Minimizing delay and delay variation in Ethernet fronthaul networks

Benefits of frame preemption

White paper

Architectural changes made in 5G to better optimize and scale the network with C-RAN have also brought strict synchronization and latency requirements into the spotlight, particularly for fronthaul networks. Standardization bodies have developed standards on how to meet these requirements. One such standard in the Time-Sensitive Networking (TSN) toolkit, is IEEE Std 802.1Qbu, which specifies frame preemption. This amendment to IEEE Std 802.1Q, along with IEEE Std 802.3br, specifies a method for transmitting time-sensitive frames in a manner that significantly reduces the delay and delay variation these frames experience. While the greatest benefit of delay reduction is realized for lower port rates, the intended benefit of frame preemption is the ability to bound the latency and frame delay variation and protect time-sensitive flows from other flows in a deterministic manner. This makes it easier to conform to the rigorous requirements imposed, for example, by the IEEE 802.1CM TSN for Fronthaul standard profile. Frame preemption requires specific hardware. This white paper explores how frame preemption works and the benefits of using it.

Contents

Background

There is a strong desire among operators to move to packet-based fronthaul to drive down cost and increase scale. In the past, fronthaul was addressed primarily via direct fiber and point-to-point WDM links, which did not cause asymmetry or delay variation. Now with the move to transport fronthaul streams over Ethernet bridged networks, achieving strict latency and jitter requirements raises the need for new standards-based mechanisms to make Ethernet deterministic.

IEEE Std 802.1CM is a standard profile for TSN and ITU-T technologies that applies to fronthaul and was developed and published by the IEEE 802.1 Working Group. The standard and the 802.1CMde amendment specify requirements for latency, synchronization, traffic control and system reliability in deterministic Ethernet bridged fronthaul networks. IEEE 802.1CM defines two bridge profiles to address queuing. Profile A prioritizes time-sensitive traffic flows over non time-sensitive flows with strict priority queuing. Profile B goes a step further by adding frame preemption to bound latency and control frame delay variation.

Ethernet bridged fronthaul networks provide converged access that can be shared among users of business services, mobile operators and emergency services, to name a few. It is, therefore, important to ensure prioritization of time-sensitive frames over non-time-sensitive frames. The term “time-sensitive” is used generically to refer to frames that are highly sensitive to either time, latency, latency variation, or all of these at once.

Even though delay is an important aspect of Quality of Service (QoS), traditional egress frame scheduling over an Ethernet link could pose a challenge in meeting very strict delay requirements or in exhibiting deterministic behavior for any mix of traffic. In Ethernet bridged networks without frame preemption, once a frame begins transmission, it cannot be interrupted to allow a time-sensitive frame to be transmitted first.

Traditional Ethernet queue management and scheduling techniques such as SPQ, RR, WRR and DWRR schedule full frames by priority or, optionally, by relative weight, but neither provide the necessary determinism needed for the transport of high-priority time-sensitive traffic. Properly bounding the transmission delay for these time-sensitive frames requires a preemption-based transmission mechanism. For this, the IEEE 802.1CM leverages IEEE 802.1Qbu in its Profile B allowing high-priority time-sensitive frames to preempt other frames, thereby bounding frame delay and mitigating variability in frame transmission delay.

