Update on Cut-Through Forwarding (CTF)

Johannes Specht

Preamble

- **This Presentation collects thoughts on cut-through forwarding (CTF) and the outcome of past discussions in IEEE 802.1.**
- **It is intended to move towards a common view in IEEE 802.1 amongst goals, needs, and operation of potential IEEE 802 standardization activities on CTF.**
- **This is an individual contribution.**

Recap

January 2020

- https://www.ieee802.org/1/files/public/docs2020/new-specht-cut-through-tech-0120-v01.pdf
- Where CTF matters, and where not
- CTF-specific issues and mitigations
- Proposed contents of a standard

December 2020

- https://www.ieee802.org/1/files/public/docs2020/new-specht-cut-through-update-1220-v02.pdf
- Approach of a WYSIWIG working document, work in progress
- Basis for discussion in 802.1 and other 802 WGs

Proposed direction of the working document

Objectives

- Demonstrate where Cut-Trough Forwarding matters
- Preview on how an IEEE 802.1 Standard (not an amendment) for Cut-Through Forwarding could look like
- Problems IEEE 802.1 cannot solve, for discussion with other IEEE 802 WGs
- Readable, comprehensible, etc.

IEEE 802[.1] Standards environment

- Fit into the IEEE 802.1 Stds environment
- Stay within the IEEE 802.1 Stds environment (layers)
- Reflect IEEE 802.1 participants

Brownfield (i.e., CTF is already implemented and used)

- Capture representative use-cases
- Representative subset of mechanisms for CTF

Working document vs. Standard

No approved IEEE SA project

No balloting process

- Development is different
- Exchange drafts/pieces with IEEE 802.1 people interested in the topic

Options and optional mechanisms in existing IEEE 802.1 Stds

- IEEE 802.1 Std: Often desirable to explore
- Document: Determine reasonable options for use-cases (while keeping compatibility in mind)

Existing CTF mechanisms (brownfield) = new mechanisms in IEEE 802.1

One possible approach:

- If motivated and within IEEE 802.1 → incorporate
- If beyond IEEE WG 802.1 → capture the problem these mechanisms solve

Navigation: Purpose of the subsequent slides

What to find in the document

1. Structural
   - Clauses
   - Content assignment
   - Relationships
2. Technical
   - Use-cases
   - Proposed technical choices
   - Open technical choices

Purpose

- Early feedback
- Pointers

Structural Overview

Top Level Structure

Structure of an IEEE SA Standard

1. Overview
2. Normative References
3. Definitions
4. Abbreviations
5. Conformance

Not now

6. Cut-Through Forwarding in Networks
7. Cut-Through Forwarding Relay
8. Managed Objects
9. YANG Modules
A. PICS

Not now

B. Bibliography
Z. Open Issues

Structure of the working document

1. Introduction
2. Cut-Through Forwarding in Networks
3. Cut-Through Forwarding Relay
Bibliography
Open Issues

Top Level Structure - Contents

1. Introduction

- *Introduction, Glue for Subsequent Clauses*
- *Not a Standard, and only subset of mechanisms/options from 802.1 Stds*

2. Cut-Through Forwarding in Networks

- High-Level Use-Cases (application independent): Topologies, Traffic Patterns/Scheduling*
- Network Structure and Elements *(Wired P2P, Extensions in Bridges, No Extensions in End Stations)*
- Performance Considerations *(a.k.a. where CTF matters, and where not)*
- QoS Maintenance *(Frame/header errors, impact, etc.)*

3. Cut-Through Forwarding Relay

- Bridge Port Transmit and Receive *(Demultiplexing, etc.)*
- Augmented Forwarding Process
- *Forwarding Process Function 1..n (Existing ones included, and new ones)*

Specific Structure

- Use-case area
- Details on next slides

Straight Forward Structure

- Bridge pipeline and operational model
- Details on next slides

Bibliography

- Standards: IEEE Std 802.3, IEEE Std 802.1 802.1AC-2016 Cor 1-2018, IEEE Std 802.1 802.1Q-2021, IEEE Std 802.1 802.1CB-2017, IEEE Std 802.1 CBcv-2021
- IEEE 802.1 contributions (http://www.ieee802.org/1/files/public/)
- External (e.g., Papers)

Open Issues

Note *: Term "Scheduling" is used in the broad sense for shaping, coordinated transmission times, TDM, etc. (i.e, not tied to what is called "Scheduled Traffic" in IEEE Std 802.1Q-2018).

2. Cut-Through Forwarding in Networks

Structure

2. Cut-Through Forwarding in Networks

2.1 Chain Networks
2.1.1 General
2.1.2 Communication Schemes
2.1.3 Quality of Service Maintenance

2.2 Ring Networks
2.2.1 General
2.2.2 Communication Schemes
2.2.3 Quality of Service Maintenance

2.3 Link Speed Transitions

Assumption:
Covers the majority of use-cases on a higher abstraction level

Placeholder/“Special topic area”:
In this case, if exclusion in 2.1/2.2 causes is not obvious

Structure

2. Cut-Through Forwarding in Networks

2.1 Chain Networks

2.1.1 General

2.1.2 Communication Schemes

2.1.3 Quality of Service Maintenance

2.2 Ring Networks

2.2.1 General

2.2.2 Communication Schemes

2.2.3 Quality of Service Maintenance

2.3 Link Speed Transitions

- Topology
- Lower layer properties
- CTF & S&F locations
- Frame structure

2.1.2.1 Overview
2.1.2.2 Uncoordinated
2.1.2.3 Coordinated Talkers
2.1.2.4 Class-based Time Division Multiplexing
2.1.2.5 Full Time Division Multiplexing

