The Parallel Redundancy Protocol for Industrial IP Networks

Markus Rentschler
Hirschmann Automation & Control GmbH
Stuttgarter Straße 45-51
72654 Neckartenzlingen, Germany
Markus.Rentschler@belden.com

Holger Heine
Siemens AG
Wernerwerkdamm 5
13629 Berlin, Germany
heine.holger@siemens.com

***Abstract*- The "Parallel Redundancy Protocol" (PRP) according to IEC 62439-3 realizes active network redundancy by packet duplication over two independent networks that operate in parallel. It has been specifically designed for industrial networks to meet highest availability requirements. In case of a single network failure, seamless redundancy is provided for data communication between PRP nodes that are connected to both networks. However, PRP was designed as a layer 2 Ethernet protocol that is transparent to higher protocol layers. In its current version, PRP is not able to support IP routing. This is because an IP router would change the source MAC address field of the Ethernet header which is used by a PRP receiver node for duplicate detection. In this work we propose and discuss a novel approach that keeps the PRP duplicate identification information across IP router boundaries.**

I. INTRODUCTION

In contrast to hot-standby switchover redundancy like the "Media Redundancy Protocol" (MRP) of IEC 62439-2 [1], the "High availability Seamless Redundancy" (HSR) and the "Parallel Redundancy Protocol" (PRP) of IEC 62439-3 [2][3] are active redundancy approaches that work without reconfiguration timeouts when a single failure in one of its two redundant network structures occurs. For PRP, this is achieved by the dual attached node (DAN) approach, which connects each end node into both networks and sends duplicated packets in both networks. For this, each DAN must be capable of PRP and discard the duplicated packet when received. The development of PRP in recent years was mainly driven for applications like process bus in power utility automation [4]. The basic network structure for PRP is depicted in Figure 1.

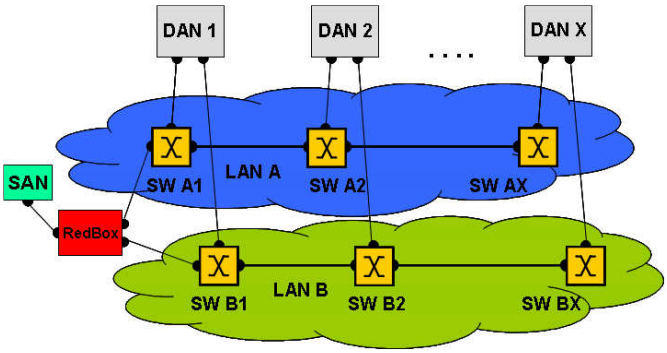

Fig. 1. Parallel Redundant Network

For better manageability larger networks with IP enabled devices are often separated in different IP address ranges. This is done by configuring the manageable network agents in different subnets on OSI layer 3. When this takes place within a flat layer 2 network with no routing devices involved, all Ethernet traffic is visible at every end node, even though they have been administratively separated by their IP address configuration. Some network designs however request to keep the traffic of the different subnets separate from each other. For this purpose, commonly IP routers are used. When such devices route packets across subnet boundaries, they replace the original source MAC address in the packet with the source MAC address of the destination router interface. But since PRP uses the originator's source MAC address for its duplicate detection mechanism, PRP will not work across such IP router boundaries as depicted in Figure 2. Due to this fact, the need for a redesigned protocol version of PRP arises that has the ability to work across router subnet boundaries. In this work, such a routing capable PRP version 2 (PRPv2) will be proposed and discussed.

