|  |
| --- |
| Proposed Revision of FFIOT Redundancy Section |
| **Date: 2018-11-12 (r03: 2019-01-09)** |
| **Author(s):** |
| **Name** | **Affiliation** | **Phone** | **email** |
| Roger Marks | Huawei | 1-802-capable | roger@ethair.net |

## Abstract

This is a contribution to the IEEE 802 Network Enhancements for the Next Decade Industry Connections Activity (Nendica). It regards the FFIOT Work Item and comment resolution regarding the Call for Comments on the FFIOT Nendica Draft Report. The contribution proposes revision of the redundancy section of that draft report. It is proposed for initial discussion at the Nendica meeting of 7 November 2018.

## Background: Revising draft to focus on wireless enhancements to TSN

During prior comment resolution regarding the Nendica Draft Report on FFIOT (802.1-18-0025-06-ICne), editorial comments have been addressed. Technical comments remain to be addressed.

Among the technical comments (802.1-18-0052-02-ICne), several suggest that the draft report should be revised to establish, as a foundation, the current applicability of IEEE 802.1 TSN standards to the wired factory environment and then to specify the challenges associated with extensions to wireless communications in the factory. In particular:

* Comment 259: “The document seems to be really focusing on wireless, so the title should probably reflect that. Did not see much said about wired use cases. Also not much is said about 802.1 TSN capabilities.”
* Comment 265: “this should be discussed in the context of TSN. At least TSN management models for factories (e.g. 802.1Qcc) should be mentioned here.”
* Comment 266: “In addition to mentioning the challenges with wireless, the document could actually start with the trend to move from proprietary wired protocols to 802-based TSN over Ethernet. Wireless is the next step.”
* Comment 267: “This overview seems a bit too narrow. Given all the 802.1TSN work and its relevance to support industrial/factory applicaitons, this initial sections should give an overview of the TSN standards. TSN adds many relevant capabilities and some are being extended to wireless already.”

* Comment 269: “this should again refer to 802.1TSN fucntions. It seems the document so far disregard all the TSN work that has already been done.”
* Comment 278: “as in a previous comment, this should be discussed within a broader context of TSN capabilities. Qat is only one of the features.”
* Comment 279: “similar to the previous comment, this should be in the TSN context, also 802.1AS over 802.11 for time sync should be mentioned.”
* Comment 281: “other wireless specific challenges that could be measured include:
	+ PHY and MAC issues in latency, reliability and redundancy similar to wired TSN are also important.
		- the coordination would still require efficient PHY/MAC solutions in order to meet very low latency use cases
	+ industrial specific channel models to enable evaluation of wireless
		- a process for deploying and planing wireless systems in a factory is important. NIST's guide for wireless system deployment should be a good reference to add.”

Based on those proposals, this contribution proposes revision of the redundancy section of the draft report to structure it as a summary of TSN redundancy and a summary of challenges involved in extension to wireless. Other IEEE 802.1 TSN standards include including time synchronization (IEEE 802.1AS), stream reservation protocol (originally IEEE 802.1Qat), time-aware traffic shaping (originally IEEE 802.1Qbv), frame-preemption (originally IEEE 802.1Qbu/802.3br), TSN configuration (originally IEEE 802.1Qcc), etc. Discussion of those technologies should likewise be introduced in other sections of the draft report.

## FFIOT Redundancy section

The draft Nendica FFIOT report includes a section on “Wireless link aggregation designed for redundancy,” on pages 27-28. The essence of the section is that duplicate packets can be transmitted on two channels. The section does not make clear how the parallel transmissions could be used, but it discusses problems of “intermittent low bandwidth” and the need to “stabilize bandwidth and latency.” It appears that the intent may be to provide two opportunities to transmit a packet without loss or without unacceptable latency. It seems that the intention is to suggest that the first packet to arrive error-free would be accepted. The draft discusses this proposal in the context of IEEE Std 802.1AX link aggregation, which, as noted, applies both links to improve throughput and does not specify duplicate transmission. IEEE Std 802.1AX link aggregation is maintained by the IEEE 802.1 TSN Task Group.

## Background: Redundancy section comments

During comment resolution (802.1-18-0052-02-ICne), two technical comments were received on “Wireless link aggregation designed for redundancy” section:

* Comment 242: “Re "two wireless ports and each wireless port", what is a wireless port?”
* Comment 243: “Re "data is transmitted redundantly through the aggregated links", the draft talks extensively about interference and spectrum resource limitation. What would be the net effect of adding redundant transmitters everywhere (and thereby doubling the interference) in an interference-limited scenario?”

