Towards Large-Scale Deterministic IP Networks: variant of CQF

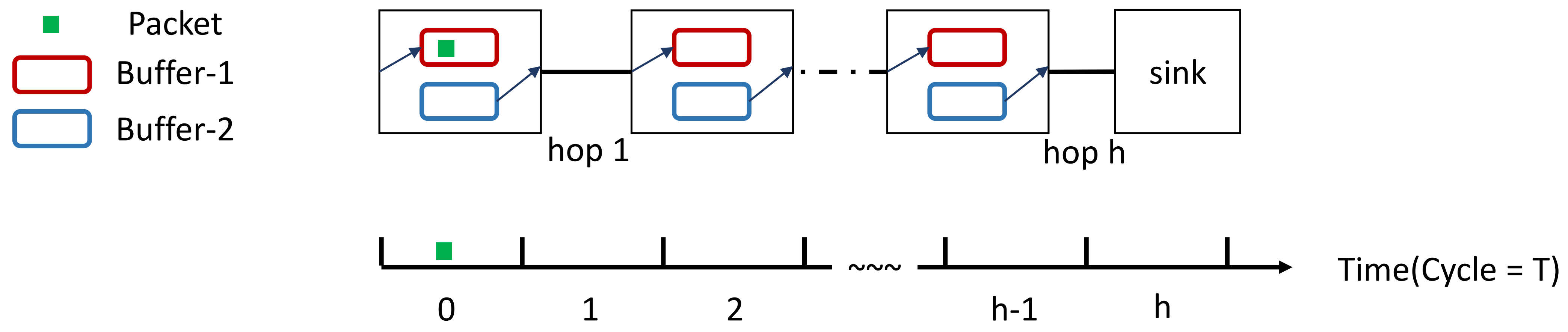
Guangpeng Li (Presenter)

https://datatracker.ietf.org/doc/html/draft-yizhou-detnet-ipv6-options-for-cqf-variant-01
https://datatracker.ietf.org/doc/html/draft-eckert-detnet-tcqf-01

July, 2022

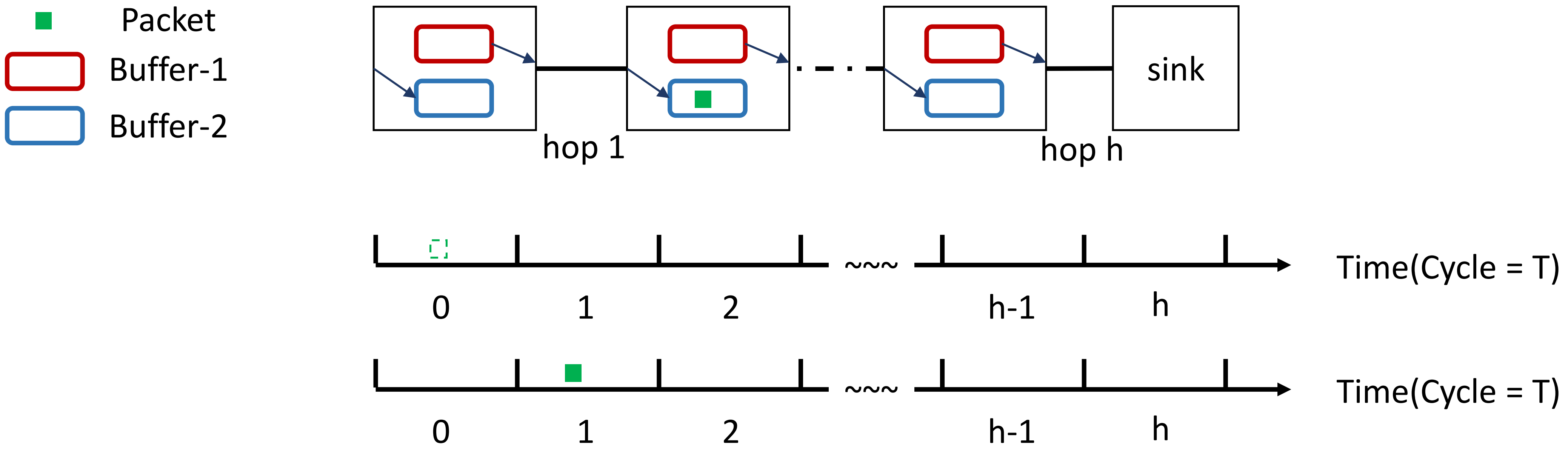
Basic CQF has attractive "simplicity" features