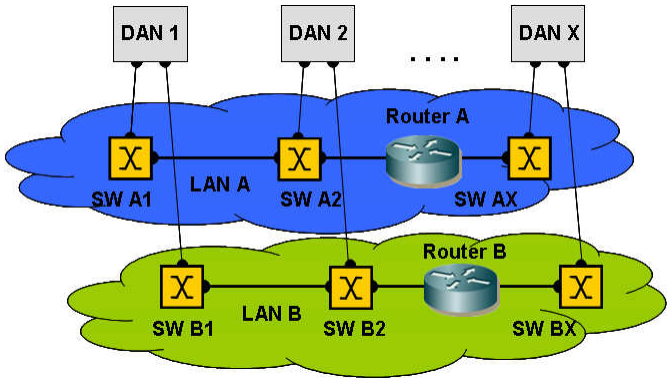

Fig. 2. Parallel Redundant Network with Routing

The paper is structured as follows: In chapter II, the operation principle of PRP will be explained, whereas chapter III briefly outlines IP routing and its impact when applied in PRP networks. Chapter IV presents possible scenarios and solutions to enable PRP for routed networks, the possible constraints are briefly mentioned. In chapter V, the identified best solution for PRPv2 is discussed in detail. Chapter VI summarizes the findings and gives an outlook to future work.

II. PARALLEL REDUNDANCY PROTOCOL

The Parallel Redundancy Protocol (PRP) according to IEC 62439-3, Clause 4 is a static network redundancy mechanism [2][3] that can compensate any single network failure. As an active redundancy scheme it does not require network reconfiguration and provides seamless failover without affecting the data transmission with packet loss.

PRP as a layer 2 redundancy operates independently of higher layer protocols. A PRP network consists of two different LANs with arbitrary, but similar topology (Figure 1). The similarity of the two networks is especially important in regard of different transmission delays which must not exceed the receiver window of the PRP nodes. A PRP node has two Ethernet interfaces, each connected to one of the two LANs, and is called a doubly attached node (DAN). Both PRP interfaces share the same MAC address. A PRP node transmits data simultaneously over the two interfaces into both networks. Each frame is tagged with a Redundancy Control Trailer (RCT) consisting of sequence number (SeqNr), LAN identifier (LanId), frame size (LSDUsize) and a fix PRP suffix (PRPsuffix). The resulting frame is illustrated in Figure 3.

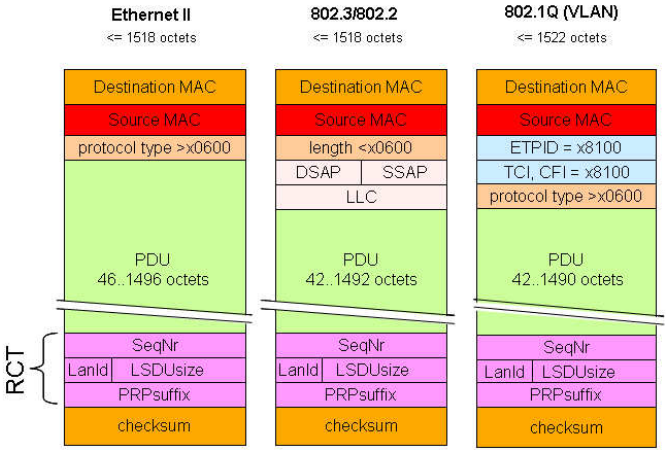

Fig. 3. Ethernet frame types with PRP Redundancy Control Trailer (RCT)

The PRP suffix allows proper identification, future extensions and coexistence with other protocols that are also using a trailer. For PRP according to [2], in the following referred to as "PRPv1", the suffix is defined with 0x88FB.

The sequence number is incremented for each frame pair sent. The first arriving frame of a pair, identified by its sequence number, is accepted by the PRP receiver node and the second frame is discarded. As long as one of the two LANs works properly, one frame of a pair always reaches its destination. Figure 4 indicates the architecture of a basic PRP communication system with two end nodes.

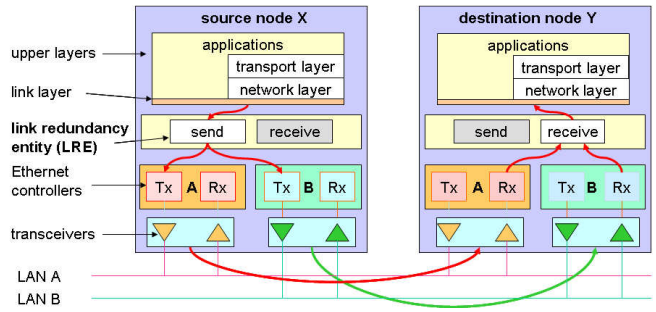

Fig. 4. Basic PRP Communication System Architecture

To use the PRP redundancy capability, non-PRP nodes must be attached through a Redundancy Box (Red Box), which is a device that behaves like a DAN.

