

# Industrial Ethernet seamless redundancy and sub-microsecond clock synchronization with IEC 62439-3 and IEC 61588

Prof. Dr. Hubert Kirrmann  
ABB Switzerland  
hubert.kirrmann@ch.abb.com  
Claudio Honegger  
claudio.honegger@ch.abb.com

Ioannis Sotiropoulos  
ABB Switzerland  
Ionanis.sotiropoulos@ch.abb.com  
Diana Ilie  
ABB Switzerland  
diana.ilie@ch.abb.com

## Abstract

**1.** IEC 62439-3 specifies two complementary seamless redundancy protocols applicable to hard real time systems, based on duplication of the transmitted information, with recognition and removal of duplicates in each node. These protocols can be applied to arbitrary topologies (PRP), but in particular to a ring, or to rings of rings (HSR). Its main application is in substation automation based on the IEC 61850 standard which references IEC 62439-3 where redundancy is required. Clock synchronisation with PTPv2 (IEC 61588-2010) allows sub-microsecond accuracy by using peer-to-peer transparent clocks. PTP cannot be treated as duplicate traffic in PRP and HSR since duplicates can take different paths. Therefore, the different paths are treated as carrying the same clock information, while preserving the pairing of PTP messages. This contribution describes the basics of IEC 62439 and the details of clock synchronization in PRP and HSR, including the connection of a PRP to an HSR network. Measurement results are shown for HSR that confirm the validity of the concept.

## 2. Introduction

The Industrial Ethernet suite IEC 61784 [2] describes a set of automation network technologies that have in common the IEEE 802.3 frames and physical layer, but differ widely on the upper protocol stack.

Outside of the automation field, similar industrial Ethernet developed for substation automation (IEC 61850) or avionics (ARINC 6)

While some solutions rely exclusively on the IP/TCP/UDP protocol suite, others may use layer 2 (link layer [3]) transmissions for time critical data, especially exploiting the broadcast or multicast ability available at level 2.

The protocol RSTP specified in [9] is widely used for this purpose, but recovery and reconfiguration cost some

time. IEC 62439-1 discloses a calculation method to estimate the worst-case RSTP reconfiguration time for different topologies, in particular rings. In typical industrial networks, recovery can take some 100 ms, which can be too long for certain time-critical applications.

One application field where actual grace times for applications are much lower is the field of power utility automation and substation communication systems as defined in [10]. IEC 61850-9-2 transmits sampled values (SV) at a rate of 4.8 kHz, an interval of 208  $\mu$ s. The loss of several samples in series can lead to a malfunction of the protection or to an inappropriate shutdown.

To address time-critical systems, IEC 62439 [4,6] specifies two redundancy protocols at layer 2 with a zero recovery time, PRP and HSR.

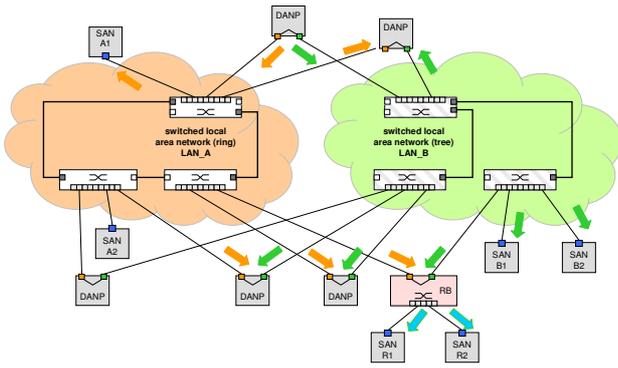
These protocols operate by sending pairs of frames over two independent paths, and letting the receiver discard the frames at the receiver. While this scheme works well for usual traffic, clock synchronization presents a challenge since the (redundant) messages carrying the time and correction can be delayed differently depending on the path they transit.

This applies especially to the precision time protocol PTP (IEC 61588).

This contribution shows how clock synchronization is handled in PRP and HSR and which practical results were achieved.

## 3. Redundancy Protocols

### 3.1. PRP Principle



**Figure 1. PRP principle.**

Figure 1 shows a typical PRP network. Redundancy-enabled nodes are called DANP (doubly attached nodes with PRP protocol).

