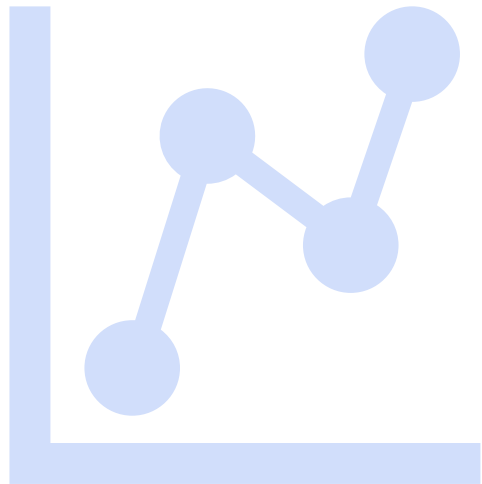


# Understanding the impact of host networking elements on traffic bursts

Erfan Sharafzadeh, Sepehr Abdous, Soudeh Ghorbani



April 2023

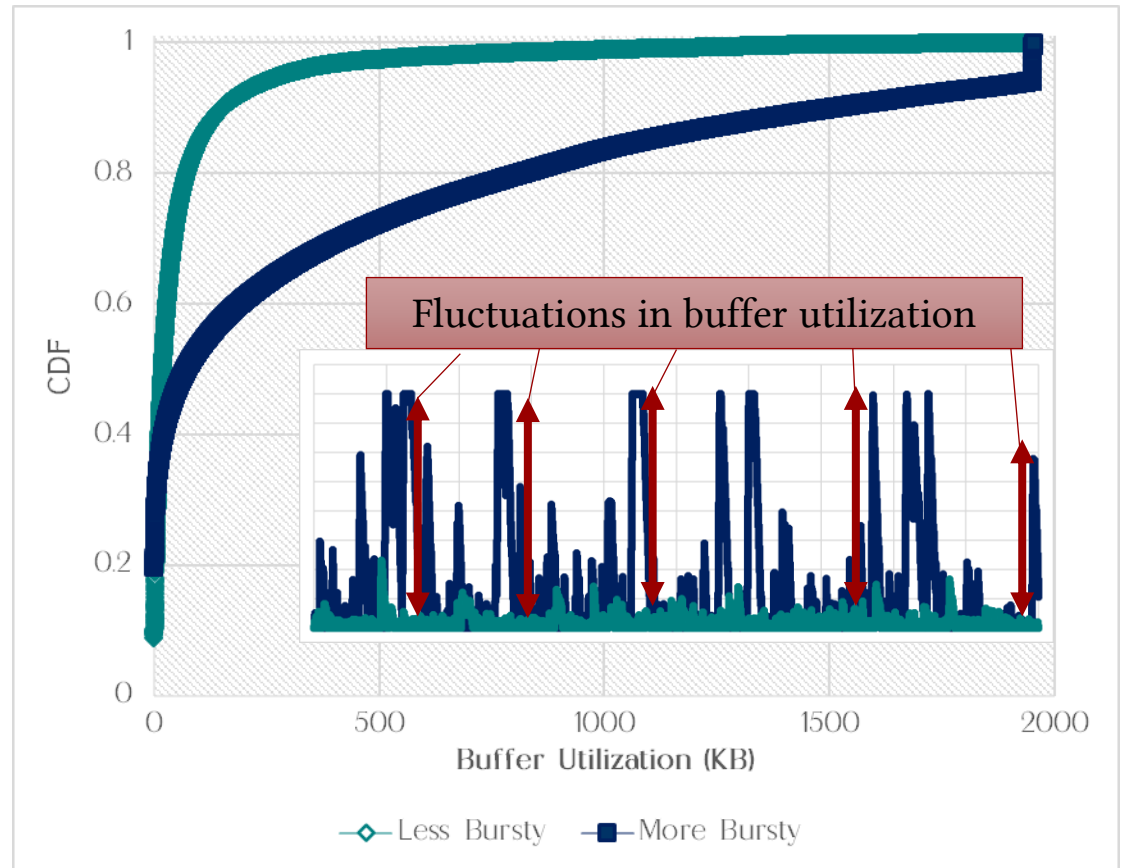
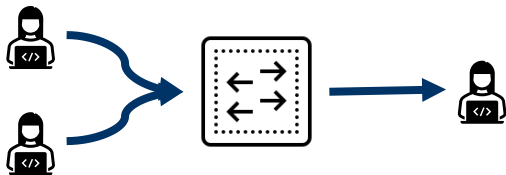


# Burstiness degrades performance

Bursts are the major cause of **queue backlogs!**

A **4X** more bursty traffic results in **10X** longer buffer utilization tails!