Singly attached nodes (SANs) without PRP capability can also be attached to the redundant LANs (see Figure 1), since packets without an RCT are transparently processed by the Link Redundancy Entity (LRE).

PRP can also be used as an easy to apply but powerful solution to improve the transmission behavior of unreliable and diverse transmission media such as wireless radio [5][6].

III. SWITCHING VS. ROUTING

The Internet Protocol (IP) [7] is a connectionless protocol for use on packet-switched link layer networks, such as Ethernet. Since it operates on a best effort delivery model, it does neither guarantee delivery nor assures proper sequencing or avoidance of duplicate delivery. These aspects, including data integrity, must be addressed by an upper layer transport protocol, such as the Transmission Control Protocol (TCP) [8]. In order to allow the flexible grouping of devices in subnets and routing of packets between them, the IP addressing scheme was introduced on OSI layer 3.

In contrast to the MAC addressing scheme on OSI layer 2, IPv4 or IPv6 addresses are individually configurable by the network operator. However, since the actual packet forwarding always takes places on layer 2, MAC and IP addresses must have a defined virtual relationship to each other. This is for IPv4 achieved through the *Address Resolution Protocol* (ARP) according to RFC 826 [9] and the *Neighbor Discovery Protocol* (NDP) for IPv6 according to RFC 4861 [10].

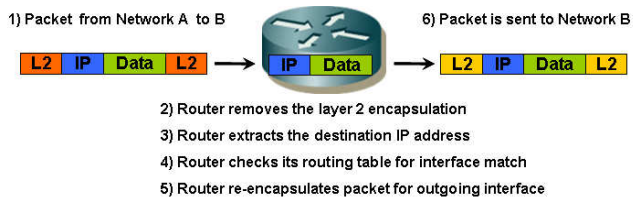

Fig. 5. Routing operation principle

"The main purpose of a router is to connect multiple networks and forward packets destined either for its own networks or other networks. A router is considered a layer 3 device because its primary forwarding decision is based on the information in the layer 3 IP packet, specifically the destination IP address. This process is known as routing. When a router receives a packet, it searches its routing table to find the best match between the destination IP address of the packet and one of the network addresses in the routing table. Once a match is found, the packet is encapsulated in the layer 2 data link frame for that outgoing interface. A router does not look into the actual data contents that the packet carries, but only at the layer 3 addresses to make a forwarding decision, plus optionally other information in the header for hints on, for example, QoS."[11].

IV. PRP AND ROUTING

In order to utilize PRP in an IP routing environment, different options will be discussed in this section:

A. Router supports PRP

PRP was designed to be a "redundancy in the nodes". It was already outlined that important PRP information gets lost when an Ethernet packet with PRP trailer is routed. One obvious solution to this would be to implement a PRP relay functionality that terminates the PRP network on either side of the router, thus containing a PRP LRE on each IP router interface (Figure 6).

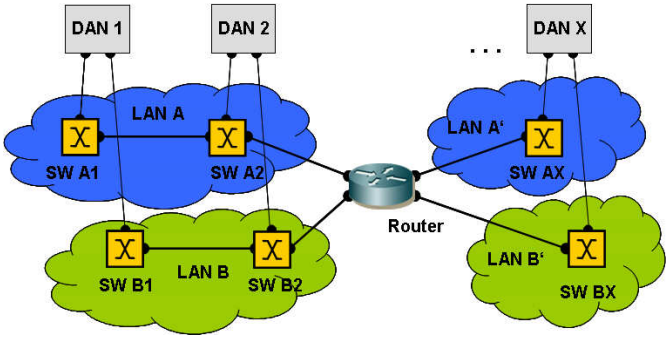

Fig. 6. Parallel Redundant Networks with PRP capable Router

Since this approach would introduce single points of failure into the network topology, and therefore contradicting the parallel redundancy paradigm, it is not leading to the desired result of having seamless redundancy in an IP network.

