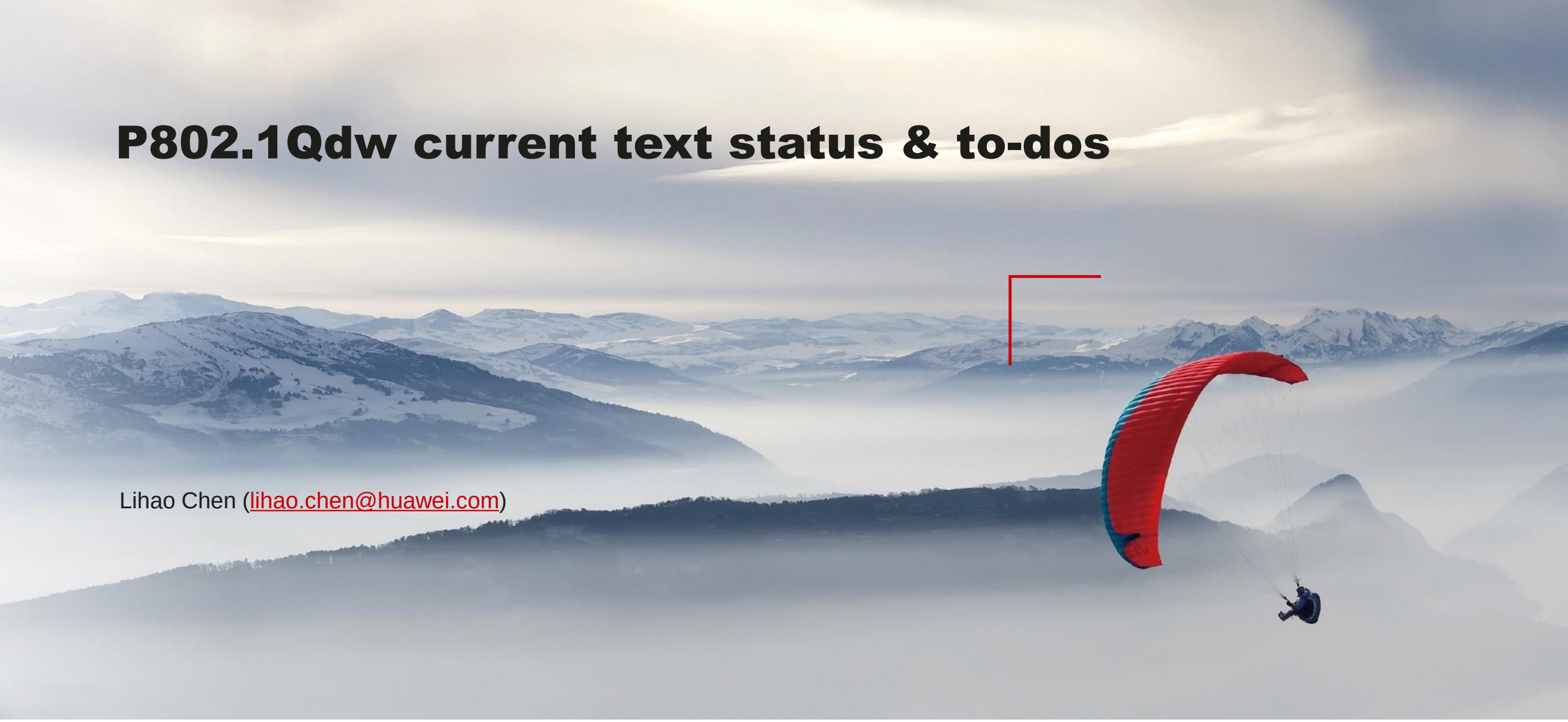
# P802.1Qdw current text status & to-dos



Lihao Chen (lihao.chen@huawei.com)

HUAWEI

# Links

- This presentation is based on the latest version of the text contribution:
  - https://www.ieee802.org/1/files/public/docs2025/dw-chen-individual-text-0325-v03.pdf