**10X longer p99 Round-Trip Times!!**

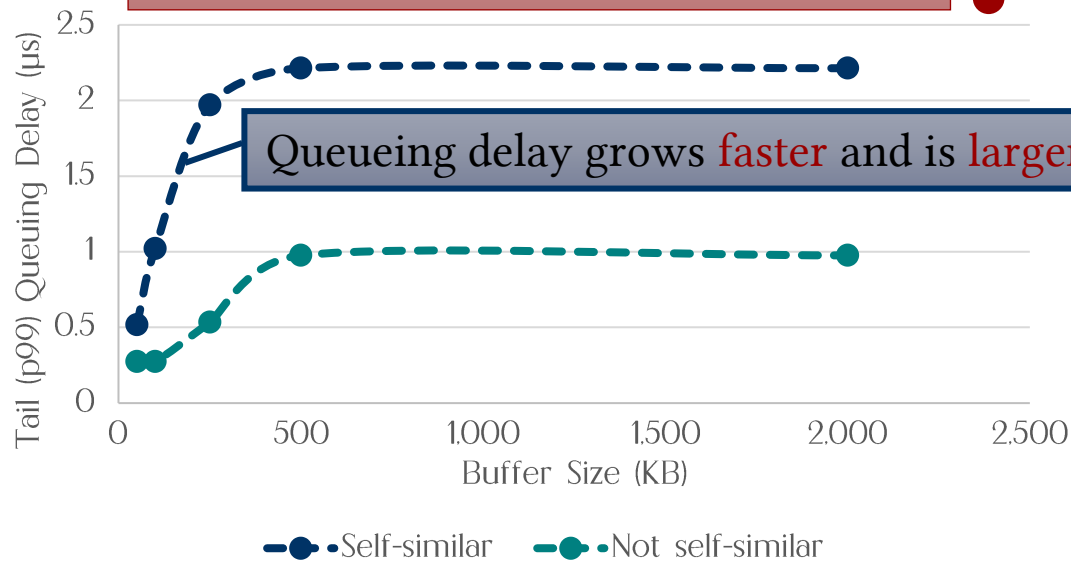


# Buffer sizing depends on burst scales

**Self-similarity:** The traffic **remains bursty** at different timescales

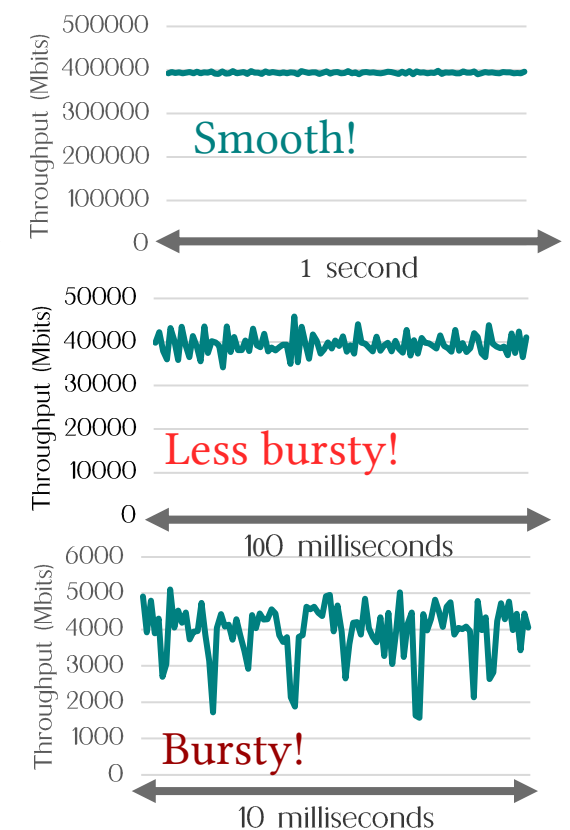
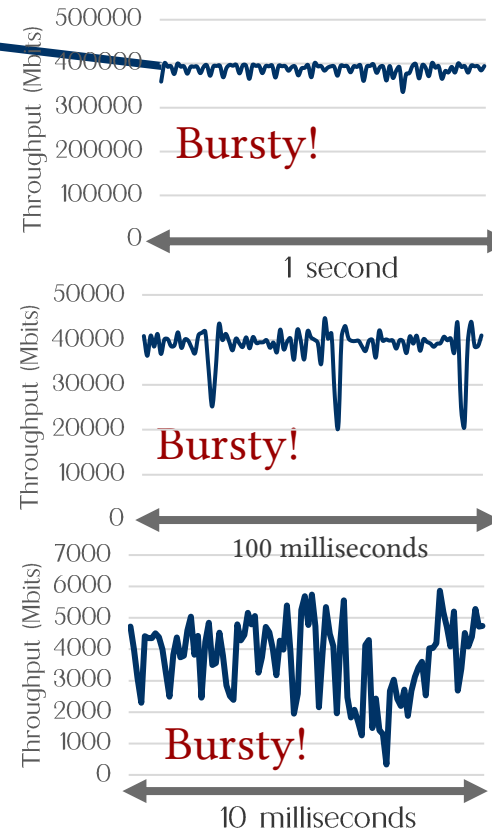
Smaller buffers are needed to control the **queueing delay!**