B. Router tunnels PRP

PRP could be tunneled over arbitrary networks via a suitable tunneling protocol [12][13][14], as shown in Figure 8, but this would not be the desired routing across IP subnets in conjunction with PRP. It would just be a way to connect plain layer 2 networks via tunneling-capable routers.

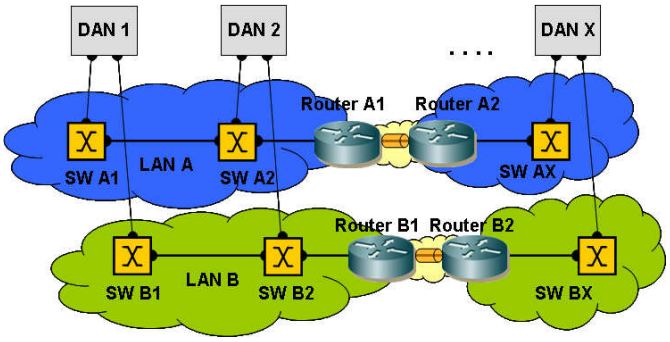

Fig. 7. Parallel Redundant Networks with PRP Tunnel via Router

C. PRP supports Routing

A suitable approach seems to be the modification or extension of the current PRP protocol definition in order to support layer 3 routing. In the following, possible solutions will be outlined.

C 1. PRP based on IP address

PRPv1 uses the pair of source MAC address and sequence number to identify duplicates. For support of parallel redundant IP routing, the source IP address and the sequence number in the IPv4 "Identification" field (Figure 8) could be utilized in a PRP like manner, albeit absolute uniqueness for the content of the "Identification" field cannot be guaranteed within a network. To realize this approach, the communication stack of the sending node simply has to send the duplicated frames via its two physical interfaces. On the receiving side, duplicated IP packets should be discarded by the standard IP stack anyway, thus additional duplicate discard processing might not even be required.

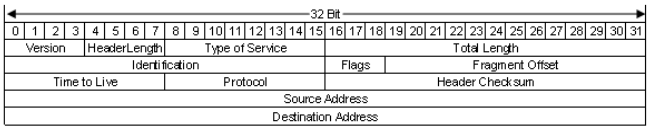

Fig. 8. IPv4 header.

For IPv6 [15], the "Identification" field has been removed, because fragmentation is handled differently. For the intended PRP-like approach, a suitable packet identifier is therefore no longer available in the IPv6 header (Figure 9).

But since IPv6 has introduced the concept of optional extension headers, this feature could be easily used to transport PRP related information in a newly defined extension header for PRP, which would perfectly resemble the PRP approach on layer 2.

Although obviously easy to realize and not even requiring a PRP trailer on layer 2, such IP based approaches would not be backward compatible to the MAC based PRP version.

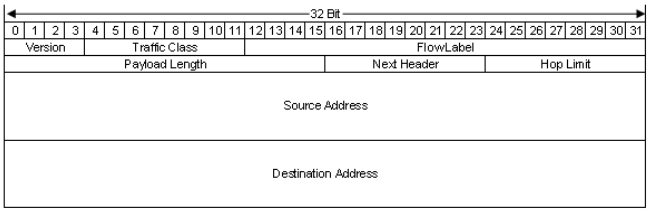

Fig. 9. IPv6 header.

C 2. PRP transports MAC address

In order to operate across any higher layer boundary, the required informational content for the operation of PRP must also be transported across this boundary. Besides the RCT with its sequence number, especially the source MAC address is required at the receiving node for duplicate detection. Since the originator source MAC address gets replaced by an intercepting IP routers destination interface MAC address, some means have to be defined to save this originator information in the routed packet. We propose the extension of the PRP trailer with a MAC address field as depicted in Figure 10.

