Multiple Cyclic Queuing and Forwarding
(slides to accompany df-finn-multiple-CQF-0919-v01)

Norman Finn
Huawei Technologies Co. Ltd
nfinn@nfinnconsulting.com
df-finn-multiple-CQF-slides-0919-v01

Why this paper?

Cyclic Queuing and Forwarding is perceived as having major flaws, which prevent its adoption:

- It is difficult to pick a cycle time:
  - A long cycle time produces long end-to-end latency.
  - A short cycle time severely restricts the number of flows handled.
  - No one cycle time serves more than a narrow range of requirements.
- Having 2-buffers per hop precludes long links.

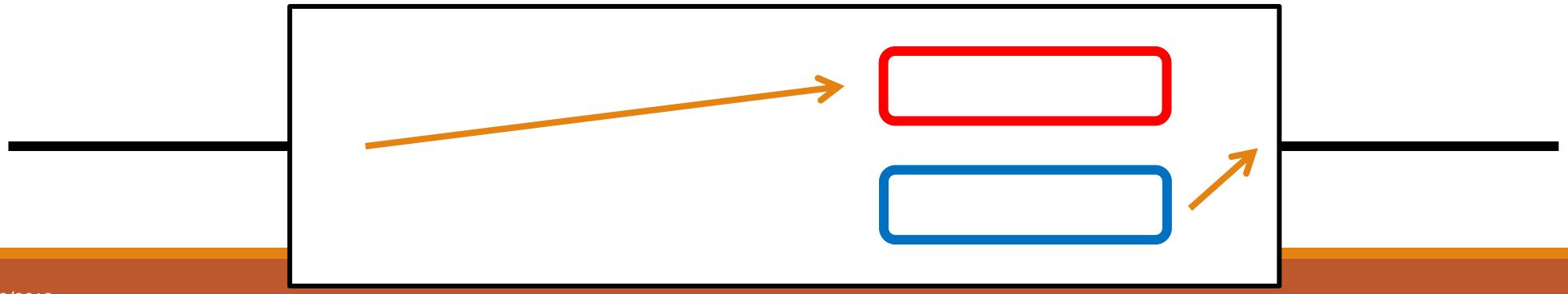
With a little imagination, these issues can be solved without adding anything to IEEE 802.1Q, so the solutions are suitable for P802.1DF.

Step 1: You can run with 3 buffers, not 2

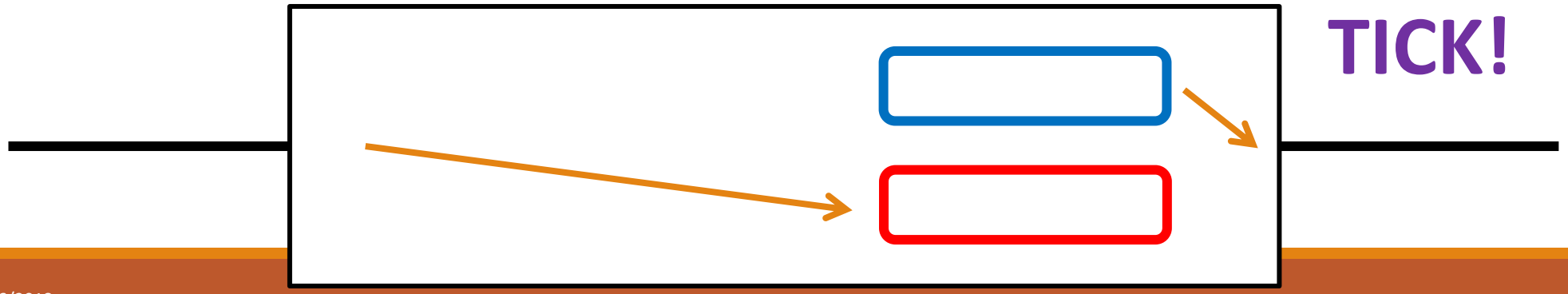
- With 2 buffers, you change input buffers at the same moment that you change output buffers.
  - Link delay and forwarding delay have to be subtracted from the useful bandwidth, to prevent filling and emptying a buffer at the same time.
- If you have 3 buffers, then you can change to the next input buffer at a different phase of the cycle from when you change to the next output buffer.
  - Each buffer spends part of a cycle idle; one cycle per hop is added to the latency.
  - But, the link delay and forwarding delay are no subtracted from the bandwidth..

