Seamless Redundancy for High Reliability Wi-Fi

Gianluca Cena*[iD], Stefano Scanzio*[iD], Dave Cavalcanti†[iD], and Valerio Frascolla‡[iD]

*National Research Council of Italy (CNR–IEIIT), Italy.
†Intel Labs, Intel Corporation, Hillsboro, OR, USA.
‡Intel Labs, Intel Deutschland, Neubiberg, Germany.
Email: gianluca.cena@cnr.it, stefano.scanzio@cnr.it, dave.cavalcanti@intel.com, valerio.frascolla@intel.com

Abstract—By removing wire harness, Wi-Fi is becoming increasingly pervasive in every aspect of our lives, in both the consumer and industrial worlds. Besides flexibility, the recent high efficiency and extremely high throughput versions managed to close the performance gap with Ethernet. However, it still lags behind Ethernet for what concerns dependability. To this aim, the ultra high reliability study group has been recently formed.

This paper reports on some preliminary ideas and proposals about the ways seamless redundancy can be exploited to make Wi-Fi more reliable, yet retaining a good degree of backward compatibility with existing network infrastructures.

I. INTRODUCTION

A common way to improve dependability of systems and networks is to exploit redundancy (space, frequency, and time) [1]. The High-Availability Seamless Redundancy (HSR) and Parallel Redundancy Protocol (PRP), defined by IEC 62439-3 [2], permit to increase availability and reliability of Ethernet-based industrial networks, also achieving fault-tolerance. These solutions were taken as the basis for defining redundancy mechanisms for time-sensitive networking (TSN). In particular, IEEE 802.1CB [3], also known as Frame Replication and Elimination for Reliability (FRER), defines a quite generic framework for redundancy, which can also describe legacy PRP and HSR. Although FRER scope is not limited to IEEE 802.3 (Ethernet), its optimal usage with other network technologies, like the wireless ones, is not readily apparent.

It must be remarked that the redundancy protocols listed above were conceived for wired networks, where frame losses are typically negligible. So, they are mainly meant to cope with failures that affect physical links (wires) and intermediate equipment (switches). Such phenomena are usually *permanent*, which means that human intervention is required to restore the system to the original state, e.g., by replacing broken hardware. In the meanwhile, the surviving redundant path ensures applications uninterrupted operation, which is essential for mission-critical systems. Unlike spanning tree (STP and RST), *seamless redundancy* achieves zero-time intervention, which makes it suitable for time-critical systems as well.

About one decade ago, the use of seamless redundancy was suggested also for high-performance wireless local area networks (WLAN), like those based on IEEE 802.11 [4]. In [5] a proposal was made to employ a pair of Wi-Fi links tuned on different channels as cable replacements, each one implemented by means of a wireless client associated to the related access point (AP). By using conventional PRP equipment (RedBoxes) on the two sides of the connection, seamless redundancy is achieved on air, which permits adoption in safety networks. In fact, although the black channel does not demand extreme reliability, relatively short outages in Wi-Fi communication may cause timeouts to expire, which in turn uselessly trigger the safety-related functions, undermining the overall system availability [6].

Subsequent works postulated that seamless redundancy could be embedded directly in IEEE 802.11 specifications, as proposed for Wi-Red [7]. By doing so, duplication avoidance mechanisms can be used to lower bandwidth consumption, which is quite useful in the $2.4\,\mathrm{GHz}$ band. In Wi-Red, a redundant STA (RSTA) associates and exchanges frames with a redundant AP (RAP). Both the RSTA and the RAP include two (or more) subSTAs (embedded in the same case), tuned on distinct bands/channels, whose behavior closely resembles legacy Wi-Fi. Frames are duplicated on the transmitting side and deduplicated on the receiving side by a link redundancy entity (LRE), which coordinates sub-STAs' operation and aborts ongoing transmissions on all channels as soon as an ACK is received for the frame being sent.

More recently, the adoption of TSN concepts in Wi-Fi has been envisaged to make it suitable for time-aware applications as well. In particular, the use of FRER was considered as a way to support seamless redundancy with no (or minimal) changes to the IEEE 802.11 specification [8]. Although Wi-Fi 6/6E is at the moment exploited in new designs for automation systems, the forthcoming Wi-Fi 7, with its Multi-Link Operation (MLO) [9], is probably the best option for implementing redundancy in reliable wireless links.

