IPv6 Options for Cyclic Queuing and Forwarding Variants

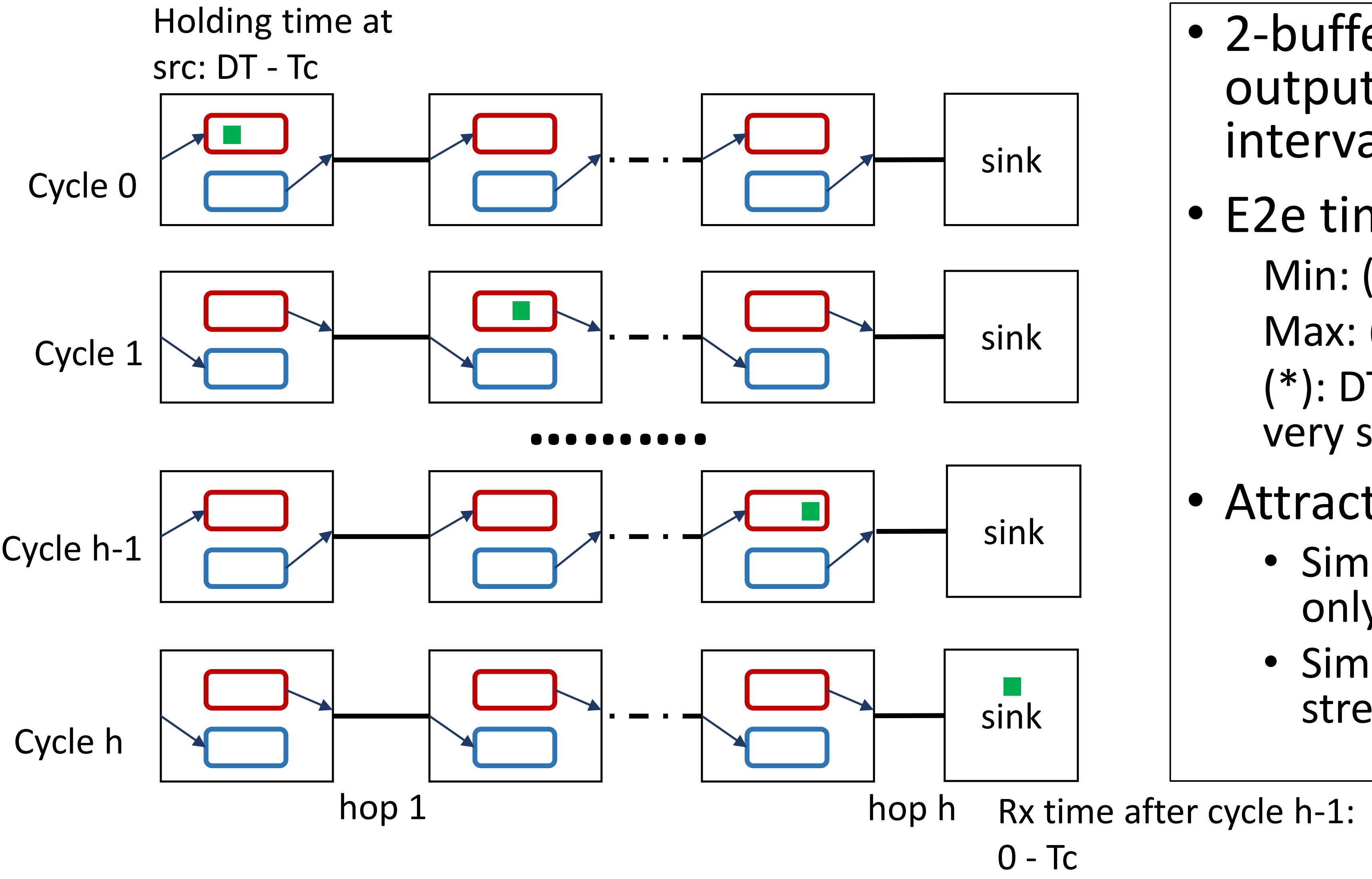
draft-yizhou-detnet-ipv6-options-for-cqf-variant-00

Yizhou Li (Presenter)

Shoushou Ren
Guangpeng Li
Fan Yang
Jeong-dong Ryoo
Peng Liu

July, 2022

Fundamental CQF has attractive "simplicity" features for wider deployments



- 2-buffer per port. Input and output swap once every cycle interval Tc.
- E2e time taken:

  Min: (h-1) Tc +DT*

  Max: (h+1) Tc

  (*): DT = dead time (revisit later). very small in fundamental CQF
- Attractive "simplicity" features:
  - Simple calculable latency bound: only relevant to Tc and h, ≈ h*Tc
  - Simple maintenance: no per-stream per-hop state maintenance