Two buffers

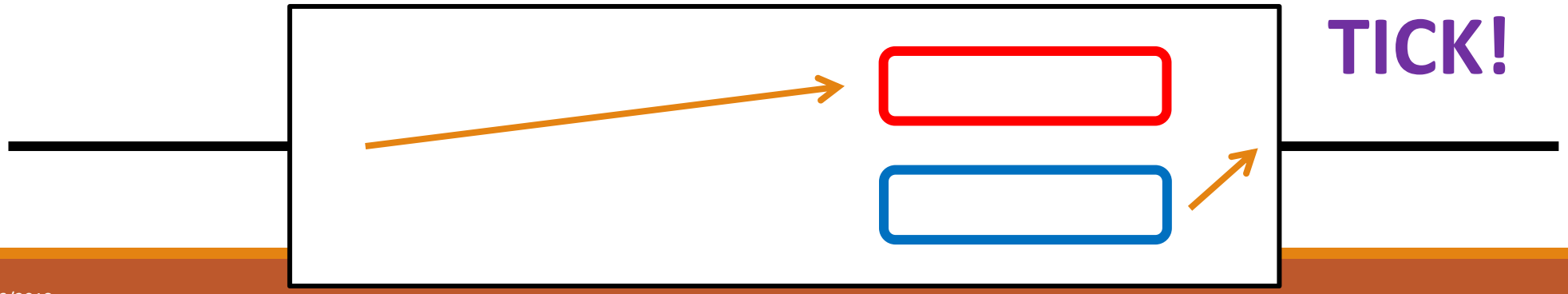
Two buffers per port. Input and output buffers swap at the same moment, once every cycle, period T_C. Small guard band to allow for link delay and forwarding delayu. All bridges are synchronized and swap buffers at the same moment. Cycle time T_C > transit time + forwarding time + clock inaccuracy + max data transmit time.