- 2 buffers per port. Input and output swap once every cycle interval T



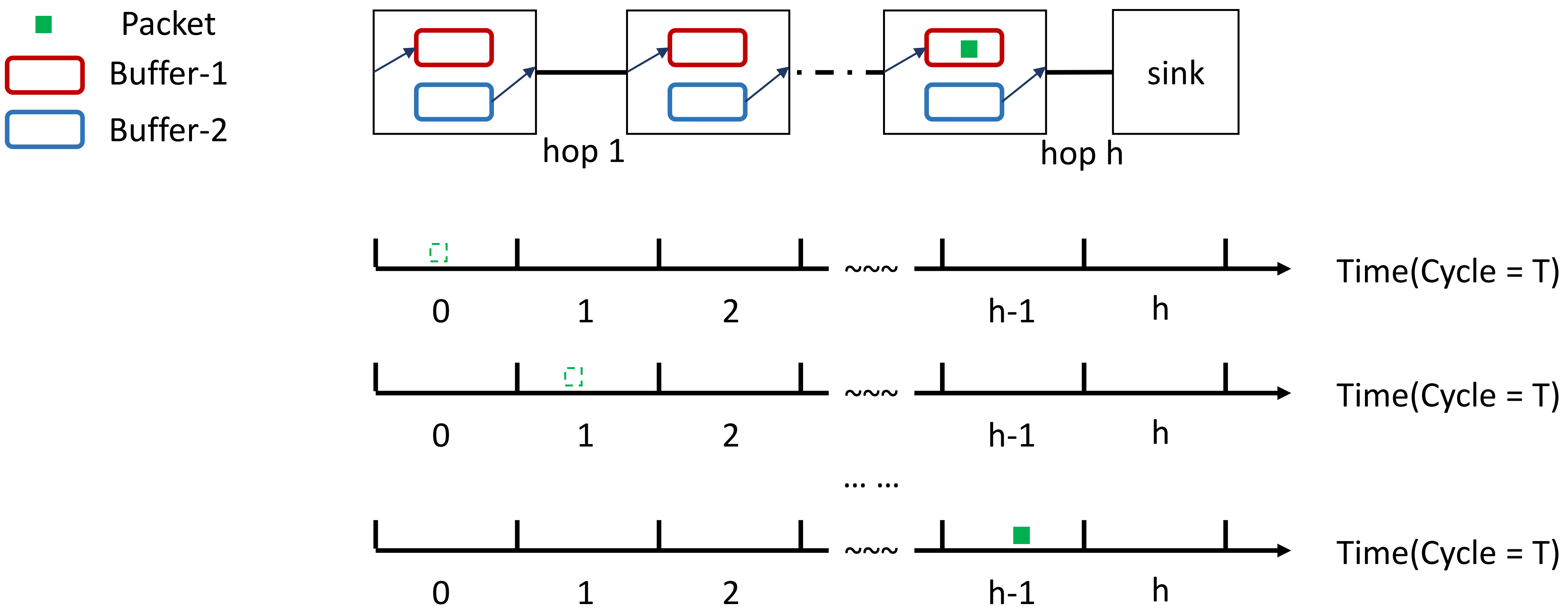
Basic CQF has attractive "simplicity" features