Queueing delay grows **faster** and is **larger**



Self-similar

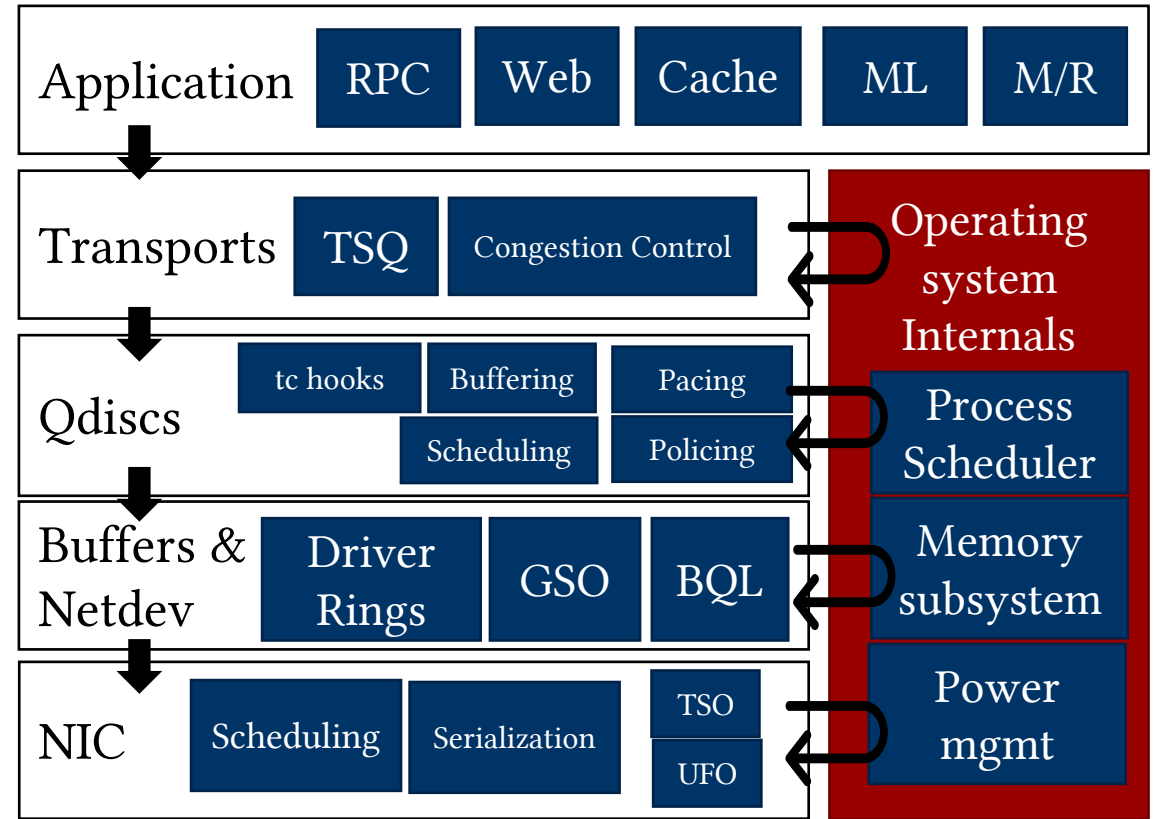
Not self-similar



# What causes the traffic to emerge in bursts at different timescales?



Valinor: A high-resolution traffic measurement framework



**On the Self-Similar Nature of Ethernet Traffic**  
 Will E. Leland<sup>†</sup> [wel@bellcore.com](mailto:wel@bellcore.com)  
 Walter Willinger<sup>‡</sup> [walter@bellcore.com](mailto:walter@bellcore.com)  
 †Bellcore, 443 South Street, Morristown, NJ 07960-6438  
 ‡Department of Mathematics, Boston University, Boston, MA 02215

**Abstract**  
 We demonstrate that Ethernet local area network (LAN) traffic is statistically self-similar, that none of the commonly used traffic models is able to capture this fractal behavior, and that such behavior has serious implications for the design, control, and analysis of high-speed call-based networks. *Introduce the*

[IEEE/ACM ToN 1994]

**On the relationship between file sizes, transport protocols, and self-similar network traffic**  
 Kihong Park<sup>\*</sup> [park@cs.purdue.edu](mailto:park@cs.purdue.edu)  
 Department of Computer Sciences, Purdue University, West Lafayette, IN 47907  
 Gitae Kim<sup>†</sup> [kgkim@cs.bu.edu](mailto:kgkim@cs.bu.edu)  
 Computer Science Department, Boston University, Boston, MA 02215  
 Mark Crovella<sup>‡</sup> [kgkim@cs.bu.edu](mailto:kgkim@cs.bu.edu)