The sending node of PRPv2 has to copy the LRE's MAC address not only to the Source MAC Field in the Ethernet

header but also to the corresponding field in the RCT. When the packet afterwards passes a Router that changes the Source MAC address in the Ethernet header, the original Source MAC address in the RCT remains unchanged. This implies an extension of PRP frames with additional 6 Bytes. Since PRPv1 already introduced oversized packets, devices have to deal with these packets in the same manner as PRPv1, but with an RCT of 12 Bytes instead of 6 Bytes.

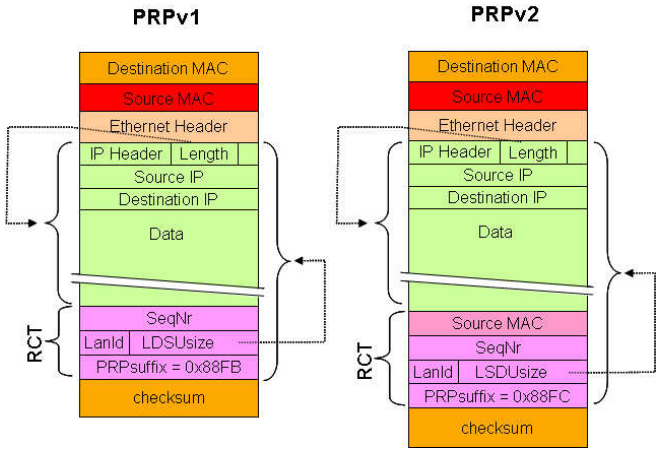

Fig. 10. PRP frame with encapsulated IP packet and extended RCT with Source MAC address field.

Additionally, the router has to add the original PRP Trailer from the incoming packet also to the outgoing packet, which requires a software extension in the router. As indicated in the routing algorithm of Figure 11, the router can even translate a PRPv1 packet into a PRPv2 packet, which is a further protocol feature.

```
1  //PRPv2 Routing Algorithm
2
3  for each Ethernet Packet do
4      if IP packet
5          if correct RCT for PRPv2 present
6              Keep RCT
7          else if correct RCT for PRPv1 present
8              Create RCT for PRPv2
9          end
10         Remove Layer 2 encapsulation
11         Extract destination IP address
12         Check routing table for outgoing interface
13         Re-encapsulate packet on Layer 2 for outgoing interface
14         Send packet on outgoing interface
15     else                                   // if Routing Switch
16         Switch packet on Layer 2 to destination port
17   end
18 end
```

Fig. 11. PRPv2 Routing Algorithm in Pseudo Code.

A receiving end node capable of PRPv2 will only rely on the Source MAC in the RCT for its duplicate discard decision.

V. Detailed Solution Discussion

In this section, the previously identified approach "PRP transports MAC address" requires an extended RCT (Figure 10) and the already mentioned routing extension for PRP in the involved routers (Figure 11). This PRPv2 will be further discussed in more detail as the best available solution.

A PRPv2-capable receiving decoder must be capable to distinguish the RCT of PRPv1 from an RCT of PRPv2, which is 6 Bytes longer. For the protocol version distinction a specific code should be employed. In the following text the suggested suffix code 0x88FC is used for PRPv2 (Figure 10). When a PRPv2 packet is received by a PRPv2 node, the receiving node can easily identify it as PRPv2 packet by this changed PRP suffix. In case the PRPv2 packet has not left the network, the Source MAC of the Ethernet header and the Source MAC in the RCT are identical.

A. Backward Compatibility Constraints

When creating PRPv2 as new protocol version of PRP, backward compatibility to the previous version PRPv1 is a key requirement. Ideally, both versions should be able to interoperate in the same network.

One advantage of PRP is the ability to include non PRP single attached nodes (SANs) in the same network. These nodes do not participate in the PRP redundancy mechanisms but can use the network infrastructure and communicate with PRP as well as non PRP nodes. Sustaining this advantage is a basic requirement of any PRP extension introducing new features and will not be violated by the additional insertion of the Source MAC address into the PRP trailer of PRPv2.