- 2 buffers per port. Input and output swap once every cycle interval T



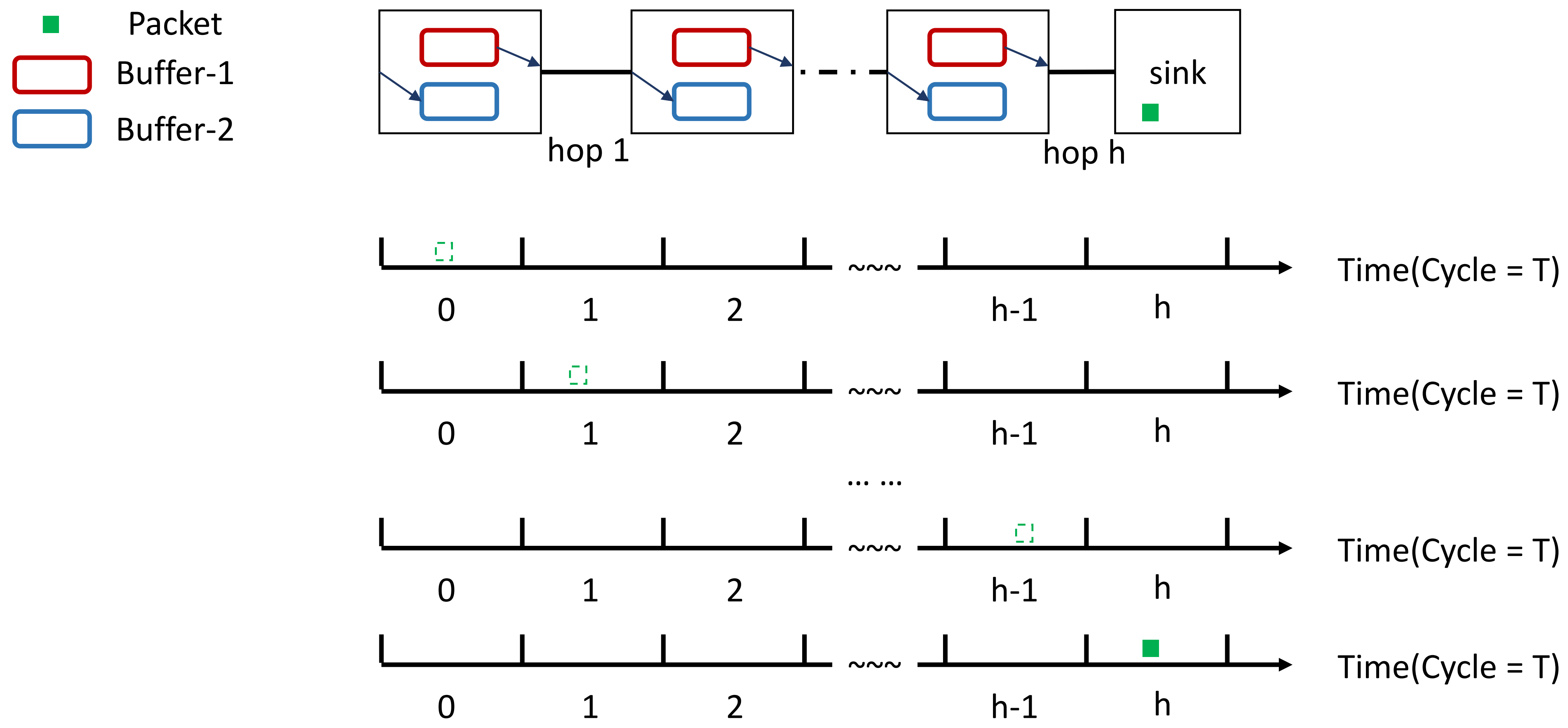
Basic CQF has attractive “simplicity” features

