Counter Rotating Dual Ring Daisy Chained Network for Real-Time Distributed Control Protocol

Yoon-Soo Lee*, ***Regular Member***

ABSTRACT

The ring topology has been in use in networking for quite some time for simplicity. However, the downside of using the ring topology is for having to worry about link or node failure which can result in network failure. Tolerating faults in the ring topology has been a popular area of research. However, it is hard to incorporate those ideas to a specialized network for distributed control networks. This work shows how the ring network can be improved in terms of performance and reliability for high performance distributed control network. Discussion will be on the ideas to increase throughput and reliability. Specifically, this work contains implementation ideas for fieldbus protocols to tolerate single point of link or node failure and prevent packet loss. Since the suggestion of this work makes it obvious on the improved throughput, the conclusion emphasizes on the reliability gain of the network system by making changes to the traditional ring network. However, the reliability of the network may decrease depending on the reliability of the nodes which consists the network.

Key Words : Fieldbus Protocol, Dual Ring, Fault Tolerant, Reliability

요 약

링 토폴로지는 네트워크 분야에서 구현의 간결함 때문에 과거로부터 많이 이용되어왔다. 하지만 링 토폴로지를 사용하는데 있어서 가장 큰 단점 중에 하나는 링크나 노드가 정상작동에 실패하여 네트워크가 죽는 것을 걱정해야 한다는 것이다. 그동안 이런 단점을 보완하기 위해 많은 연구가 이루어져 왔지만 특수하게 사용되는 분산 제어 네트워크에 적용하기는 힘들었다. 이 글에서는 과거의 연구결과들과는 다른 관점에서 접근하여 네트워크의 스루풋 및 안정성을 향상 시킬 수 있는 방법을 설명한다. 구체적으로 필드버스 프로토콜에서 단일 지점에서 발생하는 네트워크 오류를 어떻게 처리해야 하는지에 대해 서술한다. 스루풋을 향상시킬 수 있는 방법은 너무나 당연하기 때문에 그보다는 안정성 향상의 정도를 가늠하는 것으로 연구 결과의 성과에 대한 결론을 내린다. 하지만 네트워크에서 사용되는 노드들의 안정성 정도에 따라 전체적인 안정성이 오히려 감소할 수 있다.

Ⅰ. Introduction

The ring topology has been used in the past because of its simple implementation and some desirable characteristics[1]. A simple ring topology does not require a complicated routing scheme and broadcast can be done easily as the topology can ensure that a packet can reach every node as it traverses the ring. However, the communications delay is bound by the diameter of the ring and is prone to failure. Many research has been done to improve the properties of the ring topology[1, 2, 3]. They discuss about introducing another ring and using a more sophisticated topology compared to the simple ring topology. While they are all extremely valuable work, their ideas are limited to work on LAN settings and it becomes difficult to incorporate those ideas to specialized networks for real-time distributed controled systems.

Here we show how dual ring can be adapted to

* 삼성SDI(yoon-soo.lee@samsung.com)
논문번호 : 08033-0526, 접수일자 : 2008년 5월 26일

fieldbus protocols that require relatively high speed communication for increased reliability. The work is built upon the protocol developed earlier for power electronics systems. Since the method of increasing the throughput is from the original protocol, the focus will mainly be on the ideas to make the ring network fault tolerant. This work can be distinguished from previous works related to fault tolerance in the ring network as it is built based on a protocol which behaves slightly different from other protocols.

Ⅱ. Previous Work

Although this work is targeted on providing general information on implementing a reliable network protocol for specialized ring networks for real-time distributed control systems, the work which it was built upon should be mentioned to help understanding the topic and to credit work which was done previously.

This work was an augmentation to PESNet. PESNet is a fieldbus protocol for power electronics systems. It can be distinguished from many other protocols as it does not use the full protocol stack layer. Only the physical layer, link layer, and the network layer is used for high speed communications [4]. While other protocols operate above the transport layer and focuses on point to point communication at any given time, in PESNet each node serves as a router and synchronizes the action of receiving and transmitting packets at all nodes according to a global time called a network tick. The packet is consumed if it is destined to the node which received the packet and forwarded until it reaches its destination otherwise. This method is used to utilize the links as much as possible, hence getting the affect of increased throughput. While this kind of network feature may be preferable in other networks that require high speed communication, many of the previous work done may not be well adopted to protocols with such characteristics because the network behavior differs from how LAN works.

