ITU - Telecommunication Standardization Sector Temporary Document 2012-06-4A-061

STUDY GROUP 15 Original: English

Antwerp, 18-22 June, 2012

Question: 4a/SG15

SOURCE[1]: ADTRAN

TITLE: **G.fast: A flexible framing structure**

ABSTRACT

This contribution presents a flexible frame structure for G.fast that will enable dynamic adaptation to the actual needs and offered traffic load. Based on this flexible frame structure a retransmission mechanism utilizing a type II HARQ is proposed. It will provide low latency and high throughput. The flexible frame structure can also be used to save power consumption.

1. Introduction

A regular framing structure is of utmost importance for a communication system. It is required for several reasons. On the one hand, it provides synchronization, schedules time intervals for US and DS transmission, and gives implementers guidance in their design. On the other hand, a proper design of the framing structure is necessary to ensure that the system requirements are fulfilled. This includes the provision of the required data rates, US/DS asymmetry ratio, latency, and the ability to save power. The range of application scenarios (e.g. residential, business, mobile backhauling) for G.fast systems will have different requirements with respect to latency, data rates, and reliability. Therefore, optimization can't be done for all scenarios simultaneously and flexibility is needed to meet the different requirements. Furthermore, G.fast will probably be the last copper based access technology to be standardized. This means new applications with requirements not yet known might come up in future.

Adding flexibility to the framing structure could be provided by specifying a toolbox of means with tight bounds instead of dictating a static format. Other committees, e.g. 3GPP have done this very successfully. Besides giving the flexibility to optimize the frame structure for current and future applications, it will give more freedom for vendor discretionary optimizations and will enable the design of products dedicated to specific use cases. Furthermore, when the technology evolves over time new products can benefit from new algorithms (e.g. improved of accelerated decoding algorithms).

Given the above, the G.fast framing structure should provide as much flexibility as possible without losing efficiency. This contribution makes a proposal to fulfill those needs.

2. Design principles

Superframe format

It is assumed that a frame contains multiple DS (downstream) and US (upstream) transmission opportunities and that a set of frames comprises a superframe. In order to fulfill the smallest latency requirement the length of a frame shall not exceed 1ms. The number of frames per superframe is TDB but superframe duration could be in the order of 100ms. Figure 1shows the general format of a superframe consisting of n frames.

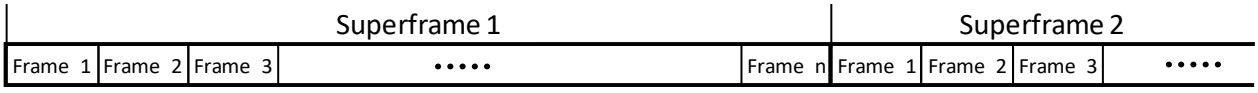

Figure 1: General superframe format

[1] Contact: Martin Kuipers
ADTRAN
T: +49 89 5159-25059
F: +49 89 5159-4425059
E: martin.kuipers@adtran.com

As mentioned in the introduction, one of the objectives is to ensure that the desired data rates and US/DS asymmetry ratio are provided to the subscribers. It is expected that the operator will be able to provision a set of parameters that might include the min and max values of US/DS asymmetry ratio and defined data rates instead of just fixed values. The portion of the frame required for US and DS transmission to fulfill these requirements depends on the offered traffic load and the loop conditions. For example, a rather long loop of a business customer subscribed to a service with an increased US data rate might require the reservation of a significant portion of the frame for its US transmission in case the allocation is fixed. In order to avoid the NEXT generated by this particular business customer, this portion of the frame is lost on all other loops connected to the same DP. On the other hand, there is usually no need to fulfill the provisioned requirements on a frame level with a granularity of e.g. 1ms. It should be sufficient to ensure that the US/DS asymmetry ratio and the data rate objectives are fulfilled on average. It is our assumption that no application will suffer from moderate variation of the rates in cases where the system provides the desired rates averaged over a superframe and the latency requirements are met.