2.1.3.1 Undetected frame errors, impact and mitigations
2.1.3.2 Filtering and policing

Structure

2. Cut-Through Forwarding in Networks

2.1 Chain Networks

2.1.1 General

2.1.2 Communication Schemes

2.1.3 Quality of Service Maintenance

2.2 Ring Networks

2.2.1 General

2.2.2 Communication Schemes

2.2.3 Quality of Service Maintenance

2.3 Link Speed Transitions

Differences to 2.1.1

- Ring = chain + (1 link)
- S&F and CTF paths in a ring
- Frame format

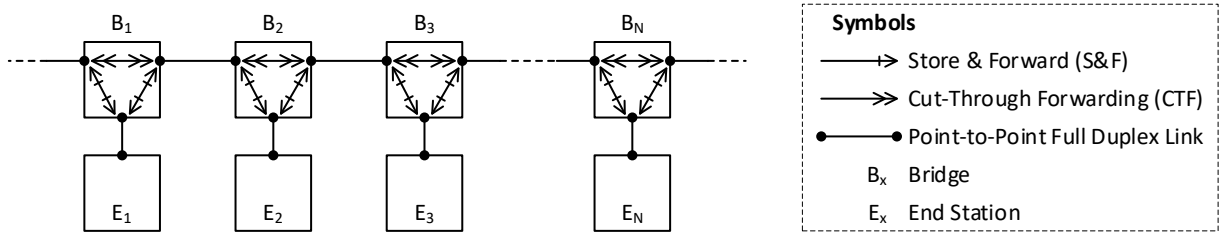
- reference to 2.1.2 (full duplex property ...)
- Redundant paths

2.2.3.1 Loop and loop prevention
(cmp. [802.1Q, 6.5.4][802.1CB, C.7])

2.2.3.2 Logical chains in ring networks

2.2.3.3 Frame shortening

Chain networks: General (2.1.1)



Lower layers

- Full-duplex point-to-point
- Same MAC type
- Identical link speeds
- Negligible signal propagation delays

TSN

- Opt. Preemption (highest priority)
- Strict priority + Opt. Tx Gates + Filtering/Policing

CTF and S&F locations

- S&F between Bridges and end stations
- CTF for high priority traffic between Bridges

C-TAG (Priority)

- Distinction by priority
  - CTF traffic (high priority)
  - S&F traffic (low priority)
- Wide support assumed
  - VLAN-aware
  - VLAN-unaware [802.1Q, 6.20]

Implies no re-tagging/tag removal & insertions

- Simplifies discussion
  - frame shortening (data stalls on transmission, etc.)
  - FCS/CRC re-computation

Chain Networks: Communication schemes overview (2.1.2.1)

Paths

- Interferences matter for quantitative comparison

Talker Transmissions

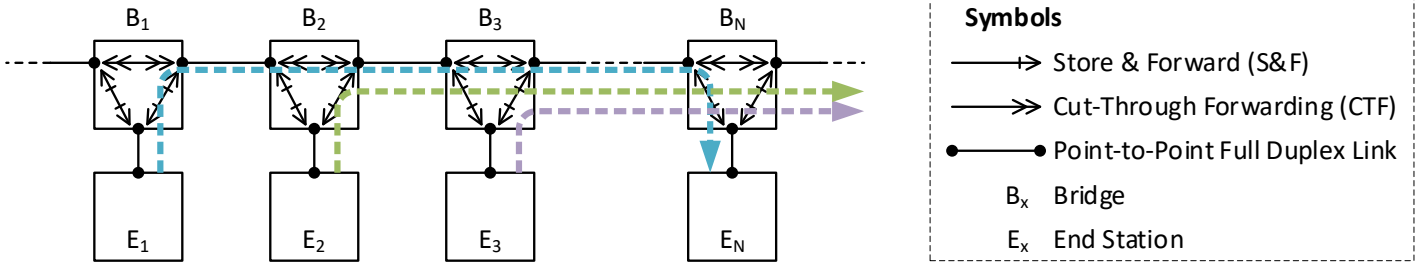
- Periodic
- Max. E2E Latency << Period
- No interference of CTF traffic by S&F on 1st hop

Per communication scheme

- Goal: Quantitative comparison, with and without cut-through
- Ordering: Incremental
  - Easiest to understand [2.1.2.2] to
  - Most latency enhancement [2.1.2.5]