**Abstract**  
 Recent measurements of local-area and wide-area traffic have shown that network traffic exhibits variability at a wide range of scales. In this paper, we examine a mechanism that gives rise to self-similar network traffic and present some of its performance implications. The mechanism we study is the transfer of files or messages whose size is drawn from a

1 Introduction  
 [ICNP 1996]

# Our study uncovers:

Lower layers of the stack can undo **TCP pacing**

In *Multi-queue NICs with Segmentation offload*, enabling or disabling TCP pacing has no effect on burst lengths.

**Congestion control** variants result in significantly different self-similarity

TCP *cubic* results in a more self-similar traffic compared to *DCTCP* and *BBR*.

**Process schedulers** with coarse time-slicing result in heavy bursts

High self-similarity when running a network application under CPU contention with *Completely Fair Scheduler*.

Smaller **buffer sizes** in the hosts can significantly reduce burstiness

Driver buffer sizing enforced by *Byte Queue Limits* algorithm is a cause for longer bursts.

# Valinor: A network traffic burst analyzer

We need to capture **timestamps!**

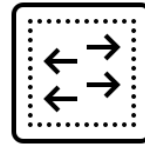
We need to collect **metadata** and **statistics!**



1) Enables observing burstiness at upper layers of stack



2) Does not need specialized HW/SW



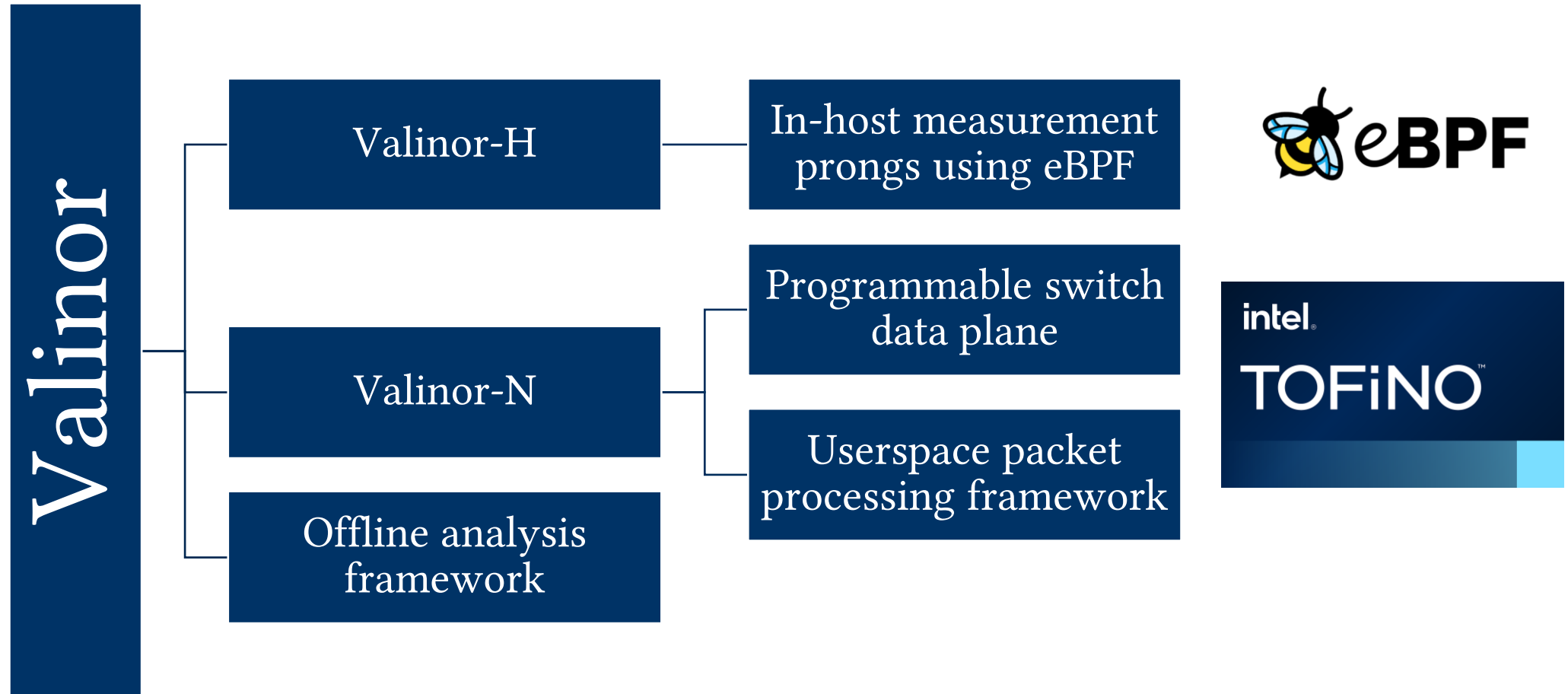
1) Enables observing bursts on the wire



2) Enables observing the aggregate behavior of bursts (queueing)