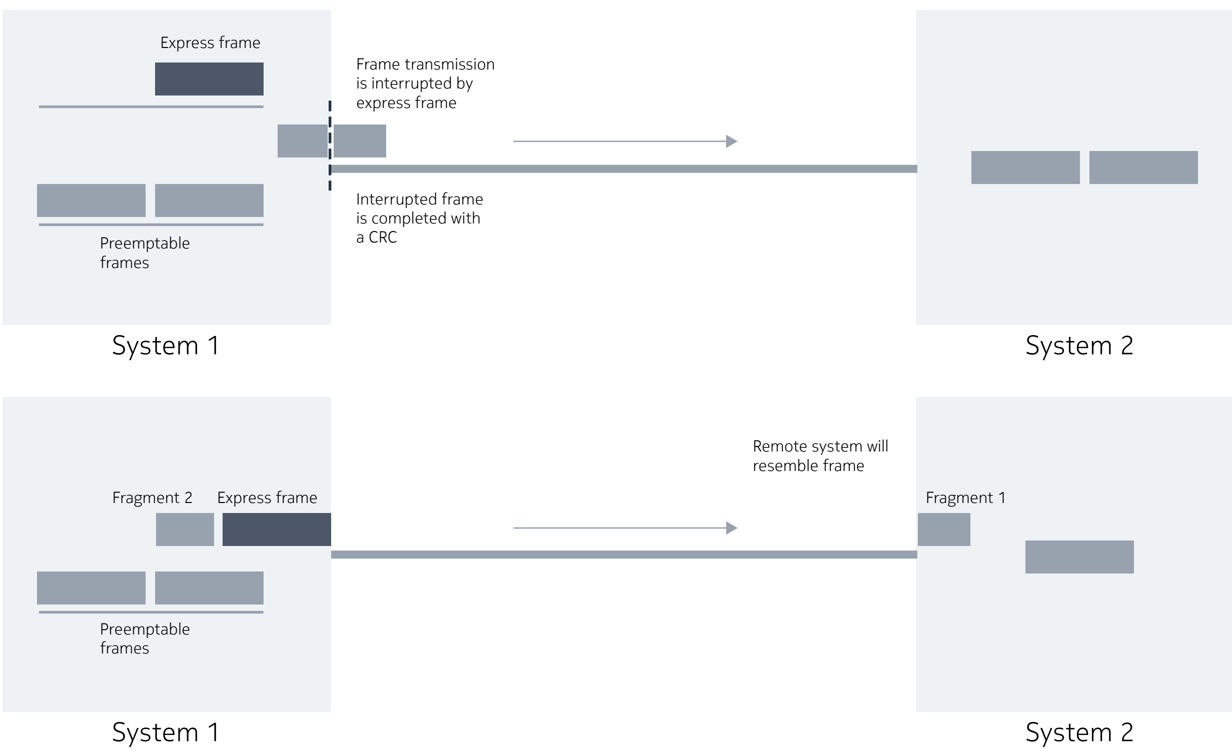
What is frame preemption?

In traditional queuing techniques like strict priority queuing (SPQ), if a high-priority frame is received while processing a lower priority frame, the high-priority frame is inserted into the highest priority queue. However, the frame still has to wait until the current lower priority frame has finished transmitting. Once complete, the high-priority frame is processed as quickly as possible due to the precedence associated with its priority. The system transmits any other high-priority frames that are in the priority queue before it resumes transmission of frames in any of the lower priority queues. In this way, time-sensitive frames experience a smaller delay and FDV than non-time sensitive frames; however, depending on the size of the frame currently being processed, these delays can get very large — in the tens of microsecond range. Note that IEEE 802.1CM Profile B does not impose a frame size limit as Profile A does. Thus, the worst-case would be when Jumbo frames (9.6k bytes) are being transmitted out of a low priority queue while a high priority frame is waiting transmission. At a very high level, this frame-by-frame processing is how a scheduler works. In practice, schedulers might transmit frames in microbursts.

Unbounded frame delay for time-sensitive frames could mean, for example, a deterioration in voice and video services and inaccurate data collection in real-time monitoring systems. With architectural changes made in 5G to better optimize and scale the network with C-RAN, frame delays and delay variations become more important to mitigate.

With IEEE 802.1Qbu, if a system receives a time-sensitive frame while a non-time sensitive frame is in the process of being transmitted, the latter is interrupted. It is preemptable because the time-sensitive frame is considered an express frame and is transmitted immediately (or after a minimum fragment of the preemptable frame has been transmitted). The system will resume transmitting the preemptable frame once the express frame has been transmitted. The receive side of the system (at the other end of the link) will re-assemble the preempted fragments into their original frames. With this in place, time-sensitive frames experience much smaller delay (marked by tight FDV). In a fronthaul application, time-critical fronthaul traffic is configured to use high-priority express queues, while other, non-time-critical traffic is configured to use preemptable queues. The figure below shows a basic example of how frame preemption works, as further discussed in the next section.

