Real-Time Performance of Cut-Through Forwarding in Hybrid Wired/Wireless Industrial Networks

Koulamas C.(1), Koubias S.(2) and Papadopoulos G.(1)
(1)Industrial Systems Institute
(2)Electrical and Computer Engineering Department
University of Patras
26500 Patras GREECE
koulamas@isi.gr, koubias@ee.upatras.gr, papadopoulos@isi.gr

Abstract – In this paper, the architecture and the operation of a cut-through forwarding device to be used in broadcasting, hybrid wired/wireless fieldbus systems are described. Analytical models of the delay overhead introduced due to frame forwarding are presented. It is shown that the usage of cut-through forwarding devices relaxes the bit-rate requirements in the radio segments, while it drastically improves the inherent advantages and reduces the drawbacks of hybrid transmission media architectures which are based on a single MAC domain.

I. INTRODUCTION

The maturity of digital radio communications and wireless LAN technologies has been already proved by the recent product evolution and their high market penetration and adoption from users in the home and office environments. In parallel with this situation, the benefits of wireless communications in other areas, like this of industrial control and automation, are well recognized, as they have already acted as the motivation behind extensive research during the last decade. The most critical differences between home/office and industrial environments which may affect an equivalent technology evolution in the latter, are either application and technology related restrictions, dealing with real-time performance and dependability aspects, or market related peculiarities, having in mind the longer expected life cycles of industrial installations. With respect to the latter, compatibility with existing infrastructures and engineering practices must be considered as strict requirements [1] for any new proposed technology or architecture. Regarding the technology related restrictions, research focus has been given to a number of directions, as to the evaluation and analysis of the impact of error prone communication links on existing industrial communications protocols [2] or to the proposal of enhancements for such protocols to compensate possible link impairments [3]. Other approaches target to the development of more robust and highly immune to noise and multipath fading radio physical layers, possibly along with accurate definitions of planning guidelines, parameter sets and limit values in order to guarantee a near-wireline communication quality and robustness [4]. Moreover, much attention has been also paid to architectural issues, resulting to an already available set of different approaches proposed, analyzed and compared, towards the construction of multi-segment, hybrid wired/wireless fieldbus systems, based either on physical layer repeaters/translators [5],[6] or on higher layer bridges/ routers [7], proxies [8], gateways [9] and wireless interconnecting backbones [10].

This paper focuses on the hybrid wired/wireless architecture defined and proposed as one of the results of the IST project RFieldbus1, especially on how the cut-through implementation of the interconnecting devices affects the performance and the real-time characteristics of the whole fieldbus system. The paper is structured as follows: section II presents the main characteristics of the proposed system and device architectures. Section III focuses on the operational and timing characteristics of a cut-through forwarding device, providing analytical figures for the calculation of the most important, performance related, timing parameters. Section IV presents the impact of such forwarding architectures in the performance of the fieldbus system and section V concludes the paper.

II. NETWORK AND DEVICE ARCHITECTURE

The two fundamental architectural characteristics of the proposed hybrid fieldbus system are the formal definition of a new wireless physical layer [6], [11] and the specification of the inter-segment communication function as a frame forwarding function at the service interface of the interconnected physical layers [11]. Consequently, the hybrid network appears with a uniform MAC protocol operation over a unified broadcasting ‘hyper-medium’ consisting of a number of forwarding functions which transparently inter-connect the different, wired or wireless, physical layer (PhL) segments.

The proposed model for the new, radio physical layer is presented in Figure 1 [12]. The model is structured in a way which provides an IEC-61158-2 [13] compatible MAC-PhL interface, while it is also influenced by the respective

1 This work was supported by the European Commission under the project name RFIELDBUS (IST-1999-11316)
specifications of the IEEE-802.11 [14] standard. It consists of two main sub-layers, namely the Data Communication Equipment (DCE) sublayer, which hides the selected modem techniques (i.e. DSSS or OFDM or other modem bursts), and the DCE Independent Sublayer (DIS), which adopts the DCE service to the service interface seen from the MAC layer and to the overall radio fieldbus system architecture [4]. The structure is completed by the addition of the Layer Management Entity (LME), which hides the management of all the parameters of the operation of the physical layer. At the DIS sublayer, an 802.11 PLCP-like frame header has been defined, consisting of 9 bytes (including two CRC bytes for header protection) followed by the physical layer service data unit (PhSDU), in a format which depends only on the PhL-User entity (i.e. the MAC or the previously mentioned forwarding function). Moreover, for the prototypes implementation, a Direct Sequence Spread Spectrum (DSSS) DCE was developed, which adds to the PhL protocol data unit (PhPDU) a total of 54usec preamble for receiver synchronization, while the supported bit-rates are programmable and up to 4Mbps.