# Valinor studies network traffic from two vantage points

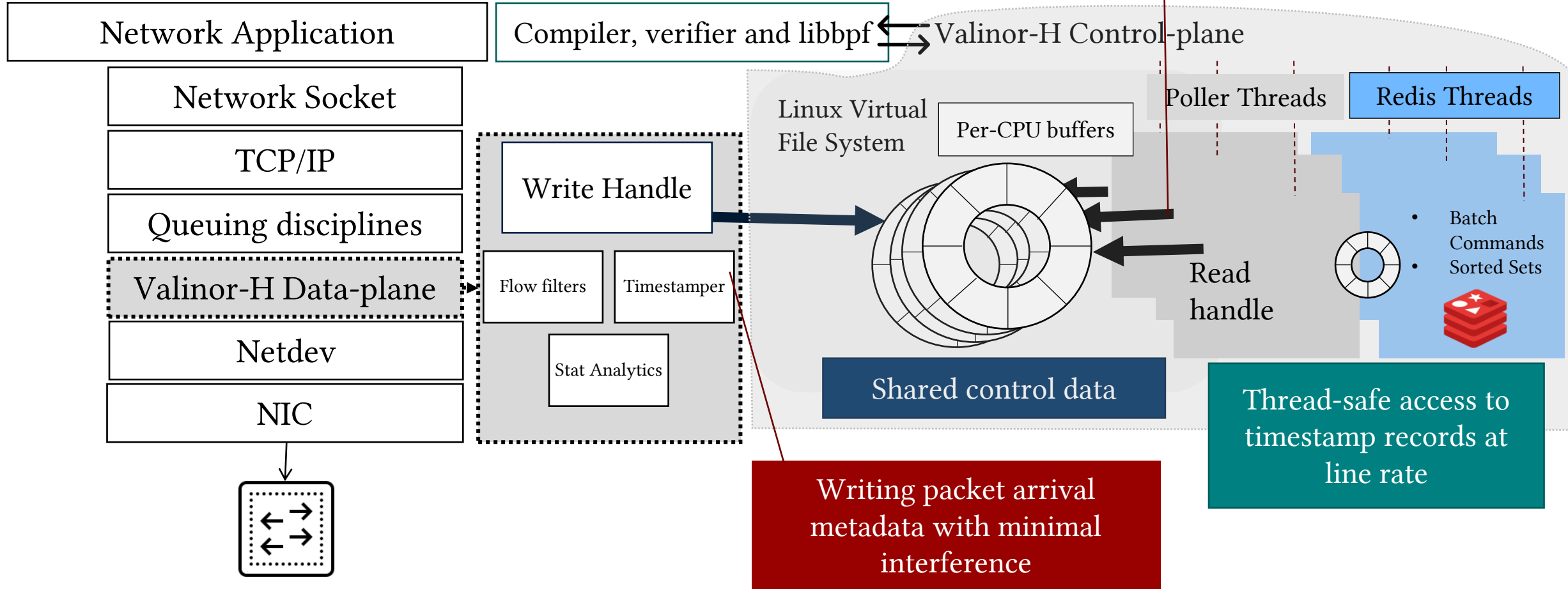




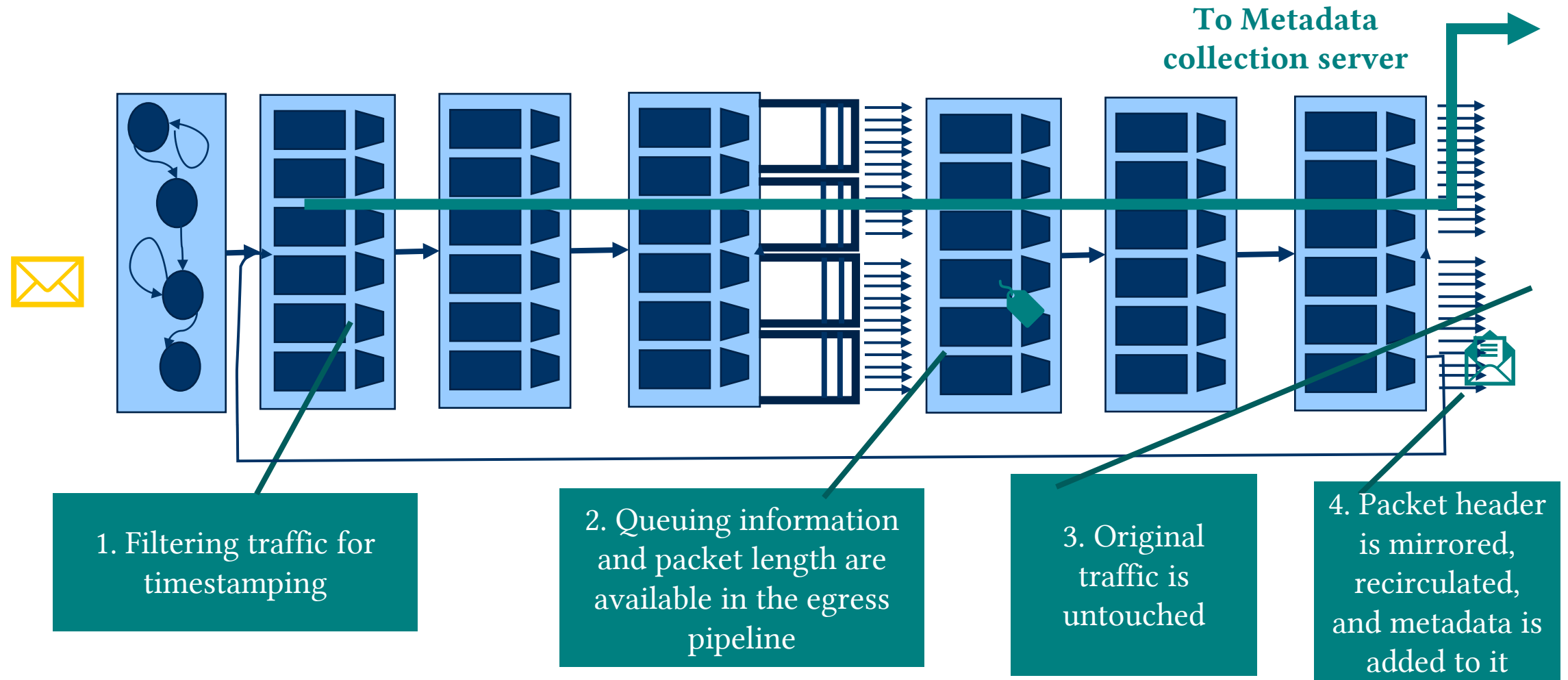
# Valinor-H: eBPF Probe

Keeping up with the fast data plane

Two-layered thread design:  
 1. Poller threads  
 2. Redis threads



# Valinor-N: Enabling In-network measurements



# Summary of Valinor findings

Lower layers of the stack can undo **TCP pacing**

In *Multi-queue NICs* with *Segmentation offload*, enabling or disabling *TCP pacing* has no effect on burst lengths.

**Congestion control** variants result in significantly different self-similarity

TCP *cubic* results in a more self-similar traffic compared to *DCTCP* and *BBR*.