- 2025 January Interim.
  - Feedback: identify the gaps in the text and make a list to drive discussions and bring contributions to fill the gaps.
  - https://www.ieee802.org/1/files/public/docs2025/dw-chen-sfc-proposal-0125-v01.pdf
- 2024 November Plenary, the SFC procedures were presented.
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-text-contribution-overview-1124-v02.pdf
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-individual-text-1124-v02.pdf
- 2024 September Interim, the whole document of the text contribution was presented.
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-text-contribution-overview-0924-v01.pdf
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-individual-text-0924-v01.pdf
- 2024 July Plenary
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-recap-restart-0724-v01.pdf

# Overall progress list

| Clause | Content | Appr. progress | Description |
|---|---|---|---|
| 1-6 | Intro., def., Abbr., Conformance, MAC. | 90% | Revise after the main text is completed and upon comments. |
| 12, 48 | Bridge management, YANG | 10% | Start after the SFC protocol specification is mostly finished (e.g., 52.5.1 Variables). |
| 52.1, 52.2 | SFC, objectives, principles | 90% | Revise after 52.5 completion and upon comments. |
| 52.3, 52.4 | SFC bridge and end station component process | 90% | Revise after 52.5 completion and upon comments. |
| 52.5.1 | Variables controlling operation | 70% | Revise upon SFC protocol specification, i.e., procedures and encodings. |
| 52.5.2 | SFCP procedures | 50% | The architecture of the SFCP procedure, as well as some of the detailed contents are already included. Though, the author plans to update this sub-clause in the next version. |
| 52.5.3 | Encoding of the SFCM PDU | 70% | 3 types of encapsulation and a basic SFCM PDU format are already included. |
| 52.5.4 | LLDP Source Flow Control TLV | 10% | To be started. |
| Annex Y | Buffer requirements for SFC | 5% | Informative, giving a delay model and a computation example. |

# Detailed TBD technical items

| Sub-Clause | Description | Gap? |
|---|---|---|
| 52.3.4 | Is the SFC Source Table really needed? Another option is to send SFCM based solely on the reported congested frame and configuration variables. Scalability vs Simplicity. | |
| 52.5.1.1.5 | Register a UDP port number? Otherwise need configuration. | |
| 52.5.2 | The whole procedure sub-clause needs refinement. | Yes. |
| 52.5.2.4 | The throttling mechanism, other than the configured sfcmMinInterval. | |
| 52.5.2.4 | Specify the calculation of the pause time (informative)? | |
| 52.5.3 | How do we know which type of SFCM to send? Currently based on identifying the type of the congested packet. | |
| 52.5.4 | LLDP and DCBX to advertise the new capability. Automatic enabling proxy mode? | Yes. |
| Annex Y | Give a delay model and a computation example, etc. | |
| / | Specify the minimum reaction time to a SFCM? | |
| / | Optionally provide means to specify congestion locators. | |
| / | May provide per-flow control including modified procedures and a new SFCM PDU types. | |

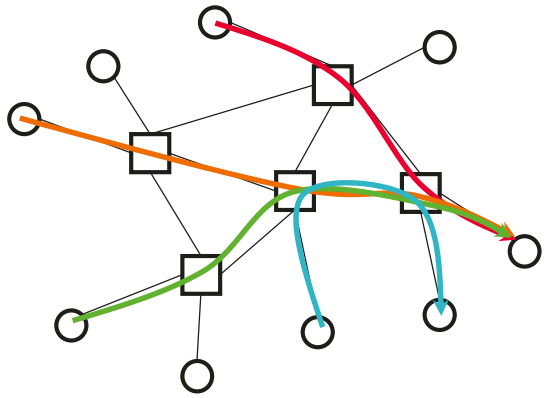
- The author believes that, based on the latest version of the text contribution, as long as the items marked as "yes" in the last column of the table are addressed, a basic yet complete SFC specification that fulfills the scope will be achieved.
- Other items are options. If no other contributions are received, the author plans to comment to address these items during the TG ballot.