Figure 1. The basics of frame preemption



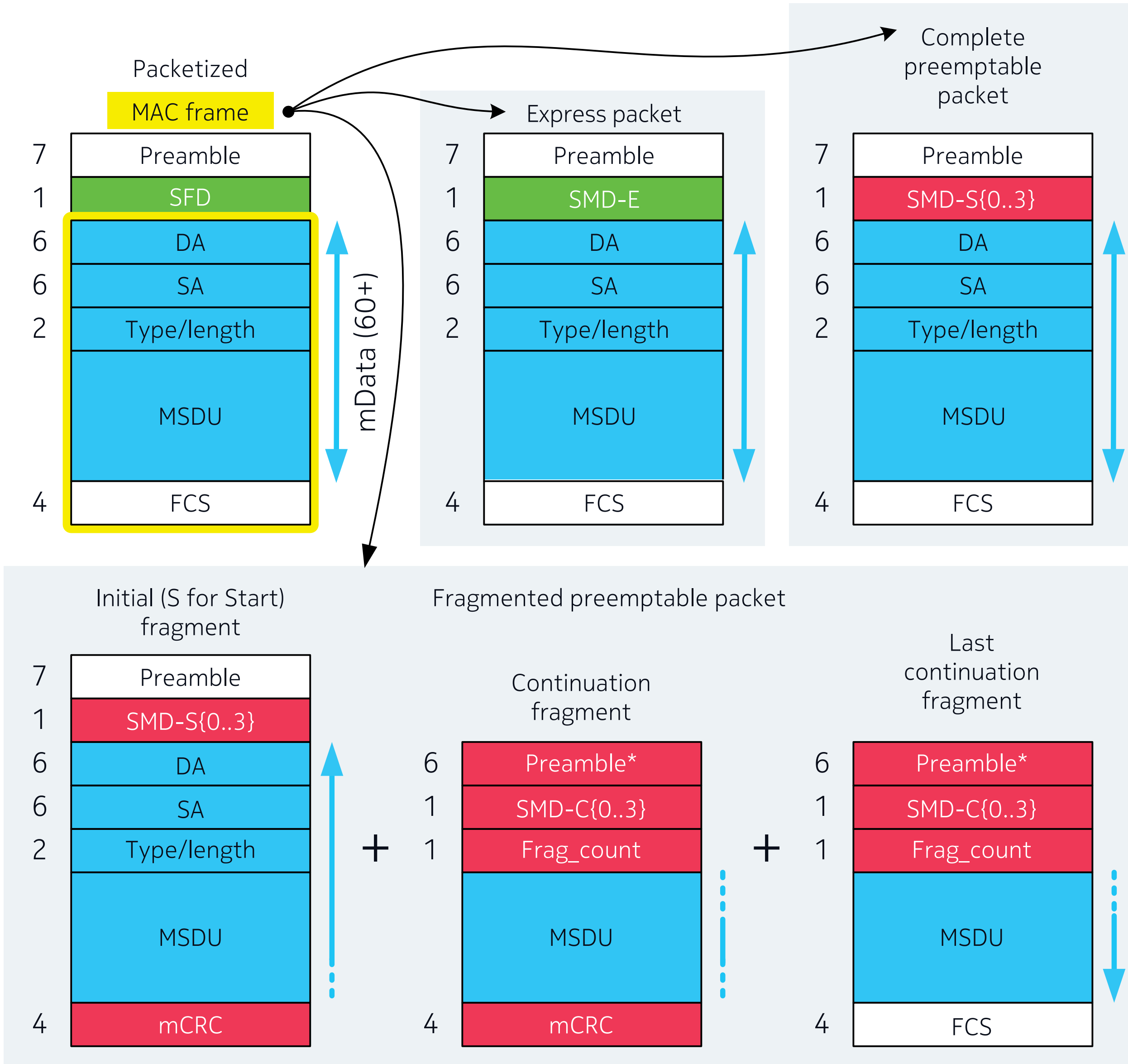
IEEE 802.1Qbu and IEEE 802.3br

Since it defines a new MAC, IEEE 802.1Qbu frame preemption requires specific hardware capabilities. For IEEE 802.1Qbu to be used, both the transmitter and the receiver on a link must enable frame preemption. It is important to note that frame preemption is disabled by default in the standard and must be enabled on all bridges in a network. To ensure that frame preemption is correctly maintained throughout the network, the values supported by priority queuing are marked in such a way that the underlying MAC layer knows which frames are express and which are preemptable. Thus, the specification of frame preemption in IEEE 802.1Qbu is complemented by the focus on aspects of the MAC layer, first standardized in IEEE Std 802.3br Interspersing Express Traffic (IET), which defined a new MAC Merge Sublayer (MMS) and MAC merge packets (mPackets). This has since been incorporated in IEEE Std 802.3-2018.

