

P802.1Qdv contribution on Cycle ID and PAR Modification Discussion

dv-yizhou-cycle-id-and-PAR-modification-discussion-0123-v01

Yizhou Li (Huawei)

Guanhua Zhuang (Huawei)

Shoushou Ren (Huawei)

Jiang Li (Huawei)

Jeong-dong Ryoo (ETRI)

Peng Liu (CMCC)

Dong Li (Shenyang Institute of Automation)

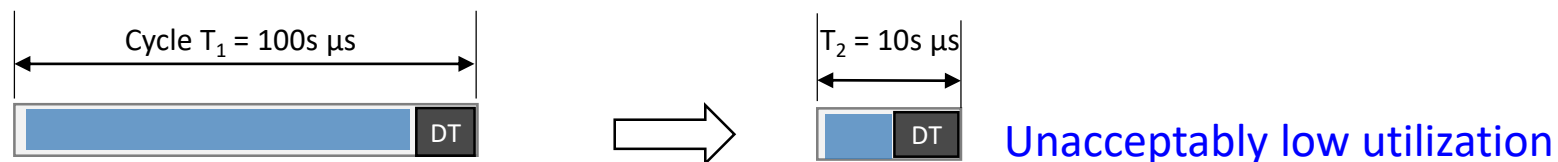