Errors excluded

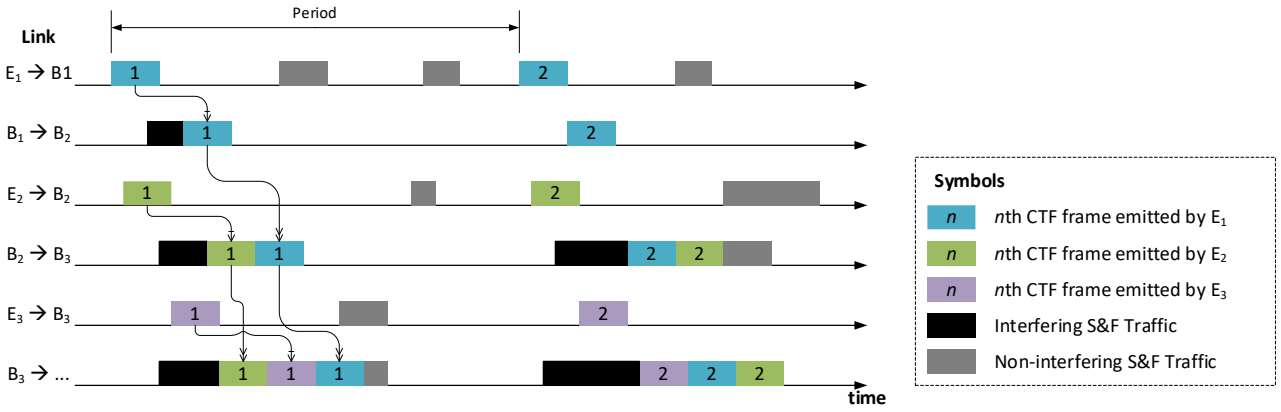
- Separate consideration in 2.1.3



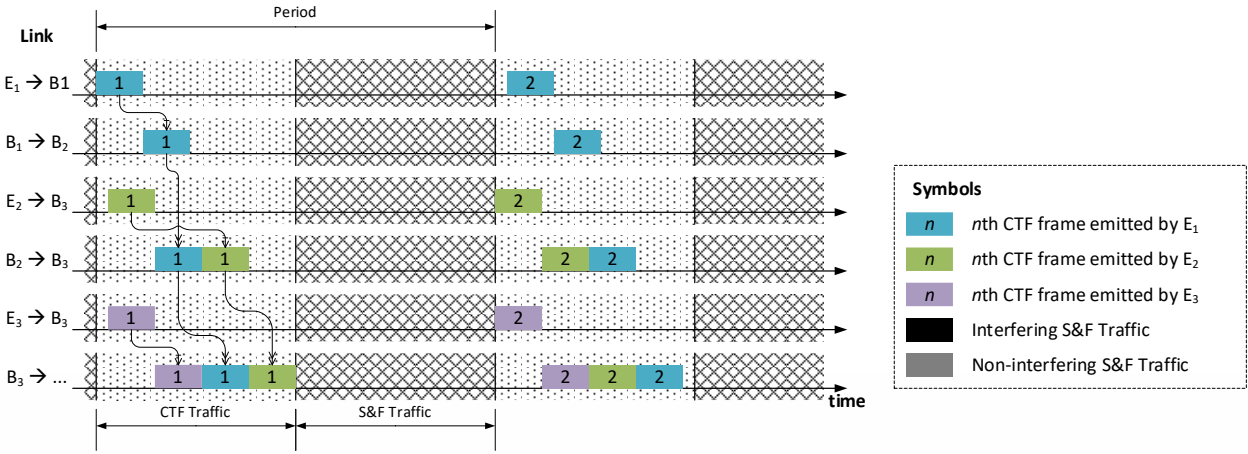
Name	Interference (Bridge transmission Ports)		Reference
	CTF Traffic ⇔ CTF Traffic	CTF Traffic ⇔ S&F Traffic	
Uncoordinated	Yes	Yes	2.1.2.2
Coordinated Talkers	No	Yes	2.1.2.3
Class-based Time Division Multiplexing	Yes	No	2.1.2.4
Full Time Division Multiplexing	No	No	2.1.2.5

Chain Networks: Communication schemes (2.1.2.2 through 2.1.2.5)

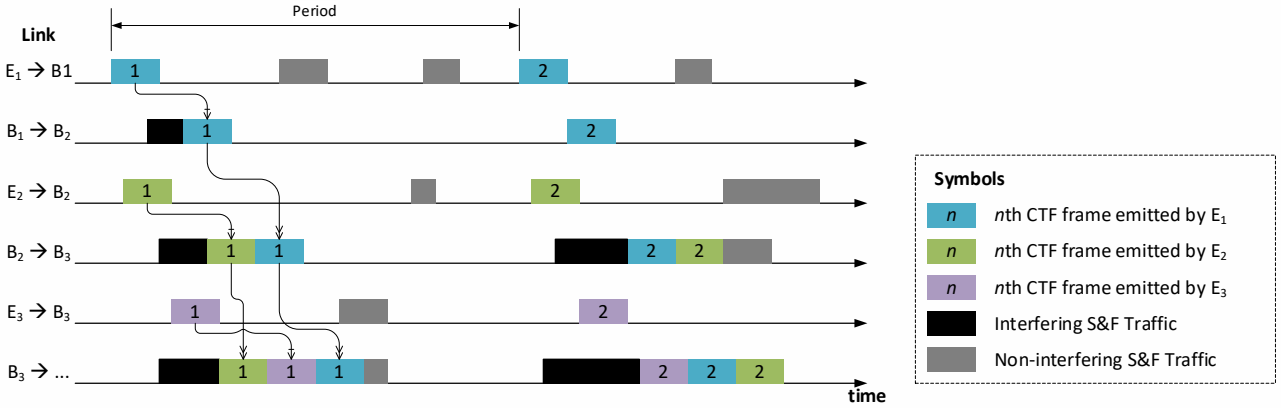
2.1.2.2 Uncoordinated



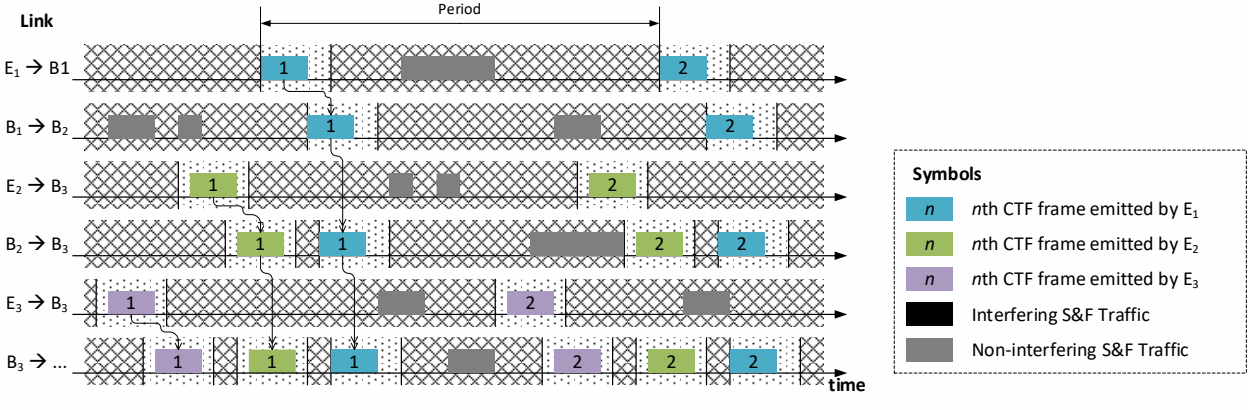
2.1.2.4 Class-based Time Division Multiplexing



2.1.2.3 Coordinated Talkers



2.1.2.5 Full Time Division Multiplexing



Chain Networks: Undetected frame errors, impact and mitigations (2.1.3.1)

Problem

Additional undetected errors by CTF (cmp. [802.1Q, 6.5.7])