## Redundancy section in view of IEEE 802.1 TSN and earlier technologies

The functionality described generally in the FFIOT draft is specified in detail in IEEE Std 802.1CB-2017. IEEE Std 802.1CB, developed by the 802.1 TSN Task Group, specifies “Frame Replication and Elimination for Reliability” (FRER). It is “one of a number of IEEE 802.1 and other standards suitable for Time-Sensitive Networking (TSN) that together have the overall goal of providing extremely low packet loss rates and finite, low, and stable end-to-end latencies.” Per the Scope, it “specifies procedures, managed objects, and protocols for bridges and end systems that provide identification and replication of packets for redundant transmission, identification of duplicate packets, and elimination of duplicate packets.” The purpose is “to increase the probability that a given packet will be delivered.” FRER “can substantially reduce the probability of packet loss due to equipment failures.” The standard emphasizes improvement in loss, rather than latency, as might be expected in a controlled-latency TSN-based network. Because duplicate packets are eliminated, the earliest one to arrive is forwarded; this works to decreases jitter when latency in the paths is not well controlled.

Similar functionality is specified in an earlier standard, (IEC 62439-3:2016, “Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR).” This is a revision of a standard initially completed in 2010. IEC 62439-3 is “applicable to high-availability automation networks based on the Ethernet technology. It is “designed to provide seamless recovery in case of single failure of an inter-bridge link or bridge in the network” based on “parallel transmission of duplicated information.”

## Prior research on redundancy for wireless in industrial settings

IEC 62439 and IEEE Std 802.1CB primarily target wired networks. However, the use of that technology in the wireless case is not excluded and has been explicitly studied. For example, Rentschler and Laukemann presented a study at the 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012) (Markus Rentschler and Per Laukemann, “Performance Analysis of Parallel Redundant WLAN,” ETFA 2012, 17-21 Sept. 2012, sponsored by IEEE Industrial Electronics Society). Industrial applications were a key target. It noted that “wireless transmission is known to be error-prone and its error characteristics behave time-variable and non-deterministic. This labels wireless communication as not very well suited for industrial applications with tight reliability requirements, such as guaranteed maximum latency times for packet transmission.” The authors indicate that they consider “reliability, latency and jitter… as the most important criteria for industrial communication systems.”

Rentschler and Laukemann applied the standardized IEC PRP protocol to two parallel wireless LANs (WLANs) based on IEEE Std 802.11n; one of the two WLANs operated in the presence of interfering WLAN traffic. Regarding latency, the paper demonstrated that the minimum latency is attained *without* PRP, because the PRP processing adds delay. However, the maximum latency is best *with* PRP, because PRP chooses the packet arriving first. PRP improved jitter (average deviation of the mean latency) by about 40% in an example. The paper reported examples of packet loss at around 0.02% per WLAN without PRP packet errors due to the unlikelihood of simultaneous loss of both packets.

The paper does not address the considerations in Comment 243: “… the draft talks extensively about interference and spectrum resource limitation. What would be the net effect of adding redundant transmitters everywhere (and thereby doubling the interference) in an interference-limited scenario?” Indeed, one of the two available wireless channels in the experiment is dedicated solely to the link. In a busy factory, this doubling of allocated resources would presumably be limited to critical applications, but the effect on less critical applications may be non-negligible.

## Proposed resolution of Comments 242 and 243

This contribution proposes the following resolution of Comments 242 and 243:

•Comment 242:

* Resolution: *Revise*
* Resolution Detail: *Take no action, since issue is superseded by Comment 243.*

•Comment 243:

* Resolution: *Revise*
* Resolution Detail: *Replace the section “Wireless link aggregation designed for redundancy” with the section “Wireless link redundancy for reliability and jitter improvement” as specified in IEEE 802. 1-18-0064-00-ICne.*

## Proposed revision of “Wireless link aggregation designed for redundancy”

The following is proposed as a replacement for the content of the section “Wireless link aggregation designed for redundancy” of the FFIOT draft:

**Wireless link redundancy** **for reliability and jitter improvement**

Beginning in around 2012, efforts began in the IEEE 802 TSN Task Group to specify seamless redundancy in conjunction with TSN streams, particularly to address Layer 2 networks in industrial control and automotive markets. Eventually, this led to the completion and publication of IEEE Std 802.1CB-2017, specifying “Frame Replication and Elimination for Reliability” (FRER). IEEE 802.1CB provides specifications “for bridges and end systems that provide identification and replication of packets for redundant transmission, identification of duplicate packets, and elimination of duplicate packets.” Essentially, packets are duplicated and transmitted along differentiated paths; copies received at the destination, following the first, are discarded. The purpose is “to increase the probability that a given packet will be delivered,” and to do so in a timely manner. FRER “can substantially reduce the probability of packet loss due to equipment failures.” [IEEE Std 802.1CB includes the following note: “The term packet is often used in this document in places where the reader of IEEE 802 standards would expect the term frame. Where the standard specifically refers to the use of IEEE 802 services, the term frame is used. Where the standard refers to more generalized instances of connectionless services, the term packet is used.”]