When a single attached node (SAN) without PRP capability receives a PRPv1 packet it cannot use the additional PRP information included in the RCT. The node simply treats the RCT as an additional padding and passes the packet to higher software layers without any discarding activities. There the packet is truncated of the RCT and treated as non PRP packet. Since a PRPv2 packet uses the same mechanisms like PRPv1 but only adds the Source MAC address and changes the PRP suffix from 0x88FB to 0x88FC, the same mechanism applies. Non-PRP nodes simply treat the PRPv2 RCT as padding and truncate the packets.

When applying the routing algorithm as described in Figure 11, line 7 and 8, the sending nodes must not even issue PRPv2 packets. It is entirely sufficient for all nodes to just send packets with PRPv1 trailer. When a node receives a PRPv2 packet, this just indicates that it has been routed before.

Nodes which have implemented the new PRPv2 are able to understand packets from PRPv1 nodes, but unfortunately not vice versa. A solution to this would be the previously mentioned approach that also PRPv2 capable nodes send only PRPv1 packets. In case of routing, the transition to PRPv2 is then performed only in involved routers. This approach allows full compatibility between PRPv2 and PRPv1 nodes at least in the same subnet.

B. Timing Constraints

PRP nodes are determined by the ability of their duplicate discard algorithm to reliably detect duplicates. Since PRP consists of two separate networks which can be build up differently a telegram and its duplicate may be received with a variable time spread. In a plain switched layer 2 network without routing or wireless elements, the maximum delay between a sending and a receiving node is well defined and rather small compared to the delay in a routed network. The time Δt_1 is defined by the transmission delay difference between both networks:

$$\Delta t_1 = \sum_{\forall SwitchesLanA} t_{delaySwitch} - \sum_{\forall SwitchesLanB} t_{delaySwitch} \quad (1)$$

A routed network may inflict additional delay and the two parallel networks can introduce a delay spread due to inhomogeneous networks. PRPv2 nodes have to deal with this possibility of asymmetric delay. The resulting time Δt_2 between a received packet and its duplicate in such a network is defined by the difference between the sum of the switch delays and the router delays in both networks, which can be unsymmetrical:

$$\Delta t_2 = \Delta t_1 + \sum_{\forall RouterLanA} t_{delayRouter} - \sum_{\forall RouterLanB} t_{delayRouter} \quad (2)$$

The delay in different networks may be rather long and unsymmetrical in routed networks compared to a plain layer 2 Ethernet network. For this, the algorithm of the duplicate discard algorithm has to be analyzed and adapted to fit this additional requirement.

C. Duplicate Discard Algorithms

Two different methods for a duplicate discard algorithm are presented in [16]. Both are shown in Figure 12 and described in this chapter. The referenced paper relates mainly to HSR [2], yet duplicate discard for PRP and HSR work similar. Therefore a HSR duplicate discard algorithm can also be used for PRP.

The first approach described in [16] is the "table of entries algorithm". As shown in Figure 12 a), each entry consists of Source MAC address and Sequence number, forming a tuple that uniquely represents a received frame. A duplicate discard decision is made by searching the table of entries. If the tuple of Source MAC Address and Sequence number of a received packet is already in the list, the packet is a duplicate and can be discarded. If the tuple is not listed, the packet is the first received and can be forwarded to the application. Additionally a new entry for the new tuple has to be generated. This tuple is then used to identify duplicates of this frame which may be received later.

For this algorithm it is important that in the time Δt_1 between a received packet and its duplicate only a certain number of packets are received. If more packets are received than there are entries in the list before the duplicate is received, the duplicate can not be detected since the original entry of the packet was already overwritten by new entries. Hence if the filter table is too short, duplicates can not be detected sufficiently. In a routed layer 3 IP network, the resulting time between a received packet and its duplicate is defined by Δt_2 and can even be unsymmetrical. PRPv2 nodes have to deal with this and if a "table of entries algorithm" is used, the number of elements in the filter has to be sufficient to fit the additional delay requirements. To meet increased timing requirements the size of the table of entries has to be increased at the cost of additional memory usage. Another way of improving the duplicate filter to meet PRPv2 requirements is to include a second table of entries. One is used for packets from the same network; thus packets not transmitted via a router. A second filter is used for packets from a distant network. Both kinds of packets can easily be distinguished in PRPv2 packets. If the Source MAC address in the Ethernet header and in the RCT is equal the packet was not transmitted through a router. If they differ, the packet was transmitted through a router since a router substitutes the Source MAC address in the Ethernet header.