It is essential to remark that seamless redundancy is applied to wired and wireless networks with quite different goals. In fact, in the latter case it is mainly meant to cope with frame losses due to disturbance, including both interference caused by nearby wireless devices and electromagnetic noise generated by high-power industrial equipment. Above phenomena are usually *temporary*, which means that no human intervention is generally required to recover from them.

In the real world, it is quite usual that about $10 \div 20\%$ of the transmission attempts performed on air fail, which is clearly unacceptable for any applications. To make communication reliable, a confirmation mechanism is customarily included in the MAC layer of every wireless protocol. On frame arrival, the recipient must return an ACK frame. If the related timeout expires before the ACK is heard, the sender transmits the frame again, until it either succeeds or

the retry limit is exceeded. Leveraging diversity, e.g., the Minstrel algorithm that dynamically selects the best modulation and coding scheme, helps further. Automatic repeat request (ARQ) mechanisms lower the amount of losses seen by datalink users dramatically, but they also increase transmission latency tangibly, making it unpredictable and often too long for control applications. Seamless redundancy (along with ARQ) improves both frame losses (it is sufficient that the frame arrives on one channel) and latency (the user is always delivered the fastest copy of every frame) at the same time.

It must be noted that, if fault tolerance has to be ensured to Wi-Fi through PRP, communication equipment must be replicated. For example, two distinct (R)APs can be deployed, to which two distinct (R)STAs are associated, as in [5]. In this way, any faults affecting a single device (a problem to the power supply, a broken antenna, etc.) do not disrupt communication. Distinguishing between reliability and fault tolerance is essential to understand what follows. The former aims at mitigating temporary disturbance, narrowing the gap between Wi-Fi and Ethernet in terms of the communication quality perceived by applications (e.g., packet loss ratio and deadline miss ratio). The latter is instead meant to counteract permanent failures, ensuring uninterrupted communication to distributed systems that are expected to be resilient. In the vast majority of the cases, what is sought from the next generation of Wi-Fi equipment is just a reliable behavior. This is certainly true for the consumer world, but it also holds in many industrial contexts. In fact, fault tolerance is expensive [10], and is only adopted when/where it is really needed.

Neglecting cost, one may argue that fault tolerance implies reliability. There is, however, another aspect to consider: PRP, HSR, and FRER are static solutions, and redundant paths are typically decided when the network architecture is designed. This is quite reasonable, since only a careful planning can provide a reasonable certainty about the effectiveness of redundancy, but it somehow betrays the plug&play nature of Ethernet and Wi-Fi, where prior configuration is not a prerequisite for communication (contrary to TSN). Typically, all it is needed to connect a node to the network backbone is related to IP (address, netmask, and default gateway) and security aspects (e.g., service set ID and WPA password). In a sizeable number of cases, it would be nice to enable reliability through redundancy without requiring the users any additional knowledge about the network. Besides multimedia communication with mobile devices, also non-safety and non-mission-critical data communication with mobile cobots (i.e., not requiring fault tolerance) could benefit from such a feature.

In the following, some preliminary ideas are presented that may help setting the path for the upcoming activities of the IEEE 802.11 Ultra High Reliability (UHR) study group, which will work on the definition of the next WLAN generation (Wi-Fi 8). Following this introduction, some relevant concepts about redundancy are briefly discussed in Sect. II, while the concept of High Reliability STA (HR STA) will be proposed in Sect. III, where its operation will be also sketched.

II. REDUNDANT COMMUNICATION BASICS

Below, two standard solutions aimed at supporting redundancy in IEEE 802 networks are briefly reviewed.

A. Wi-Fi MLO

One of the key features introduced by the IEEE 802.11be task group (Wi-Fi 7) is multi-link operation (MLO). In particular, a multi-link device (MLD), either STA (non-AP MLD) or AP (AP MLD), may send its traffic on two (or more) *affiliated* STAs, also denoted Low MAC (L-MAC), whose behavior is mostly independent (unlike channel bonding, where transmission/reception take place exactly at the same time on all the involved channels). We will focus on simultaneous transmit and receive multi-link multi-radio (STR MLMR) solutions, where transmission of different frames by the affiliated STAs (on distinct bands/channels) may overlap partially or totally.

In the current (draft) version of the specification, MLO is mainly intended to increase throughput and decrease latency, by letting a single entity in the MLD, the Upper MAC (U-MAC), decide at runtime which one among underlying L-MACs must deal with the transmission of any buffered frame. A multi-link element (MLE) is defined, included in Beacon and Association frames, that encodes the information needed to deal with this function.