With the characteristics explained above, PESNet started as a simple protocol to work on a simple single ring network. However, changes had to be made to PESNet to improve the reliability. As a result, a proposal to use counter rotating dual ring daisy chained network to gain reliability was made[5]. The proposal includes the idea of ring wrap operations on failure. Only one ring, the primary ring, is normally used, but the other ring is utilized when a failure occurs to get around the point of failure. This idea was not a new idea[3, 6], however, it began as a starting point of this research.

The proposed idea only contains conceptual ideas which can be adapted to PESNet. Work on which how to initiate the ring-wrap operation and detecting faults were not addressed in the proposal. Also, considerations about packet loss issues were not stated.

The following section will first search for a better network topology which can decrease the network delay. An attempt to search for a better choice of the network topology is made because the delay can be as large as the diameter of the ring and it can become worse by a factor of 2 when operating in a ring-wrapped state when using a simple counter rotating dual ring topology. Then a description on how to carry out the ring-wrap operation when faults are introduced to the network and its behavior for when the failure is recovered will be present.

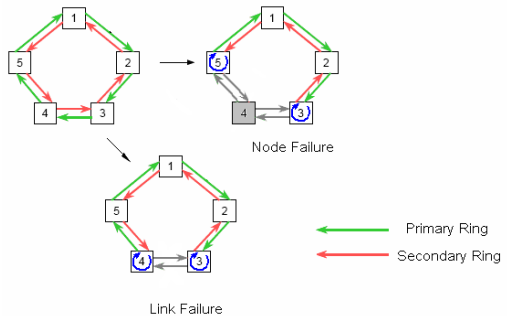

Fig. 1. Example of the use of counter rotating dual ring network through ring-wrap operation when existence of failure.

Ⅲ. A Search for a better Topology

The search for a better topology was inspired by the work of Raghavendra[7]. While one ring, the forward loop, is constructed by interconnecting neighboring nodes, the other loop is constructed by interconnecting nodes that are several nodes(h) apart. There is no definite rule for choosing the value for

h. But, it is said that the loop may become optimal if *h* is chosen to be $\lfloor\sqrt{n}\rfloor$ where *n* is the number of nodes. For example, a network with 15 nodes will become optimal if *h* is set to be $\lfloor\sqrt{15}\rfloor$ = 3 as shown in Figure 2.

Unfortunately, the optimal loop topology was not able to be adopted as one of the important requirements for fieldbus protocols is to always have a path to all nodes in the network in the form of a closed loop.This property is important in a sense that all traffic may have to be observed as the control is distributed across several nodes. This constraint is the main reason why the ring topology was used initially.

In order to suffice this requirement, the network must form a hamilton circuit at all times. In order to have a Hamilton circuit, every vertex in the graph that is equivalent to the network must have degree two. Each node is expressed with two vertices for each loop that it is connected to and the links can be depicted as edges. When a fault occurs, the two closed loops become segmented. In order to construct a graph with a Hamilton circuit with multiple segments, the number of additional edges required is the number of segments in the resulting graph after the failure occurred. Also the segments must not contain any circuits. If a link fails in an optimal loop topology, the other loop which the failed link is not involved in remains closed. If we remove an edge on the remaining loop to have two segments that do not contain a circuit, we can construct a Hamilton circuit. However, an edge that connects two vertices that do not belong to the same node must be added which is not possible without physically connecting them with a new communication link. If a node fails in the optimal loop topology, we lose two vertices and four edges in the graph. The graph becomes segmented without any circuits. However, again, an edge that connects two vertices that do not belong to the same node must be added in order to construct a Hamilton circuit. Therefore, the optimal loop topology cannot be used. Therefore, the topology was determined to stay the same.

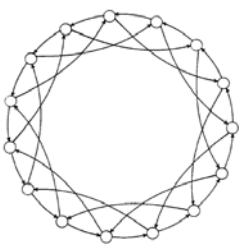

Fig. 2. Optimal Loop Topology with 15 Nodes, Skip distance 3

Ⅳ. Carrying out the Ring-Wrap Operation

In order to carry out the ring-wrap operation, there must be a fault detection mechanism. Detecting the fault becomes a challenge as fault can only be detected on the receiving side of the link. The problem is that the ring-wrap operation must be carried out on both ends of the point of failure while the fault can only be detected on one side. However, FDDI also has a similar functionality[6] for reliability and the idea was borrowed from it. A failure can be detected on the receiving side of the node by observing the signal frequency. The receivers will flag the Low Frequency Indicator flag. when the flag is flagged, the transmission in the opposite direction on the other ring is disabled. As a result, the node at the opposite side of the point of failure will be automatically notified to perform a ring-wrap operation simultaneously. The process of performing a ring-wrap operation at both ends of the point of failure is shown in Figure 3. In case of a node failure, the two neighboring nodes to the failed node will not receive any signal from it. As a result, the ring-wrap operation can be carried out simultaneously and independently at the two nodes.