With respect to the forwarding function (FF), a general model of the corresponding Forwarding Device (FD) in which it is hosted, is presented in Fig. 2. Depending on the characteristics of the FF, the FD may be one of three different types, namely either a Link Station (LS), in which, frames received from one PhL segment are forwarded - in a cut-through or in a store&forward manner - to the other PhL segment, or a Base Station (BS), in which, frames received from a PhL segment are repeated to the same segment - obviously always in store&forward - or finally, a Link Base Station (LBS), which is actually a combined LS and BS device. The repeater functionality of the BS and LBS devices has been mainly defined for infrastructure topologies, where broadcast communication is achieved through central points, (the BS/LBS), in order to alleviate the hidden terminal problem of direct link (ad-hoc) structures and to support inter-cell mobility, as this mechanism is described in [6],[15].

Moreover, as it is also depicted in Fig. 2, each FD has to implement part of the hybrid system’s DLL, either for framing reasons or in order to implement a service access point in which the station management entity may receive or transmit remote management PDUs. It must be noticed that, not all MAC and DLL protocols, among these defined in [16],[17], are equivalently suitable to such hybrid architectures, while in the context of the RFieldbus project the Type 3 (Profibus) protocol was selected as the enabling platform for the prototypes development and the pilot installations. The selection of a particular protocol type among these in [16] automatically defines the available wire-line physical layer combinations, bit-rates and device types. In the Profibus case, and for high speed traffic, the asynchronous transmission frame format over RS485 lines is used, while Fig.3 presents an example of such a hybrid topology, depicting also the two different types of stations defined, namely the master (M) and slave (S) stations.

Although both wire-line and wireless terminal stations (M or S) have been defined in the proposed architecture [5], Fig.3 depicts a hybrid system which consists only of standard (i.e. existing) wire-line Profibus nodes connected to standard RS485 asynchronous wire-line PhL segments which, in turn, are interconnected through the new radio PhL by the usage of FDs (LSs in particular). Recalling the long expected life time of industrial installations, it is foreseen that such topologies will present the initial roadmap for faster deployment of wireless extensions to existing installations, having also in mind that the forwarding nodes are more general products in contrast to complete Profibus nodes which usually include additional application related components (specific sensor, actuator, PLC etc). Nevertheless, the analysis presented in section III is equivalently applicable to any valid hybrid topology, according to [5].

III. FORWARDING DEVICE OPERATION

A. Ground Definitions

In general, the transmission time \( T_F \) of a frame with a PhSDU size of ‘N’ bytes is calculated as

\[
T_F = T_{POV} + T_{PSDU} = T_{PSST} + N \cdot B_F + T_{TR} = T_{FR} + T_{HD} + N \cdot O_F \cdot T_{BIT} + T_{PTR}
\]  

(1)

where, \( O_F \) equals to the number of bits used for the transmission of a byte of information, \( T_{BIT} \) equals to the transmission time of 1 bit in the medium, \( T_{PR} \) corresponds to the preamble transmission time possibly needed from the selected PhL technology and \( T_{HD} \), \( T_{PTR} \) correspond to the transmission time of additional bytes in a possible header and trailer respectively, in the PhPDU. Moreover, \( B_F \) stands for the transmission time of one byte of information (\( B_F = O_F \cdot T_{BIT} \)), \( T_{PSST} \) stands for the physical layer overhead prior to the PhSDU part (\( T_{PSST} = T_{PR} + T_{HD} \)) and \( T_{POV} \) stands for the overall PhL overhead over the PhSDU (\( T_{POV} = T_{PSST} + T_{PTR} \)).
B. Forwarding Operation Timing