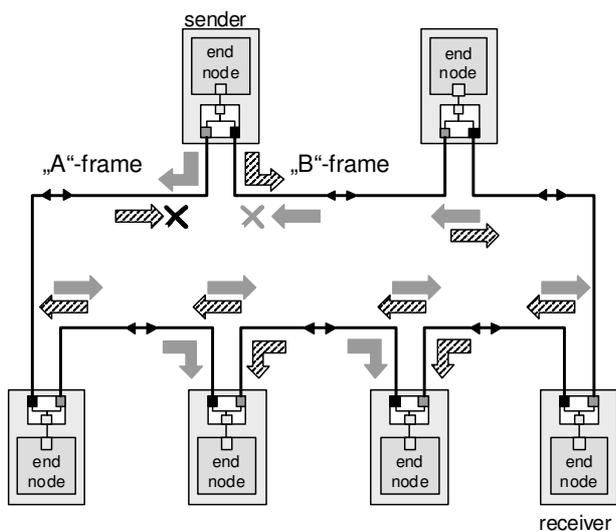
The source of a frame sends the same frame over two different, independent networks. The two frames travel simultaneously. The DANP receiver(s) accept the first frame and discard the duplicate.

To allow detection of duplicates, each frame carries a sequence number. This sequence number is appended within a redundancy control trailer to the frames, so it remains transparent to the application. This allows to connect usual devices such as printers as SANs (Singly attached nodes) to one LAN only, without any modification.

### 3.2. HSR principle

HSR (IEC 62439-3, Clause 5) [8] also provides seamless redundancy, but with only one additional link. It relies on the same principle of duplication of the frames and discarding of duplicates.

HSR is explained with the ring structure shown in Figure 2.



**Figure 2. HSR principle.**

Each node has at least two Ethernet ports of similar capabilities; this is called a DANH (doubly attached node for HSR). The nodes are daisy-chained into a ring.

Indeed, the network topology in automation is mostly a combination of linear segments rather than a meshed star topology network. A structure of rings is therefore the preferred redundancy structure. As nodes with two Ethernet ports and an integrated MAC relay are becoming common, the basis for redundant rings is already there.

Each node implements the forwarding capability described in [9].

The sender of an Ethernet frame sends two copies of the same frame (nearly) simultaneously through each port in both directions in the ring. The two frames of a pair travel in opposite direction. Each node in the ring forwards the frames, except if it is the destination of the frame.

In the fault-free state of the ring, each destination node receives two identical frames, passes the first frame of a pair to its upper layers and discards the duplicate.

In case of loss of a frame, of a link or of a node, only one frame is lost. The application on the destination node operates with the remaining frame undisturbed, but may signal the absence of the second frame of a pair for diagnostics purposes.

When the frame is a multicast or a broadcast frame, the node that injected the frame into the ring does not forward it any more when it comes back to it.

To cope with a situation in which the originator of the frame was removed from the ring or in which a frame is truncated or so badly damaged that its source cannot recognize it: every node is responsible to detect duplicates and remove them from the ring. Thus, an erroneous package will be discarded at the beginning of the second round as it has been already been seen by the ring devices.

The penalty of HSR is that overall traffic is doubled on the ring, but communication can sustain the loss of any frame, link or node that is neither source nor destination. .

The worst case delay over the network is however, if the network is in a healthy state, half that of a normal ring with RSTP.

If the network has sustained one failure and is in a degraded state, the worst case delay is that of a similar RSTP network, as a frame will potentially have to travel through the whole network.

### 3.3. Node structure

Figure 3 shows the architectural view as introduced by IEEE 802.1 (the so called baggy pants model) [9]. HSR is a component that is embedded in the forwarding process and has an additional protocol entity between the protocols using HSR and the forwarding process. The redundant forwarding extension represents a alternate

forwarding method compared to the standard forwarding method.

Figure 3. HSR in the IEEE 802 context

Figure 4 shows the basic structure of an HSR node.

Each node in HSR has two Ethernet network adapters. These two network adapters form a single interface with the same MAC address and present the same IP address(es); therefore, HSR is a layer 2 redundancy. This allows all protocols using IEEE 802 link layer to operate as usual and simplifies engineering.

