Latency Guarantee with Stateless Fair Queuing (C-SCORE)

draft-joung-detnet-stateless-fair-queuing-05

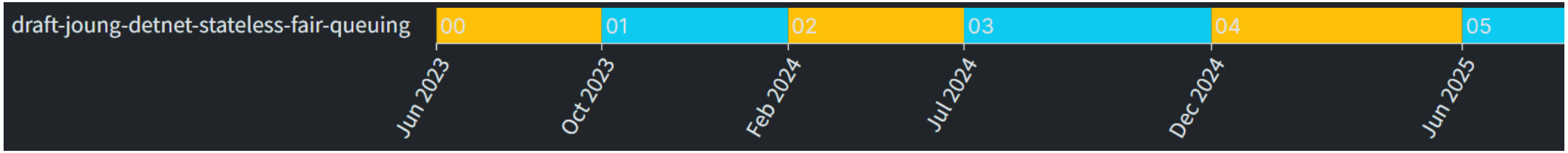
Jinoo Joung	**Jeong-dong Ryoo**	**Tae-sik Cheung**	**Yizhou Li**	**Peng Liu**
Sangmyung University	ETRI	ETRI	Huawei	China Mobile

IETF 123, DetNet WG, July 2025

History of C-SCORE

- Included in “Asynchronous Deterministic Networking Framework for Large-Scale Networks” draft-joung-detnet-asynch-detnet-framework
- as one of the async. data plane solutions, from June 2022.

- Separated to an independent draft with the requests from WG, on June 2023.



- Has been through intensive debates, discussions, and revisions until IETF 119.

Revisions in the 05 version

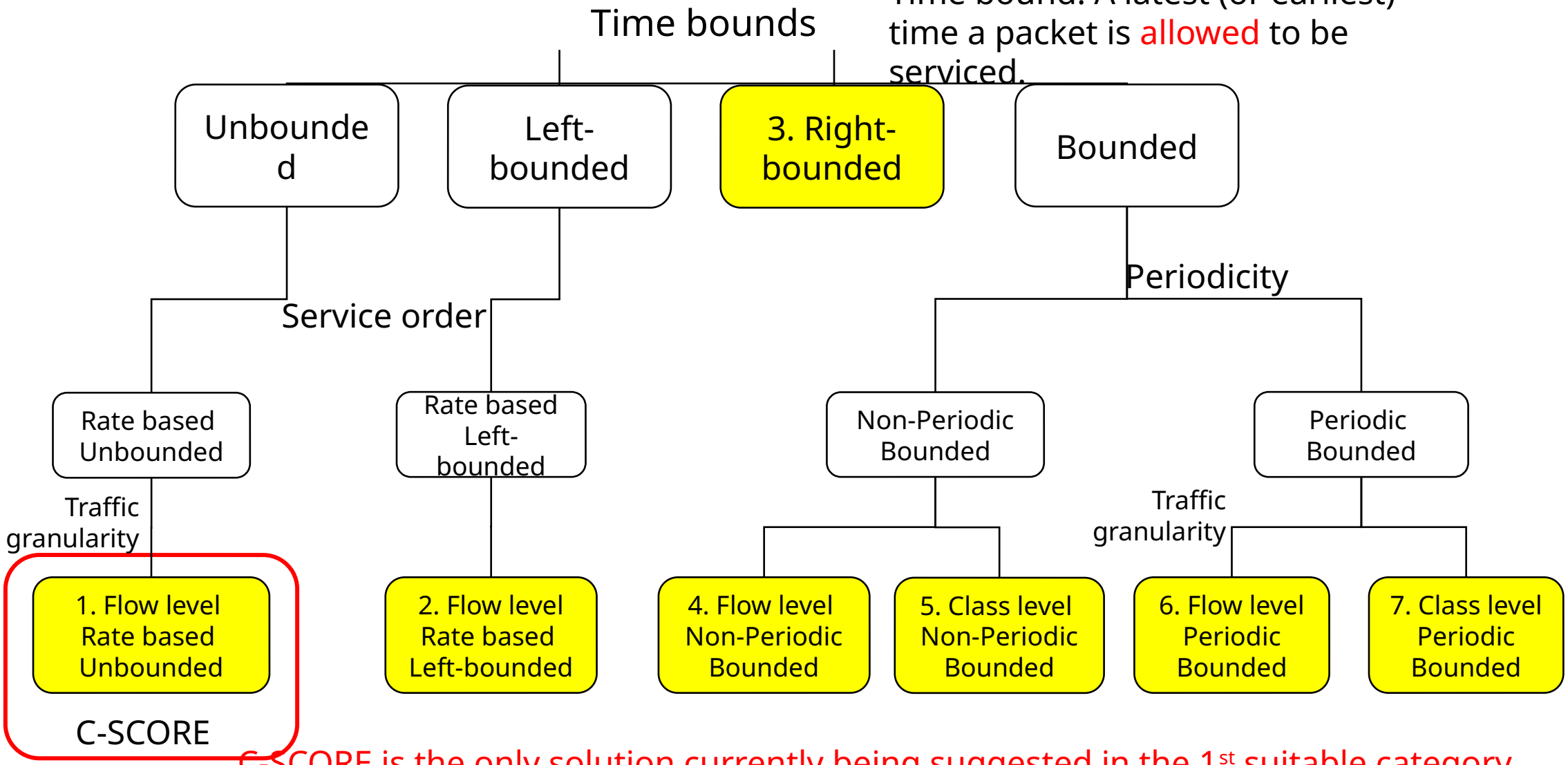
- More elaboration on the characteristics