To define the context of the analysis presented in the following sub-section, a snapshot of a general forwarding operation is depicted in Fig. 4. At time zero, one of the ports of the FD becomes active at the medium, indicating the starting point of a frame reception. The time elapsed - due to implementation related processing - from the appearance of such a line event up to the activation of the appropriate service primitive at the PhL-FF interface (see Fig. 2) is defined as \( T_{PD} \). Due to this event, the relevant port becomes the “input” port (I) while the other becomes the “output” port (O) of the FD. For a number of restrictions discussed below, this point in time does not necessarily define the activation time of the forwarding transmission in the output port. The time instance in which all the relevant requirements due to such restrictions are satisfied, is defined as the triggering time instance, \( T_{TRG} \), in which a triggering event is assumed to be generated inside the FF, translated to a service request primitive to the output port’s PhL interface. In response to this request, the device has to initiate the transmission in the output port, while this time component - due to initialization processing - from \( T_{TRG} \) up to the activation of the output medium is defined as \( T_{PS} \). The objective is to calculate the optimal \( T_{TRG} \) time instance which, in turn, will define the optimal time overhead \( T_{FWD,OVHD} \) of the forwarding operation, qualitatively defined as the time elapsed from the end of the frame in the input port (I) to the end of the frame in the output port (O), and quantified according to Fig. 4 by the following equation:

\[
T_{FWD,OVHD} = T_{TRG} + T_{PS} + (T_{F(O)} - T_{F(I)})
\]  

(2)

It is evident that the overall frame forwarding transaction lasts:

\[
T_{F,FWD} = T_{F(I)} + T_{FWD,OVHD} = T_{TRG} + T_{PS} + T_{F(O)}
\]  

(3)

C. Calculation of \( T_{TRG} \)

Getting into the inherent restrictions of a cut-through forwarding operation, the first should be mentioned is that after a frame transmission starts, the point in time for each byte’s transmission is accurately defined for a known bit-rate in the medium, since no gaps are allowed between bytes in any transmission technique used in the physical layers of [13]. Therefore, after transmission initiation, there are real-time requirements concerning the availability deadline for each byte which is to be transmitted, as it is evident that at the point in time of a byte’s transmission in the output port, that byte has to be already received by the FF, or else a transmitter under-run occurs and the frame transmission fails. Assuming that incoming data bytes are made available to the FF with the respective PhL indication primitive then this signaling, according to the previous sub-section, appears after \( T_{PD} \). The quantitative description of this requirement is

\[
T_{PS} + n \cdot T_{PD} + T_{TRG} + T_{IPST} + n \cdot T_{IPST} 
\]

(4)

Moreover, since the decision upon a correctly or erroneously received frame is made at the very end of any frame’s reception and there are decisions and actions that should be taken on certain events and CRC forwarding policies (i.e. pass unchanged, reconstruct etc.), there may be an additional requirement due to such a processing overhead \( T_{PC} \) after the end of the frame reception and prior to sending the last byte in the forwarding retransmission. Therefore, it is derived that the time offset between the end of the whole frame reception in the input port and the starting point of the transmission of the last PhSDU byte in the output port must never fall below \( T_{PC} + T_{PD} \). This is described as

\[
T_{PS} + n \cdot T_{PD} + T_{TRG} + T_{IPST} + T_{IPST} + (n - 1) \cdot T_{PD} + T_{PC} 
\]

(4)
which, actually, is a further constraint for (4), and can be easily merged with it by adding \( T_{PR} + T_{PC} \) in the left part and for \( nN \).

From (4) & (5) it is clear that for an FD to decide for its \( T_{TRG} \) time, the size information of the frame must be available. According to the PhL service primitive description [11],[13], this length information is made available to the forward control entity either with the RX START service primitive, if it is encoded in the PhL header, or after the reception of \( L \) bytes of the PhSDU, assuming that the length is encoded, directly or indirectly, in the \( L \)-th byte of the DLL PDU (PhL SDU). In the latter case it is evident that the forwarding entity has to decode, in a degree, the DLL protocol transfer rules. Concluding, it holds

\[
\min T_{TRG} = T_{PS} + L^* B_{F(I)} + T_{PD}
\]

where \( L \) may be 0, to cover the first case described above.