An additional layer is introduced in the (otherwise unmodified) forwarding component, the LRE (Link Redundancy Entity), that handles two Ethernet controllers and presents the same interface towards the upper layers as a single Ethernet interface.

In a sending node, the LRE duplicates the frame it received from the above layer and sends the frames over both adapters at nearly the same time

In a receiving node that is the destination, the LRE receives the same frame from both interfaces, within some time skew since one frame may come from the neighbor node and the other has to propagate through the whole ring. The LRE discards the second frame of a pair to offload the application.

In a forwarding node, the LRE passes the frame from one ring port to the other, except if it already forwarded the same frame: the duplicate detection is used to discover and remove circulating frames.

If a link or a port is damaged, the LRE will still receive frames over the other path. So, in case of failure of one network component, data keep flowing over the other path without influence on the application.

The LRE could be implemented in software, like in PRP (IEC 62439-3 Clause 4). However, in a software implementation, the forwarding delay in each node is quite high (in the order of a millisecond) and the sum of the propagation times along the ring can exceed the required response time. Therefore, a hardware implementation of the LRE directly on the physical layer is preferable. This allows operating in the cut-through mode, i.e. a frame is forwarded before it is entirely received. Cut-through reduces the average forwarding delay, but it does not improve the worst case propagation delay, which happens when every node is sending an own frame at the same time, and delays the forwarded frames. Therefore, if an Ethernet protocol for a hard real-time communication system utilizing HSR wants to fully exploit its cut-through properties, it has to constrain the nodes to send their frames in a pre-allocated time window (TDMA), for instance using a common frame generator (such as Ethercat) or using a common precision clock (see 5).

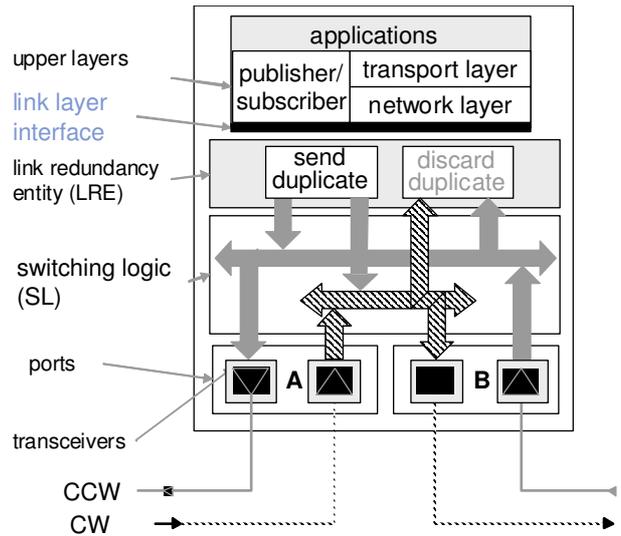
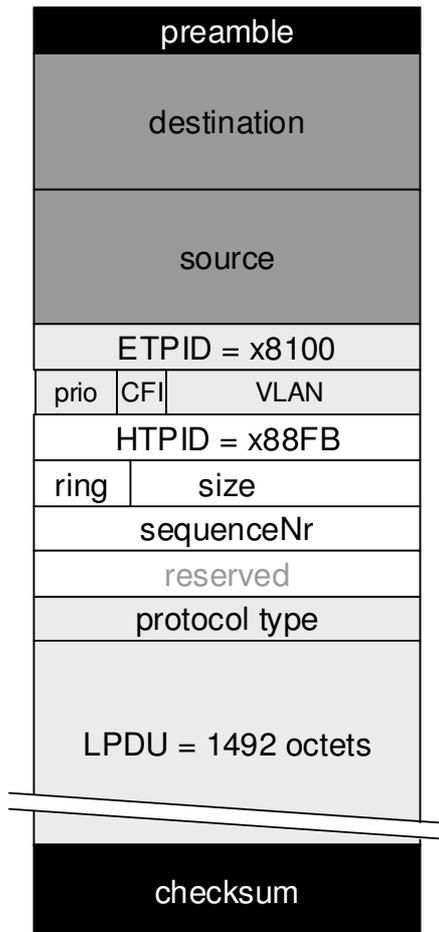


Figure 4. Node structure

### 3.4. Duplicate identification

The LRE is responsible to detect and discard duplicates. The LRE should rather accept a duplicate than discard it. The application normally can handle duplicates, but less well loss of frames.