CQF has potentials for wider deployments - 1

- Wider deployment requires supporting one or combination of the followings:
  - Smaller e2e latency bound (1)
  - Larger number of hops (2)
  - Longer links (3)
  - Larger processing time variance as node type diversity increases (4)
- Recall that CQF latency bound ≈ h*Tc
- Higher speed link provides the potential to reduce Tc, even with greater value of h
  - allow at least one 1500B/max size packet to be sent within Tc
  - With increasing of link speed, the same amount of data can be transmitted within a smaller cycle time
  - Counteract larger h

Cycle Time (µs)	**Buffer Size per Cycle (Byte)**		
	Link bandwidth		
	100Mbps	1Gbps	10Gbps
1	12.5	125	1250
1.2	15	150	1500
2	25	250	2500
4	50	500	5000
10	125	1250	12500
12	150	1500	15000
120	1500	15000	150000

Cycle time decreasing:
100x µs -> 10x µs -> few µs

- Potentials for item (1) and (2), next page for item (3) & (4)

Fundamental CQF support req (3) &(4) but with low utilization

```
  ------------------------------------------------------> Time

          |             |             |             |
Node A    |  cycle i-1  |   cycle i   |  cycle i+1  |
          |             |             |             |
Sending  ---------------+---------------------------------------
          |+------+     |+------+     |+------+     |
          ||//////|     ||//////|     ||//////|     |
          |+------+     |+------+     |+------+     |
          |  buf_1      |  buf_2      |  buf_1      |
          |       |     |       |     |       |     |
          |       | DT  |       | DT  |       | DT  |
Node B    |       |<--->|       |<--->|       |<--->|
          |             |             |             |
Receiving-------------------------------------------------------
          |     +------+|     +------+|     +------+|
          |     |//////||     |//////||     |//////||
          |     +------+|     +------+|     +------+|
          |       buf_1 |       buf_2 |       buf_1 |
          |             |             |             |
          |             |             |             |
Node B    |             |             |             |
          |             |             |             |
Sending  -------------------------------------------------------
          |             |+------+     |+------+     |
          |             ||//////|     ||//////|     |
          |             |+------+     |+------+     |
          |             |  buf_1      |  buf_2

                                                  DT=Dead Time

                Figure 2: Fundamental Two Buffer CQF
```

- Revisit DT (dead time): the last byte sent by node A in cycle (i-1) has to be ready for sending at node B before the start of cycle i.
- DT is at least: max propagation delay + max processing delay at the next node + max other time variations.
- The longer the propagation or processing delay, the larger the DT.
- DT eats up cycle interval Tc when Tc is small (both values < 1ms): result in low utilization or impractical in extreme case (consider prop delay > Tc)
- Hard for fundamental CQF:
  - Shorter Tc for lower e2e latency bound
  - Larger DT for longer link and/or processing time
  - Smaller ratio of DT/Tc for better utilization

CQF Variant (>2 buffer) has the potential to support (3) & (4)

```
  --------------------------------------------------------> Time

          |             |             |             |
Node A    |  cycle i-1  |   cycle i   |  cycle i+1  |
          |             |             |             |
Sending  ---------------+-----------------------------------
          |+----------+ |+----------+ |+----------+ |
          ||//////////| ||//////////| ||//////////| |
          |+----------+ |+----------+ |+----------+ |
          |  buf_1      |  buf_2         buf_3      |
          |           | |           | |           | |
          |         ->| |<-       ->| |<-       ->| |<-
          |            DT            DT            DT
          |
          ------------------------------------------------
Node B    |     +-----------+ +-----------+ +-----------+
          |     |///////////| |///////////| |///////////|
Receiving |     +-----------+ +-----------+ +-----------+
          |       buf_1 |       buf_2 |       buf_3 |
          |             |             |             |
          |             |             |             |
          |             |             |             |
          |             |             |             |
         ---------------|-----------------------------------
Node B    |             |             |+----------+ |+----------+
          |             |             ||//////////| ||//////////|
Sending   |             |             |+----------+ |+----------+
          |             |               buf_1         buf_2

                                                  DT=Dead Time
```

Figure 3: Three Buffer CQF

- 3 buffer works in rotation manner
- A straightforward variant to fundamental 2-buffer CQF:
    - Configuration is similar
    - Can easily deduce from fundamental CQF without the rigid requirement to produce new standard
- More than 3 buffer is required when the receiving time spans over two cycle interval boundaries.
- In general, it is feasible.