Recall that the protocol which is being discussed transmits and receives packets simultaneously at all nodes every network tick. Suppose that a fault was detected during a packet transmission and the ring-wrap operation was completed before the end of the network tick. If the packet is not retransmitted after the ring-wrap operation during the next network tick, messages can get lost. In order to prevent message lost, a temporary buffer must be present for retransmitting the packet which was failed to be delivered to the next node. Figure 4 is a block diagram on how the protocol must handle packet forwarding.

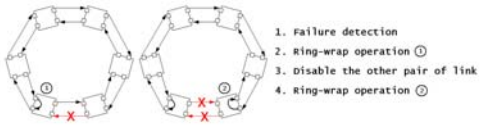

Fig. 3. Ring-wrap operation upon fault detection.

As a packet is received at each receiver, the data is stored in separate buffers. According to the Low Frequency Indicator (LFI) signal from each receiver, the node knows when a break condition has occurred on either of the ring. The LFI signal is low when the optical signal on the corresponding link is sufficient to properly receive the packets. When LFI becomes high for either ring, transmission on the other ring is disabled immediately to force the LFI signal to rise on the other side of the point of failure. This ensures that the ring-wrap operation will also be carried out at the node residing one the other end of the link pair.

In order to formalize the action of forwarding process, the condition of the network is divided into three situations. They are normal operation mode, ring-wrapped mode, and healing mode.

At all times, the destination of the packet is checked to see whether the node should consume the packet or forward it on to the next node when the packet is placed in the primary buffer. During normal operation, the packet that is placed in the Primary Buffer is transmitted on the primary ring and the packet that is placed in the Secondary Buffer is transmitted on the secondary ring.

When the network is in ring-wrapped mode, the nodes which performed the ring-wrap will forward the packet to the other ring depending on their ring-wrapped state. The packets to be forwarded to the primary ring is embedded with the information of the node that is thought to have failed, which is the previous node. This action is required to determine whether the failure has resulted from a link or node failure. In order to distinguish between a link failure and a node failure, a test is required at the other end of the point of failure and cannot be distinguished until the first packet that contains the information of the node that is thought to have failed reaches the other end of point of failure. When a packet in the primary buffer is forwarded to the secondary ring, the information of the node to be thought as faulty is examined. If the information matches the network address of the current node, the failure is determined to be a link failure. Otherwise, it indicates a node failure has occurred.

After the ring-wrap operation, the disabled

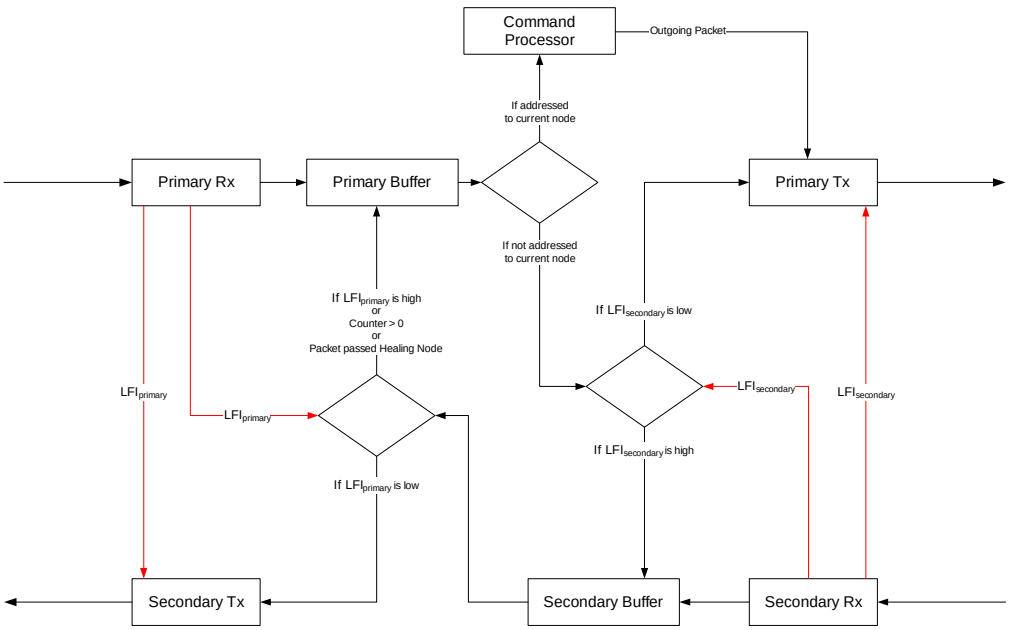

Fig. 4. Block diagram of the forwarding process on each node.

transmitter resumes operation for reestablishing connection. The network will remains in a ring-wrap state until the two nodes complete a hand-shake. A hand-shake is required to ensure that both links between them are operational to enter normal operation mode again.

