RLP : An Efficient HSR Traffic Reduction Algorithm

Saad Allawi NSAIF[*], Je Hyun JUN[**], Sang Heon SHIN[***], Jong Myung RHEE[****]

ABSTRACT

In this paper, we present an algorithm called redundant logical paths (RLP) for efficient HSR traffic reduction. It creates redundant logical paths between each HSR node and all the other nodes. Eventually, a logical full-meshed network or paths will be established among all HSR node types, except the Quadbox type, which is used only for interconnection. The logical full-meshed network will be used instead of using the standard HSR protocol that depends on the concepts of the duplication and forwarding of the received frame until it reaches the destination node. The RLP algorithm results in significantly less frame traffic because there is no random forwarding as in the standard HSR protocol. For the sample network in this paper simulation results showed a 61.5 - 80% reduction in network frame traffic compared to the standard HSR. Our algorithm will avoid latency issues in the network and even network congestion, thus improving network efficiency.

Key Words : IEC 61850, IEC 62439-3, HSR, RLP, Traffic performance, Unicasting, Redundancy.

I. Introduction

High-availability seamless redundancy (HSR) is a redundancy protocol standardized as IEC 62439-3, Clause 5 [1]. Its principles depend on the replication of the frame over both sides of the HSR node in order to make sure that at least one frame copy will reach the required destination. If both copies reach the destination node, then it will take the fastest copy and discard the other copy. HSR thus provides zero recovery time in case one of the frame copies is lost. However, this principle divides the available network bandwidth for network traffic roughly in half [2]. The HSR protocol is very useful for time-stringent, real-time and mission-critical systems, such as automation substation networks, nuclear stations, and military command, smart grid system and control centers that are located within the same site. Therefore, instead of using traditional Ethernet network switches that utilize the rapid spanning tree protocol (RSTP), HSR is one of the best candidate solutions for redundancy management. However, RSTP is not suitable for real-time applications. The reason is that once a link or a switch fails, the network undergoes reconfiguration to rebuild the logical path using the RSTP protocol, which usually takes quite a long time (typically, one second) [3]. In contrast, HSR provides two independent FTE paths and provides a zero switchover-time characteristic in the case of failure. Although the Ethernet has become a de facto choice for open control network strategies, it was not designed to handle network faults, such as hub/switch failure or link failure. An obvious solution to fault tolerance is to equip a fault-tolerant Ethernet (FTE) protocol that is designed with specific goals: 1) No single point of network failure shall take down or cause loss of communication to more than one node on the network; 2) Any failure must be detected and recovered within a specific time [4].

As documented in IEC 62439, various FTE protocols for IEC 61850-based substation networks have been proposed, such as media redundancy protocol (MPR), parallel redundancy protocol (PRP), cross-network redundancy protocol (CRP), beacon redundancy protocol (BRP), and high-availability seamless redundancy (HSR). Recently, the HSR, standardized as IEC 62439-3 Clause 5 [1], became part of the IEC 61850 standard in addition to the PRP. However, IEC 61850-90-4, a network- engineering guideline for communication networks and systems in substations, recommends IEEE RSTP, PRP, and HSR [5]. Tan and Luan analyzed substation automation system (SAS) architecture designs using these three protocols [3]. They also proposed hybrid configurations with

---

※이 논문은 2012년도 삼성탈레스(주)의 재원을 지원 받아 수행된 연구임

[*] 명지대학교 정보통신공학과 Ubiquitous&Convergence 연구실 (saad.allawi1@gmail.com),
[**], [***] 삼성탈레스 C4I 연구소 (jehyun.jun@samsung.com[**], sangheon.shin@samsung.com[***])
[****] 명지대학교 정보통신공학과 (jmr77@mju.ac.kr), 교신저자 : 이종명

접수일자 : 2012년 11월 8일, 수정완료일자 : 2012년 11월 12일, 최종게재확정일자 : 2012년 12월 3일

combinations of these protocols. Furthermore, they concluded that HSR-based implementation would prevail in future power-related applications because of its excellent performance and affordable cost. The HSR protocol is used mainly for ring topologies, although redundant connections to other networks are also possible, such as connected rings or mesh topologies.