# Questions?

# The SFC benefits for IEEE 802 congestion and flow control

- The differences and similarities compare to CN (Qau), PFC (Qbb), and CI (Qcz):
  - Page 4 and 5 in https://www.ieee802.org/1/files/public/docs2022/new-congdon-SFC-proposal-0322-v01.pdf
  - Clause 52.2.5, 52.2.6, and 52.2.7 in https://www.ieee802.org/1/files/public/docs2024/dw-chen-individual-text-1124-v02.pdf
- A comparison of PFC, DCQCN, SFC in P18 https://www.ieee802.org/1/files/public/docs2022/new-bottorff-sfc-0322-v6.pdf
- Besides the commonly used end-to-end congestion control algorithms and hop-by-hop flow control (i.e., PFC), a new edge-to-edge flow control scheme is needed.
  - Page 4 in https://mentor.ieee.org/802.1/dcn/21/1-21-0055-00-ICne-source-flow-control.pdf

- Key features:
  - **Fastest reaction** to stop Reaction Point Egress.
    - Crucial for ultra-high-bandwidth Ethernet.
    - “Precise PFC” mitigates PFC side effects, e.g., PFC storm, HoL, deadlock.
  - **Non-scenario-specific**. Congestion cased by many-to-one traffic is very common.
  - **Easy adoption**. No mandatory math, and end stations don't necessarily need to be modified.



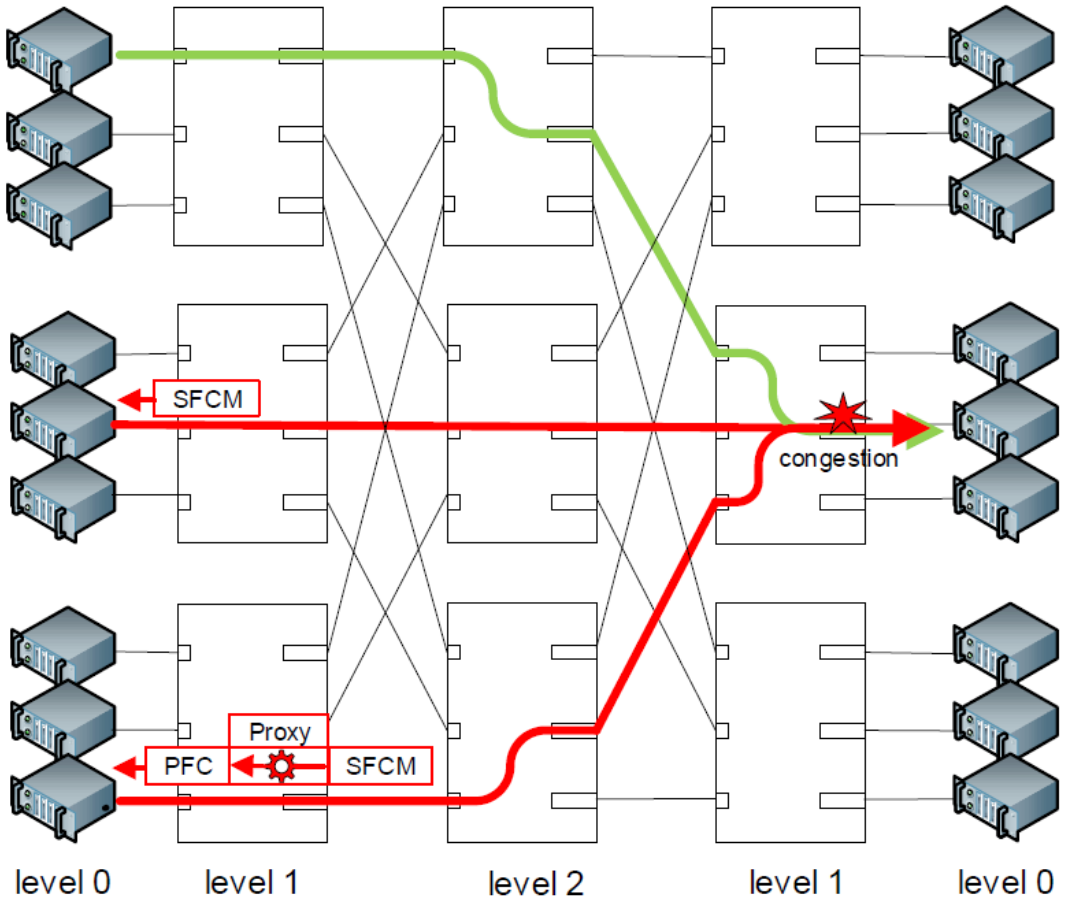
# Proposal for the text - Structure aspects