Frame format

For the purpose of this contribution, a "symbol" corresponds to an OFDM/DMT symbol, and a "timeslot" corresponds to one or more consecutive symbols in a given upstream or downstream direction. A frame will start with a symbol sent by the FTU-O in the DS direction that comprises multiple functions. It will be easily detectable to function as a SYNC symbol and provide means for direct and crosstalk channel estimation. For this purpose, the full capacity of the first symbol will not be needed. The measurements on all subcarriers may be spread over multiple frames providing more space for other purposes since it will not be required to do measurements on all subcarriers in each frame. This can be achieved by using a rotating pattern of subcarriers that will provide a different subset of subcarriers for subsequent frames. In an alternative implementation a fixed pattern (e.g. fixed subcarrier spacing) is used in all frames and the values between measured subcarriers will be interpolated.

The first symbol will additionally contain a MAP (Media Access Plan). This MAP will provide the indication of which symbols in the present frame are US and which are DS. It will therefore instruct the FTU-R at what time to listen to DS transmission and the interval to be used for US transmission. The MAP may change from frame to frame according to the needs.

The overhead introduced by a multifunction SYNC symbol should be very small. The functionality of a SYNC symbol to provide synchronization and channel estimation is required, and this SYNC symbol can be sent once per superframe or alternatively distributed over multiple frames. For the MAP the amount of data needed is negligible. It needs to indicate for each symbol in the frame whether if it is used for US, DS, or unused. This can be easily realized with two bitmaps (one for DS and one for US) indicating for each symbol if it is used for the particular direction or not. In total two bits for every symbol in the frame would be sufficient. For a 50 symbol frame a total of 100 bits would be required. The MAP may be even omitted in case time slot allocation does not change from the previous frame but it should be transmitted at least at the beginning of every superframe. The rest of the first symbol could be used for DS data transmission or other management purposes.

In addition to a single DS period a second DS period might be scheduled at the end of the frame to provide in-frame retransmission information. This will be explained more in detail below.

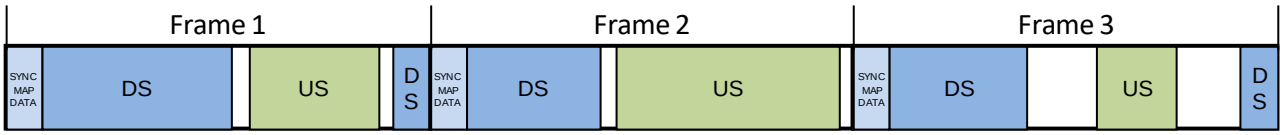

Figure 2: Frame format examples

When constructing the MAP the FTU-O needs to ensure that DS and US will not overlap on lines connected to the same DP. Figure 2 shows an example of three frames that could be part of the same superframe on the same line. The transmission starts always with the first symbol in DS direction containing SYNC and MAP information. Usually, the DS transmission continues on the next symbol unless otherwise indicated by the MAP. After a necessary gap to switch between reception and transmission the US starts. This gap may be longer as necessary in case the transceiver does not have enough data to send in DS but US can't start due to an ongoing transmission on another line. In other words, if there is no data to send, no symbol will be transmitted on the line. The third frame of Figure 2 indicates this. A second DS period is scheduled for HARQ information at the end of the first and third frame.

3. Type II Hybrid ARQ

Current retransmission schemes as specified in G.998.4 [2] are based on the full retransmission of the corrupted DTUs (Data Transfer Units). In case of a DTU received in error, the errored DTU is thrown away and an identical copy of this DTU is transmitted. This is done despite the fact that the original transmission contains valuable information which could be used in the subsequent decoding process. Such a scheme that solely relies on the retransmitted DTU is also denoted as Type I Hybrid ARQ (HARQ).

More advanced techniques are denoted as type II Hybrid ARQ. These more reliable techniques were already mentioned in [3]. One of the type II Hybrid ARQ approaches is called Chase Combining (CC). Using CC the original transmission and the retransmission are combined before decoding. This can be done in different ways. The simplest approach is just to add both transmissions mathematically according to the relation of their SNRs before decoding (this assumes the exact same modulation is used for the original DTU and the retransmitted DTU). Other schemes based on soft-bits in case of different modulation schemes for the two transmissions exist. CC is a receiver only technique and no extra complexity is required at the transmitter side.

