part.

4. PERFORMANCE EVALUATION

Ethernet based real time control network (ERCNet) is a RMS developed as the distributed control system (DCS) for industrial environment. The ERCnet adopts ring topology for the ease of maintainability. Also, it uses fast ethernet and fiber optic cable for physical layer. As for medium access control (MAC), token pass scheme is used to have real time property. Every nodes has its own RM and the RM maps all node's data space on the same address space. Figure 7 is showing the architecture of the ERCnet. Field network controls attached controllers and transfers updated data to the control network. The control network shares the data updated from the field networks using the RM concept. Information network observes the status of communication.

Fig. 6 Topology Conversion Switch

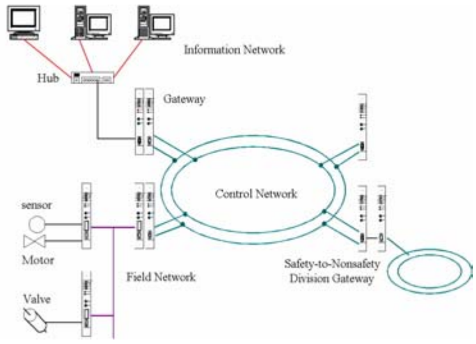

Fig. 7 Structure of ERCnet

The proposed architecture is adopted to construct the control network. The nodes are connected to the ring network by way of the TCS.

The ERCnet uses 100 Mbps ethernet for physical layer and therefore the shared data are transmitted by the ethernet frame. According to fast ethernet standard, the maximum frame size is 1518 bytes and the minimum frame size is 64 bytes. If the size of the shared data is larger than 1500 bytes, then the data are fragmented into 1500 bytes frames and transmitted. In addition, even if the size is less then 64 bytes, the size of the data frame should be larger than 64 bytes.

In order to evaluate the performance of the proposed architecture properly, it is evaluated when the ERCnet configures it's inter network with the TCS and then the same procedure is repeated in the same system which is constructed without the TCS. Four nodes are connected in the ring network, and the time for a certain size of data is rotated through out the ring is measured. This time is called data rotation time (DRT). The DRT is also measured repeatedly with different size of data transmission.

As shown figure 8, the network using the TCS takes 147us delay to share 1500 bytes data. However 607.8 us is needed to share the same size of data without the TCS. The time difference between two graphs and the gradient of each graph are almost identical. This means that about 560us delay can be saved by using the proposed architecture. The most of time delay occurring when the data are updated to all nodes is due to the buffering and retrasmitting process in each node. Therefore the SDTT of the RMS can be efficiently improved when the RMS constructs the inter network using the proposed network architecture.

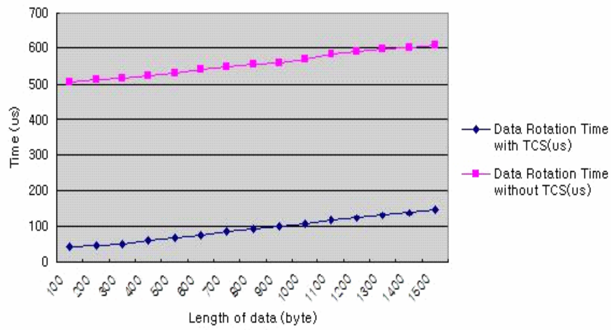

Fig. 8 Data rotation time

5. CONCLUSION

In this paper, the important factors which influence on the performance of the RMS are presented. To overcome drawbacks of the existing results and improve the performance for the RMS, the broadcasting based ring network architecture is proposed. Using this architecture, the time for broadcasting a shared-data to all nodes can be significantly shortened and therefore the RMS can share more data at high speed. The TCS is introduced to implement the architecture. The performance is evaluated by adopting the implemented TCS to the ERCnet. There was about 560us delay reduction when 1500 bytes data were shared by using the proposed architecture. This experiment verifies that the proposed broadcasting based ring network architecture can improve the performance of the RMS as the number of nodes and size of data increases.

REFERENCES

[1] M. Jovanovic, V. Milutinovic, "An Overview of Relfective Memory Systems", IEEE Concurrency, vol. 7, 1999, pp. 56-64.
[2] Ramanjuan R.S., Bonney J.C., Thurber K.J., ``Network shared memory: a new approach for clustering workstations for parallel processing", Proceedings of the Fourth IEEE International Symposium, 1995, pp. 48-56.
[3] Systran Corporation, ``SCRAMNet+ Overview ", www.systran.com
[4] VME Microsystems Corp., ``VMIC's Reflective Memory Network", www.vmic.com
[5] Chia Shen, Ichiro Mizunuma, "RT-CRM: Real-Time Channel-Based Reflective Memory", IEEE Transaction on Computers, vol 49, no 11, November 2000, pp 1202-1214.
[6] R. Bisani et al., "Plus: A distributed shared memory system", in Proc. 17th Ann, Int'l Symp. Computer Architecture, 1995.
[7] D. Chaiken et al. "Software extended coherent shared memory: performance and cost", in Proc. 21th Ann. Int'l Symp. Computer Architecture, 1990.
[8] A. Wilson et al. "Hardware assist for distributed shared memory", in Proc. 13th Int'l Conf. On Distributed Computing Systems, 1993.