**Process schedulers** with coarse time-slicing result in heavy bursts

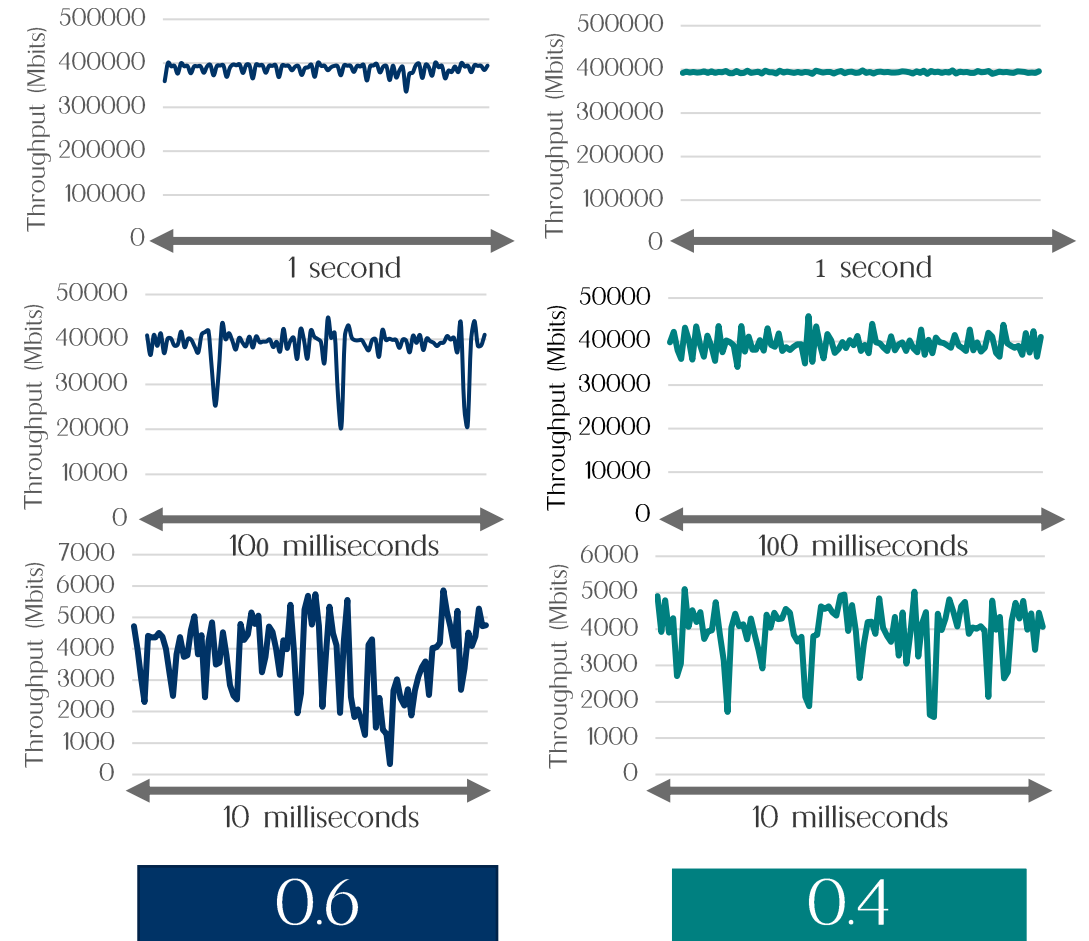
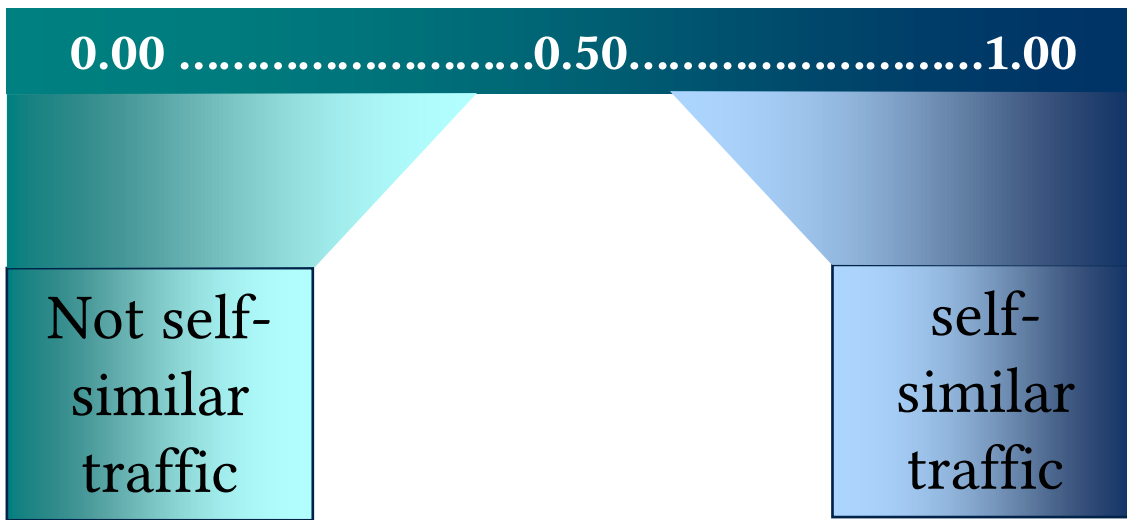
High self-similarity when running a network application under CPU contention with *Completely Fair Scheduler*.