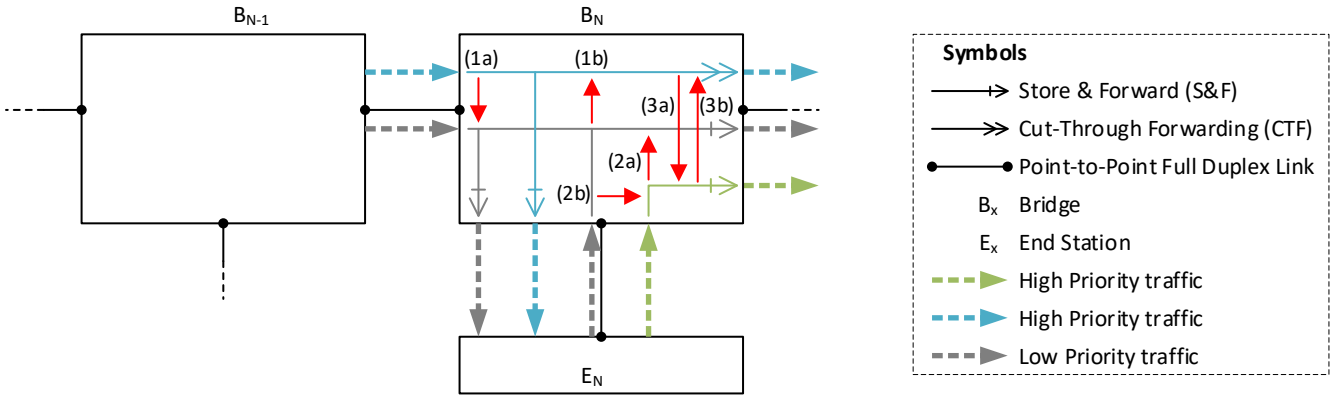
Impact

Additional congestion due to:

1. Wrong transmission port selection
2. Wrong traffic class selection

Further Reduction

- Low priority S&F traffic **from bridge to bridge** classified as high priority CTF traffic

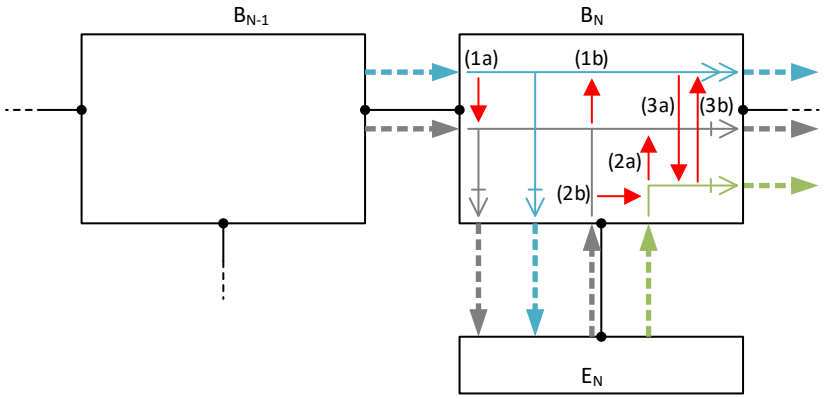


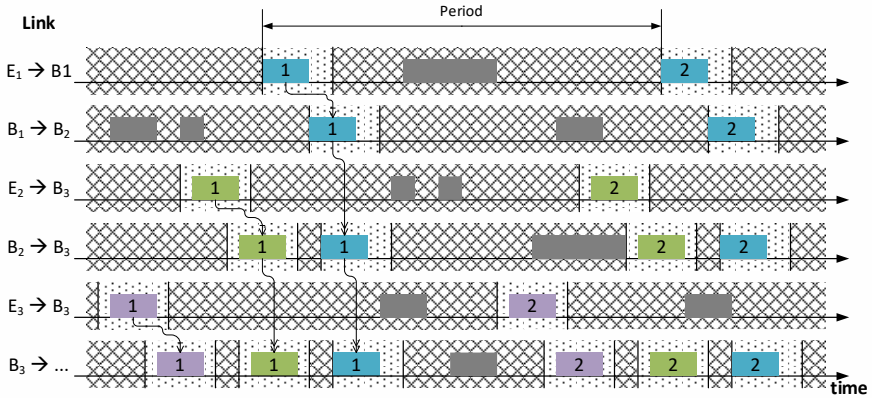
Case	Description	Detected before transmission	Mitigation	Reference
(1a)	High priority CTF traffic from B_{N-1} classified as low priority S&F traffic.	Yes	N/A	2.1.1
(1b)	Low priority S&F traffic classified as high priority CTF traffic.	No	Filtering and Policing	2.1.3.2
(2a)	High priority CTF traffic from E_N classified as low priority S&F traffic.	Yes	N/A	2.1.1
(2b)	Low priority S&F traffic from E_N classified as high priority CTF traffic from E_N.	Yes	N/A	2.1.1
(3a)	High priority CTF traffic from B_{N-1} classified as high priority S&F traffic from E_N.	Yes	N/A	2.1.1
(3b)	High priority S&F traffic from E_N classified as high priority CTF traffic from B_{N-1}.	Yes	N/A	2.1.1

Chain Networks: Filtering and policing (2.1.3.2)

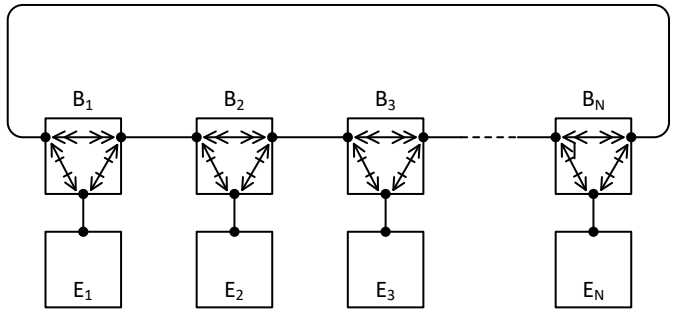
Communication scheme dependency

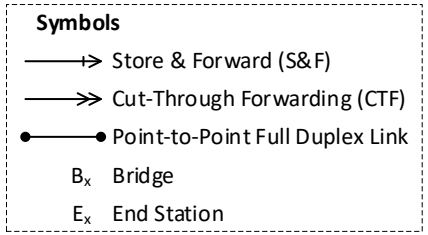
- TDM schemes (2.1.2.3 and 2.1.2.4)
  → Maximum SDU size filtering (3.7.3) + stream gating (3.7.4)
- Asynchronous schemes (2.1.2.1 and 2.1.2.2)
  → Maximum SDU size filtering (3.7.3) + flow metering (3.7.5)





Ring Networks: General (2.2.1)





