TSN Background

Background

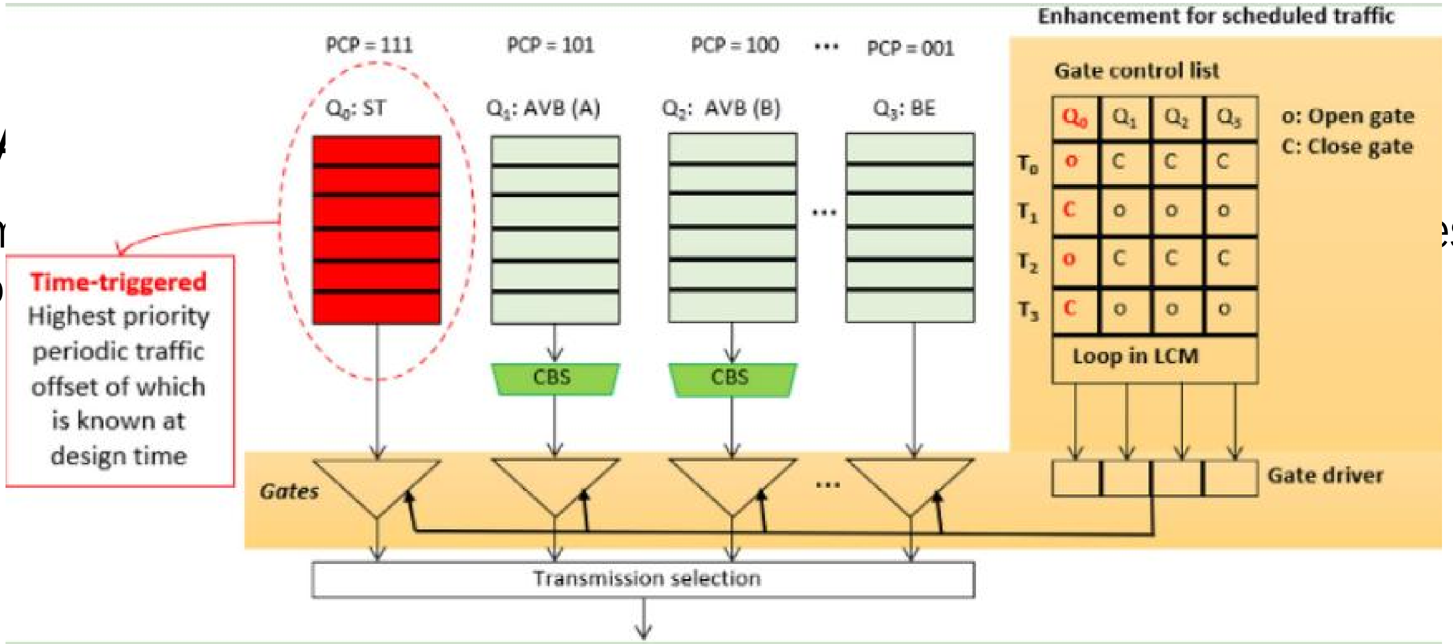
- **Time Sensitive Networking (TSN)** is a set of data link layer protocol specifications developed by the IEEE 802.1 task group to build a more reliable, low-latency, and low-jitter Ethernet.

- TSN's data scheduling is the basis for ensuring time sensitivity. Its core idea is to perform flow control in different application scenarios **based on different shapers**.

- According to the IEEE 802.1Q standard, traffic categories are divided into three categories: **scheduled traffic ST, reserved traffic RT, and best-effort BE traffic**.
- **ST**: industrial automation and control traffic requiring limited latency and zero congestion loss.
- **BE**: consists of generic Ethernet traffic with no specific Quality of Service (QoS) requirements.
- **RT**: frames allocated in different time slots with specified bandwidth reservations. (class A and B)

- # Qav
- **Credit-Based Shaper CBS**
- The goal is to ensure that **the maximum bandwidth required for audio/video transmission** is provided over time series without causing significant disruption to concurrently transmitted best-effort data traffic
- Assign a "credit value" to different queues to schedule data transmission
  - The "credit values" of different transmission queues will change as the data is transmitted.
  - As long as the send credit value is in the positive range (⩾0), the bandwidth-reserved data frames can be transmitted with higher priority (marked blue).
  - Each time a transfer is prioritized, the send credit decreases until it eventually reaches negative territory. When transmit credits are in the negative range, data frames that reserve bandwidth may no longer be transmitted.
  - Therefore, best effort frames in the transmission queue can be processed at this time.
  - If the transmission of a data frame with reserved bandwidth is therefore delayed, the send credit of the corresponding data flow is increased.