Smaller **buffer sizes** in the hosts can significantly reduce burstiness

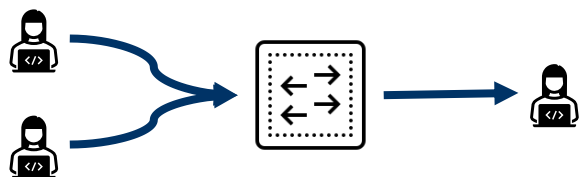
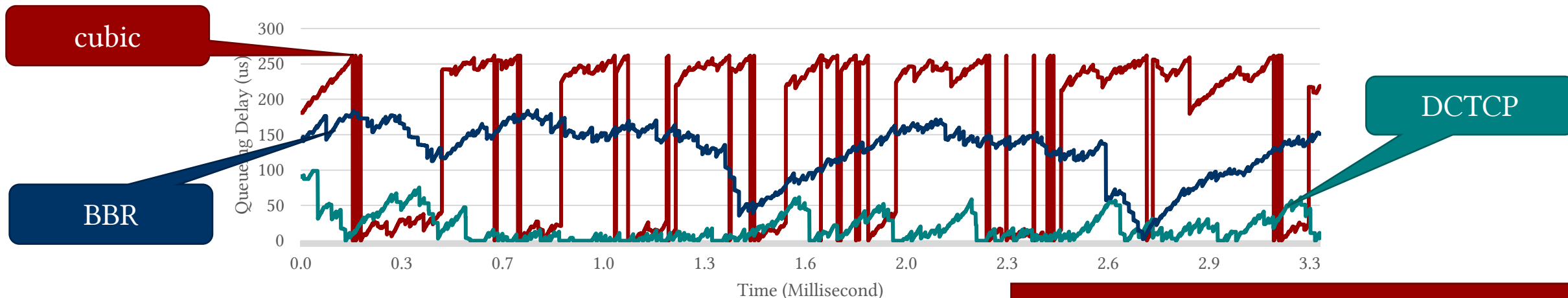
Driver buffer sizing enforced by *Byte Queue Limits* algorithm is a cause for longer bursts.