- Refer to previous standards and text contributions.
  - IEEE 802.1Q-2022 Clause 30-33 (Quantized) Congestion Notification.
  - IEEE 802.1Qcz-2023 Congestion Isolation.
  - https://www.ieee802.org/1/files/public/docs2022/dw-congdon-individual-text-1122-v01.pdf
  - https://www.ieee802.org/1/files/public/docs2024/dw-chen-individual-text-1124-v02.pdf

| Contents | QCN (in 802.1Q-2022) | CI (in 802.1Qcz-2023) | SFC (in individual text) |
|---|---|---|---|
| Objectives | Clause 30 | Clause 49.1 | Clause 52.1 |
| Principles | Clause 30 | Clause 49.2 | Clause 52.2 |
| Entity operation | Clause 31 | Clause 49.3 | Clause 52.3 and 52.4 |
| Protocol (variables, procedures, encodings) | Clause 32 | Clause 49.4 | Clause 52.5 |

# Proposal for the text - High level technical aspects

- Follow the scope to provide a straightforward solution.
  - This amendment specifies procedures, managed objects, and a YANG data model for **the signaling and remote invocation of flow control at the source of transmission** in a data center network. This amendment specifies enhancements to the Data Center Bridging Capability (DCBX) protocol to advertise the new capability. This amendment specifies the optional use of existing stream filters to allow **bridges at the edge of the network to intercept and convert signaling messages to existing Priority-based Flow Control (PFC) frames**. This amendment also addresses errors and omissions in the description of existing IEEE Std 802.1Q functionality.