Generally, as a redundancy protocol for an Ethernet-based network, as previously mentioned here, the HSR provides duplicated frame copies for each frame sent. In other words, HSR provides two frame copies on separate physical paths, pursing zero-fault recovery time. This means that even in the case of a node or a link failure, there is no stoppage of network operations. Moreover, the HSR protocol is a layer 2 protocol and independent of the upper layer protocols, thus returning high flexibility in a wide range of different applications on HSR networks. The HSR network has several types of nodes, as listed below [1]:

A. Doubly attached node for HSR (DANH)
B. Singly attached node (SAN)
C. Redundancy box (Redbox)
D. Quadruple port device (Quadbox)

The other HSR principle and its implementation are examined and discussed in [1] and [5-8].

However, the main drawback of HSR is the unnecessary traffic caused by the duplicated frames that are generated and circulated inside the network. The traffic then will nearly double if the standard HSR is applied. This downside will degrade network performance and may cause network congestion or even stoppage. However, despite the potential advantages of HSR, most research on HSR has been dedicated to its implementation [7,8]. Few studies have reported on HSR traffic issues or traffic reduction techniques except our recent works [9,10].

In this paper, we present a new algorithm called redundant logical paths (RLP), which is suited to any closed-loop network topology. This algorithm will solve the unnecessary traffic problem that is caused by duplication and random forwarding. The idea of our algorithm is to establish two logical paths between each HSR node and all the other nodes, which eventually will establish a logical, full-meshed network among all the nodes, except the Quadbox node type, which is used only for interconnecting purposes. Instead of using the standard HSR protocol, which depends on duplication and forwarding, these logical paths will be used for the unicast traffic type. The RLP algorithm will significantly reduce the frame traffic in HSR networks, which will result in enhanced traffic performance.

The rest of the paper is organized as follows. In Section II, we briefly present our RLP algorithm and its concept. In Section III, we compare different cases of traffic performance between RLP and the standard HSR. The network simulator OMNeT++ V4.2.2 is used for this purpose. Finally, in Section IV, we describe our conclusions and future work.

Ⅱ. RLP Algorithm

A. RLP Algorithm

Our novel algorithm, redundant logical paths (RLP), will enable DANH, SAN, and Redbox nodes to communicate with each other through a point-to-point connection (pair connection) by establishing the fastest two logical paths between each pair. These paths will be established automatically after completing the learning process of the Quadbox nodes, which will make them “smart.” That is, they will know how to forward the received frame to the proper destination through one of their ports instead of rebroadcasting the frame through all the ports with the exception of the port from which the frame has been received.

The Quadbox node will learn how to interconnect the path segments that pass through it. Eventually, these paths will establish a full logical mesh network among the DANH, SAN, and Redbox nodes, which may be available within any network.

Thereafter, when the full-mesh network is established and a data frame is required to be sent to a certain destination, the sending node a DANH, SAN, or even a Redbox node will duplicate and send the frame copies, one from each direction. Later, the Quadbox nodes located in the middle between the source and the destination nodes will guide each frame copy to the destination node through two paths until they reach the destination. The destination node will then receive both copies, one from each direction, take the fastest one, and discard the other. However, any Quadbox node will not establish paths originating from it to other nodes. It does not generate or originate data frames to the other nodes because it is used only for interconnection purposes. In addition, the other HSR nodes do not need to send any direct data frames to any Quadbox node because this type does not have any attached end users.

For simplicity, and to avoid any confusion or recursion in this paper, the DANH, SAN and the Redbox nodes will be called the operation nodes.

B. RLP Frame Format

During the process of establishing the paths, the RLP algorithm will use several frame types in order to complete the process. In order to differentiate between these frame types, we suggest inserting a field with a unique code for each frame type inside the standard HSR frame. The location of this field will be between the HSR tag and the payload fields, as shown in Fig.1. This field is called the RLP frame type (RFT), and its size is two octets. The RFT size portion will be cut from the LPDU portion in the payload field in order to keep the whole size of the HSR frame within the same standard size instead of adding new two octets to its tag.

The RLP frame types will have several reliable frames, which have acknowledgment frames with a unique RFT code.