A closer look at the CQF variant: a time ambiguity window exists

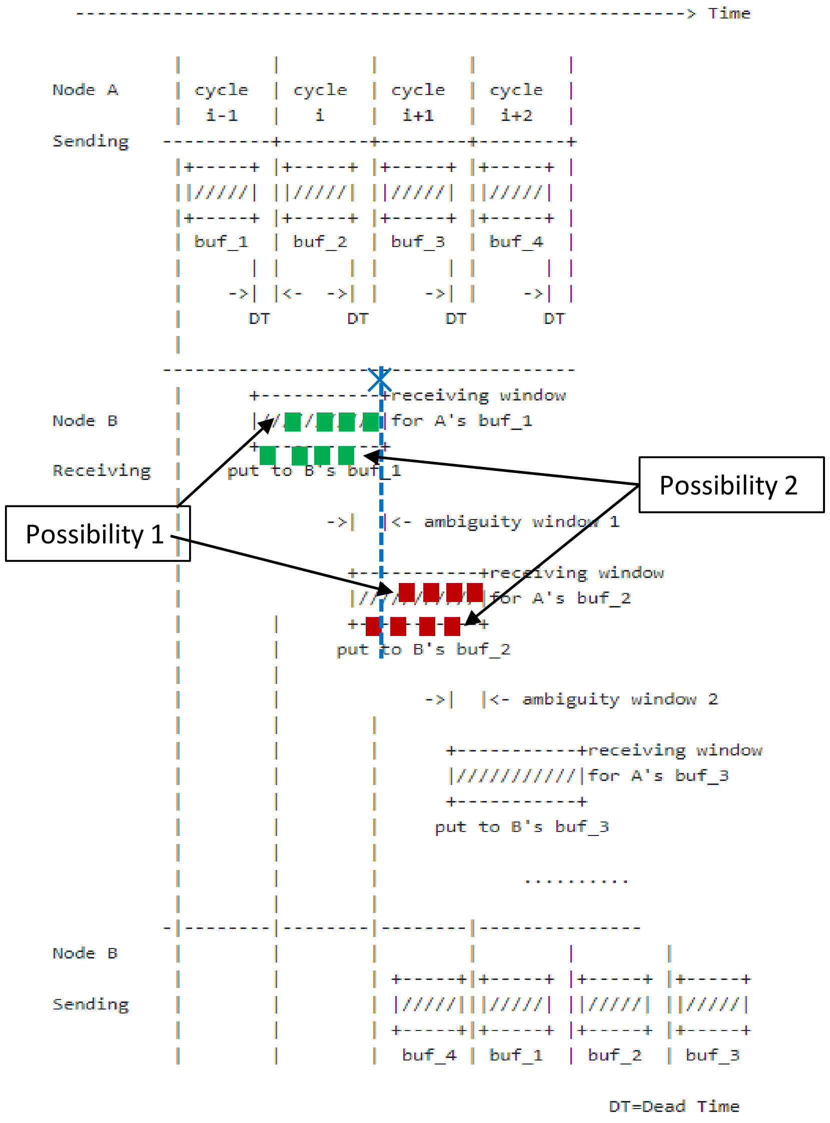


Figure 4: Ambiguity of time based cycle demarcation in CQF

- Receiving time window swells when the (processing) time variance increases
- Keep DT small
- Time ambiguity window exists for two consecutive cycles
- The larger the time variance and/or the smaller the DT, the larger the ambiguity window
- So setting the time demarcation to differentiate pkts from two consecutive cycles is impractical (see left)
- Way out: pkt carry cycle id metadata at output to help the downstream node determine the correct buffer to put it in

Summary

- CQF has attractive features and potentials for wider deployments
- CQF variant is a straightforward extension from fundamental CQF:
  - use more than two buffers
  - some extra configurations would be required
  - Other variants may exist
- A missing part in current CQF variant: remove the ambiguity when identifying the packets from the upstream's two consecutive cycles
- IPv6 options to carry cycle id metadata is proposed.

```
+---------------------------------+
|          DetNet IP Packet       |
+---------------------------------+
|            other EHs            |
+---------------------------------+
|       IPv6 Hop-by-Hop Ex Hdr    |
|         (CQF-Variant Option)    |
+---------------------------------+
|            IPv6 Header          |
+---------------------------------+
|             Data-Link           |
+---------------------------------+
|             Physical            |
+---------------------------------+
```

Figure 6: Example 1: CQF-Variant Option Encapsulated in HbH

```
+---------------------------------+
|          DetNet IP Packet       |
+---------------------------------+
|      other EHs including RH     |
+---------------------------------+
|      IPv6 Destination Ex Hdr    |
|        (CQF-Variant Option)     |
+---------------------------------+
|            IPv6 Header          |
+---------------------------------+
|             Data-Link           |
+---------------------------------+
|             Physical            |
+---------------------------------+
```

Figure 7: Example 2: CQF-Variant Option Encapsulated in DOH

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Option Type  | Opt Data Len |E|     Flags     |    Cycle Id   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
~         (64-bit extension if flag E-bit is 1)                 ~
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: CQF-Variant Option Format Example

Solicit feedback

- Is it a good way to address the "ambiguity" issue in order to facilitate the increasing demand to use CQF and its variants in the wider scenarios?
- IPv6 options, whether and/or how to collaborate with other WG (6man)