Wenbin Dai (Shanghai Jiao Tong University)

Agenda

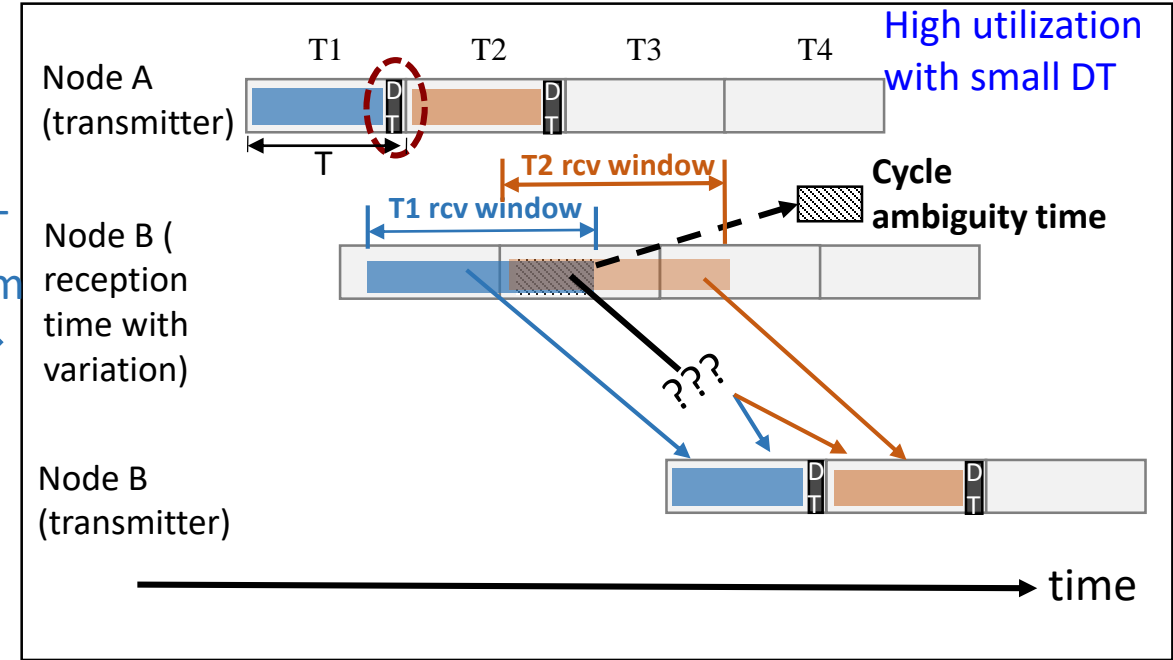
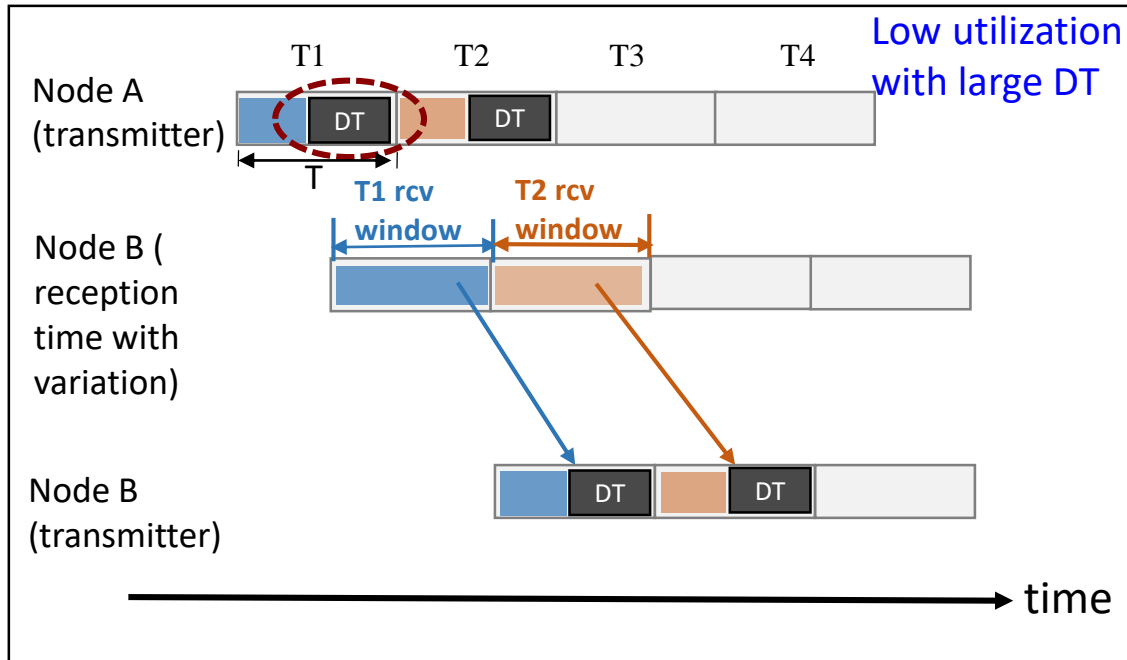
- A summary of Cycle ID proposal to P802.1Qdv
 - Goal and problems to be solved
 - Proposal to add a data plane cycle id tag
 - History of contributions
- Discussion on the potential P802.1Qdv PAR modification required