C. RLP Implementation Process

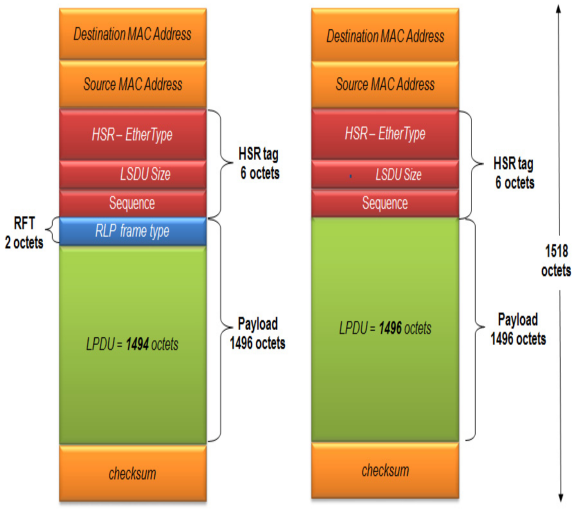
All nodes of the network will use the standard HSR protocol on start up until they complete the RLP implementation process.

The implementation process for the RLP algorithm will consist of three stages;

1. Declaration Stage

This stage consists of three steps summarized as follows:

a) Each operation node will declare itself by broadcasting a frame called the declaration (DeC) frame. This frame will be duplicated at each node and spread through the whole network until it reaches the source node or passes twice through the same node and the same direction in order to be discarded.



a) RLP frame with RFT field

b) Standard HSR frame

Fig 1. Frame structure for both the RLP algorithm and the standard HSR protocol

b) Each node, including the Quadbox nodes, will read the DeC frames that pass through it and build a table called neighbors (Ne). The Ne table contains the source MAC address and the sequence number of each DeC frame passing through it. The Ne table will list all the operation nodes with their MAC addresses; hence, all the operation nodes are capable of sending unicast frames to each other instead of broadcasting them.

c) The time interval needed for building the Ne table is called the neighbors table-building (Ne.B) interval. The Ne.B will be estimated as follows:

- Each node could introduce a 125-μ sec delay at 100 Mbit/s [1].
- Assume all operation nodes send a DeC frame, simultaneously.
- Assume each DeC frame passes through 250 nodes; the total transmission delay per one DeC frame is then equal to 125 μ x 250 = 31.25 msec.
- 125 μ sec processing time is required per each node for each received DeC frame; therefore 125 μ x 250 = 31.25 msec is the total processing time for processing 250 DeC frames at each node.
- The estimated Ne.B interval = 31.25 +31.25 = 62.5 msec.
- For margin purposes, we assumed that the total estimated Ne.B interval will equal 1 sec.

After the Ne.B expires, each node will move to the second stage.

2. Paths Establishment Stage

In this stage, all the Quadbox nodes will learn how to forward or guide the passing frames to their sources or destinations without the need to rebroadcast them at each Quadbox node. The activities within this stage can be summarized as follows:

a) Each operation node will send a unicast frame called the path selection (PaS) frame to each operation node listed in its Ne table. In other words, each source and destination node of each connection pair will send a PaS frame to each other. The PaS frames will be used to build a table called the pre-path table (PrP) at each Quadbox node type, as will be shown later.

The PaS frame that will be sent from a source node of a certain connection pair will make the Quadbox nodes in the middle-way of that connection pair know how to select the fastest path to connect the source with that destination node. The PaS frame that will be sent from the destination node will make the Quadbox nodes in the middle-way of that connection pair know how to select the fastest path to connect the destination with that source node. Therefore, each connection pair will have two PaS frames, one from the source and one from the destination node.

b) As soon as all the PaS frames are sent and start passing randomly from node to node towards their destinations, each Quadbox mid-way node will perform the following steps:

- Select the fastest PaS frame per each connection pair that reaches it and discard the other copies.
- Read the input port number that was entered from it, the sequence number, source MAC address, and the destination MAC address.
- All information read earlier will be arranged and tabulated into the PrP table and all the Quadbox nodes will know from their PrP tables from which the port the fastest PaS frame has been delivered in order to use it the next time this node needs to forward a data frame to the source node of that PaS frame.
- Rebroadcast the frames from the other ports.