Two buffers



Two buffers

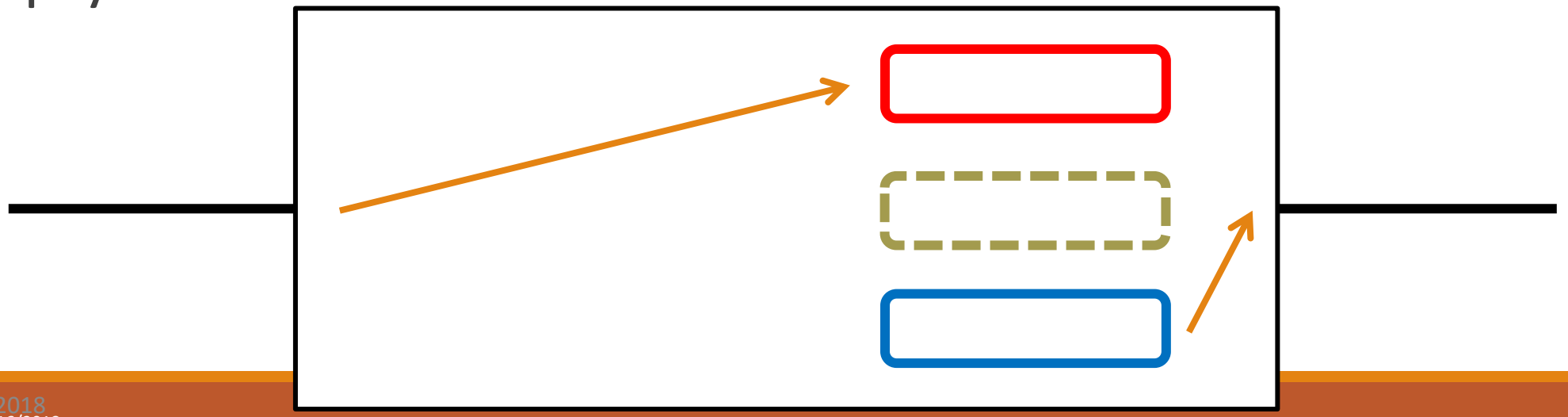


Three buffers

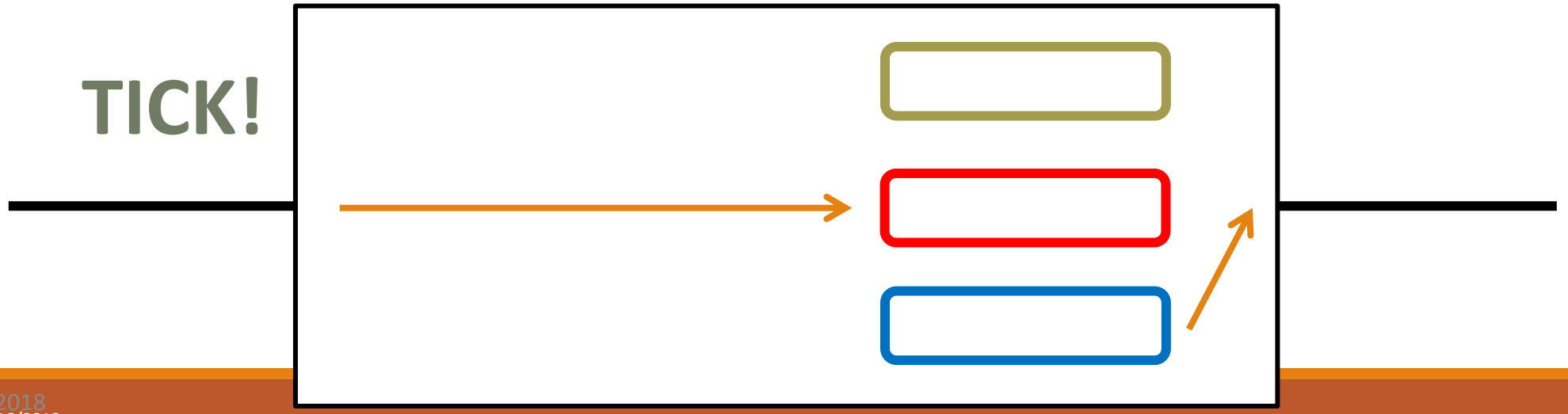
Input buffer swap is out-of-phase with output buffer swap to allow for arbitrary link delay.

No guard band needed for link/forwarding delay.

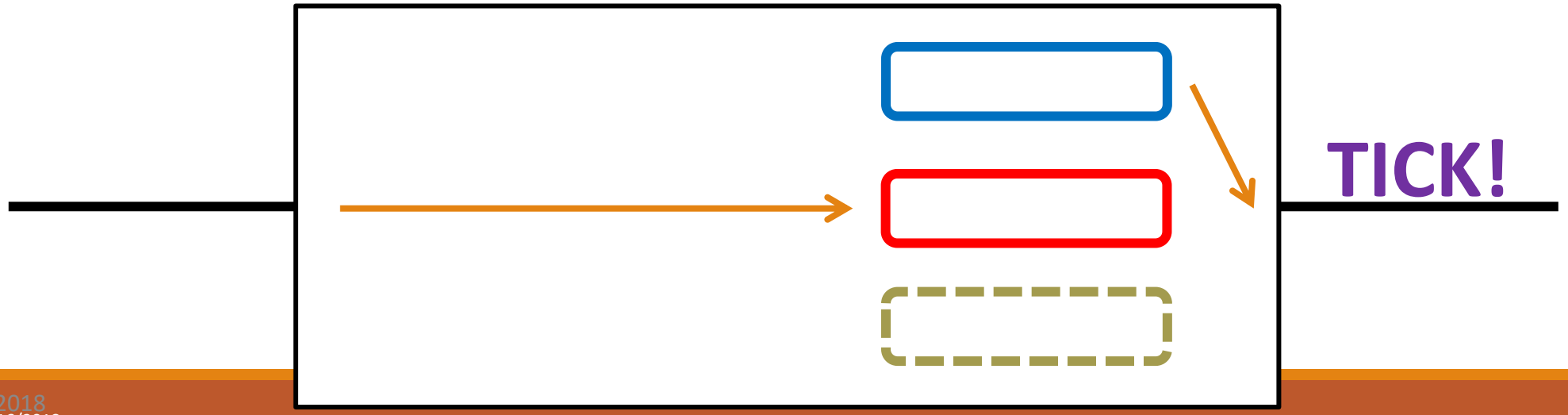
Each buffer cycles through four states: filling, full, draining, empty