In case both DTUs have the same SNR the resulting SNR is doubled in linear scale providing a combining gain of 3dB. Using the simulation results from [1] the benefit become evident. Figure 11 of [1] shows the LDPC decoding performance for a 1024 QAM of 120 bytes blocks. In case we consider a code rate of 20/21 the required SNR to reach a BLER of 10^{-5} is about 33.8 dB. To reach the same decoding performance it would be sufficient if the original and the retransmitted block would reach 30.8 dB. When we look up this number in the simulation results it follow that both transmissions would have a BLER of > 0.5 and the retransmission would fail with a high probability.

A variant of CC is called Partial Chase Combining (PCC). This scheme is based on the fact that it is not required to retransmit the DTU completely all the time when very little information is missing to decode it correctly. In this case only a part of the DTU may be retransmitted while it is not necessary to identify the actual defect sections. It is only required to have knowledge about the amount of information missing. In this case the retransmission can be much shorter (see section 2.5 of [1]).

Another type II HARQ scheme is called incremental redundancy (IR). In this case the DTU is encoded using a FEC code with a rather low code rate (e.g. 1/2). The original transmission contains only a part of the redundancy (higher code rate, e.g. 5/6) obtained by puncturing. In case decoding fails only the additional redundancy that was punctured away will be transmitted. This can be done in multiple variants. The complete remaining redundancy can be sent in a single or in multiple steps. In the later case the system will send additional redundancy until decoding was successfully done. In another variant, the receiver signals how much information is missing to decode the DTU correctly. The transmitter will then forward at least as much redundancy as required to ensure successful decoding. To get an idea about the amount of information that is retransmitted in case of IR the following example should be given. To achieve performance parity with chase combining under the assumption that all transmissions have been received with the same SNR an additional coding gain of 3 dB is required. Assuming Gaussian noise and going back to the example given above from Figure 11 of [1], we notice that the difference in coding gain of the two LDPC codes with rates of 20/21 and 5/6 is about 4 dB. The difference in coding rate between 20/21 and 5/6 is 11.9%. For a 120 byte block the additional redundancy would be 18 bytes.

4. Efficient usage of the flexible frame format

In case we combine the proposed flexible frame format with one of the type II HARQ techniques a very efficient scheme can be constructed. As given above the allocation of symbols to DS or US can change from frame to frame. This will require some knowledge about the actual offered traffic load in DS and US direction. In DS direction the FTU-O has direct knowledge about the status of the transmit buffers. The FTU-O can use this information easily to plan for the DS transmission. To do the same for the US, it is therefore envisaged that the FTU-Rs signal in each frame the actual fill status of their transmit buffers including the age of the oldest packet. In an alternative implementation the FTU-R can signal the desired capacity it would like to use in the US direction for the upcoming frame or frames.

The DP will collect the information of all lines connected to it. That information is used together with the provided configuration (desired rates, US/DS asymmetry ratio, etc) and status of all other loops to partition the next frame between DS and US. The system can profit from rather small gaps in transmissions and give lines that have significant higher or different demands more capacity (e.g. in the US) for a short period. Due to the NEXT issue the DP need to ensure that this is synchronized between lines.

There are many examples where such a scheme will be beneficial. For example, one line would like to do a large upload while there is not a lot of traffic currently ongoing on other lines. Furthermore, many applications generate traffic that is bursty due to the applied coding (e.g. packets on a 20ms basis). If the system is able to transmit this within a single frame instead of distributing this over multiple, the latency might be reduced and the system will gain flexibility for the remaining frames.

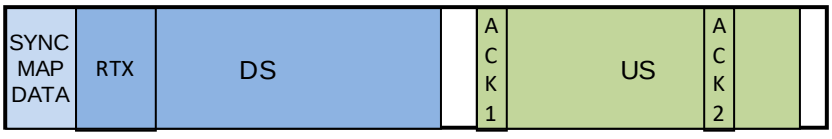

Figure 3: Example frame with acknowledgements

Examples of such frames are given in Figure 3. The first symbol contains SYNC and MAP and also DS data. In case there are still packets to be retransmitted, these retransmissions should be done first. Then the remaining DS traffic is then transmitted. The FTU-R can start transmission after an obligatory gap as scheduled for US in the MAP. The FTU-O will ensure by proper MAP definition that there will be no US transmission overlapping with DS operations on other loops.