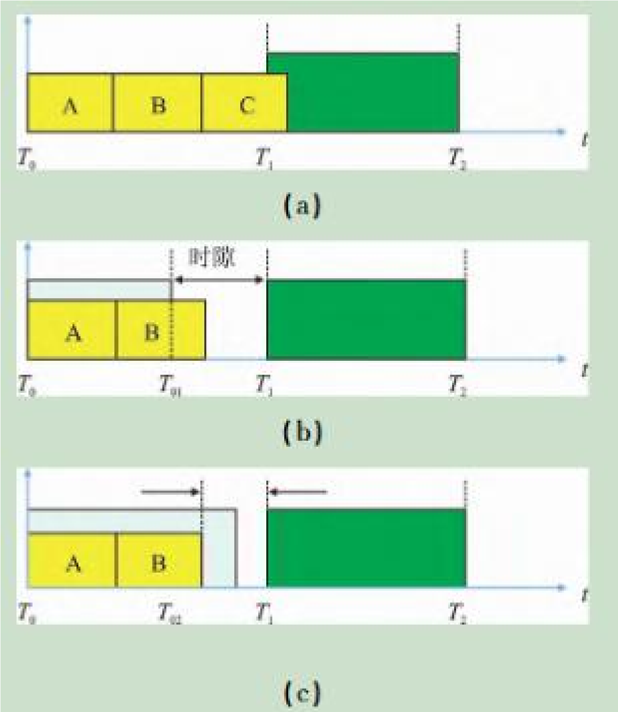
- # Qbv
- **Non-preemptive TA**
- TAS introduces a Tin es the assignment of transmission slots to



- Periodically control the opening/closing of doors through the **Gate Control List (GCL).** Gate control lists determine which traffic queues are allowed to transmit at specific points in the cycle.
- Network packets are queued by their arrival time (that is, FIFO queuing) and transmitted in a non-preemptive manner
- Data streams (**ST**) that require real-time transmission are usually scheduled for transmission first and need to be determined in advance during time scheduling configuration.
- Whether it is periodic data or aperiodic reserved data, reserved channels are required(**AVB/RT**). There are also some other data in the TSN network.
- These data are also aperiodic, but there are not enough reserved channels. Then at this time The data transmission is the "Best-effort" data scheduling(**BE**).