Lower layers

- As in 2.1.1

CTF & S&F locations

- As in 2.1.1

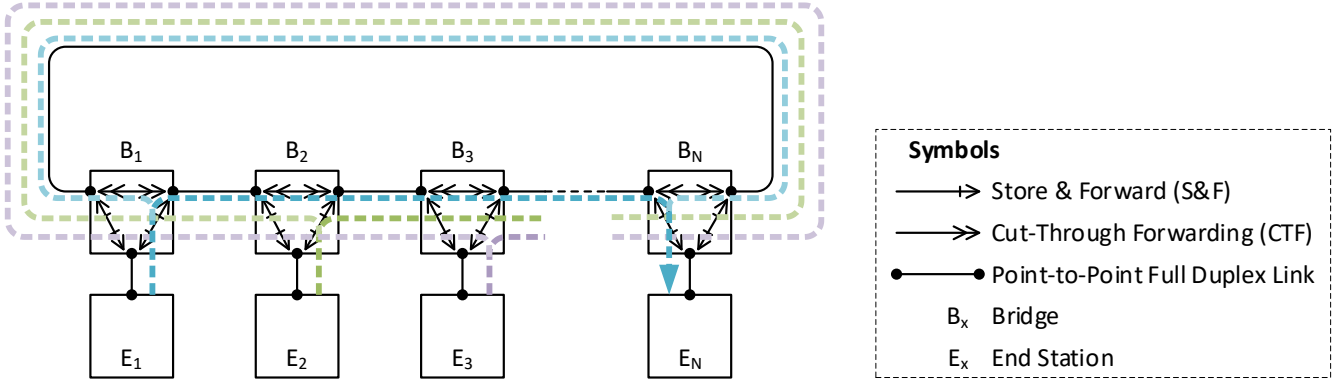
C-TAG/Priority-tagged

- As in 2.1.1

R-TAG

- **E_x → B_x:** Splitting to both directions of the ring
- **B_x → E_x:** Sequence recovery
- *Yet, a reasonable choice out of the options of IEEE Std 802.1CB-2017 is required …*

Ring Networks: Communication (2.2.2)



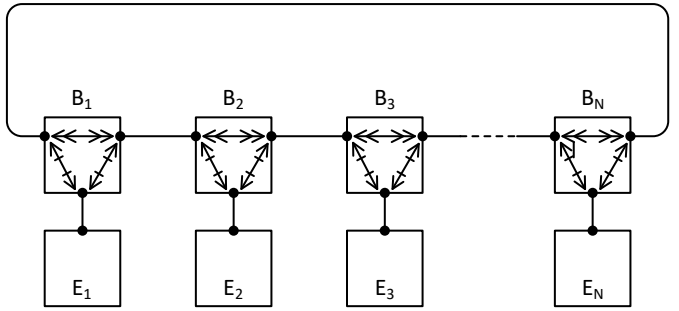
Comparison to 2.1.2

- Identical assumptions on talker transmissions and errors
- Treating “long path” and “short path” separately (full duplex links)
- Quantitative consideration is as in 2.1.2 (i.e., separation does not affect quantitative comparison of CTF with S&F)

Ring Networks: Quality of Service Maintenance (2.2.3.1)

Problem

- Same problem (additional undetected errors)
- Different Impact



Impact

- Loops/circulating frames
- Cmp. [802.1Q, 6.5.4] and [802.1CB, C.7]

Reduction

- Again, treating “long path” and “short path” separately (full duplex links)

Goal definition

Once a frame became erroneous, this frame is removed in a ring network with N Bridges after at most N hops.

Mitigations

Logical chains in ring networks (2.2.3.2)

- Force S&F for all traffic in one bridge.
- Satisfies the goal
  - First error after this S&F bridge
  - Discovered and removed after N hops (reaching this bridge again).
- Increases latency for all high priority traffic.

Frame shortening (2.2.3.3)

- Force S&F for all traffic in one bridge.
- Satisfies the goal, but under constraints
  - Erroneous frames cannot exceed a maximum size S (max. SDU size filtering).
  - Bridge B_x shortens erroneous frames by at least $T_{min}(B_x)$.
  - Frame removed after N hops or earlier:

$$S \leq \sum_{x=1}^{N} T_{min}(B_x)$$

Note: More details are found in https://www.ieee802.org/1/files/public/docs2020/new-specht-cut-through-tech-0120-v01.pdf

3. Cut-Through Forwarding Relay

High-level model: *Transient Frames, Complete Frames, Stalls,* and *Late Discarding*

Properties of *Transient Frames*

- Device Internal
- Content visible can change over time
- *Late discarding* (e.g., FCS errors)
- Only for Relaying path
- Relay stages *stalled* until enough content is available

Distinction in Descriptions

- *Transient Frames* v.s.
- *Complete Frames* (just “Frame” in IEEE 802.1Q)

Transient Frames v.s. Complete Frames

- Receive Timing
  - *Transient Frames*: At Frame Start from the Wire
  - *Complete Frames*: After Frame End from the Wire
- Transient Frames can be completed
  - Become Complete Frames (e.g., if FCS ok)
  - *Late discarding* (e.g., if FCS is not ok)