Several methods can be applied to identify duplicates, for instance the duplicate discard algorithm used in PRP. To avoid checking complete frames, a special Ethertype (0x88FB) has been reserved for HSR, which allows introducing a tag between the link layer header and the payload. Frame with such a tag are handled like any other frame in an RSTP switch, if such are used in the network.



**Figure 5. Frame format (with VLAN tag)**

The HSR tag consists of four fields:

- The HSR tag (0x88FB)
- The size of the LSDU
- A sequence number
- A reserved field

A frame is uniquely identified by its source address, destination address, VLAN tag and sequence number. This means that the source of the frame is responsible for maintaining and incrementing the sequence numbers for each connection (unicast or multicast) it operates.

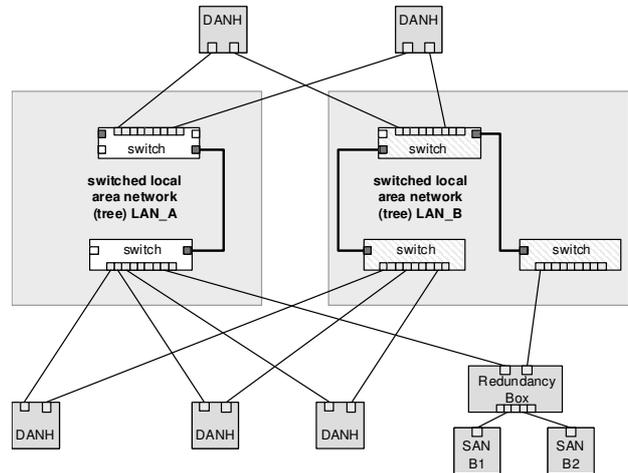
Contrarily to PRP, the LSDU size is not used to identify the HSR frames, since the Ethertype provides this function, but the size is a help for a hardware implementation.

The identification of the duplicates is then reduced to a comparison similar to the MAC address tables used in commercial switches.

## 4. Extension to general networks.

### 4.1. Network topology

HSR is not restricted to a ring topology. The same principle can be used to operate on two separate networks, like in the case of PRP (IEC 62439-3 Clause 4), see Figure 6. Thus, HSR is an universal approach that can be used in various topologies.



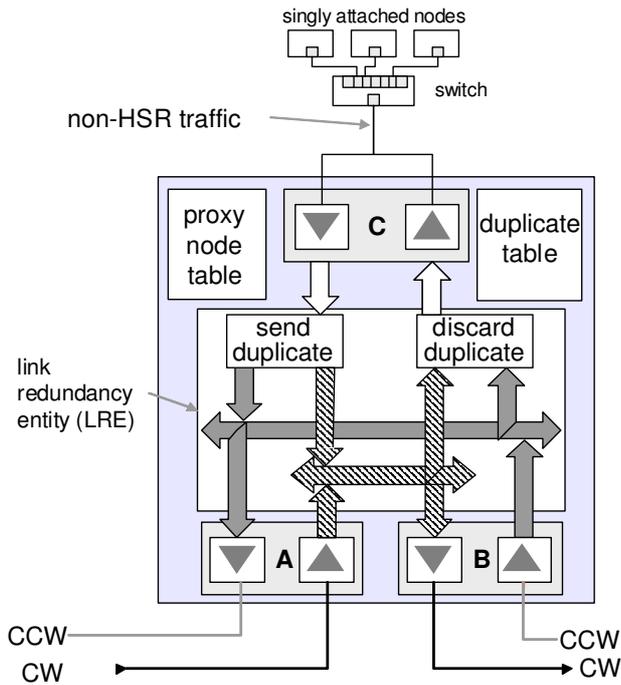
**Figure 6. Example of HSR Topology**

Each HSR DANH is attached to two independent LANs of similar topology. The networks are completely separated and are assumed to be fail-independent. The networks operate in parallel, thus providing a seamless recovery and allowing checking redundancy continuously to avoid lurking failures. Singly attached devices (SAN = non-HSR nodes such as a PC or a printer) can be attached to one network, but a SAN attached to one network cannot communicate with a SAN attached to the second network.