When a node exits the ring-wrap state, the network would operate as if it were in normal operation mode except that it must complete the healing process. The healing mode is a state of the network where the meaningful packets that are traversing the secondary ring are forwarded back to the primary ring. The process is done in an opportunistic fashion. As the network enters healing mode, the node that was forwarding packets from the secondary ring to the primary ring becomes the Healing Node. As a node becomes a Healing Node, a counter is set to the number of nodes in the network. The counter is used as an indication that the node has become a Healing Node. The counter is decremented after each network tick and the node serves as a Healing Node until the counter reaches 0. The Healing Node attempts to forward the packet received from the secondary ring to the primary ring by placing the packet in the Primary Buffer. However, if the Primary Buffer is already occupied by a meaningful packet, the attempt fails. If an attempt fails, the packet is forwarded to the next node on the secondary ring for another attempt on that node. An attempt to forward a packet that was supposed to be forwarded to the primary ring to the primary ring is continuously made until success as it propagates in the secondary ring.

V. Preventing Packet Loss

Primary and secondary buffers are used to prevent packet loss that might occur due to a link failure during packet transmission. However, packets may get lost due to node failures. Packet loss caused by a node failure cannot be prevented using the same technique for link failures. In general, a lost packet in a network can be detected by the node that generated the packet when it fails to receive an acknowledgement within a given time frame. Detecting packet loss becomes very trivial with the assumption that all nodes are aware of the number of nodes in the network and the acknowledgement is made as soon as a packet is received. With these assumptions, a node can detect a lost packet when it does not receive an acknowledgement for a packet for $2n$ network ticks where n is the number of nodes in the network.

While the method for using response time for detecting packet loss is simple, the drawback of it is the indefinite amount of resource needed. Each node must keep a list of packets that were generated and transmitted in the last $2n$ network ticks. Although each nodes can be built to have enough resource to store the information of the packets that it generates during the past $2n$ network ticks, the required amount of resource for the nodes to have should not be a dynamic property affected by the communications protocol and the property of the network. Also, packet loss can be detected as early as when a node failure is detected. Since network delay is a very sensitive property in real-time applications, it would be better not to wait $2n$ networks ticks to retransmit the lost packet, but retransmit as soon as a node failure has been detected.

A detection scheme for node failures has been already specified in the previous section. The challenge in retransmitting the lost packets lies in the fact that failure detection does not necessarily indicate a node failure. Another transmission of a packet should only occur when the failure is determined to be a node failure. The determination takes approximately n network ticks after a failure has been detected. In order to retransmit the lost packet, there must be some mechanism to preserve the packet that potentially has been lost due to the node failure until a node failure has been confirmed. Preservation of a packet can be done by saving the packet that has been sent until a new packet is finished being transmitted to the next node on the primary ring. Therefore, a packet that has been transmitted at a node is preserved for a network tick. The process of saving a packet stops as a failure is detected on the secondary ring. By the time a failure is detected on the secondary ring, the packet being preserved at the node might have been lost due to a node failure. Recall that a node that detects a failure on the receiver of the primary ring will start