Three buffers

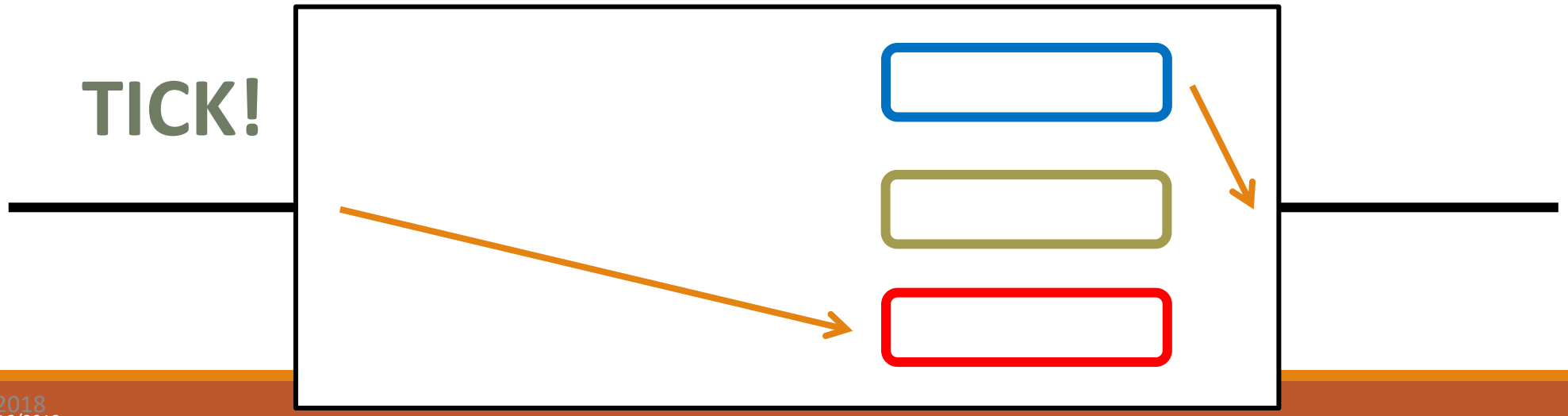


Three buffers



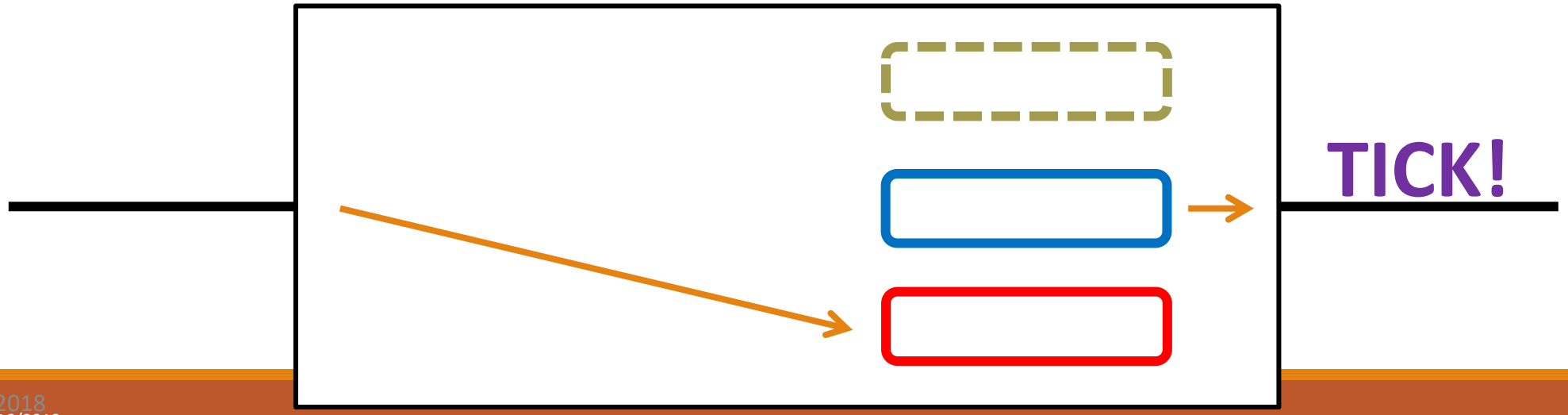
7/16/2018
9

Three buffers