6.4. Characteristics 10
6.4.1. Taxonomy 10
6.4.2. Strengths 11

- Specified that C-SCORE belongs to the one of the “Suitable categories” defined in the Taxonomy draft [draft-ietf-detnet-dataplane-taxonomy], reflecting the Chair's request.

Category of C-SCORE [draft-ietf-detnet-dataplane-taxonomy]

* Time bound: A latest (or earliest) time a packet is allowed to be serviced.

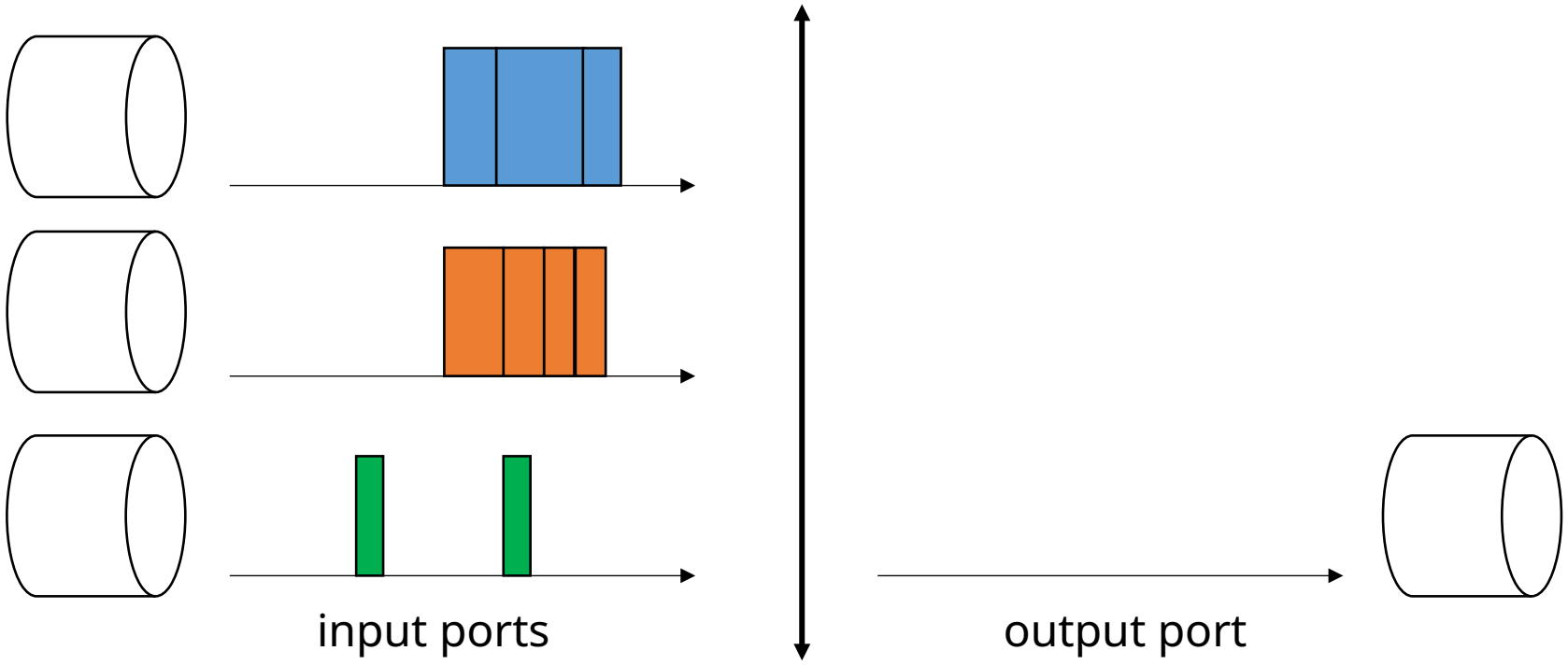


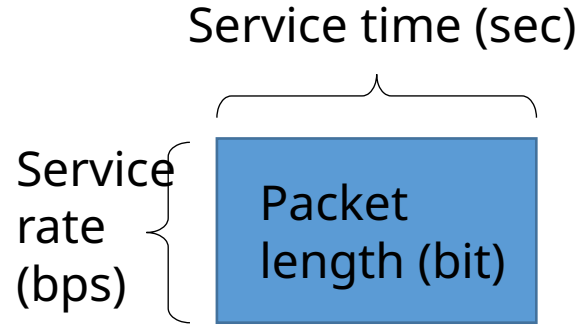
C-SCORE is the only solution currently being suggested in the 1st suitable category.

The key idea of C-SCORE

It boils down to “how to resolve the bursts collision”

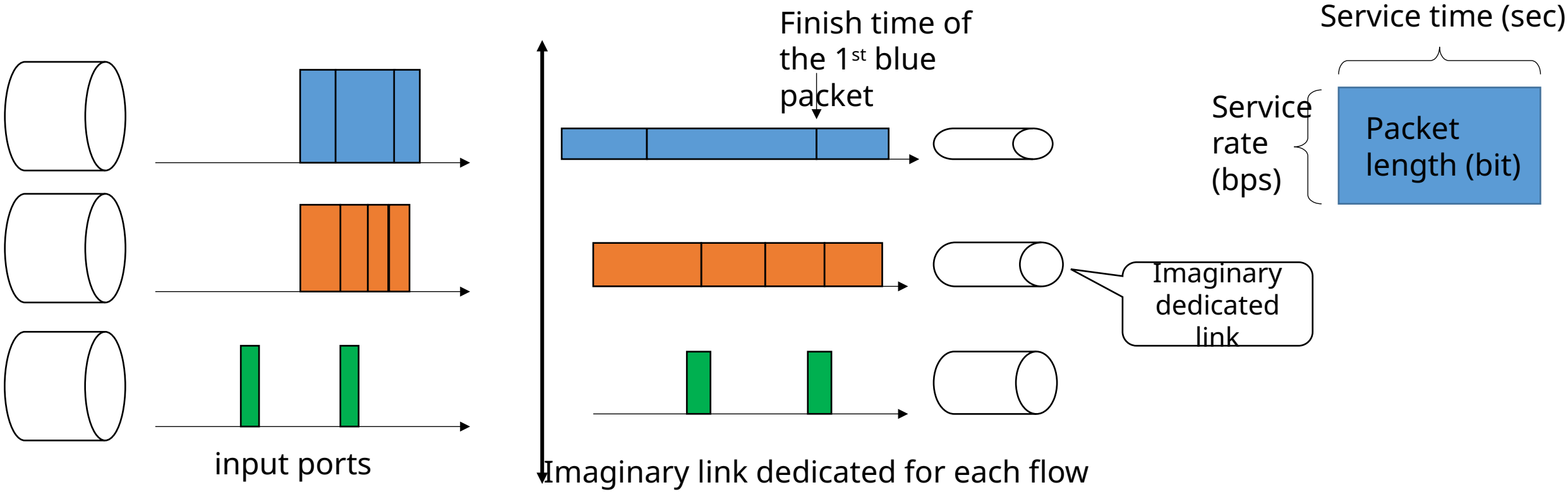
- Three flows arrive from input ports.
- Two of them are bursty.
- How will these packets be scheduled?