Since (6) gives the absolute minimum possible time instance of the \( T_{TRG} \), and since (4) & (5) solved for \( T_{TRG} \) may give negative (i.e. unreal) results, the definition of the \( \Delta T_{TRG} \) and the identification of the prerequisites for the \( \Delta T_{TRG} \) to exist, that is, to be positive, seem more appropriate. It is defined

\[
T_{TRG} = \min T_{TRG} + \Delta T_{TRG} = T_{PS} + L^* B_{F(I)} + T_{PD} + \Delta T_{TRG}
\]

After merging (4) & (5) and substituting in (7), it finally holds

\[
\Delta T_{TRG} = \max\left\{ \begin{align*}
& n^* \Delta B_F + B_{F(O)} - P_2 & \forall n, 1 \leq n \leq (N-1) \\
& N^* \Delta B_F + B_{F(O)} - P_2 + P_1 \\
& 0
\end{align*}\right. \quad (8)
\]

where,

\[
P_1 = T_{PR} + T_{PC}, \\
P_2 = T_{PS} + T_{PS} + L^* B_{F(I)}, \\
\Delta B_F = B_{F(I)} - B_{F(O)}
\]

Equation (8) shows that, if \( \Delta B_F = 0 \) (i.e. \( B_{F(O)} = B_{F(O)} \)) then the \( \Delta T_{TRG} \) for a frame is derived for \( nN \), since \( P_1 \) & \( P_2 \) are always zero or positive numbers, and \( \Delta T_{TRG} > 0 \) happens if and only if it is \( B_{F(O)} + P_1 > P_2 \). Moreover, if \( \Delta B_F > 0 \) (i.e. \( B_{F(O)} > B_{F(O)} \)) then the \( \Delta T_{TRG} \) increases with ‘\( n \)’, therefore the worst case is again for \( nN \) - since \( P_1 \) is always zero or positive - that is, defined from the length of the frame. Then, \( \Delta T_{TRG} \) is positive if and only if

\[
N > N_{FWD_{-1}} = \frac{P_2 - (B_{F(O)} + P_1)}{\Delta B_F} \quad (10)
\]

which happens for all sizes \( N \) if \( B_{F(O)} + P_1 > P_2 \). Finally, if \( \Delta B_F < 0 \) (i.e. \( B_{F(O)} < B_{F(O)} \)) then the \( \Delta T_{TRG} \) increases when ‘\( n \)’ decreases, and the actual \( \Delta T_{TRG} \) for a frame forwarding is defined by either the frame’s length \( N \) or by its first byte (i.e. \( n=1 \)). The first byte’s case dominate when

\[
(\Delta B_F + B_{F(O)} - P_2) - (N^* \Delta B_F + B_{F(O)} - P_2 + P_1) > 0
\]

which leads to

\[
N > N_{FWD_{-2}} = \frac{P_1}{-\Delta B_F} + 1 \quad (11)
\]

and \( \Delta T_{TRG} \) is positive only when \( \Delta B_F + B_{F(O)} > P_2 \) or, equivalently, \( B_{F(O)} > P_2 \). For the latter, it may be easily proven that it holds only for \( L = 0 \) and \( B_{F(O)} > T_{PS} + T_{PS} \). If (11) does not hold, then the second branch of (8) applies, and consequently (10) with the inequality reversed (since we divide with \( \Delta B_F \) which is negative) for the \( \Delta T_{TRG} \) to be positive. Concluding, \( \Delta T_{TRG} \) is finally described by the following:

\[
\Delta T_{TRG} = \begin{cases}
B_{F(O)} + T_{PR} + T_{PC} \\
B_F - P_2 + P_1 \\
0
\end{cases}
\]

\[
\text{if } \{(B_{F(I)} > B_{F(O)}) \& (N > N_{FWD_{-1}})\} \\
\text{or if } \{(B_{F(I)} < B_{F(O)}) \& (N \leq N_{FWD_{-1}})\} \\
\{B_{F(I)} \& (N > N_{FWD_{-1}})\}
\]

\[
D. Calculation of \( T_{FWD_{-OVHD}} \)