c) Later the destination node will receive the two copies of the sent PaS frame, one from each direction.

d) The building process for the PrP tables at all the Quadbox nodes will be approximately equal to the double Ne.B interval that we calculated earlier. This interval is called the PrP building (PrPB) interval, and it is calculated from the time of sending the PaS frames until the building process of the PrP table is complete. The PrP table size will increase with respect to the number of operation nodes in the network, which results in more processing time within each table. Therefore, we considered that the PrPB interval is doubled as the Ne.B.

e) During the PrPB interval, all the nodes listed in the Ne tables send their PaS frames to each other in order to establish two paths among them. After completing the building process for the PrP table, the network will move to the next step.

f) The PrP table shows the gateways or the port numbers of the fastest received PaS frames or we can say the fastest path for each connection pair that may pass through it.

g) Each Quadbox node will deduce another table from its PrP table in order to use it for forwarding the received frames in both directions. This table is called the final paths (FP) table. The FP table will be built by taking the source and the destination MAC address columns from the PrP table and flipping their contents between them. It will also use the input port column as the output port column. The FP table does not need the sequence number column because its role is finished. Note that the building process for the FP table will start directly after completing the building process for the PrP table.

h) After completing the FP tables, each Quadbox node will delete its PrP table because its purpose is finished after deducing the FP table.

i) In this stage, each destination node of each connection pair will reply with a frame called the path confirmed (PaC) frame as soon as it receives the two copies of the PaS frame. Each copy of the PaC frame will travel through one of the two logical paths that have been established previously. During the trip of each copy towards the source node, each Quadbox mid-way node will confirm the entries of their FP table for that connection pair, whereas the nodes that the PaC frame will not pass through will delete their entries of FP tables of that connection pair because they will not be included in the selected paths. Hence, the size of the FP tables will remain small and will not contain any non-useful entries. The same process will occur for the source node when it receives the PaS frame from the destination.

3. Final Stage

After completing the building process of the FP tables, the RLP algorithm is ready and the next frames sent will be forwarded or guided within the two logical paths of each connection pair. We can summarize this stage as follows:

a) Each Quadbox node will be able to forward the received data frames to their destination or source nodes by using its FP table, instead of using the standard HSR transmission process. This will be done according to the following steps:

- Read the received frame and determine its source and destination MAC addresses.
- Look at the FP table for the source and the destination MAC addresses (connection pair) in order to identify and send out the received frame from the proper port. For the operation node, it will just forward it from the opposite port.
- If a Quadbox node could not find the proper source-destination pair for the received frame inside its FP table, then it will use the standard HSR transmission process (duplication and forward). Therefore, no data frame will be lost.

b) If the Quadbox nodes receive a broadcast or a multicast data frame, then it will use the standard HSR transmission process of duplication and forwarding.

D. Path Re-Establishment

As we mentioned earlier, sometimes it is necessary to re-establish the paths for a certain connection pair. To achieve this task, the following procedure is followed:

1. Using the example node A and node B, if node A did not receive a PaS frame from another operation node that is listed in its PrP table, then node A will send a frame called the resend-PaS (RePaS) to node B.
2. During the travelling of the RePaS frame, all the mid-way nodes will duplicate and forward it. The duplication and forwarding process is very important for the RePaS frame because there are no two paths for the connection pair A-B. Therefore, the frame will be spread out through the network in order to reach node B.
3. When node B receives the RePaS frame, it will re-send the PaS frame to node A.
4. All the nodes that are located in the re-established paths will delete the previous entries in their FP tables of that connection pair as soon as the nodes receive the new PaS frame, then from the new PaS frame information will create a new entry instead of the deleted one.

Ⅲ. Traffic Performance

To compare network-frame traffic performance between the standard HSR protocol and our RLP algorithm, under various scenarios, simulations have been made for the network sample shown in Fig. 2. Assuming that node 1 sends one frame to node 2, we simulate the frame traffic performance under healthy and faulty cases. In Fig. 2 all nodes 1, 2, 3 and 4 are DANH, while nodes 5, 6, 7 and 8 are Quadbox. The network simulator OMNeT++ V4.2.2 is used for this purpose.