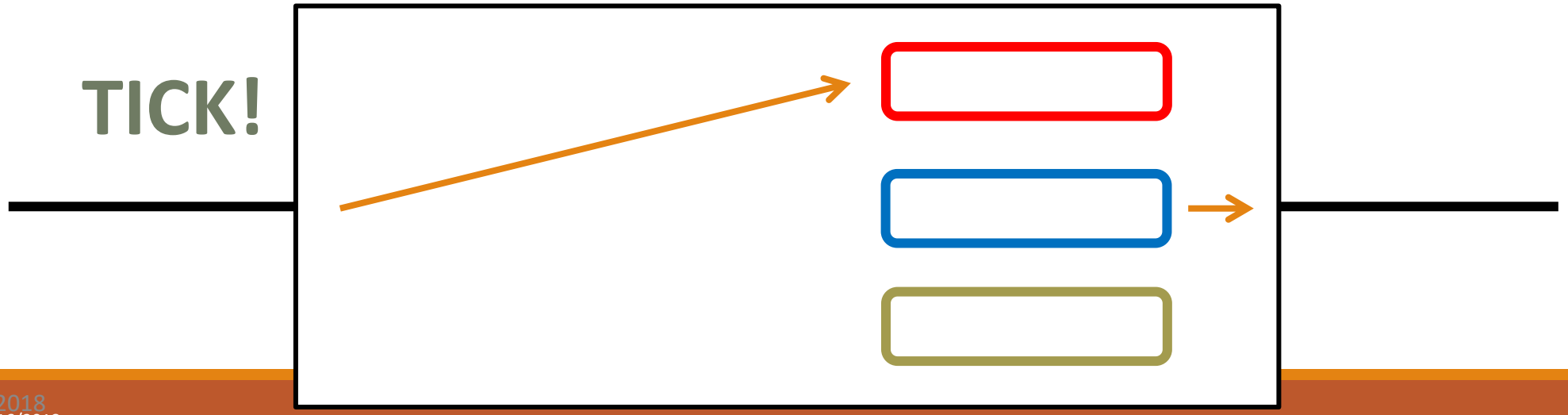
7/16/2018
10

Three buffers



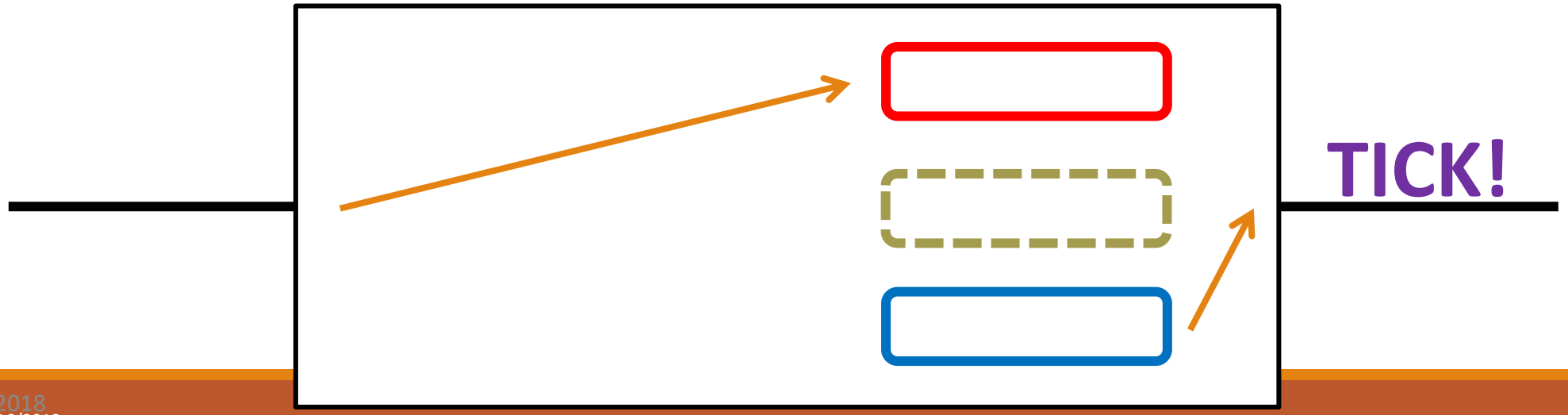
7/16/2018
11

Three buffers



7/16/2018
12