This means that there exist a traffic in the networks that is not HSR, and therefore usual network components such as RSTP switches are required to break possible loops. While an HSR node can understand non-HSR traffic, it must send frames without HSR tag to reach the SANs. The LRE does this distinction depending on the protocol tag. Note that the redundant networks do not carry the same traffic, but this is permitted.

### 4.2. Redundancy Box

SAN that should be connected to the two redundant networks must be attached through a “Red Box”, a device that behaves like a DANH toward the HSR devices and like a normal device towards the SAN. A RedBox can act as a proxy for a number of SAN devices, as Figure 7 shows.



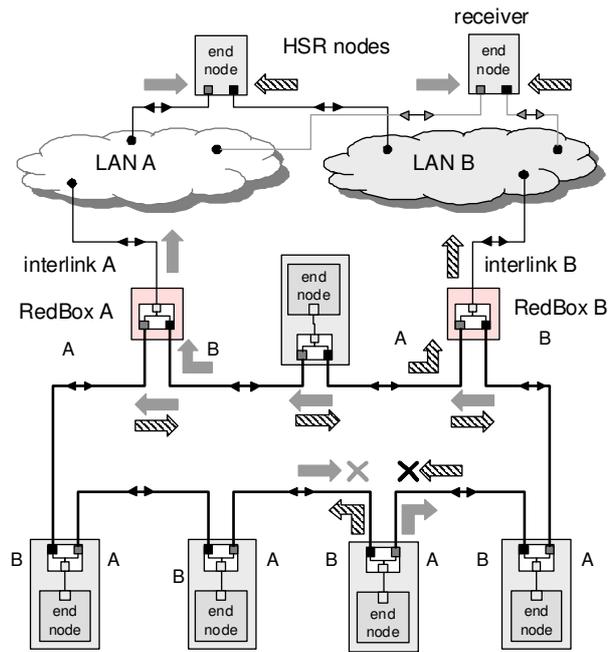
**Figure 7. RedBox**

The RedBox keeps a table of SAN, removes the frames that it injected into the ring and removes the HSR tag and duplicate frames.

In fact, HSR nodes will also perform to some extent the function of a RedBox, since they often have more than one Ethernet port, for instance for maintenance. The difference is that a RedBox supports a large number of SANs and can perform at the same time the role of a switch.

#### 4.3. Hierarchical connection.

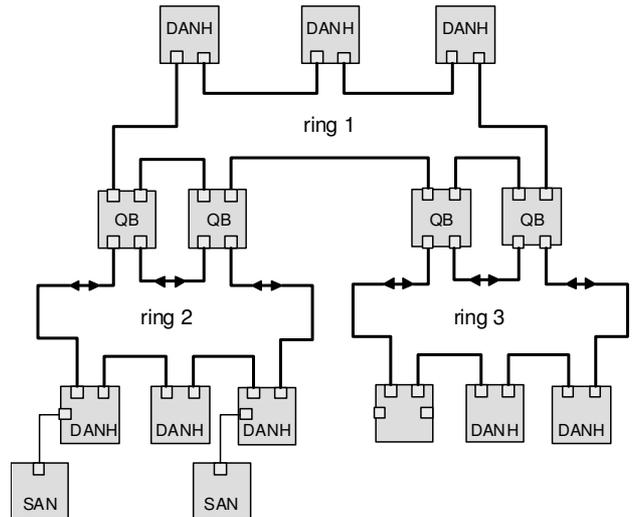
A ring can be connected to a duplicated network according to PRP as Figure 8 shows. Two RedBoxes are needed to prevent a single point of failure. Since each Redbox injects two frames into the ring, four copy of a frame will circulate, but as soon as a frame reaches a port that has already sent a copy, it will be discarded, so that the overall traffic consists just of the pair of frames. On reception, a RedBox only forwards one copy.



**Figure 8. Example of HSR Topology**

#### 4.4. Rings of rings

HSR allows building rings of rings, a preferred solution in substation automation.