Traditionally, the IEEE 802.1 MAC layer only supported a single MAC service interface as the concept of preemptable frames did not exist. With IEEE 802.3br, the MAC layer leverages the MMS to support two MAC service interfaces: express MAC (eMAC) and preemptable MAC (pMAC). Frames associated with the express MAC, which are packetized as express packets, can interrupt the transmission of frames associated with the preemptable MAC. This results in frame fragments as the express frame(s) is interspersed between an initial fragment and a continuation fragment(s) of the same preemptable frame, which is transmitted after the express packet. When all fragments of a preemptable frame have been received on the next system, they are reassembled by the MMS into a complete preemptable packet that can be de-packetized to reveal the original preemptable frame. This fragmentation and re-assembly process are the reason why frame preemption must be supported at both ends of the link.

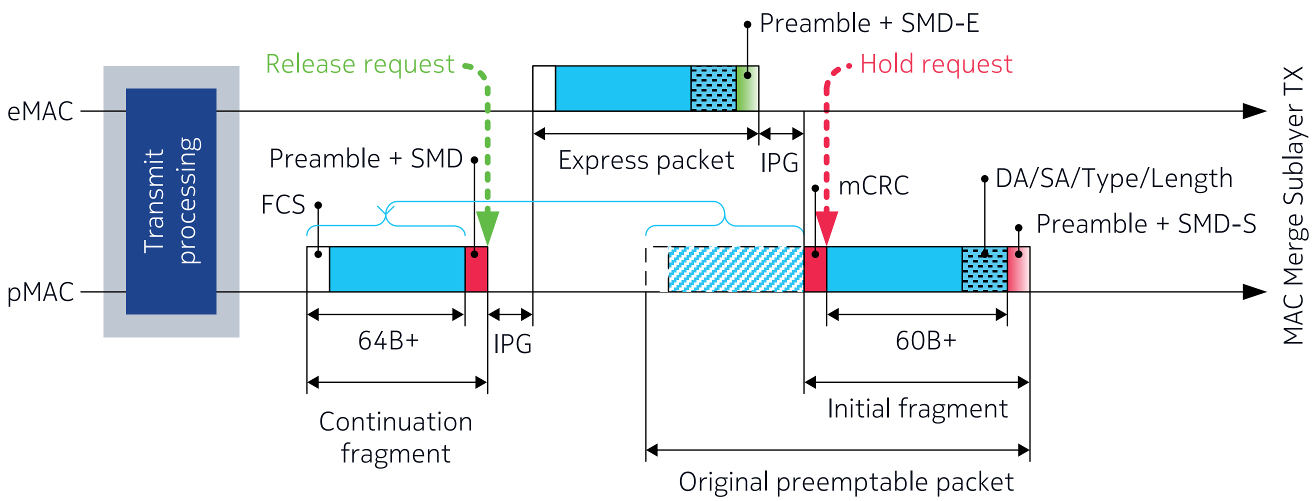
The packetization and de-packetization leverages mPackets to packetize express frames and preemptable frames. The Start Frame Delimiter for all mPackets is a new Start mPacket Delimiter (SMD). It can be an SMD-E for express frames. Since only pMAC frames can be interrupted and thus fragmented, the SMD for packetized preemptable frames includes an SMD-S for the start fragment and an SMD-C for any continuation frames. Except for the last one, each fragment is appended with a dedicated frame check sequence, called the mCRC, to allows its content to be verified at destination. A fragment count protects against mPacket reassembly errors by detecting a loss of up to three fragments. The figure below shows the formats for express and preemptable packets as well as the fragments of a preempted frame.

Figure 2. Only preemptable packets can be fragmented



Preemptable frames are preempted if a hold request is received by the MAC client or if an express frame needs to be transmitted; however, before preemption can take place, at least 60 bytes of the preemptable frame must have been transmitted and at least 64 bytes (including 4 bytes for the CRC) remain to be transmitted. This is illustrated in figure 3 below. The worst-case frame delay and frame delay variation correspond to the transmission of a 123-byte fragment.

Figure 3. The requirements for preempting a packet



Frame preemption configuration

Frame preemption operates on a per-link basis. Fragments of frames are only transmitted to the next system on the link where they must be reassembled. There is no support for end-to-end fragmentation. In other words, all interconnected bridges must support frame preemption. The Link Layer Discovery Protocol (LLDP) is used to communicate between systems to advertise frame preemption capability and configuration. Once a system that supports frame preemption receives an LLDP notification verifying that its link partner supports frame preemption, it turns on its own frame preemption capabilities. The system then safely assumes that its link partner understands the IEEE 802.3br-specified frame formats required to support fragmented frames.