By distinguishing the two major cases of the \( \Delta T_{TRG} \) value - i.e. being or not a function of the frame size - substitution from (12) to (2) leads to a more expressive form for the \( T_{FWD_{-OVHD}} \) definition, which is

\[
T_{FWD_{-OVHD}} = \begin{cases}
B_{F(O)} + T_{PR} + T_{PC} \\
\text{min } T_{TRG} + \Delta T_{TRG} + T_{PS} \\
+ \left(T_{NOV(I)} - T_{NOV(I)}\right) + N^* (B_{F(O)} - B_{F(I)})
\end{cases}
\]

\[
\text{otherwise}
\]

Equation (13) shows that in the cases where the \( T_{TRG} \) time instance is constant, the overhead introduced from the FD is
a function of the frame size, increasing with a size increase if \( B_{\text{WD}} > B_{\text{SD}} \), decreasing otherwise – see the second branch of (13), covering the last three, constant value \( \Delta T_{\text{TRG}} \) branches of (12). Moreover, the opposite applies, which is that whenever the \( T_{\text{TRG}} \) time instance is a function of the frame size, the resulting forwarding overhead is constant. The latter constant value, given from the first branch of (13) is also the minimum achievable time overhead in any frame forwarding.

It must be also noticed that (13) is a general equation describing the time overhead added when a frame passes through a forwarding device to a separate physical layer segment. It applies equivalently to cut-through and store-and-forward forwarders, when for the latter case the second branch of (13) is used, and \( \min T_{\text{TRG}} + \Delta T_{\text{TRG}} \) is substituted by the time in which the whole input frame is received and checked, that is \( T_{\text{FWD}} + T_{\text{PD}} + T_{\text{PC}} \). In this case (i.e. store&forward), (13) becomes

\[
T_{\text{FWD-OVHD}} = T_{\text{PS}} + T_{\text{PD}} + T_{\text{PC}} + T_{\text{F(OUT)}} \tag{14}
\]

which is the rather obvious result that in a store-and-forward operation the overhead is the total duration of the retransmission of the frame in the output port plus any processing and data transfer overhead delay inside the forwarder.

Getting into the specific usage of the high-speed Profibus standard in the wire-line segments, accurate plots of the time overheads of such operations can be given. In the proposed and implemented radio PhL the related parameters are \( T_{\text{PR}} = 54\,\mu s \), \( T_{\text{HD}} = 9B_{\text{F}} \) (since the physical layer frame header consists of 9 bytes), \( O_{\text{F}} = 8 \), \( L = 0 \) (since the length of the frame is encoded in the physical layer header). In Profibus PhL Type 1 (RS485) the parameters are \( T_{\text{PR}} = T_{\text{HD}} = T_{\text{PST}} = T_{\text{PTR}} = 0 \) and \( O_{\text{F}} = 11 \). For the length parameter, different options exist, since there are different frame formats defined. For the fixed size frame formats, the length is indirectly encoded in the frame type which is defined by the first byte of the frame, therefore it holds \( L = B_{\text{F}} \). For the variable size frame format (SD2), it holds \( L = 2B_{\text{F}} \). Concerning Profibus Type 1 PhL, there is another issue with respect to the different frame formats defined. There are two from the five frame types, namely the one-byte fast acknowledge (SC) and the three-byte token frame (SD4) that are not protected by the two byte closing sequence consisting of the Frame Check Sequence (FCS) and the end of frame (ED) bytes, leaving as the only transmission error detection mechanism the parity bits of the asynchronous byte transmission technique. Since in the wireless frame the byte is transmitted as a 8-bit entity, the forwarder shall add for these two frames the missing error protection, which has been decided to be equivalent with the same protocol (Profibus) but different frame encoding (Type 2), resulting to two additional CRC bytes. For these frames, (12) & (13) are again valid if this \( 2B_{\text{F(WIRELESS)}} \) additional transmission time is – artificially – considered as a trailing overhead, therefore encoded in the \( T_{\text{PTR}} \) component.
In conclusion, for topologies equivalent to that of Fig.3, it was shown that:

- For a given bitrate in the wired segments of a hybrid wired/wireless system there is an optimal (worst-case) bitrate for the wireless (radio) segments, which depend on the maximum frame size allowed in the system, while an increase of this bitrate leads to the degradation of the system performance.
- In this optimal bitrate, the end-to-end delay overhead due to frame forwarding from one wireline segment to another through a wireless link is constant and independent from its size.
IV. IMPACT ON NETWORK PERFORMANCE