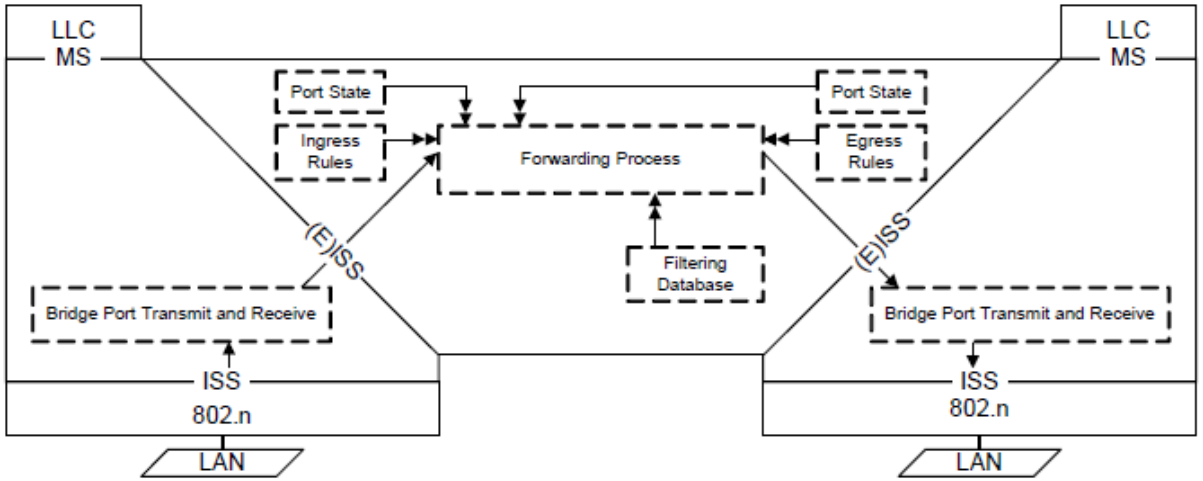


Figure 8-4—Relaying MAC frames

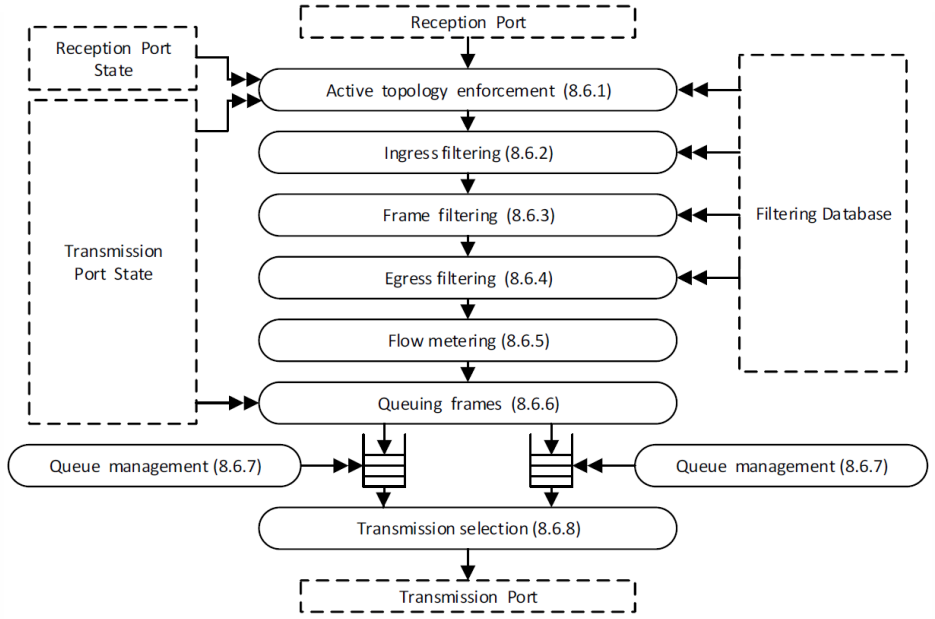


Figure 8-12—Forwarding process functions

Matching into the structure

- Relaying of transient frames
- Introduce:
  - “diff”-clause concept
  - Essentially only for transient frames
  - Absent 802.1* functions: unsupported
  - Pipeline stages/stalls/late discarding

- Initial List (subclauses added/removed over time)
- Case-by-case diff to 802.1Q
  - Min: “As described in A.B.C of 802.1Q-20XX.”
  - Typical: Different handling of transient frames/late discarding
  - Max: New stages (not illustrated)