**Figure 9. Ring of rings**

The rings are connected by Quadboxes, which consist conceptually of two Redboxes connected in series. Each ring in this network is an independent redundancy domain, which means that each ring can tolerate one single fault. If the Quad Box is realized in one single physical entity, the loss of one Quad Box puts both rings it interconnects into the degraded state.

#### 4.5. Generalized topology

HSR can be used in an arbitrary topology, provided each node has more than three connections and has the

capability to store enough connection identifiers to reject reliably the duplicates.

## 5. Clocks in PRP

### 5.1. HSR and clocks

Clock synchronization in HSR relies on IEEE 1588 V2 [14]. This standard compensates propagation delays in the network by sending a correction field together with the synchronization trigger frame. A pair of frames is often used to do this as the manipulation of the correction field has to be done while sending the frame.. If two-step transparent clocks are used, propagation delays in the network are compensated by sending the time through pairs of frames. The basic assumption is that the two frames of a sync pair take the same path. This condition is not guaranteed in HSR since one frame of a sync pair can propagate in one direction and the other take a different path. Therefore, clock synchronization in HSR relies on hybrid clocks, that is, a combination of transparent clock and ordinary clock. This way, synchronization always takes place on a node-to-node basis and the clock is inherently redundant. Alternatively, one-step transparent clocks can be used. In this mode of operation, no follow up messages with additional correction factors are needed as the precise correction values are inserted into the PTP frame on the fly while the frame is assembled and transmitted at wire speed. In this case the clock forwarding can be done as usual on both HSR paths.

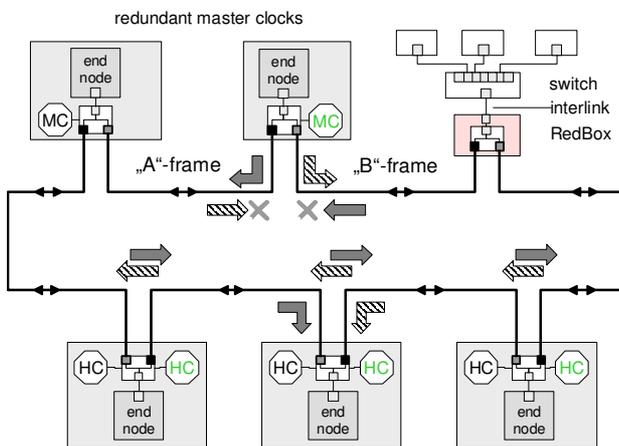


Figure 10. Clocks in HSR

This can be used as an additional means to synchronize e.g. IEC 61850 compliant merging units. A synchronization over the network with IEEE 1588v2 could augment or even replace current methods with IRIG-B or PPS signals.

## 6. Clock synchronization with PTP

Figure 11 shows the principle of the PTP communication.

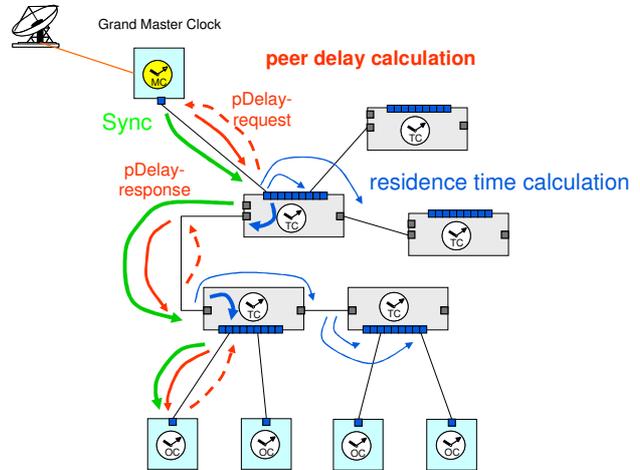


Figure 11. PTP principle

## 7. Conclusions

In order to fulfill the requirements of modern Ethernet networks with hard real time behavior and a near-zero frame loss tolerance, the HSR protocol is proposed. Where traditional, proven redundancy concepts and protocols fail to deliver redundancy on par with the requirements and constraints of e.g. IEC 61850 substation automation, process automation or even motion control applications, HSR provides seamless redundancy and a means of communication that suffices the needs of these demanding applications.