Using MLOs for sending replicated copies of the frames is in theory possible, even though a clear description about the way this can be done is not included in the draft specification.

B. TSN FRER

Among the many features TSN adds to Ethernet, frame replication and elimination (FRER) aims at improving reliability by providing an agreed way to set redundant communication paths between end-points. The concept of compound stream is introduced, made up of several member streams. Streams can be split by using frame replication, or joined by using frame elimination. To enable these operations, frames in a stream are identified by sequence numbers, encoded in a suitable R-TAG. FRER functions are implemented in both end-nodes and TSN switches. By suitably configuring them, redundant paths can be set, which provide fault tolerance.

III. APPLICATION SCENARIOS FOR RELIABLE WI-FI

We consider contexts where infrastructure Wi-Fi networks are connected to a wired Ethernet backbone and possibly to the Internet through routers. We will assume that at least one of the two communication end-points relies on Wi-Fi. The case where both communicating devices use Wi-Fi (associated to either the same or distinct APs) is not irrelevant. A Wi-Fi node that supports seamless redundancy will be denoted for short HR STA or HR AP, depending on its role. Their behavior resembles RSTA and RAP in [7], respectively. To comply with Wi-Fi 7, they are implemented as multi-link devices (MLD), and basically consist of a pair of affiliated STAs. Every frame for which reliability is demanded is sent by the HR STA/AP on all its affiliated links at the same time.

Similarly to the access category (AC) foreseen by Wi-Fi to specify the frame QoS, a *reliability category* (RC) must
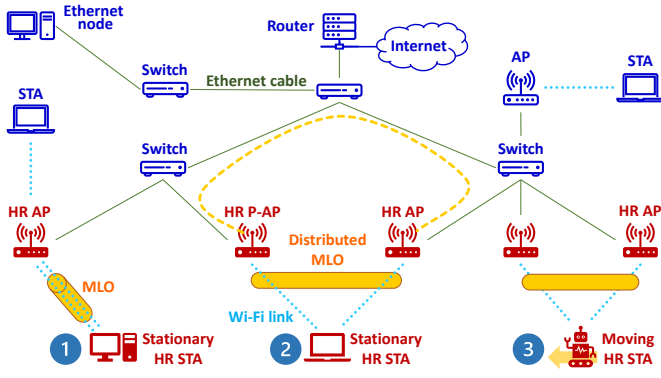
Fig. 1. Three relevant scenarios for HR Wi-Fi: 1) Stationary – Single AP; 2) Stationary – Multiple AP; and 3) Moving – Multiple AP.

be additionally defined that tells whether or not redundancy is exploited. Its exact definition is still to be decided: in fact, more-than-duplex redundancy is possible (also thanks to the new 6 GHz band), which means that different degrees of reliability can be envisaged.

The very first aspect to be taken into account when exploiting seamless redundancy to enhance reliability is where the frame duplication and deduplication functions have to be located. Below we will consider three scenarios, characterized by increasing complexity. What we seek in perspective is a single, unified solution, to be incorporated in Wi-Fi 8. Only reliable frame exchanges are considered, as those that do not need UHR service follow the default Wi-Fi behavior.

A. Stationary Wireless Device – Single AP

In the first scenario, depicted on the left side of Fig. 1 ①, all the affiliated STAs of an HR STA are associated to the same HR AP. By exploiting the advertising MLO functions, based on multi-link elements (MLE), both end-points of the redundant wireless link are completely aware of what is needed to support seamless redundancy operations. If required, the MLE format could be augmented to include additional information.

All duplication-deduplication operations are performed in the MLD of the HR STA and the associated HR AP. This means that any HR STA is seen by the other devices in the network as a conventional node, and this holds for legacy Ethernet and Wi-Fi devices, as well as for other HR STAs associated to either the same or other HR APs. They will be unaware of seamless redundancy and just see a more reliable end-to-end path. Every frame sent by the HR STA is deduplicated by the HR AP before being relayed on Ethernet or to another associated STA. To make redundancy completely transparent, one of the STAs in the HR STA is selected as the *primary STA*, and all frames exiting the associated HR AP include its MAC address as the source address (SA).

When the target is another HR STA associated to the same HR AP, the latter will duplicate the frame again on the related multi-link. Similarly, every frame arriving from Ethernet (or from another STA) is duplicated by the HR AP before being sent on the multi-link to which the target HR STA is

associated. The destination address (DA) of incoming Ethernet frames coincides with the MAC address of the primary STA of the HR STA, but in the relay process on Wi-Fi it is changed to match the MAC address of the actual receivers (affiliated STAs).