- 2 buffers per port. Input and output swap once every cycle interval T



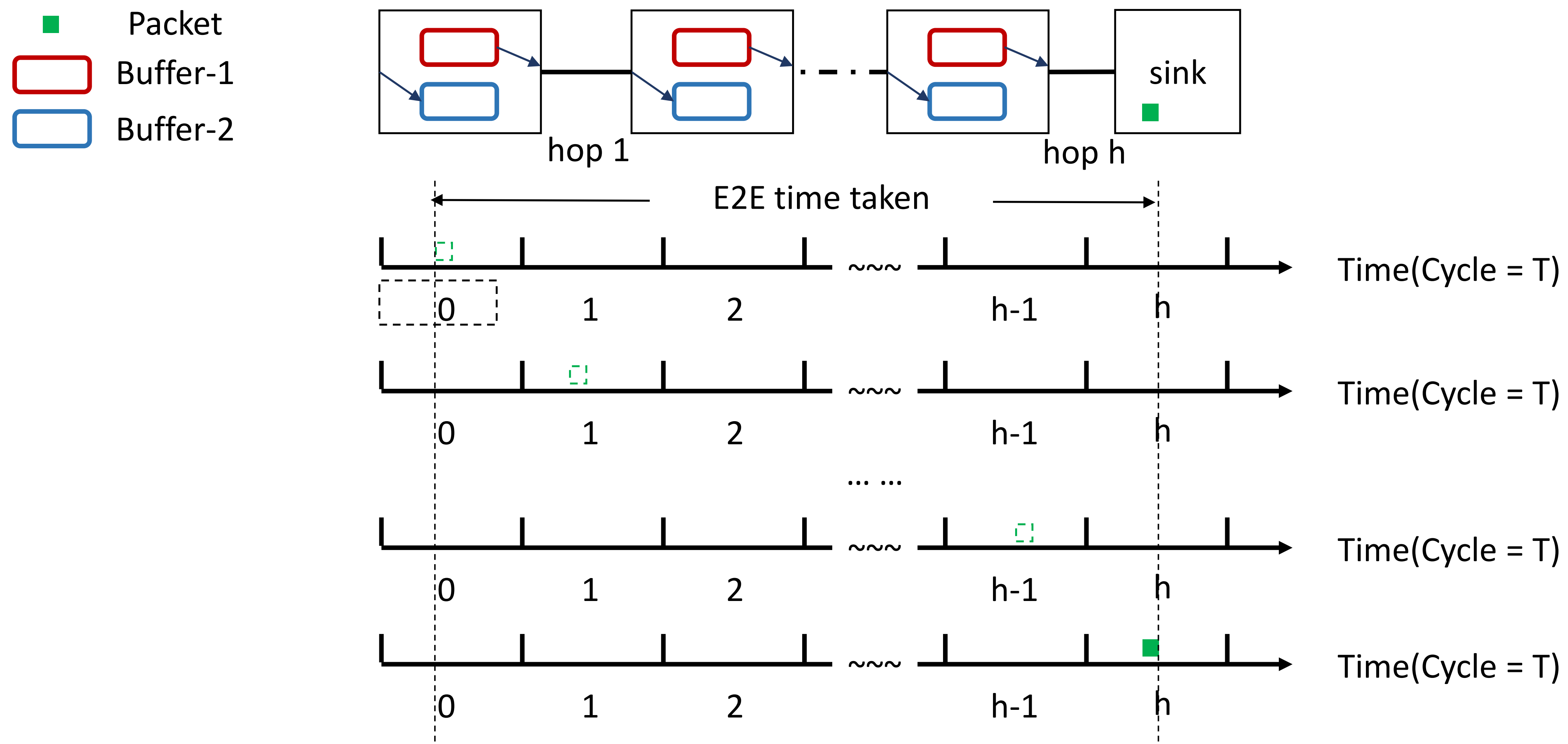
Basic CQF has attractive "simplicity" features