Problem to be solved – Low bandwidth utilization when Dead Time (DT) eats Cycle Time (T) in cyclic forwarding

- E2e latency and jitter are proportional to cycle time T, and # of hops, h
 - E2e delay $\approx h * T$; jitter = 2T
- Desire to use smaller T for better latency bound and jitter
 - To achieve e2e latency (e.g. < 1ms) or large # of hops (e.g. 10+)
 - A known app cycle 31.25 μ s; cycle time T < 31.25 would be desired
 - Should support T in the order of ~ 10 s μ s
- DT is imposed in each cycle to absorb the time variation between two neighbor nodes
 - DT is in the order of ~ 10 s μ s as well
- DT and T are in the same order of magnitude, causing low bandwidth utilization
 - DT eats T, e.g. $\approx 50\%$ when T=40 μ s & DT=20 μ s

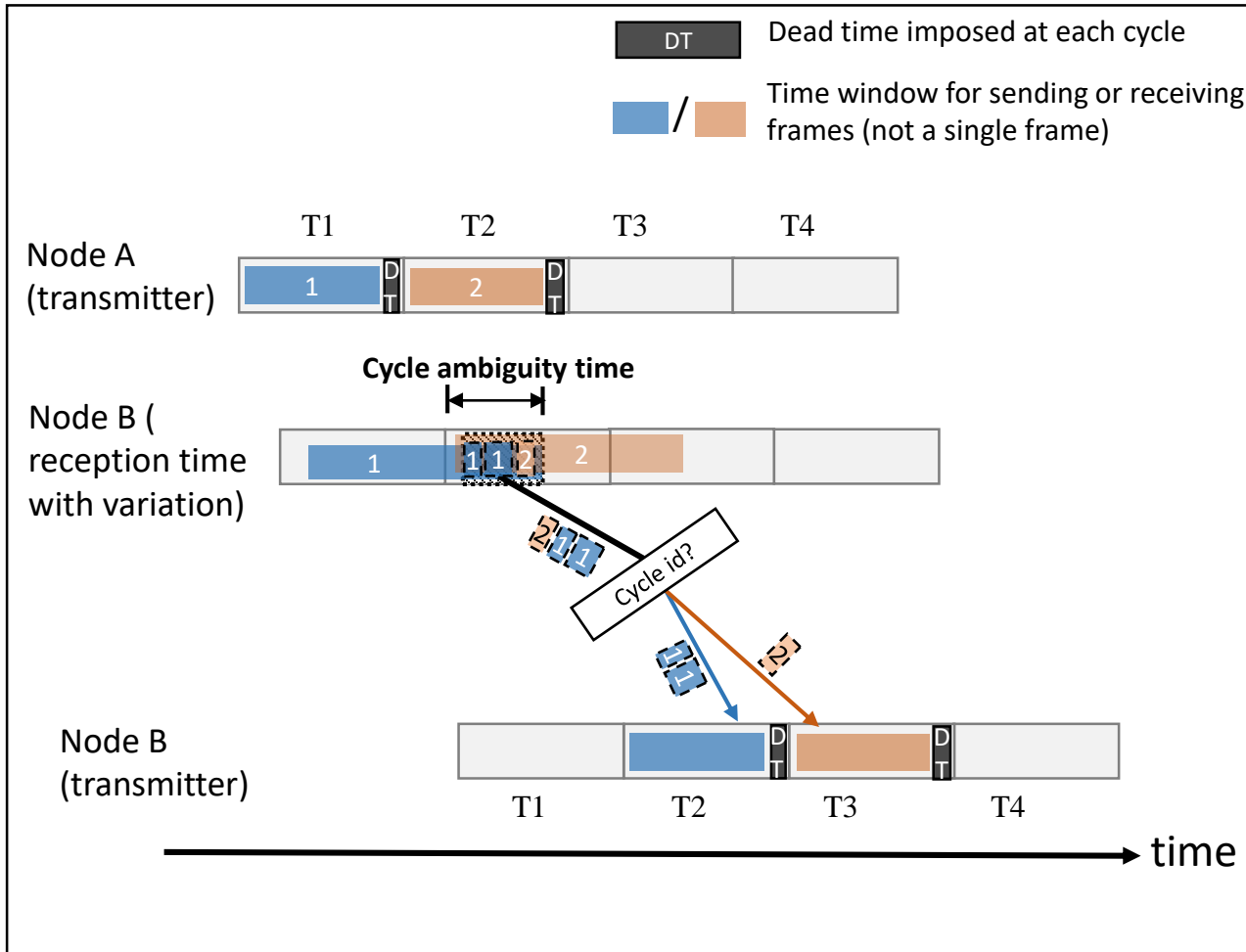


Goal – Make DT minimum to improve the b/w utilization in small T



- **Why low utilization?** DT takes a significant part of T.
- **A straightforward way to improve utilization:** make DT minimum
- **Output buffer (cycle) determination was made based on the reception time of the packet, i.e. the assumption was there was a clear time demarcation for packets from the consecutive cycles**
- **Such an assumption may not hold**
 - DT is minimum so that it does not cover the full time variation between adjacent nodes
 - The reception time of the packets experience variations themselves, e.g. due to the variable residence time before frame started being processed by PBA (packet bus arbiter), or the storing of the variable frame size
- **Cycle ID tries to solve the cycle ambiguity when the packet reception time for output cycle determination includes the variation**

Propose to use the explicit cycle id at data plane



- Carry cycle id and change per hop
- Cycle id is an implicit indicator of pivot reception time (i.e. without variation) for output buffer/cycle determination instead of the packet reception time with variation
- Remove the output cycle determination ambiguity
- Achieve the good utilization in small cycles by making DT minimum

How to carry a cycle id

1. R-tag (defined in 802.1CB)

Ethertype (F1-C1)	Reserved (16-bit)	Sequence number (16-bit)
-------------------	-------------------	--------------------------

To define a subtype flag and use the last 5-bit in Reserved field for cycle id.