Encoding information about RC into the frames sent on air, e.g., as an MLE, could make operation easier. This is not out of question, as UHR is planned for the next Wi-Fi version, which has still to be defined. However, this goes beyond the level of detail foreseen for this paper.

B. Stationary Wireless Device – Multiple AP

In this second, static scenario, shown in the middle of Fig. 1 ②, the STAs affiliated to the HR STA are associated to different HR APs. As a consequence, duplication/deduplication can not be performed directly by such APs, as they only see part of the frame copies. This implies that, unlike the previous case, MLO can not be exploited to support redundancy. For this reason a "distributed MLO" concept has been conceived.

Redundancy as per TSN could be exploited, by leveraging FRER to perform duplication and deduplication in some point of the network located between the two HR APs and the destination node. To limit bandwidth consumption, FRER should be implemented in the nearest common ancestor of the different HR APs associated to the given HR STA. To this purpose, an R-TAG is added by the involved HR APs to the frames sent on Ethernet, which permits them to suitably travel on the wired portion of the network. Doing so is not trivial, especially in the case when wireless devices are moving, as it requires that: 1) the wired portion of the network employs TSN-compliant switches (which is hardly feasible/simple in the brownfield), and 2) FRER paths are configured in advance for all possible communication end-points.

A more practical solution, which preserves plug&play behavior, is to have frame replication and elimination performed entirely by a single AP, chosen among those affiliated to the involved HR APs, that will be termed *primary AP* (P-AP) for the distributed MLO. A sensible choice is to select as P-AP the AP associated to the primary STA of the HR STA. To forward the replicated copies to the P-AP for processing, a new *relay* tag (Y-TAG) is defined that implements sort of a source routing for redundancy in Ethernet. Besides a reserved value for the EtherType (on 2 B), which permits to detect the Y-TAG, it includes the next destination address (NA, on 6 B) to which the frame must be forwarded in the subsequent relay, a sequence number to manage elimination of replicated copies (on 2 B), plus (possibly) ancillary information, similar to what is needed to deal with MLO redundancy.

When a copy of the frame is received (on air) from the associated HR STA, all the involved non-primary HR APs relay it to the P-AP over Ethernet using the Y-TAG. Upon reception of one of these frames, the P-AP removes the Y-TAG and reconstructs the original frame by using NA as the new DA. It has all the information needed to understand that this is a copy that employs distributed MLO. Thus, deduplication can
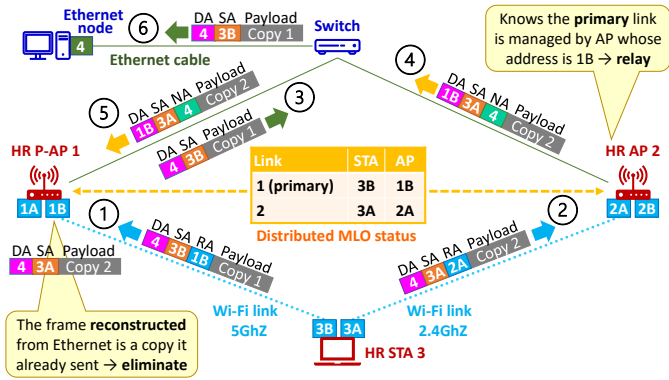
Fig. 2. Sample HR distributed multi-link operation: all copies are relayed to the primary AP (using the Y-TAG), which performs elimination.

be performed in the same way as if the frame were received on air from the HR STA associated to the P-AP. If the frame is not discarded, it is relayed to destination.

An example of distributed MLO is sketched in Fig. 2. HR STA 3 is associated to two distinct HR APs (1 and 2). A frame for which reliable service is demanded is sent contextually by both its affiliated STAs (3A and 3B). The copy on the left (from 3B) is sent on air to the P-AP (1B). When arriving at the non-primary AP (2A), the copy on the right (from 3A) is relayed to the P-AP (1B) on Ethernet. By exploiting the Y-TAG, where NA points to the final destination (4), the original frame can be reconstructed. The copy that arrives first to the P-AP (the one on the left, in the example) is relayed to destination, the other is discarded.

The reverse direction (from an Ethernet node to an HR STA associated to multiple HR APs) is managed similarly, with the P-AP that relays copies of the frame to the other involved HR APs for parallel transmission (on air) to the target HR STA.