While HSR has been designed with the needs of these applications in focus and while the actual protocol implementation focuses on the fulfillment of their requirements, an HSR network structure is fully compatible with standard 802.3 Ethernet. The protocol is described in the International Standard IEC 62439-3 where actual implementations are based on. The protocol itself works on OSI Layer 2 and is transparent to the higher layer applications. To set up redundant network paths, the usage of switches is not necessary, as the HSR devices can be daisy-chained and interconnected to form a redundant network without the use of switches.

If in an application scenario, additional Ethernet switches are necessary; these network components can be low cost "commodity of the shelf" devices. In addition to this fact, the actual implementation on each device is simple and cost effective, as soon as dedicated hardware components implementing the dual Ethernet network interface and protocol stack become available.

With this outlook, HSR promises to be a means of creating network redundancy in application fields where

the use of Ethernet was not feasible up to now because of the restrictions, hard real time applications pose on the communication infrastructure. Now with HSR, Ethernet becomes a viable means on communication in demanding application scenarios aside the (proprietary) fieldbus systems.

In summary, HSR is a redundancy method that:

- Can be used with any Industrial Ethernet
- Can be used with any topology (tree, ring,...)
- Avoids the use of switches
- Is transparent to the application
- Tolerates any single network component failure
- Achieves zero delay recovery
- Is suitable for the most time-critical processes.
- Does not rely on higher layer protocols
- Uses off-the shelf network components (tools, controllers, switches and links)
- Uses a redundancy box to attach nodes not equipped for redundancy (SAN).
- Is standardized in IEC 62439-3(2010)
- Has the potential to emerge to the universal seamless redundancy method of industrial networks.

## References

- [1] The Institute of Electrical and Electronic Engineers, “CSMA/CD access method and physical layer specifications. IEEE Std 802.3”, 2005.
- [2] International Electrotechnical Commission, Geneva “IEC61784-2, Additional profiles for ISO/IEC 8802.3 based communication networks in real-time applications”, 2006.
- [3] The Institute of Electrical and Electronic Engineers, “IEEE Std 802.2 Logical Link Control”, 1985.
- [4] Kirrmann, H; Hoga, C.; Kleineberg, O; Weibel, H; “HSR: Industrial Ethernet redundancy with zero recovery time and low cost (High availability Seamless Ring, IEC 62439-3)”, *ETFA 2009*, 2009
- [5] Kirrmann, H.; Dzung, D.; “Selecting a Standard Redundancy Method for Highly Available Industrial Networks”, *2006 IEEE International Workshop on Factory Communication Systems*, June 27, 2006 Page(s):386 – 390
- [6] International Electrotechnical Commission, Geneva “IEC 62439 High Availability Automation Networks”, 2008
- [7] Kirrmann Hubert, Hansson Mats, Müri Peter, “IEC 62439 PRP: Bumpless Recovery for Highly Available, Hard Real-Time Industrial Networks”, *ETFA 2007*.
- [8] International Electrotechnical Commission, Geneva “IEC FDIS 62439-3 “High Availability Automation Networks – PRP & HSR”, April 2009
- [9] The Institute of Electrical and Electronic Engineers, “ANSI/IEEE Std 801.2D, Media Access Control (MAC) Bridges”, 2004.
- [10] IEC 61850: “Communication networks and systems for power utility automation”, available at [www.iec.ch](http://www.iec.ch)
- [11] IEC 61158-5-10: “Industrial communication networks – Fieldbus specifications – Part 5–10: Application layer service specification – Type 10 elements”, available at [www.iec.ch](http://www.iec.ch)
- [12] IEC 61158-6-10: “Industrial communication networks – Fieldbus specifications – Part 6–10: Application layer protocol specification – Type 10 elements”, available at [www.iec.ch](http://www.iec.ch)
- [13] IEC 61784-2: “Industrial communication networks – Profiles - Part 2: Additional fieldbus profiles for real-time networks based on ISO/IEC 8802-3”, available at [www.iec.ch](http://www.iec.ch)
- [14] The Institute of Electrical and Electronic Engineers, “ANSI/IEEE Std 1588, Precision Time Protocol (PTP), version 2, 2008.