In the US transmission the FTU-R has to send ACK or NACK for DTUs transmitted in the DS earlier in the frame. This should be done as soon as the decoding process has finished (e.g. directly at the beginning of the US transmission, ACK1). In case the decoding process was not complete for all DTUs at the beginning of the US transmission, the FTU-R will insert the missing ACKs at a later stage during (e.g. ACK2).

Due to its superior performance it is assumed that a type II HARQ scheme with Incremental Redundancy (IR) is applied. In case a DTU can't be decoded correctly the redundancy that was punctured away will be sent instead of a complete retransmission. In order to better estimate the amount of additional redundancy that is needed the FTU-R will not only provide ACK/NACK as binary information. The decoding process can either provide the amount of information missing (received coded bit information, RBIR [4]) or can directly signal the amount of additional redundancy it wishes to receive. Compared to retransmitting a complete DTU much less data is needed. The additional signaling information is negligible. It consists of only a few more bits for each incorrectly received DTU.
In the example given above it was shown that sending additional 18 bytes for a 120 byte block will provide an additional 4 dB coding gain. To give a few more numbers the additional redundancy and resulting coding will be derived for the simulation results given in Figure 9 of [1] for 1024QAM and 540 bytes blocks at a BELR of 10^{-5}:

Original code rate	New code rate	IR size	Add. coding gain
20/21	5/6	81 bytes	4 dB
5/6	2/3	162 bytes	4 dB
2/3	1/2	270 bytes	3.5 dB

The benefit of IR will be not that significant for sporadic retransmissions caused by Gaussian noise. In many situations error events are clustered due to impulses or other kinds of sporadic noises. If many DTUs are affected, the complete retransmission of all DTU might cause a significant reduction of the error free throughput and introduces additional jitter.

Since the amount of retransmission data sent with IR is rather small, a single symbol will be sufficient to carry the additional data for many DTUs. Therefore, a second DS period is scheduled in Figure 4 at the end of the frame. By doing this, the incorrectly received DTUs can be corrected within the same frame. This will on the one hand save time and keeps the latency small. On the other hand it will save transmission bandwidth because a retransmission of the complete DTU is not necessary.

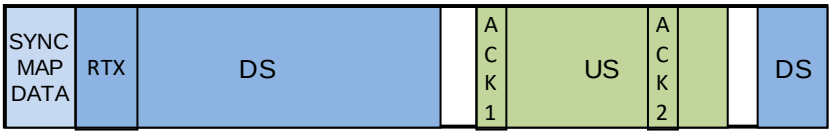

Figure 4: Example frame with second DS timeslot

In case a successful decoding is not possible using IR the complete DTU will be retransmitted in the next frame. This complete retransmission can be Chase combined with the previously received data to improve decoding probability. Please note that two retransmissions are possible within the same period that would usually allow only one retransmission.

For the scheduling of the second DS timeslot multiple options exists. One option is to permanently schedule a second DS timeslot in advance. In case no transmission error occurred the symbol could be used for other purposes in DS direction. Alternatively, the second DS timeslot will be scheduled on demand in case US will not use the complete rest of the frame or low latency and high reliability is required by the kind of application or customer.

In a potential alternative solution the second DS period will not be scheduled and the last symbol of a frame will carry the IR tentatively. In case the FTU-R has received all data correctly it will use the US period completely. In that case, the last symbol at the end of the frame will act as a transmission gap between US and next DS. If there are transmission errors that are signaled in the US the FTU-R will expect HARQ information in the last symbol. FTU-R will end the US transmission early to have time to switch to reception before the last symbol starts. The FTU-O will use the last symbol for transmission of HARQ information on that particular line. NEXT will not be an issue since loops not having transmission errors will be silent during this symbol period, see Figure 6.