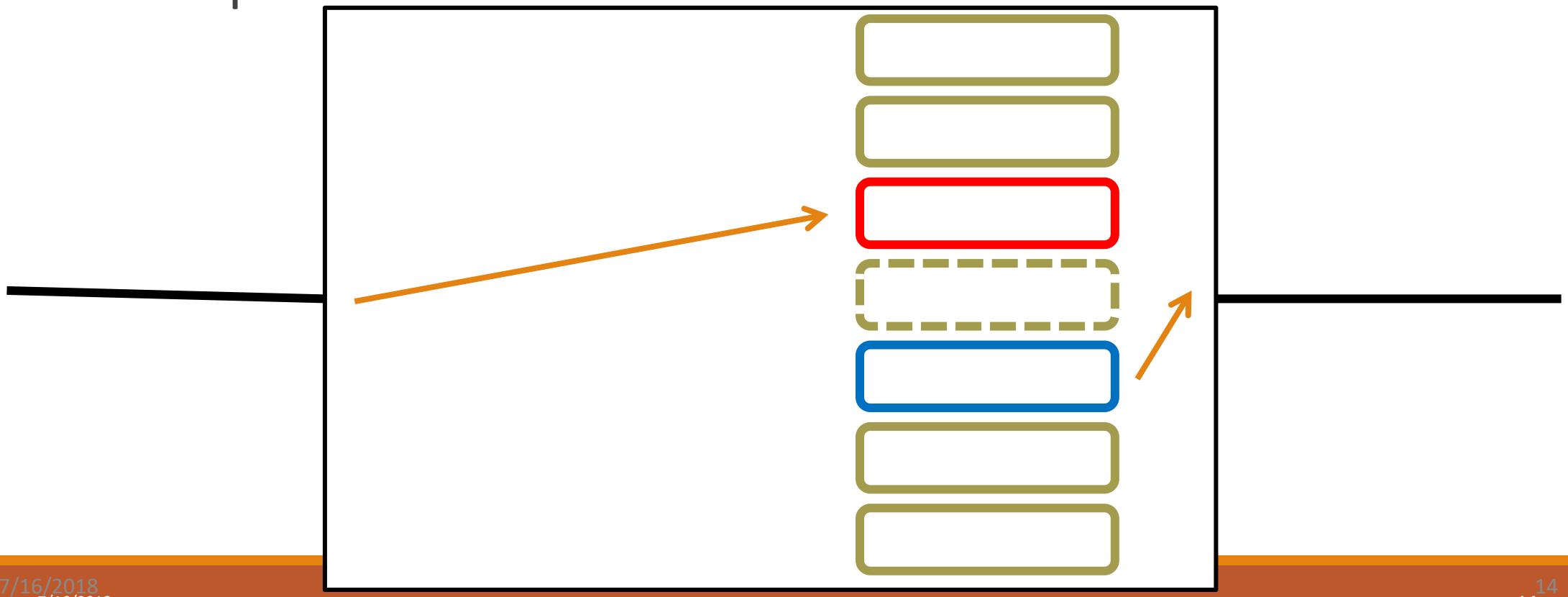
Three buffers



7/16/2018
13

N buffers

Any number of buffers per port. Useful for delay matching on different paths.