FRER emphasizes improvement in loss, rather than latency. FRER is built upon earlier TSN standards and groups and, accordingly, presumes that frames are parts of a stream carried along a provisioned reservation. Accordingly, the latency of the reservation may be determined and presumed bounded; the bounds, however, depend on the reliability of the network along the reserved path. For some applications, this reliability limitation is insufficient. FRER can, in effect, provide instantaneous backup of each frame. This dramatically reduces the likelihood frame loss rate due to independent failure of identical equipment, roughly squaring it. For example, if each link experiences a frame loss rate of ε, FRER would be expected to have a frame loss rate of ε2. The difference may be highly significant in practice.

FRER is specified to apply only to frames carried in TSN streams. Not all streams in a network need to be subject to FRER; it can be limited to mission-critical streams only.

The concept of frame duplication and duplicate elimination preceded TSN discussions toward IEEE Std 802.1CB. In fact, the concept was standardized as early as 2010 in IEC 62439-3:2010, “Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR).” The standard supports the use of Ethernet in industrial applications. It is not based on TSN technologies and accordingly does not support the flexibility to sequence frames per stream. A number of industrial applications of PRP have followed.

The use of PRP wireless networks is not excluded and has been explicitly studied. This case is similar in principle but may be qualitatively different because the wireless link may be far more variable that the typical industrial wire link. As a result, a frame may be delayed significantly and unpredictably on a link without equipment failure. One implication is that, in the wireless environment, PRP may be more prominently used for jitter reduction rather than simply for frame loss.

Rentschler and Laukemann presented a study at the 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012) regarding PRP and wireless LAN (WLAN) [Markus Rentschler and Per Laukemann, “Performance Analysis of Parallel Redundant WLAN,” ETFA 2012, 17-21 Sept. 2012, sponsored by IEEE Industrial Electronics Society]. Industrial applications were a key target. It noted that “wireless transmission is known to be error-prone and its error characteristics behave time-variable and non-deterministic. This labels wireless communication as not very well suited for industrial applications with tight reliability requirements, such as guaranteed maximum latency times for packet transmission.” The authors indicate that they consider “reliability, latency and jitter… as the most important criteria for industrial communication systems.”

Rentschler and Laukemann applied the standardized IEC PRP protocol to two parallel wireless LANs (WLANs) based on IEEE Std 802.11n; one of the two WLANs operated in the presence of interfering WLAN traffic. Regarding latency, the paper demonstrated that the minimum latency is attained *without* PRP, because the PRP processing adds delay. However, the maximum latency is attained *with* PRP, because PRP chooses the frame arriving first. PRP improved jitter (average deviation of the mean latency) by about 40% in an example. The paper reported examples in which frame loss was around 0.02% per individual WLAN but in which frame errors were not observed using PRP due to the unlikelihood of simultaneous loss of both packets.

Rentschler and Laukemann study do not address the resource requirements necessary to implement PRP. In the wired case, whether PRP or FRER, the additional bandwidth resources to support redundancy may be supported by a cable and some switch ports. However, in the wireless case, the primary resource is a radio channel. As noted, one of the two available wireless channels in the Rentschler and Laukemann experiment was dedicated solely to the link. However, as discussed throughout this report, spectrum resources are limited in the factory environment. Each duplicated frame consumes twice the spectral resource of a single frame. If interference and channel availability are limiting factors, transmitting each packet in duplicate seems likely to be counter-productive. However, in some circumstances, such as for low-bandwidth mission-critical control messaging, duplicate wireless transmission might prove effective.

Another issue that needs to be considered regarding the application of PRP or FRER duplication in the wireless setting is the degree to which the pair of wireless channels is independent. For many realistic scenarios, such independence is a reasonable assumption in many wired networks. In the wireless case, the LAN elements may be physically separate, but the wireless environments may nevertheless be correlated. Operating the two links in different radio channels, or better yet different radio bands, can help to separate the interference conditions. However, even then, it is easy to imagine scenarios that would result in simultaneous degeneration of both links. One example might be a broadband noise source that affects both channels. Another is example is that of large moving machinery, such as a moving truck discussed earlier in this report, that blocks the direct line-of-sight of two antennas.

A number of WLAN applications of PRP have since been discussed in the literature, and wireless industrial applications of PRP have been introduced in the market, primarily regarding WLAN. However, no wireless applications of IEEE Std 802.1CB have been identified in the literature. Perhaps the best explanation is that 802.1 TSN is rarely implemented in wireless networks and wireless traffic is rarely carried in TSN stream reservations, and therefore 802.1CB FRER is inapplicable. Should 802.1 TSN functionality, including TSN streams, become introduced into wireless networks, techniques like FRER could be considered. However, it appears that some additional complications could arise. For example, FRER relies on sequence numbering in which the number of bits required depends on the maximum possible path latency difference that needs to be accommodated. In the wireless case, given the expected difficulty in ascertaining a tight latency bound, that number could be difficult to assign or could be impractically large without improvements in network control and management.

Concepts like FRER may find application in contributing to improved reliability and jitter in wireless factory networks. However, some of the challenges discussed will first need to be addressed and resolved.