a) "**Table of Entries"**

Src MAC Addr. 5	Seq. Nr. 4
Src MAC Addr. 1	Seq. Nr. 3
Src MAC Addr. 3	Seq. Nr. 4
Src MAC Addr. 1	Seq. Nr. 7
Src MAC Addr. n	Seq. Nr. 1
Src MAC Addr. 7	Seq. Nr. 9
⋮	⋮
Src MAC Addr. 1	Seq. Nr. 8

b) "**Sliding Window**"

Src MAC Addr. 1	Seq. Nr. $_{max}$	Window
Src MAC Addr. 2	Seq. Nr. $_{max}$	Window
Src MAC Addr. 3	Seq. Nr. $_{max}$	Window
Src MAC Addr. 4	Seq. Nr. $_{max}$	Window
Src MAC Addr. 5	Seq. Nr. $_{max}$	Window
Src MAC Addr. 6	Seq. Nr. $_{max}$	Window
⋮	⋮	⋮
Src MAC Addr. n	Seq. Nr. $_{max}$	Window

Fig. 12. Duplicate Discard Algorithms

The second approach for a duplicate discard filter described in [16] is the "sliding window algorithm" shown in Figure 12 b). This algorithm is based on the knowledge that every PRP sender increments the sequence number with every transmitted packet. A receiver sorts incoming packets according to their Source MAC address and spans a window over the highest received sequence number and previously received lower sequence numbers of a specific Source MAC address (Figure 12 b)). If a new packet with a higher sequence number than the highest sequence number in the window is received, the sliding window algorithm shifts the window's highest sequence to the newly received packet sequence number. If a packet with a smaller sequence number is received and marked in the sliding window algorithm this packet is a duplicate and can be discarded.

The advantage of the "sliding window algorithm" lies in the fact that it does not rely on the depth of a list of entries. An unlimited number of packets can be received between the reception of a packet and it's duplicate. However this leads to a higher implementation effort. Another restriction is the limited number of entries in a realization of the sliding window table. Since every source node requires such an entry, only a limited number of source nodes can be supervised.

Additionally, the window size is essential, because the window has to deal with packets that are received out of order and with the time difference Δt_1 or Δt_2 between the reception of a packet and its duplicate. The window must be large enough to reliably detect duplicates even if routers adding additional delay Δt_2 in PRPv2. A way of assuring this could be by the limitation of the number of packets a device may exchange with another device connected via router. This could be easily found out by comparing the Ethernet header Source MAC address with the PRPv2 Source MAC address in the RCT. If they differ, a router has exchanged the Source MAC address. Another way would be increasing the window size to meet increased timing requirements. Since the "sliding window algorithm" needs only one additional bit for each additional packet to monitor, the additional memory usage is rather small.

Considering advantages and disadvantages of both approaches, either duplicate discard filter algorithm can cope with the possibly increased timing requirements of routable PRP. Small improvements in the algorithms may be made to improve performance and reliability of standard PRPv1 traffic in an PRPv2 environment. Thus both algorithms could fulfill the needs for PRPv2. However, due to its higher tolerance regarding delay, the "sliding window algorithm" seems to be better suited for routed PRP.

VI. CONCLUSION

The "Parallel Redundancy Protocol" (PRP) according to IEC 62439-3 was designed for plain layer 2 networks to meet highest availability requirements due to its seamless redundancy. Applications like process bus in power utility automation [4] or industrial motion control have these demands, which can not be fulfilled by standard reconfiguration redundancy technologies like MRP. Additionally the support of single attached non-PRP devices (SAN's) in the PRP network without additional components makes PRP more flexible in contrast to HSR. It allows the use of low cost devices for uncritical communication without building another separate network infrastructure for these non PRP devices.