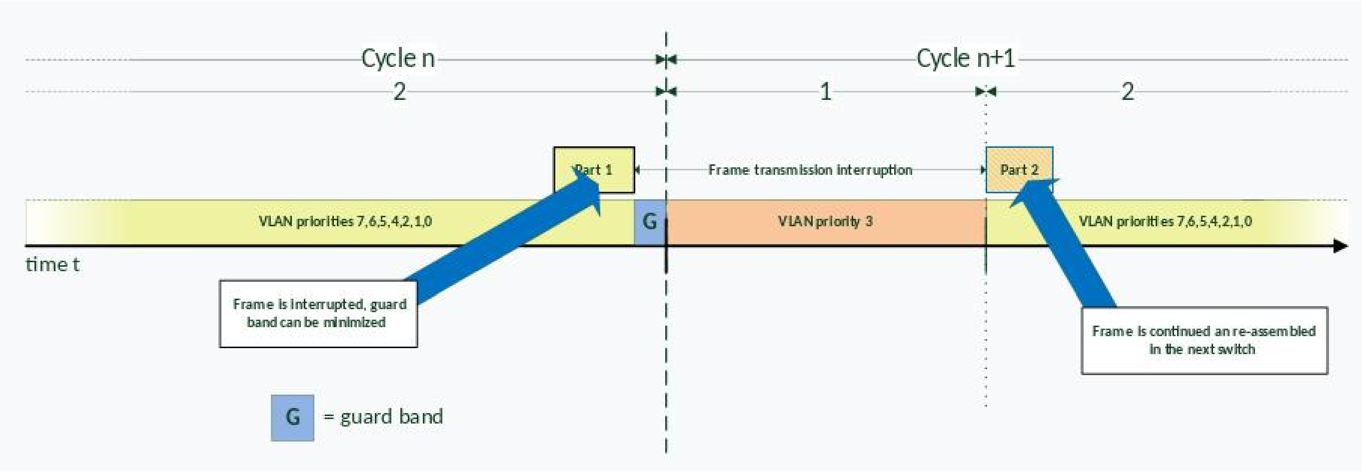
• Qbv

- Set a specified time period before the end of each cycle, called the **guard band**.
- This operation ensures that no frames are being transmitted during cycle switching.
- It is stipulated that within the guard band, no new frames are allowed to start transmission,
- but frames that have already started to be transmitted before the guard band can continue to complete the transmission within the guard band.
- If the port cannot confirm the transmission time of the next frame, the guard band should be long enough to cover the transmission time of the longest frame on the current link.

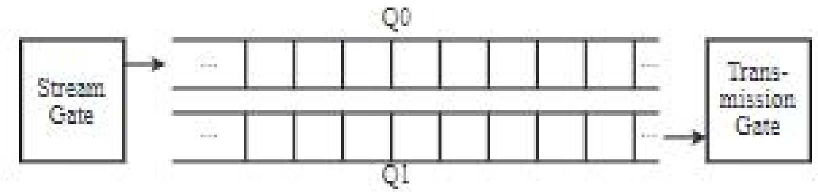


• Qbu

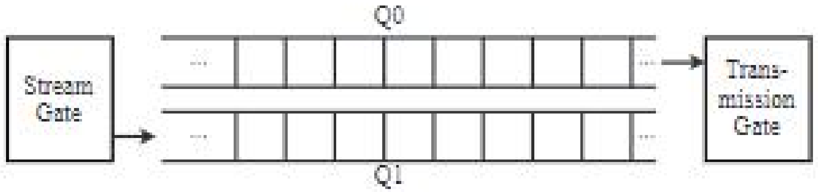
- **Preemptive TAS**
- Frames smaller than 123 bytes cannot be preempted
- In frame preemption, frames that are sensitive to delay requirements are called high-speed frames (**express**),
- the remaining frames are called low-speed frames or preemptable frames (**preemptable**).
- The principle of frame preemption can be compared to the preemptive CPU scheduling in computer operating systems.
- Express traffic can preempt preemptible traffic, but it cannot itself be preempted. Preemption support can be combined with CBS and gate mechanisms.
  - For example, the ST queue can be set to express, while other queues are preemptible.



- **QCH**
- The network is divided into consecutive equal-length time slots, i, i+1..., i+N, and the traffic is controlled by alternating two ping-pong queues.
- In time slot i, queue Q0 can receive tasks but cannot transmit, and Q1 can transmit tasks but tasks cannot be queued. (The queue control at time slot i+1 is opposite to that at time slot i.)
- So the maximum delay is (h+1)l, and the minimum delay is (h-1)l.
  - h is the number of hops passed by the data frame, and l is the time slot length.
- WCRT=(h+1)l, Depends on slot length and hop count



(a) Slot i



(b) Slot i+1

- Qcr
- Data flows are shaped and queues are allocated to data frames based on token bucket.
- Data frames reaching the head of the queue are output together with other non-shaped data flows after evaluating their eligibility time and priority selection.

- For shaping queues, it is necessary to adhere to the rules of queue allocation. Data frames in the following situations should not be assigned to the same shaping queue :
- P1: Frames from different transmitters.
- P2: Frames from the same transmitter but with different priorities.
- P3: Frames from the same transmitter with the same priority, but have different priorities in the same receiver.