3. Cut-Through Forwarding Relay

3.1 General

3.2 Bridge Port Receive

3.3 Active Topology Enforcement

3.4 Ingress Filtering

3.5 Frame Filtering

3.6 Egress Filtering

3.7 Flow Classification and Metering

3.8 Queuing Frames

3.9 Queue Management

3.10 Transmission Selection

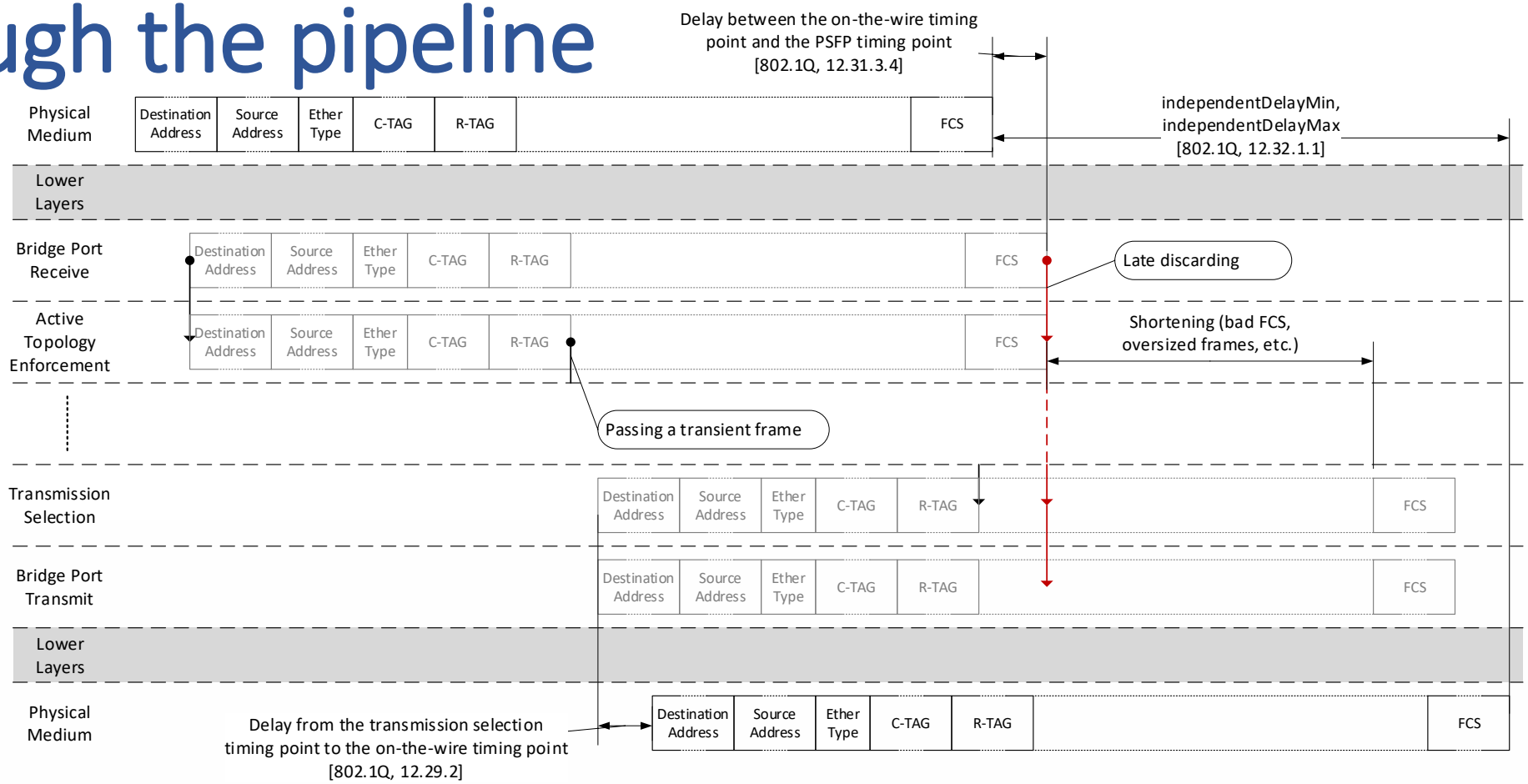
3.11 Bridge Port Transmit

- Relationship:
  - Transient Frames v.s.
  - Complete Frames v.s.
  - M_UNITDATA.indication
- Transient frames not sent to higher layer entities in Bridges [802.1Q, 8.5]
- Stage may stall:
  - Any stage could do so, if transition to Complete Frame required

- Multiplexing:
  - Higher Layer PDUs,
  - Transient Frames,
  - Complete Frames
- Handling late discarding of transient frames

Note: Arrangement of *Bridge Port Receive* and *Bridge Port Transmit* is in a pipeline manner, not in a combined/layered manner (like in IEEE Std 802.1Q). This proposal is a trade-off: On the one hand, the pipelined manner appeared more readable to the author, and it’s no obligation to organize contents identical to IEEE Std 802.1Q. On the other hand, symmetry might be helpful for readers familiar with IEEE Std 802.1Q.

Path through the pipeline



A possible illustration

The externally visible behavior matters → most timing properties don't need to be exposed.

Close to IEEE Std 802.1Q

Multiple relevant timing elements standardized, at most two new proposed external visible timing elements:

(cmp. https://www.ieee802.org/1/files/public/docs2019/60802-Ademaj-et-al-CutThrough-0919-v11.pdf)

1. Shortening timing
2. Optional initial delay in Bridge Port Receive (not illustrated above)

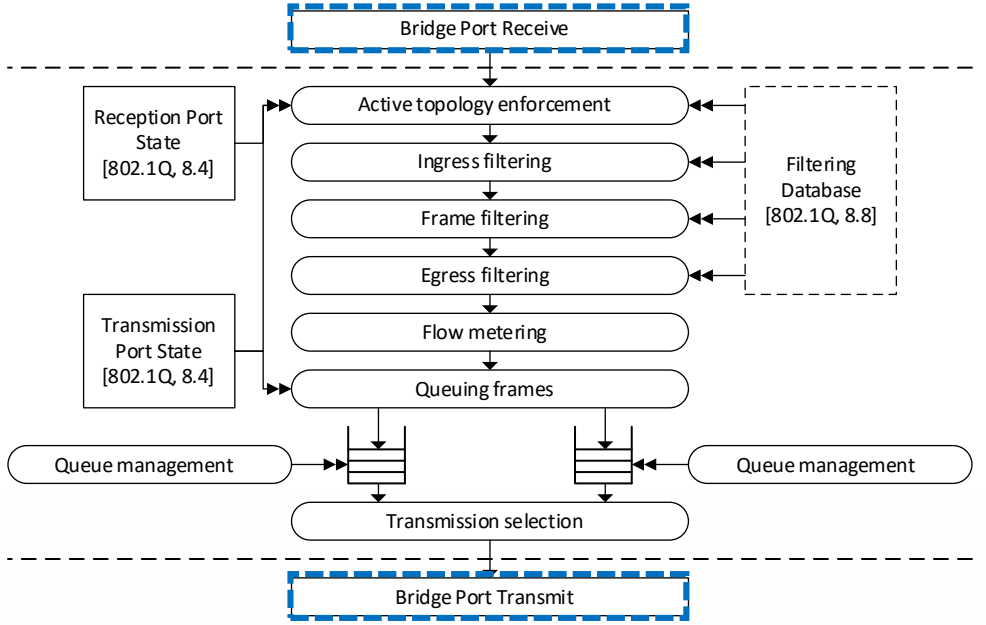
Interacting with Lower Layers

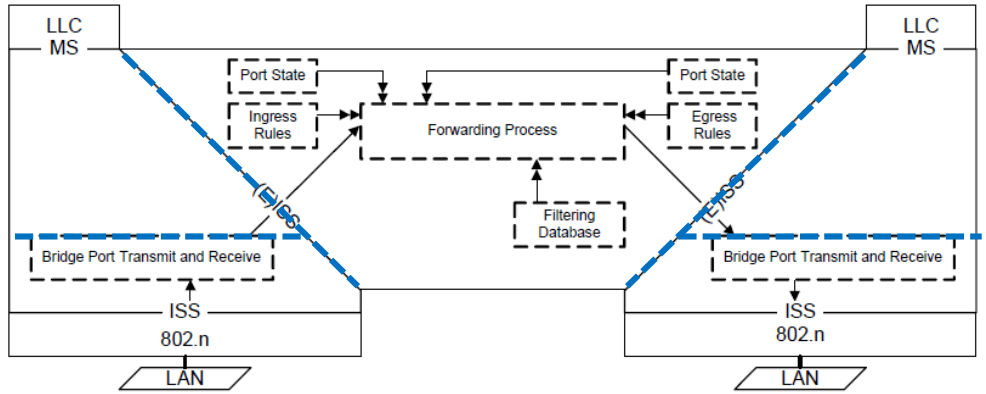
Situation

- S&F: Standardized
  - MAC [802.3]
  - (E)ISS and support functions [802.1Q, 6.6 ff.]
  - MAC Services/Translations [802.1AC]
  - Stream identification, Sequencing recovery/decoding/encoding [802.1CB, 8.1 & 8.2]
- CTF: None of these

Making Assumptions (not particular solutions)

- Description in a Relay boundary oriented manner
- Information elements used in 3.3 through 3.10
  - Which ones
    destination address, source address, drop eligible, priority, stream handle, frame check sequence, current length received[, frame start][, service data unit][, sequence number]
  - Encoding/Decoding
    References into 802.1 Stds
  - When
    - Assumed association with physical frame contents
    - assignment/update to transient frames





For illustration in this slide set.