HUAWEI

# Proposal for the text - Overview of components and procedures

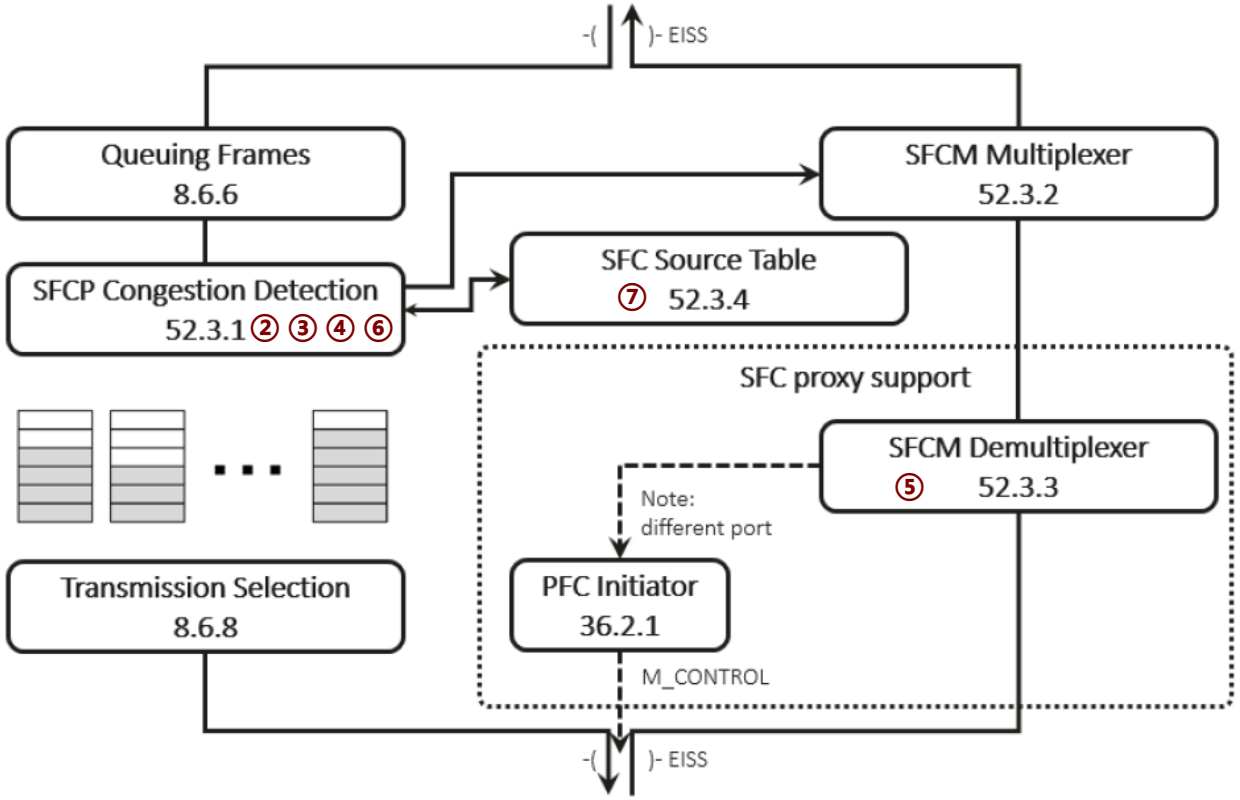



Figure 52-2—Bridge component SFC reference diagram

① sfcInitialize()
② EM_UNITDATA.request()
③ condTransmitSfcmPdu()
④ buildAndSendSfcm()
⑤ processSfcmPdu()
⑥ addSfcSource()
⑦ periodicTableCleanup()

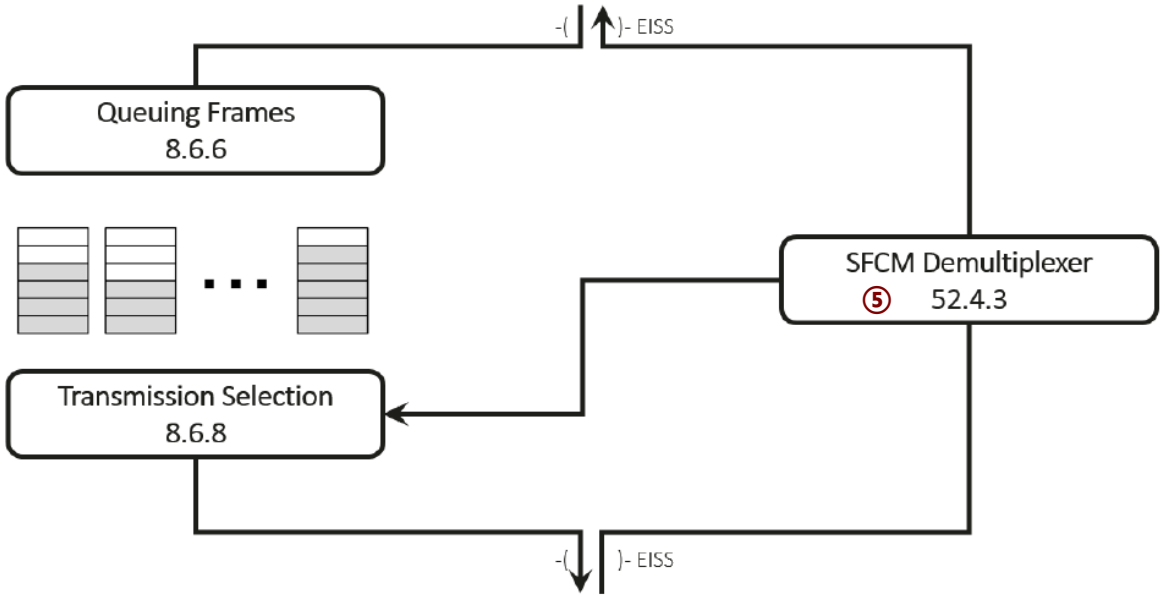



Figure 52-3—End station SFC reference diagram



HUAWEI