How to resolve the collision: Decision of Finish Time (FT)

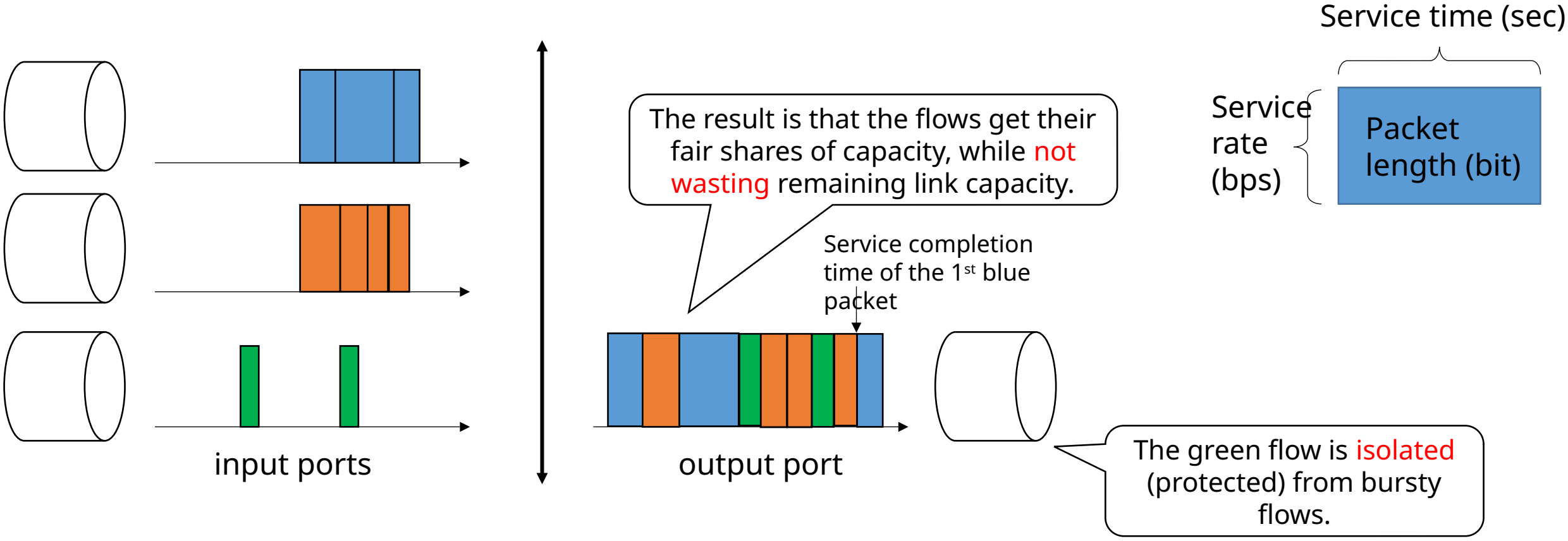
- Allocate an imaginary dedicated link to each flow, whose capacity equals to the flows' service rate.
- The service completion time of a packet in that imaginary link is its Finish time (FT).
- The only necessary condition is that the sum of service rates are less than the link capacity.