Ethertype (F1-C1)	Reserved (16-bit)			Sequence number (16-bit)
	flag(1)	Rsvd (10)	Cycle id (5)	

2. Define a new cycle-tag

Ethertype (cycle-tag)	Subtype (4-bit)	Reserved (4-bit)	Cycle ID (8-bit)
--------------------------	--------------------	---------------------	---------------------

How to carry a cycle id (cont'd) – Not a good choice based on previous discussion

~~3. Use vlan stacking + vlan mapping function~~

- Inner vlan is used for cycle id, use ACL to map from ingress cycle id to egress cycle id
- Outer vlan is used as normal vlan based mac learning and forwarding
- Used in a controlled domain, may not be compatible with some existing s-vlan + c-vlan usage

Ethertype (s-vlan)	vlan-tag (16-bit)	Ethertype (c-vlan)	Cycle id (16-bit)
--------------------	-------------------	--------------------	-------------------

ACL example: `if-match cvlan-id cycle-id-in`

`remark cvlan-id cycle-id-out`

History

- Further details on Cycle ID proposal to answer the questions on
 - How the initialization and cycle ID mapping relation establishment are performed
 - How to compute cycle and buffer IDs as they have different ID spaces?
 - How the cycle ID mapping works when the #of bins(buffers) are different on two neighbor nodes?
 - How large should the cycle ID space be? What are the considerations when choosing the field length of it?
- **dv-yizhou-cycle-identification-details-1122-v02**
 - Added clarifications for checksum re-compute, multiple cycle time and computation delay.
 - Added explanation on cycle id = implicit pivot reception time
 - Editorial changes from -v01
- **dv-yizhou-cycle-identification-details-1022-v01**
 - Provided further details on the cycle ID based implicit time determination approach
 - Focused on the initialization, especially how to establish the cycle ID mapping relations for the port pairs.
 - Suggested to use 5 bit for cycle ID
- **Cycle ID proposal (dv-yizhou-cycle-identification-0922-v02)**
 - Showed the goal to improve the bandwidth utilization in small cycle
 - Discussed the output cycle ambiguity problem when making dead time (DT) minimum to improve the utilization
 - Proposed to use implicit pivot reception time indicated by cycle id to determine to which bin a frame goes
 - Provided three options to carry cycle id

Takeaways

- Concept:
 - Cycle id conceptually is an implicit pivot reception time of the frame. Not impacted by the time variation experienced by the data frame. Reliable for output cycle determination.
- Goal:
 - It helps to remove the ambiguity of time based output cycle determination and improve the bandwidth utilizations when packet reception time experiences time variation.
- Details have been provided to show the feasibility and to answer the technical questions received
 - Most of the work is done at the provisioning stage. The calculations for initialization and mapping relation establishment are feasible.
 - Per-frame data plane cycle ID mapping is straightforward.

Discussions regarding the current PAR

- Do we need to modify the current PAR of IEEE 802.1Qdv if the cycle ID based output bin determination is included?
 - Seems so from last meeting as a data plane new tag or modification to the current tagged field is required to be defined.
- Shall we start the discussion to make PAR and CSD modifications of the P802.1Qdv?

Current PAR scope:

- 5.2.b Scope of the project: This amendment specifies procedures, protocols and managed objects to enhance Cyclic Queuing and Forwarding, comprising: a transmission selection procedure that organizes frames in a traffic class output queue into logical bins that are output in strict rotation at a constant frequency; a procedure for storing received frames into bins based on the time of reception **or the cycle ID** of the frame; **data field definition to carry cycle ID information**; a procedure for storing received frames into bins based on per-flow octet counters; a protocol for determining the phase relationship between a transmitter's and a receiver's bin boundaries in time; managed objects, Management Information Base (MIB), and YANG modules for controlling these procedures; and an informative annex to provide guidance for applying these procedures. This amendment also addresses errors and omissions in the description of existing IEEE Std 802.1Q functionality.