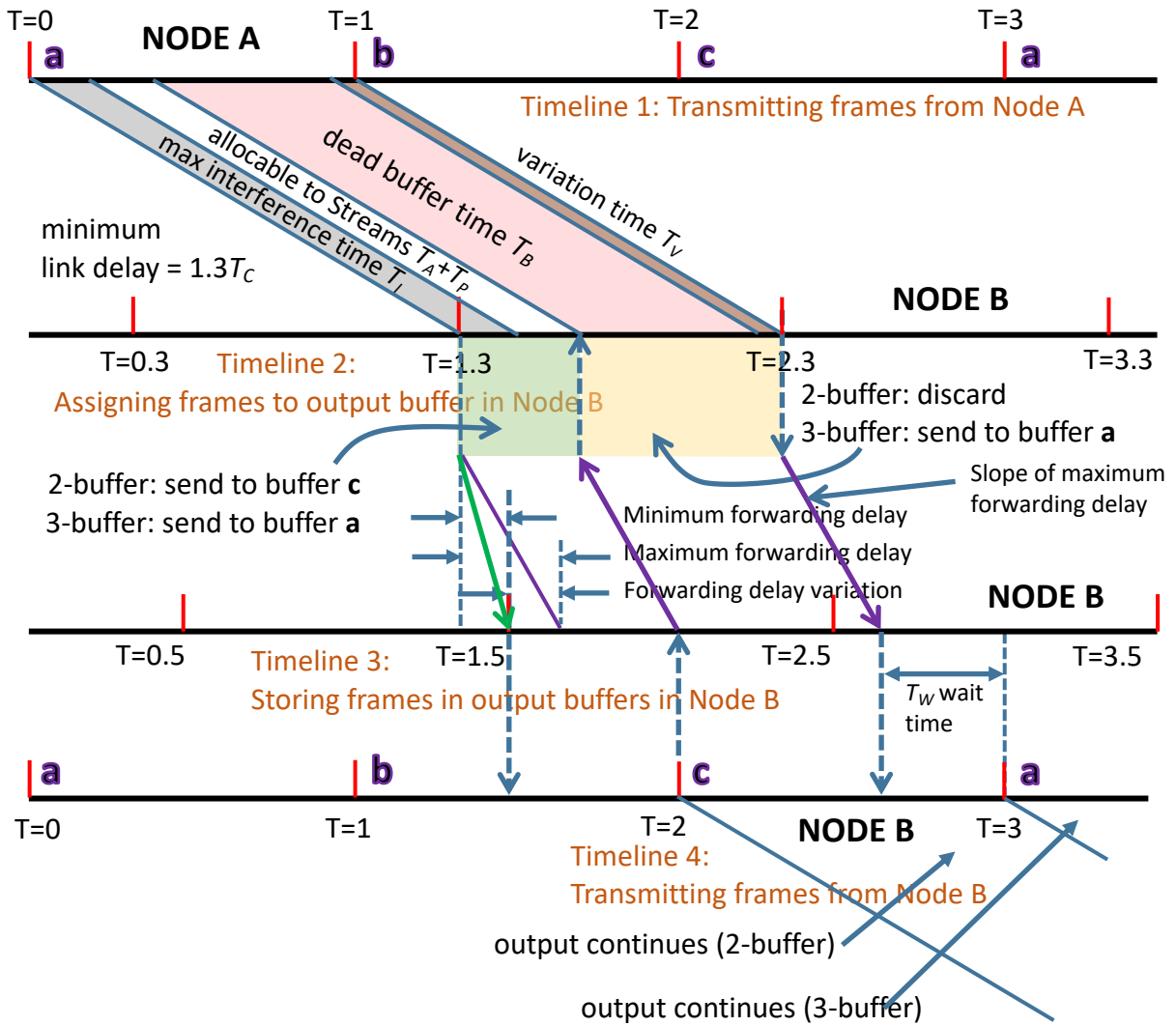


7/16/2018
14

The number of buffers used by a stream

- Different output ports on one system can have different numbers of buffers.
- Different input ports feeding the same output port can each use a different number of the buffers on that output port.
- One Stream can use a different number of buffers at each output port along the path.