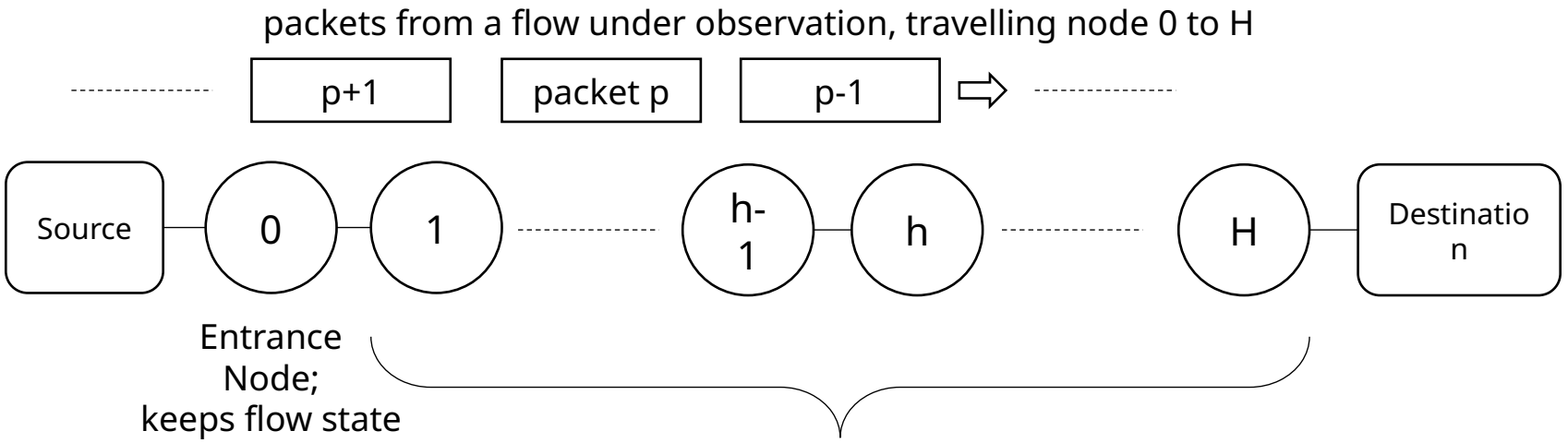
Flow isolation from other bursts: FT as the service order

- Then, serve those packets in the order of their FTs.
- The service completion time of packets are guaranteed to be less than their FTs.

 This is the basis of E2E latency guarantee.



More than that: Stateless core nodes



Entrance Node;
keeps flow state $F_0(p-1)$, L/r;
calculates $F_0(p)$ & $F_1(p)$;
marks $F_1(p)$, L/r as packet metadata;
serves packets in ascending order of $F_0(p)$.

Core Nodes; stateless.

Node h simply updates metadata $F_h(p)$ to $F_{h+1}(p)$;

serves packets in ascending order of $F_h(p)$.

Symbol	Definition
$F_h(p)$	'Finish time' of packet p at node h
$A_0(p)$	Arrival time of p at node 0
L(p)	Length of p
L	Max Packet Length of p's flow
	Arrival rate of flow j
Bj	Max burst of flow j
r	Service rate of p's flow
	Service rate of flow j at node h
$Lmax_h$	Max Packet Length at node h
R_h	Link capacity of h
f(h)	Set of flows in node h

Advantages

- C-SCORE has
  - a simple admission control procedure.
  - an excellent flow isolation (thus protection) property. Once admitted, a flow is not affected by other flows' fluctuation (their joins/leaves/mal-behaviors).
  - a flexibility to adjust E2E latency bounds, by setting the assigned service rates.
  - an excellent statistical performance, due to its work conserving nature. This is important to average buffer occupancy in the network. Refer to Little's law; $B = \lambda * D$ (Buffer occupancy = arrival rate * delay, in average).