C. Moving Wireless Device – Multiple AP

The third scenario, on the right side of Fig. 1 ③, closely resembles the second one, but moving nodes are now present. Ensuring reliable interconnection on air to roaming devices makes the link transition between different states of the former two kinds, and implies that associations to APs change over the time. So that communication quality is not impaired, reassociation must be performed one link at a time. In other words, when this process starts on one of the STAs of the HR STA, the other affiliated STAs are not allowed to do so until the ongoing reassociation finishes. Eventually, all STAs will be associated to a new HR AP.

Generally speaking, three types of transitions may occur while roaming: the former turns a single AP scenario into a multiple AP one (MLO to distributed MLO), while the second performs the reverse operation (MLO is managed more effectively than distributed MLO, and hence the HR STA should do its best to return to the single AP scenario). A third type of transition occurs when switching between different multiple AP scenarios. Overall, the protocol must deal with such transitions smoothly and flawlessly. To this aim, the use of artificial intelligence to select the best time/way for

reassociation is likely to bring tangible benefits. Instead, single AP to single AP transitions are not possible, as reassociation is performed one link at a time.

IV. CONCLUSION

High reliability is the next Holy Grail for Wi-Fi, as witnessed by the formation of the UHR study group. The solution we propose focuses on enhancing the Wi-Fi operation without requiring any changes to the already deployed Ethernet networks. Conversely, frame replication is performed only on the wireless portion of the paths, where losses due to disturbance actually occur. We deem that this is advantageous for non-critical systems, where fault tolerance is not needed.

A new kind of Wi-Fi devices is envisaged, denoted HR STAs and HR APs, each one including two (or more) affiliated STAs. Qualitative analysis was performed for three relevant scenarios. When the HR STA is associated to a single HR AP, MLO can be used to support seamless redundancy. On the contrary, when it is associated to multiple HR APs, a mechanism denoted distributed MLO has been introduced that adheres to the plug&play philosophy of Ethernet/Wi-Fi and makes the involved APs transparently manage duplication and deduplication, without requiring any prior configuration. The case where HR STAs are allowed to move relies on the previous two mechanisms, by defining reliable reassociation procedures. Our proposal is clearly in a very early stage, and requires a formal definition, a careful investigation about achievable performance, and some proof of correctness.

REFERENCES

[1] I. Álvarez, A. Ballesteros, M. Barranco, D. Gessner, S. Djerasevic, and J. Proenza, "Fault Tolerance in Highly Reliable Ethernet-Based Industrial Systems," *Proc. IEEE*, vol. 107, no. 6, pp. 977–1010, 2019.

[2] International Electrotechnical Commission, "Industrial communication networks - High availability automation networks - Part 3: Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR)," International Electrotechnical Commission, Geneva, CH, Standard EN IEC 62439-3:2022, 2022.

[3] "IEEE Standard for Local and metropolitan area networks–Frame Replication and Elimination for Reliability," *IEEE Std 802.1CB-2017*, pp. 1–102, 2017.

[4] "IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)*, pp. 1–4379, 2021.

[5] M. Rentschler and P. Laukemann, "Towards a reliable parallel redundant WLAN black channel," in *2012 9th IEEE International Workshop on Factory Communication Systems*, 2012, pp. 255–264.

[6] G. Cena, M. Cereia, and A. Valenzano, "Security aspects of safety networks," in *ETFA2011*, 2011, pp. 1–4.

[7] G. Cena, S. Scanzio, and A. Valenzano, "Seamless Link-Level Redundancy to Improve Reliability of Industrial Wi-Fi Networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 608–620, April 2016.

[8] J. Fang, S. Sudhakaran, D. Cavalcanti, C. Cordeiro, and C. Chen, "Wireless TSN with Multi-Radio Wi-Fi," in *2021 IEEE Conf. on Standards for Communications and Networking (CSCN)*, 2021, pp. 105–110.

[9] B. Bellalta, M. Carrascosa, L. Galati-Giordano, and G. Geraci, "Delay Analysis of IEEE 802.11be Multi-link Operation under Finite Load," *IEEE Wireless Communications Letters*, pp. 1–1, 2023.

[10] G. Ditzel, "The Comparison/Contrast of TSN Frame Replication and Elimination for Reliability (FRER) and Parallel Redundancy Protocol (PRP)," in *2020 Industry Conference & 20th Annual Meeting*, 2020. [Online]. Available: https://www.odva.org/wp-content/uploads/2020/05/2020-ODVA-Conference_TSN_FRER_PRP_Ditzel_Final.pdf