df-finn-multiple-CQF-0919-v01 Figure 1 Reference timelines for time-based CQF



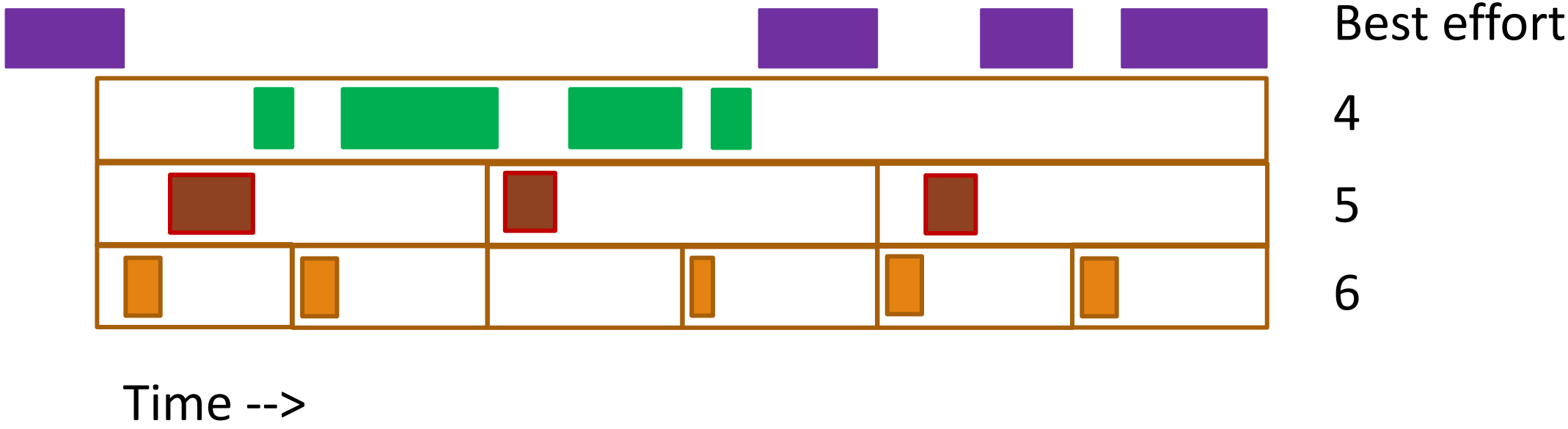
Example of Multiple cycle times on one output port

- Four cycle times at four priorities 3-6.
- 1 priority 3 cycle, buffers h, i (not shown)
- 3 priority 4 cycles, buffers f, g
- 6 priority 5 cycles, buffers d, e
- 24 priority 6 cycles, buffers a, b, and c.

<table>
<tr><td colspan="24">h</td><td>3</td></tr>
<tr><td colspan="12">f</td><td colspan="8">g</td><td colspan="4">f</td><td>4</td></tr>
<tr><td colspan="4">d</td><td colspan="4">e</td><td colspan="4">d</td><td colspan="4">e</td><td colspan="4">d</td><td colspan="4">e</td><td>5</td></tr>
<tr><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>6</td></tr>
</table>

Transmission order

- The total bandwidth allocated over all CQF priorities cannot exceed the link bandwidth.
- Draining of CQF buffers are by strict priority, fastest cycles first
- As long as every cycle is an integer multiple of the next-faster cycle, every buffer will empty before the end of its cycle.



More ways to improve the options available to CQF.

- Preemption:
  - Lowers the interference from lower-priority queues, especially best-effort
  - **CQF queues can be preemptable**, because the preemption penalty is bounded.
- Packing varied frame sizes into cycles can require overprovision:
  - "Sausage making," packing the customers' packets into fixed-sized encapsulations, requires much less overprovision, and can be done at the edges of a Service Provider network.
- In multiple CQF, deliberate overprovision of a Stream (assigning it to a too-fast cycle) reduces its latency.

The multiple-CQF value proposition

- CQF, as described here, requires network time synchronization.
- CQF requires time-aware state machines on input/output ports.

but,

- Except, perhaps, for the ingress bridge, CQF requires no per-Stream configuration, no per-Stream state machines, no per-Stream knowledge at any node in the network.
- CQF allows a network controller to perform admission control with no per-flow conversations with any (except perhaps the ingress) nodes in the network.

To come …

There are many more improvements that can be made to TSN, many of which require no new standards. I hope to supply more papers and presentations in this vein.

Thank you