Complexities in data & control plane

- Currently, the E2E jitter bound is the same with the latency bound.
- Data plane complexity:
  - A priority queue is necessary to sort the packet according to their FT.
  - This renders about a half the packet processing speed compared to a FIFO. [Bhagwan]
    - i.e. at the same cost, a C-SCORE chip can run at 100Gbps, while a FIFO chip can run at 200Gbps.
- Control plane complexity:
  - Admission control procedure is necessary to check if the maximum total service rate exceeds the link capacity.
  - No more control procedure, other than the admission control.

Conformance to the requirements for scaling

Requirements for scaling [draft-ietf-detnet-scaling-requirements-07]

1. Tolerate Time Asynchrony; due to 1) difficulties in synchronization across multiple domains & 2) non-periodic traffic
2. Support Large Single-hop Propagation Latency; which makes the cyclic operations difficult
3. Accommodate the Higher Link Speed (or the discrepancy of the link speeds); which also makes the cyclic operations difficult
4. Be Scalable to A Large Number of Flows and Tolerate High Utilization of Bandwidth
5. Tolerate Failures of Links or Nodes and Topology Changes
6. Prevent Flow Fluctuation (Flows join/leaves & Bursty flows)
7. Be Scalable to a Large Number of Hops with Complex Topology
8. Support Multi-Mechanisms in Single Domain and Multi-Domains

Restructured Requirements

- Overcome the inflexibility of synchronization and cyclic operation
  1. Tolerate Time Asynchrony; due to 1) difficulties in synchronization across multiple domains & 2) non-periodic traffic
  2. Support Large Single-hop Propagation Latency; which makes the cyclic operations difficult
  3. Accommodate the Higher Link Speed (or the discrepancy of the link speeds); which also makes the cyclic operations difficult
- Overcome the complexity due to large network size and high utilization
  4. Be Scalable to A Large Number of Flows and Tolerate High Utilization of Bandwidth
  7. Be Scalable to a Large Number of Hops with Complex Topology
- Overcome the network dynamicity
  5. Tolerate Failures of Links or Nodes and Topology Changes
  6. Prevent Flow Fluctuation (Flows join/leaves & Bursty flows)
  8. Support Multi-Mechanisms in Single Domain and Multi-Domains

How C-SCORE meets the Requirements

- Overcome the inflexibility of synchronization and cyclic operation
  - → No need for synchronization or cyclic operation
- Overcome the complexity due to network size and high utilization
  - → C-SCORE's work conserving nature optimizes the statistical performance in highly utilized large-scale networks.
  - → Flow isolation capability guarantees E2E latency in a highly loaded network with a large number of hops.
- Overcome the network dynamicity
  - → Simple admission control mechanism
  - → Easy E2E latency bound calculation
  - → A service rate given to a flow can be unaltered even with network changes
    → E2E latency bounds unaltered

Future Plan

- The draft is now stable.
- WG adoption after this meeting, hopefully.

Thank you

- Please take a look at

  https://datatracker.ietf.org/doc/draft-joung-detnet-stateless-fair-queuing/

- [Kaur] Jasleen Kaur, and Harrick M. Vin. "Core stateless guaranteed rate scheduling algorithms." In Proceedings IEEE INFOCOM 2001.
- [Bhagwan] Ranjita Bhagwan and Bill Lin, “Fast and Scalable Priority Queue Architecture for High Speed Network Switches", IEEE INFOCOM 2000.