All these advantages lead to the assumption that the seamless redundancy methods of IEC 62439-3 will soon gain widespread acceptance, since in the near future an increasing number of PRP and HSR devices will be available and used in industrial installations.

However, the current focus of PRP on plain layer 2 networks may be unsatisfactory for a range of applications. In this paper some network scenarios and related issues regarding PRP and IP Routing were highlighted. Different proposals of how PRP can also work between routed subnets were outlined. A proposal for a new PRPv2 was discussed in more detail and some light shed on its interoperability with PRPv1. To implement the proposed version, only a slight extension to the protocol trailer and some software extensions in the router are required. In fact, the PRPv2 capable router must become aware of PRP trailers on incoming packets and simply keep them with the routed packet on the outgoing interface without disturbing other protocols.

Since these requirements for PRPv2 are not very demanding to implement, it would be very beneficial to include the proposed extensions in standardization and future implementations.

REFERENCES

[1] IEC 62439-2: "*Industrial communication networks: High availability automation networks*" – Part 2: Media Redundancy Protocol, available at www.iec.ch

[2] IEC 62439-3 (2012): "*Industrial communication networks: High availability automation networks*" – Part 3: Parallel Redundancy Protocol (PRP) and High Availability Seamless Redundancy (HSR), available at www.iec.ch

[3] H. Kirrmann, M. Hansson, P. Muri; "*IEC 62439 PRP: Bumpless recovery for highly available, hard real-time industrial networks*"; ETFA 2007, Patras, Greece

[4] D. Baigent, M. Adamiak and R. Mackiewicz, *"IEC 61850 Communication Networks and Systems in Substations: An Overview for Users"*. SIPSEP 2004, Monterrey, Mexico, available at http://www.gedigitalenergy.com/multilin/journals/issues/Spring09/IEC 61850.pdf

[5] M. Rentschler, P. Laukemann; "*Towards a Reliable Parallel Redundant WLAN Black Channel*"; WFCS 2012, Lemgo, Germany

[6] M. Rentschler, P. Laukemann; "*Performance Analysis of Parallel Redundant WLAN*"; ETFA 2012, Krakow, Poland

[7] RFC 791, "*Internet Protocol – DARPA Internet Program Protocol Specification*", J. Postel, September 1981, available at http://tools.ietf.org/html/rfc791

[8] RFC 793, "Transmission Control *Protocol – DARPA Internet Program Protocol Specification*", September 1981, available at http://tools.ietf.org/html/rfc793

[9] RFC 826, "*An Ethernet Address Resolution Protocol -- or -- Converting Network Protocol Addresses*", C. Plummer, November 1982, available at http://tools.ietf.org/html/rfc826

[10] RFC 4861, "*Neighbor Discovery for IP version 6 (IPv6)*", T. Narten et al., September 2007, available at http://tools.ietf.org/html/rfc4861.

[11] Wikipedia online article, "Routing (computing)", available at http://en.wikipedia.org/wiki/Router_(computing), retrieved 09.08.2012.

[12] RFC 1702, "*Generic Routing Encapsulation over IPv4 networks*", S. Hanks et al., October 1994, available at http://tools.ietf.org/html/rfc1702.

[13] RFC 2516, "*A Method for Transmitting PPP Over Ethernet (PPPoE)*", L. Mamakous et al., February 1999, available at http://tools.ietf.org/html/rfc2516.

[14] RFC 2661, "*Layer Two Tunneling Protocol (L2TP)*", T. Narten et al., September 2007, available at http://tools.ietf.org/html/rfc2661.

[15] RFC 1752, "*The Recommendation for the IP Next Generation Protocol"*, S. Bradner, A. Mankin, January 1995, available at http://tools.ietf.org/html/rfc1752.

[16] H. Heine, O. Kleineberg; "*The High-Availibility Seamless redundancy protocol (HSR): Robust fault-tolerant networking and loop prevention through duplicate discard",* WFCS 2012, Lemgo, Germany