Benefits of frame preemption

Skeptics, often protecting product limitations, will question whether frame preemption is really necessary, arguing that the greatest impact of frame preemption on latency is seen at lower port rates and that impact decreases as port rates increase. Earlier drafts of IEEE Std 802.1CM-2018 (up to D1.0) included an example in Annex B.2.2 (deleted in the final standard). It showed examples of latency calculations in a TSN bridged network composed of four bridges. Frame preemption saves ~5 µs on the total bridge delay over a total budget of 100 µs, which equates to a gain of ~1 km of distance (from 13 to 14 km). The difference between the two is the queuing delay for fronthaul data traffic corresponding to the worst-case frame preemption delay (1240-bit times).

It is true that the real benefit of frame preemption in terms of latency reduction decreases as rates increase, however:

- Testing reveals that there are certain conditions with high fan-in of client ports competing for the same line port, with transient congestion, where even at high rates, frame preemption shows a real benefit by saving multiple microseconds of frame delay, which is more than the nodal delay through the TSN bridge

- In practice, even in ASSP/ASICs, schedulers are not perfect and have inherent characteristics that make FDV a major problem for time-sensitive frames. Schedulers may transmit in microbursts instead of frame by frame allowing for features such as look-ahead scheduling. Frames in between the scheduler and the MAC are considered inflight frames and can negatively impact FDV. Once a frame has left the scheduler, it cannot be interrupted.

As testimony, the standard has been enhanced to reflect this. The initial published version of IEEE Std 802.1CM-2018 only asked to support frame preemption on ports whose data rate is not higher than 10 Gbps, and several products in the industry still exhibit this restriction. This restriction has in the meantime been removed in IEEE P802.1CMde. Frame preemption is now specified for all rates.

However, the latency reduction is not the most important aspect. The real value of frame preemption is to decrease the effects of non-fronthaul traffic on fronthaul traffic and to make network design for fronthaul traffic deterministic by bounding its latency and latency variation in the queued bridged network. This becomes crucial when mixing fronthaul and non-fronthaul services within a C-RAN fronthaul network offering converged multiservice access.

Frame preemption adds determinism to traditional Ethernet by enabling control of queueing and latency. This eliminates the need for connection-oriented TDM-like technologies, which are not optimized for fronthaul.

Summary

The evolution of networks to accommodate a wide variety of 5G-related services brings in a mix of traffic within the same network segment where the impact of non-time-critical traffic on time-critical traffic can be deleterious. A proven methodology to ensure that time-critical frames are transmitted with the smallest delay and delay variation is frame preemption. Frame preemption not only reduces the delay experienced by time-critical frames, it also reduces the delay variation. This makes the transmission of low latency services deterministic and reliable.

Glossary

ASSP	Application-specific standard product
ASIC	Application-specific integrated circuit
C-RAN	Cloud-RAN
CRC	Cyclic redundancy check
DWRR	Deficit-weighted round robin
eMAC	Express MAC
FDV	Frame delay variation
IET	Interspersing express traffic
IEEE	Institute of Electrical and Electronics Engineers
ITU-T	International Telecommunication Union – Telecommunications standardization
LLDP	Link layer discovery protocol
MAC	Media access control

MMS	MAC merge sublayer
mPackets	MAC merge packets
pMAC	Preemptable MAC
QoS	Quality of service
RAN	Radio access network
RR	Round robin
SMD	Start mPacket delimiter
SPQ	Strict priority queuing
TDM	Time-division multiplexing
TSN	Time-sensitive networking
WDM	Wavelength-division multiplexing
WRR	Weighted round robin

References

1. IEEE Std 802.1CM-2018, IEEE Standard for Local and metropolitan area networks – Time-Sensitive Networking for Fronthaul.
2. IEEE Std 802.1Qbu-2016, IEEE Standard for Local and metropolitan area networks – Bridges and Bridged Networks – Amendment 26: Frame Preemption.
3. IEEE Std 802.3br-2016, IEEE Standard for Ethernet Amendment 5: Specification and Management Parameters for Interspersing Express Traffic.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: SR2006044411EN (July) CID207583