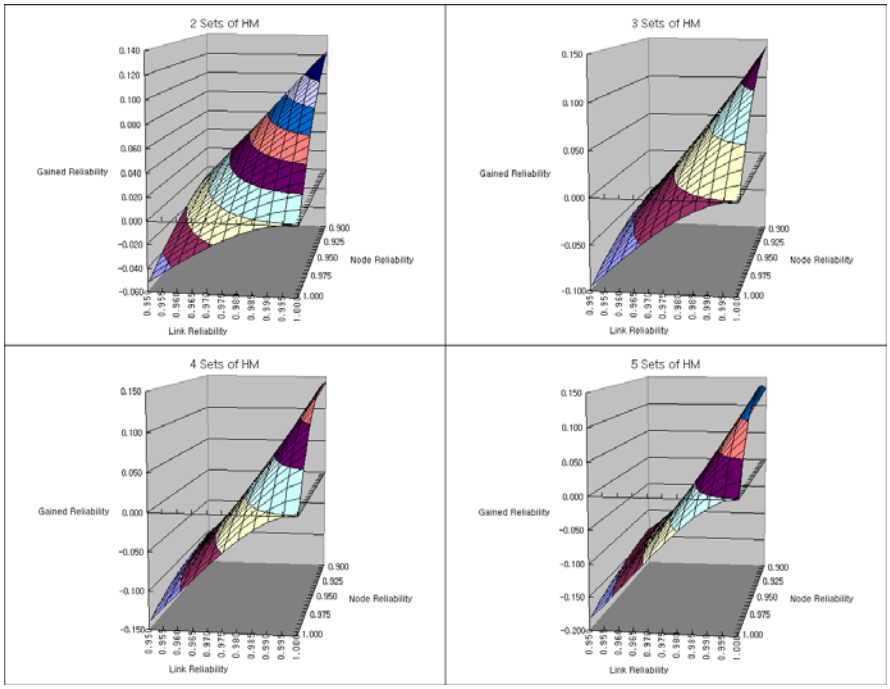

Fig. 5. Gained reliability

to send information about the node that has possibly failed in every packet it transmits. Whether the node has failed or not can be determined when the packet containing the supposedly faulty node's address arrives at the other end of failure, which is the node that detects a failure on the secondary ring. If the fault is determined to result from a node failure, the preserved packet is retransmitted and begins to preserve incoming packets for a network tick again.

Ⅵ. Conclusions

This work was conducted to incorporate the ideas of reliable ring topology techniques to specialized network protocol for real-time distributed control systems. Figure 5 shows the reliability that can be gained by being able to tolerate faults. The numbers are mathematically calculated rather than simulated or tested. The reliability gain which it delivers may not be so significant. However, there is a slight reliability gain by allowing single point of link or node failure as expected. Also, the results show that presence of unreliable nodes may decrease the overall reliability as the number of nodes increases. This is obvious because the possibility of failures increases as more nodes are introduced. However, many people might have neglected to expect the outcome.

Problems still remain for network segmentation when there are multiple points of failure in the network. However, this can be another starting point for making more improvements. The biggest challenge of making such protocol specifically designed for high throughput relies in the fact that it does not make use of the entire protocol stack layer. Nonetheless, there exists a tradeoff between reliability and performance. Before more research is done to achieve reliability of the network without using the full protocol stack layer, such protocols should be designed carefully depending on the required performance and the degree of reliability that must be reached.

References

〔1〕 Raghavendra, C. S. and Gerla, M. 1981. "Optimal loop topologies for distributed systems." *In Proceedings of the Seventh Symposium on Data Communications,* (Mexico City, Mexico, October 27 - 29, 1981). SIGCOMM '81. ACM, New York,

NY, 218-223.

〔2〕 Peha, J.M., Tobagi, F.A., "Analyzing the fault tolerance of double-loop networks," *Networking, IEEE/ACM Transactions on , vol.2, no.4*, pp.363-373, Aug., 1994.

〔3〕 Rom, R.; Shacham, N., "A reconfiguration algorithm for a double-loop token ring local area network," *Computers, IEEE Transactions on , vol.37, no.2, pp.182-189*, Feb., 1988.

〔4〕 Milosavljevic, I., "Power Electronics System Communications," in Electrical Engineering. 1999, Virginia Tech: Blacksburg, Virginia.

〔5〕 Jerry Francis, J.G., and Stephen H. Edwards, "Protocol Design of Dual Ring PESNet (DRPESNet)," in CPES 2002 Power Electronics Seminar and NSF/Industry Annual Review. April, 2002.

〔6〕 Ross, F.E., "An overview of FDDI:the fiber distributed data interface." *Selected Areas in Communications, IEEE Journalon*, 1989.**7**(7): p.1043-1051.

〔7〕 Raghavendra, C.S.; Gerla, M.; Avizienis, A., "Reliable Loop Topologies for Large Local Computer Networks," *Computers, IEEE Transactions on, vol.C-34, no.1, pp.46-55*, Jan., 1985.

이 윤 수 (Yoon-Soo Lee) 정회원

2004년 5월 Virginia Tech Computer Science 졸업

2006년 7월 Virginia Tech Computer Science 석사

2006년 9월~현재 삼성 SDI PDP 개발팀 근무

<관심분야> Software Design/ Architecture, Software Platform, System Software, Embedded System