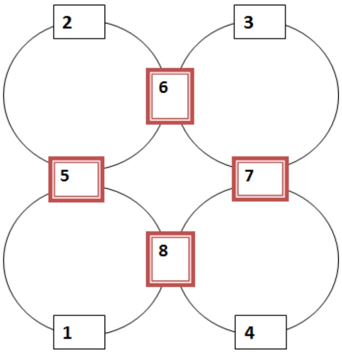


Fig 2. An example of HSR network

A. Healthy Case

In this case, all the nodes and the links will work properly without failure or faulty issues.

The network-frame traffic of the network shown in Fig. 2 for both the standard HSR and the RLP algorithm is listed in table 1 below:

Table 1. The network frame traffic under both Std. HSR and RLP-Healthy cases

#	Frames-Standard HSR	Frames - RLP	Frame Reduction Percentage
1	21	6	71.4%

Note that the two logical paths of our RLP algorithm for this case are

- The primary path: node 1- node 5-node 2
- The secondary path: node 1-node 8-node 7-node 6-node 2

In addition, the percentage of the transmission interval time (TIT) required to complete the transmission of one frame compared with the standard HSR protocol will be

$$TIT = (1 - \frac{4\,clocks}{5clocks}) \times 100\% = 20\% \qquad (1)$$

B. Faulty Cases

1. Assume that the link between node 1 and node 5 is failed; then the results will be

Table 2. The network frame traffic under both Std. HSR and RLP-link failure case

#	Frames-Standard HSR	Frames - RLP	Frame Reduction Percentage
1	20	4	80%

And TIT = 20%

In this case, the frame-reduction percentage is increased, but one of the two logical paths for the connection pair between node 1 and 2 is lost.

2. Assume node 5 is failed; then the results will be

Table 3. The network frame traffic under both Std. HSR and RLP-node 5 failure case

#	Frames -Standard HSR	Frames - RLP	Frame Reduction Percentage
1	13	5	61.5%

And TIT = 20%

3. Assume node 4 is failed; then the results will be

Table 4. The network frame traffic under both Std. HSR and RLP-node 4 failure case

#	Frames -Standard HSR	Frames - RLP	Frame Reduction Percentage
1	16	6	62.5%

And TIT = 0%

Furthermore, the RLP algorithm can also be applied to a ring topology that has no Quadbox nodes, but it will be useful because the network performance will be the same of the standard HSR protocol and the Quadbox nodes will need bigger size of memory.

Ⅳ. Conclusions

In this paper, we present a new algorithm called redundant logical paths (RLP), which is suited to any closed-loop network topology. This algorithm solves the unnecessary traffic problem in the standard HSR that is caused by duplication and random forwarding. The idea of our algorithm is to establish two logical paths between each HSR node and all the other nodes, which eventually will establish a logical, full-meshed network among all the nodes, except the Quadbox node type, which is used only for interconnecting purposes. Furthermore, the RLP algorithm can also be applied to a ring topology that has no Quadbox nodes, but it will be useful because the network performance will be the same of the standard HSR protocol and the Quadbox nodes will need bigger size of memory. Instead of using the standard HSR protocol, which depends on duplication and forwarding, these logical paths will be used for the unicasting traffic type. Thus the RLP algorithm will significantly reduce the frame traffic in HSR networks, which will result in enhanced traffic performance. For the sample network in this paper, simulation results show 61.5 - 80% reduction in network frame traffic compared to the standard HSR as presented in tables 1, 2, 3 and 4.

The RLP algorithm is suitable for the unicasting traffic type within any closed loop network topology. In our future research, we will expand our algorithm to cover multi/broadcasting traffic.

References

[1] IEC 62439-3, High-available Automation Networks, Part:3 Parallel Redundancy (PRP) and High-available Seamless Redundancy (HSR), pp.3036-3054, 2010.

[2] Jean-Charles Tournier, Karl Weber, and Clemens Hoga, Precise Time Synchronization on a Highly Available Redundant Ring Protocol, IEEE International Symposium on Precise Clock Synchronization, Brescia, Italy, Sep. 2009.