- 2 buffers per port. Input and output swap once every cycle interval T

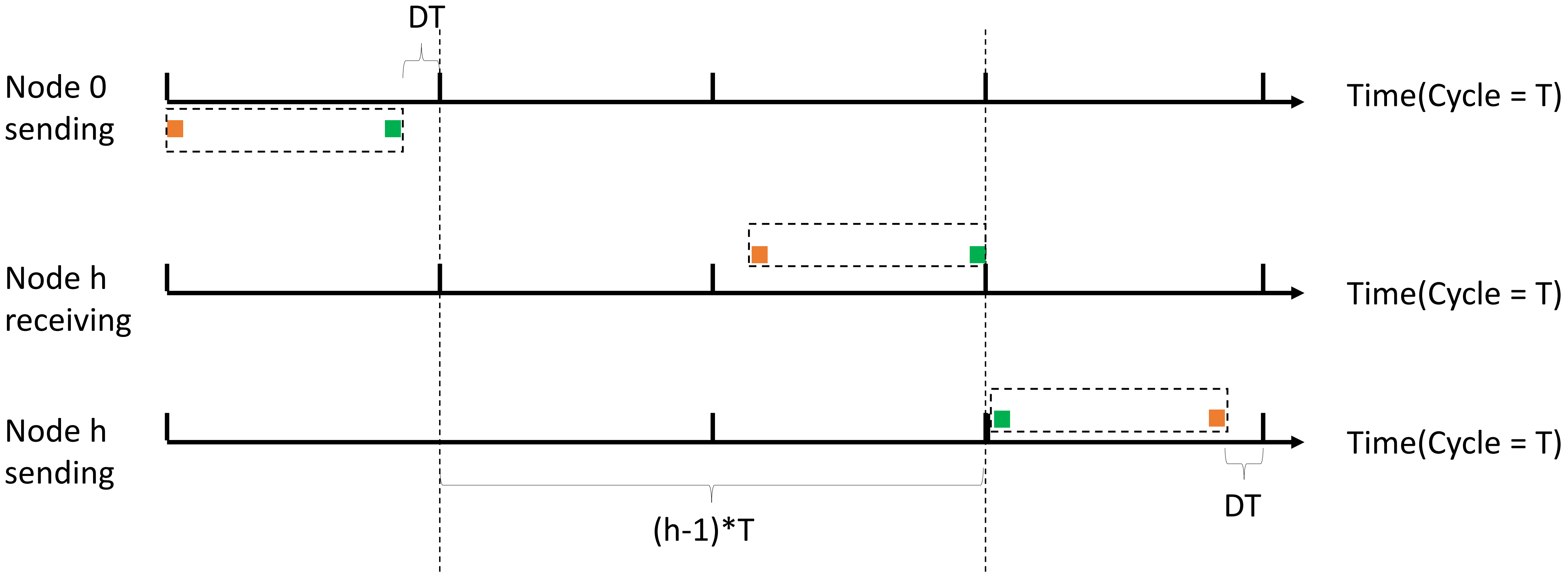


Basic CQF has attractive "simplicity" features

- 2 buffers per port. Input and output swap once every cycle interval T



Basic CQF's bounded latency



Total queueing time:

- Min: (h-1) Tc +DT*
- Max: (h+1) Tc -DT
- (*): DT = Dead Time

DT is at least: ↑
max propagation delay ↑
+ max processing delay at the next node
+ max other time variations.

Lower utilization or impractical with larger DT/T

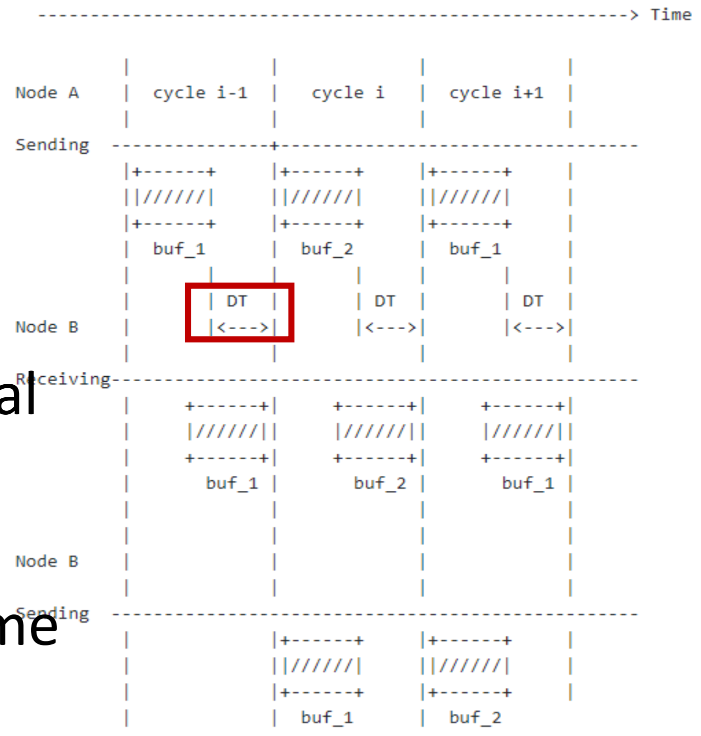
Cycle time decreasing:
100x µs -> 10x µs -> few µs

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

- allow at least one 1500B/max size packet to be sent within T
- With increasing of link speed, the same amount of data can be transmitted within a smaller cycle time
- Counteract larger h

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.