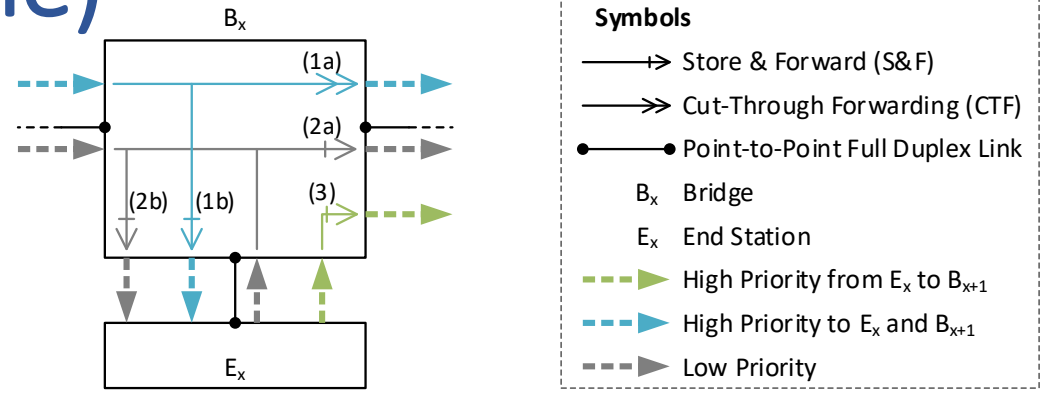
Differentiating CTF and S&F frames (and non-preemptible & preemptible)

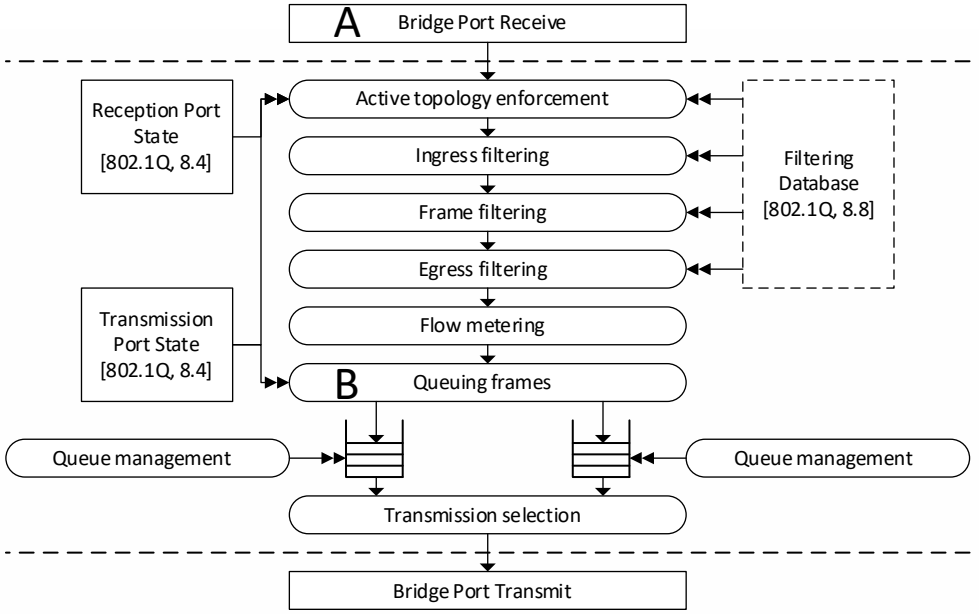
Considering 2.1 and 2.2

- A: Per reception port
  - Covers (2a), (2b) and (3)
  - Insufficient to distinguish (1a) from (2a), (2b) and (3)
- B: Per output port per class
  - Can distinguish (1a) from (2a)
  - Insufficient to distinguish (1a) from (3)
- Combination of both, A+B
  - A: Earliest stage is Bridge port receive [3.1]
  - B: Earliest stage is Queueing frames [3.8]

Possible constraints on B

- Less than 8 classes
  (2 appear sufficient for 2.1 and 2.2)
- Not all classes support CTF





New CTF functions affecting lower layers

Executing Frame Shortening

(cmp. https://www.ieee802.org/1/files/public/docs2019/60802-Ademaj-et-al-CutThrough-0919-v11.pdf)

- **Purpose:**
  Abort ongoing transmission of a transient frame → don't add a valid FCS → mark this frame "seen invalid"
- **Options to discuss:**
  - Describe the problem (open issues)
    Based on 2.2.3.2 (frame shortening in ring networks), late discarding and gaps in clause 3.
  - Think about a special FCS?
    Available in the relay, though it rather seems like a lower layer topic.
  - Other?

Header CRCs

(December 21st 2020 discussion)

- **Purpose:**
  Generic tool against errors causing wrong priority assignment and wrong output port selection of CTF frames.
- **Options to discuss:**
  - Describe the problem (open issues)
    The approaches in 2.1.3 and 2.2.3 are specialized and add complexity. In addition, it seems possible to reduce end station S&F hops.
  - Skip for now
    There are at least approaches in 2.1.3 and 2.2.3
  - Other?

Thank you for your Attention!

Questions, Opinions, Ideas?

Johannes Specht
Dipl.-Inform. (FH)

Kurfürstenwall 2
45657 Recklinghausen
North Rhine-Westphalia
GERMANY
M +49 (0)170 718-4422
johannes.specht.standards@gmail.com

Backup

Discussion points from December 21st 2020: Summary

Call it *Bridge*, or different?

→ For now, [CTF] Bridge works, at least for me

Conformance clause in the working document?

→ Under discussion

Criteria to stall transient frames until completion/discarding

→ #1: Fallback to the standardized operation during the forwarding process

→ #2: Enforce S&F at selected points (minimal proposal on a later slide)

Preemption: CTF only for non-preemptible traffic assumed

→ OK

Late Discarding (supported/unsupported relay functions)

→ Most parts of the forwarding process that can discard may result in late discarding

Configuration: Static only assumed, not dynamic ([R]STP)

→ Appears wrong, though it seems to require no special consideration