[3] Jian-cheng Tan and Wenpeng Luan, IEC61850 Based Substation Automation System Architecture Design, IEEE-Power and Energy Society General Meeting, July 2011.

[4] D.H. Lee, Y.Z. Cho, H.A. Pham, J.M. Rhee, and Y.S. Ryu, “SAFE: A Scalable Autonomous Fault-tolerant Ethernet Scheme for Large-scale Star Networks,”IEICE Trans. Communications, vol.E95-B, no.10, pp 3158-3167, Oct. 2012.

[5] IEC 61850-90-4, Network Engineering Guideline for Communication Networks and Systems in Substations, 2011.

[6] Hubert Kirrmann, HSR - High Availability Seamless Redundancy, Fault-tolerance in Ethernet networks IEC62439, Presentation slides available on http://wenku.baidu.com/view/f32178f2f61fb7360b4c65e8.html, 2012.

[7] Hubert Kirrmann and Oliver Kleineberg, Seamless and Low-Cost Redundancy for Substation Automation Systems (High availability Seamless Redundancy, HSR), IEEE Power and Energy Society General Meeting, 2011.

[8] Xiaozhuo Jiang, High Availability Seamless Ring Protocol Implementation in FPGA, Master Thesis, Computer Engineering and Networks Laboratory, ETH Zurich, 2009.

[9] S.A. NSAIF and J.M. RHEE, Improvement of High-availability Seamless Redundancy (HSR) Traffic Performance, ICACT IEEE 2012.

[10] S.A. NSAIF and J.M. RHEE, Improvement of High-availability Seamless Redundancy (HSR) Traffic Performance for Smart Grid Communications, Journal of Communications and Networks, (forthcoming).

AUTHORS

Saad Allawi NSAIF

- 1999년 6월 : Univ. of Baghdad Iraq B.Sc in Electrical Eng.
- 2002년 6월 : Univ. of Baghdad Iraq M.Sc in Computers and control Eng.
- 2002년 ~ 2004년 : Assistant lecturer, Univ. of Baghdad
- 2004년 ~ 2011년 : Director of the Command and Control Systems (C2) in Iraqi Ministry of Defense.
- 2011년 9월 ~ 현재 : 명지대학교 정보통신공학과 박사과정

<관심분야> : Military communications, ubiquitous networks and smart grid communications.

전 제 현(Jehyun Jun)

- 1990년 2월 : 인하대학교 전자계산학과 학사졸업
- 1990년 1월 ~ 현재 : 삼성탈레스 C4I 연구소 수석연구원

<관심분야> : 군 이동 무선 통신 시스템 및 프로토콜 설계/구현, QoS, Routing 알고리즘

신 상 헌(Sangheon Shin)

- 1998년 2월 : 영남대학교 전자공학과 학사졸업
- 2000년 2월 : 영남대학교 정보통신공학과 석사졸업
- 2004년 2월 : 영남대학교 정보통신공학과 박사졸업
- 2004년 ~ 2005년 : 미국 NIST 초청 연구원
- 2005년 ~ 2007년 : 인텔 코리아 R&D 센터 과장
- 2007년 ~ 2009년 : POSDATA Flyvo 연구소 차장
- 2009년 ~ 현재 : 삼성탈레스 C4I 연구소 전문연구원

<관심분야> : WLAN, WiMAX, QoS, 군 전술통신망, 네트워크 M&S

이 종 명(Jong Myung Rhee) **정회원**

- 1976년 2월 : 서울대학교 전자공학과 (공학사)
- 1978년 2월 : 서울대학교 전자공학과 (공학석사)
- 1987년 12월 : North Carolina State Univ. ECE Dept. (공학박사)
- 1978년 ~ 1997년 : 국방과학연구소 책임연구원
- 1997년 ~ 1999년 : 데이콤 연구소 부소장
- 1999년 ~ 2005년 : 하나로텔레콤 CTO (부사장)
- 2006년 9월 ~ 현재 : 명지대학교 정보통신공학과 교수

<주 관심분야> : Military Communication, Fault Tolerant System, Ad-hoc, Data Link, Convergence, Smart Grid Communications