- The longer the propagation or processing delay, the larger the DT.
- DT eats up cycle interval T when T is small
- result in low utilization or impractical in extreme case (consider DT > T)



DT=Dead Time

Figure 2: Fundamental Two Buffer CQF

Requirements of deploying in LDN

- Wider deployment requires supporting one or combination of the followings:
  - Smaller e2e latency bound, relating to h and T (1)
  - Larger number of hops (2)
  - Longer links introducing longer propagation delay (3)
  - Larger processing time variance as node type diversity increases (4)

- Hard for fundamental CQF:
  - Shorter T 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 buffers) 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

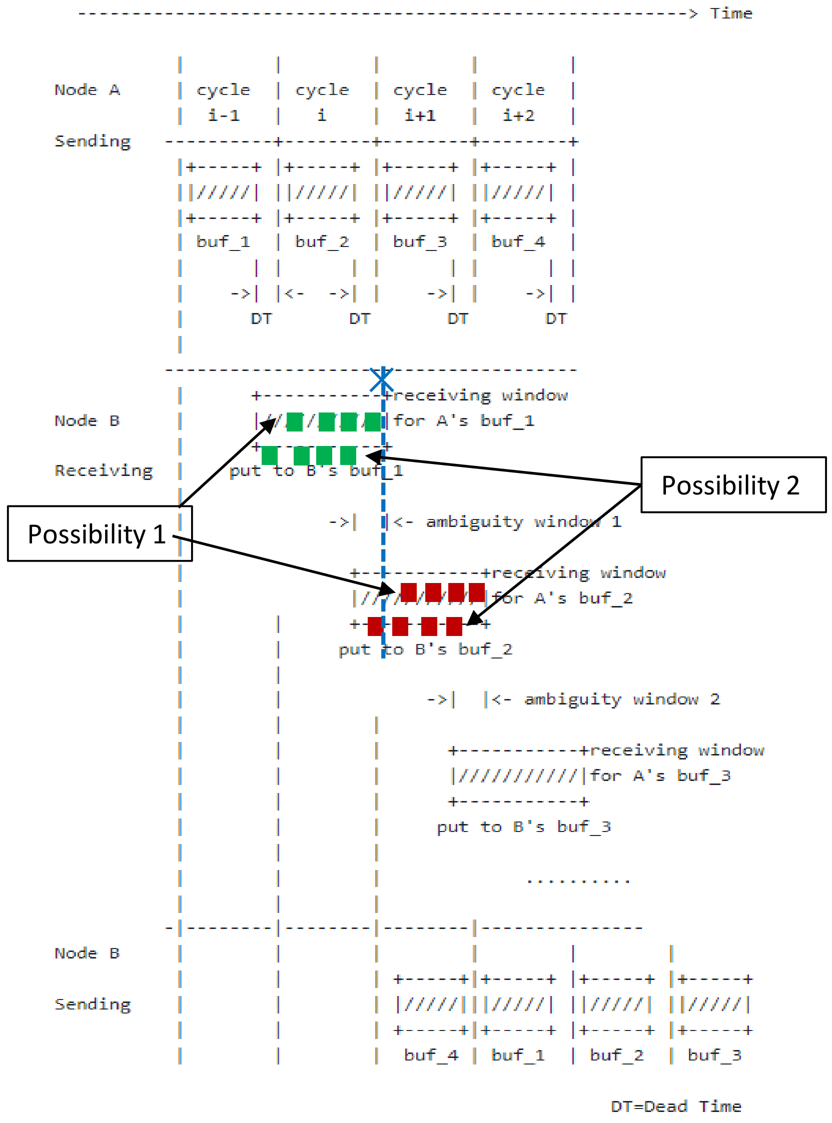


Figure 4: Ambiguity of time based cycle demarcation in CQF

- 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

Simulation Results in LDN

TABLE I
TRAFFIC PATTERNS IN THE PoC TEST.