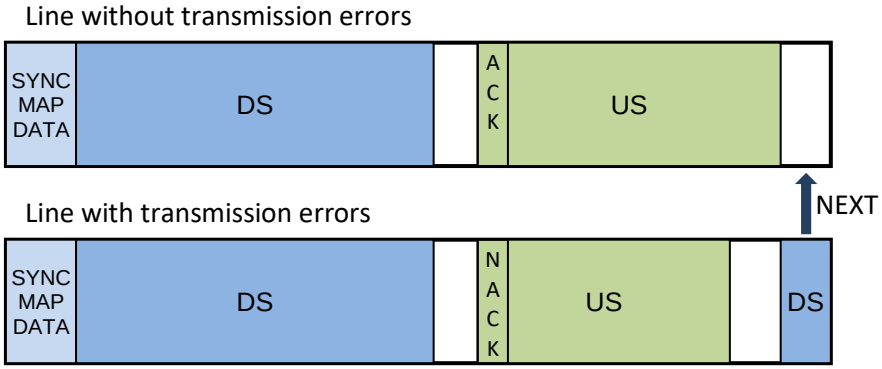

Figure 5: Example of dynamic usage of the second DS

It the obvious that the examples given here for the case of IR will also work for PCC and CC.

5. Power saving capabilities

The flexible framer can also be used to reduce power consumption. By receiving the MAP the FTU-Rs know exactly when to receive and when to transmit. This can be used to plan and shutdown unused parts of the chipset. In case there is a need to further reduce the power consumption (e.g. due to thermal restrictions or available power) this can be done by shortening the DS and US time periods. Figure 6 is providing an example of normal and power saving operation.

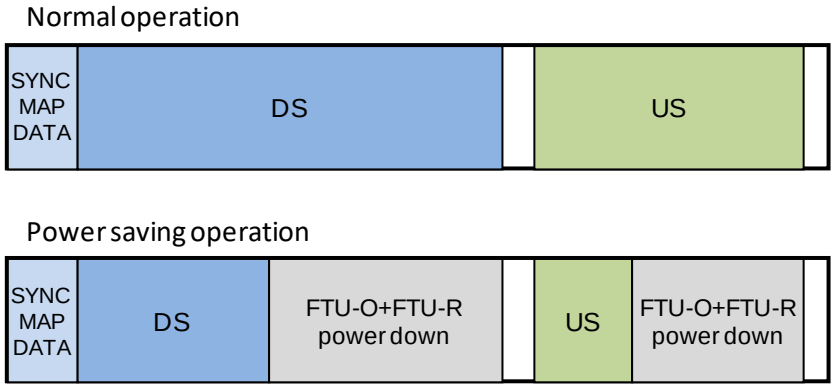

Figure 6: Power saving example

6. Summary

We believe that a flexible frame structure as presented in this paper together with a type II HARQ scheme will enable the implementation of efficient and future proof systems that will fulfill the requirements of different applications. In addition to that this frame structure provides also means to save power

This paper shall be presented at the G.fast session.

It is proposed that the following new issues be agreed:

6.3.x	Open	Should a flexible frame structure be defined for G.fast that allows changing the allocation of US/DS timeslots dynamically?	2012-06-4A-061
6.3.x	Open	Should a flexible frame structure that enables in-frame HARQ with a second downstream time slot for the retransmission be defined for G.fast?	2012-06-4A-061
6.1.x	Open	Should G.fast utilize a type II HARQ retransmission scheme with soft ACK	2012-05-4A-052

7. References

[1] Marvell, “G.fast: LDPC-BC proposal as the FEC method for G.fast”, ITU-T SG15 Q4a contribution 2012-05-4A-054

[2] ITU-T Recommendation G.998.4 (2010), Improved Impulse Noise Protection (INP) for DSL Transceivers.

[3] Sckipio, “G.fast: Retransmission scheme for G.fast”, ITU-T SG15 Q4a contribution 2012-05-4A-052.

[4] Xiang Chen, et al., “Prediction of Hybrid-ARQ Based on Mutual Information Model for LDPC coded OFDM system”, 2008, 11th IEEE International Conference on Communication Technology Proceedings.