The impact of the introduction of forwarding devices, additional physical layer segments and additional delays in a fieldbus network is that a number of parameters which coordinate each specific protocol operation have to be reconsidered. This especially applies to timeouts, either between a request and a response frame or in physical medium idle times which have to be prolonged for the normal operation of the hybrid network (see Fig. 8, message cycle between M1 and S2, with respect to the topology of Fig. 3). Moreover, some kind of message rate control has to be applied in cases that two or more stations connected to the same physical medium of a hybrid system are communicating between each other or a single station is continuously transmitting unacknowledged frames, as proposed in [18] (see in Fig. 8, the propagation of message cycles between M1 and S1).

The advantages and drawbacks of broadcasting hybrid architectures based on a single MAC operation over the whole set of PhL segments in the system, as the one described in Section II, have been discussed in [7] and [8]. Among the most crucial drawbacks of such architectures seem to be the low network utilization achieved, due to higher medium idle times while waiting for responses (i.e. the maxT_{SDR} parameter in Profibus) and the system responsiveness to failures, as this is governed by timeout values (the T_{SLOT} parameter, regarding the Profibus protocol, which is also based on maxT_{SDR}). Considering the snapshot depicted in Fig. 8, the T_{SDR} of S2 in response to a request from M1 (first message cycle) is

\[
\text{max } T_{SDR} = \text{max } T_{FWD\_OVHD(REQ)} + \text{min } T_{SDR} + \\
+ \text{max } \sum_{i=1}^{N_{ED}} (T_{TRG(i)} + T_{PS})_{(RESP)}
\]

(18)

which, if the assumptions for (10) hold, becomes

\[
\text{max } T_{SDR} = \text{max } T_{FWD\_OVHD(REQ)} + \text{min } T_{SDR} + \\
+ \text{max } T_{FWD\_OVHD(RES)}
\]

(19)

following analogous curves with these of Fig. 7.

Moreover, among the most important advantages of single MAC hybrid networks are the straightforward and simplified deployment, since they are not different in principle from the standard, existing wire-line solutions, as well as the preservation of the notion of the global, system-wide time, which is important to control applications. In bridged/routed architectures, system-wide SYNC-type commands (e.g. input FREEZE in Profibus DP) capture or drive the environment conditions with a worst-case time skew equal to the sum of the maximum latencies in the intermediate segments’ MAC operation, which for multi-master Profibus segments are in the order of several msec (target rotation times). In contrast, in single MAC, multi-physical segment hybrid networks this worst-case time-skew equals to the maximum T_{FWD\_OVHD} delay over the longer path in the system, which may be in the order of usec, since it is calculated for a single message cycle.

In conclusion, since most crucial network parameters which affect the performance of the described hybrid fieldbus system depend on the forwarding overhead delay, the reduction of the latter by the usage of cut-through forwarding devices elevates the advantages and downgrades the drawbacks of such hybrid architectures. In order to point

![Figure 9. Comparison of Store&Forward (SF) and Cut-Through (CT) forwarding overhead delays](image)
out the improvement, comparing store-&-forward and cut-through forwarding devices, Fig.9 depicts the ratio of the end-to-end $T_{FWD\_OVHD}$ delays which correspond to the 2-LS path scenario described by Fig.7, that is, the ratio $T_{FWD\_OVHD\_SF} / T_{FWD\_OVHD\_CT}$. Fig. 9 clearly shows that in the optimal radio PhL bitrate and for the maximum size Profibus telegram the overhead is more than 25 times less when using cut-through forwarding instead of store-&-forward.

V. CONCLUSION

In this paper, the architecture and operation of a cut-through forwarding device and a corresponding network architecture of broadcasting, hybrid wired/wireless fieldbus systems were described. Analytical models of the delay introduced during a frame forwarding and propagation were also presented. It was proved that there are certain bitrate relationships between the various physical layer segments for which the real-time communication characteristics are optimal, while it was shown that the usage of cut-through forwarding lowers the bitrate requirements in the radio segments, improves the inherent advantages and reduces the drawbacks of hybrid architectures which are based on a single MAC domain.

VI. REFERENCES