Parameter	HP flows	BE flows
Rate (Mbps)	600	600
Burst size (KB)	1.5	192

T = 10μs

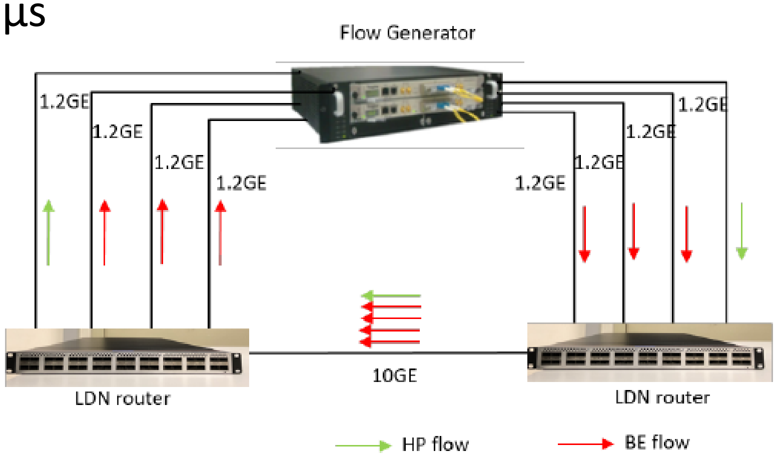


Fig. 5. Test topology for the PoC implementation.

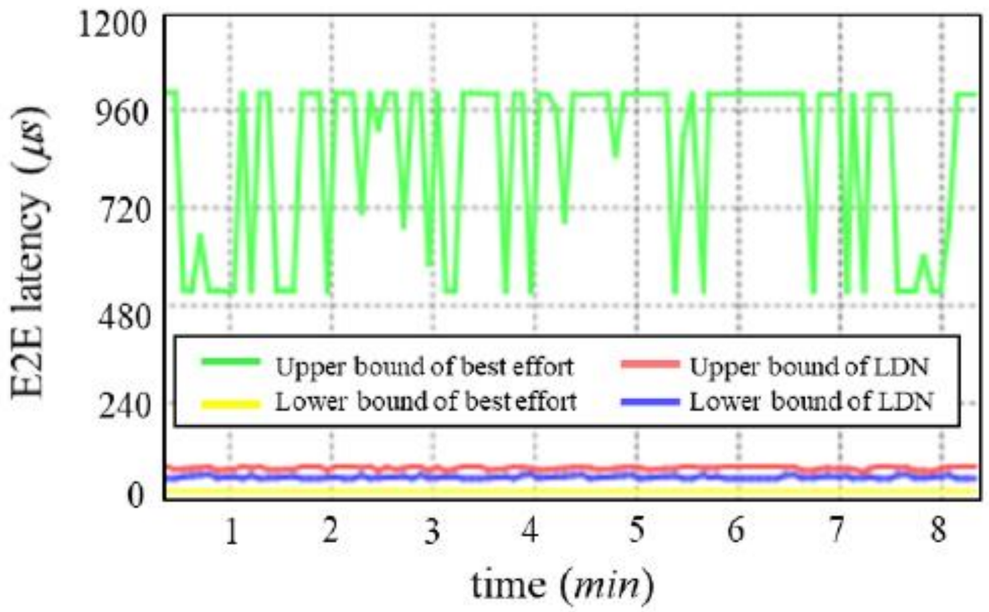


Fig. 6. Test results over the PoC implementation.

The largest value of the worst case E2E queuing delay is 67:046 μs, while the smallest value of the best case delay is 49:886 μs.
By consequence, the largest jitter experienced by the flows is 67:046 μs-49:886 μs = 17:16 μs, smaller than 2T=20 μs [1]

[1] Liu B, Ren S, Wang C, et al. Towards large-scale deterministic IP networks[C]//2021 IFIP Networking Conference (IFIP Networking). IEEE, 2021: 1-9.

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

Ask for comments and collaboration

- Remove the ambiguity when identifying the packets from the upstream’s two consecutive cycles
- Ways to carry cycle id metadata is needed to be discussed.