# Hurst exponent: A measure of self-similarity

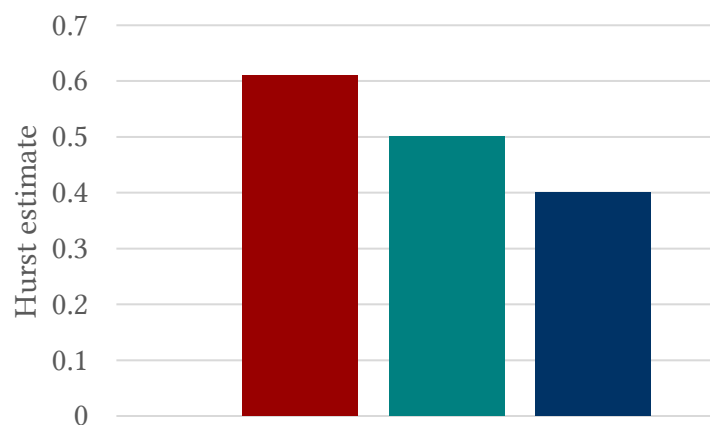
 Measure of how bursts are preserved at scales



# Observing TCP congestion control variants



- 2-1 Incast
- Map-Reduce flow size distribution
- 40Gbps offered load

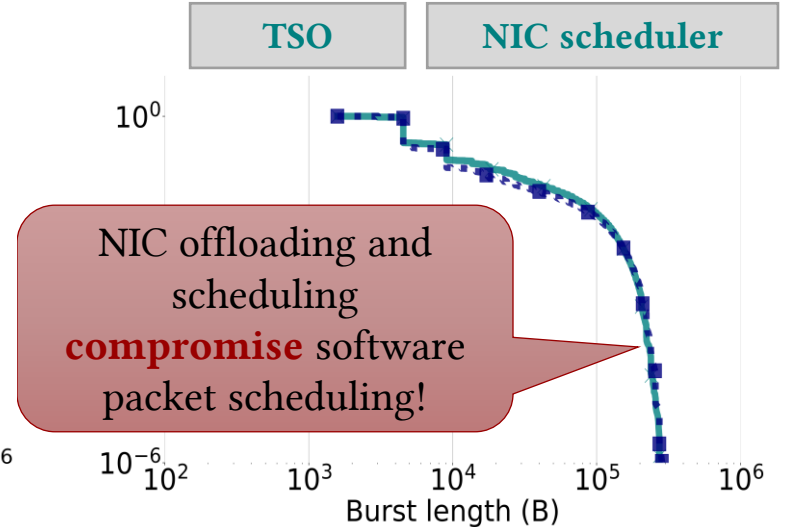
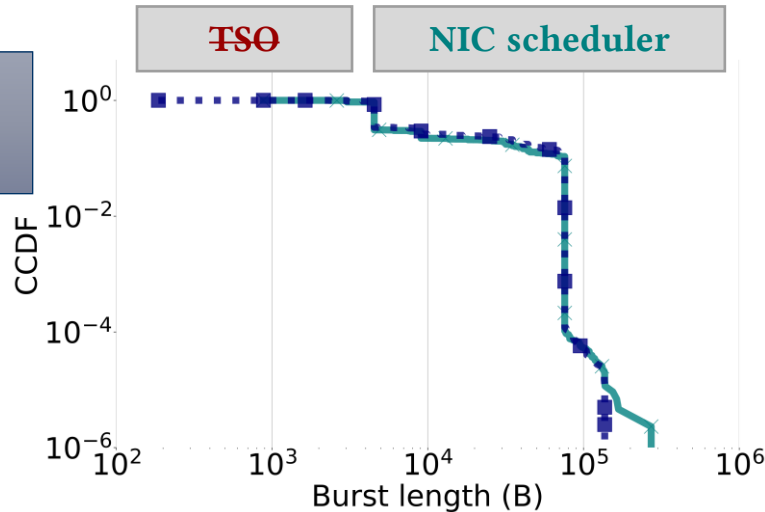
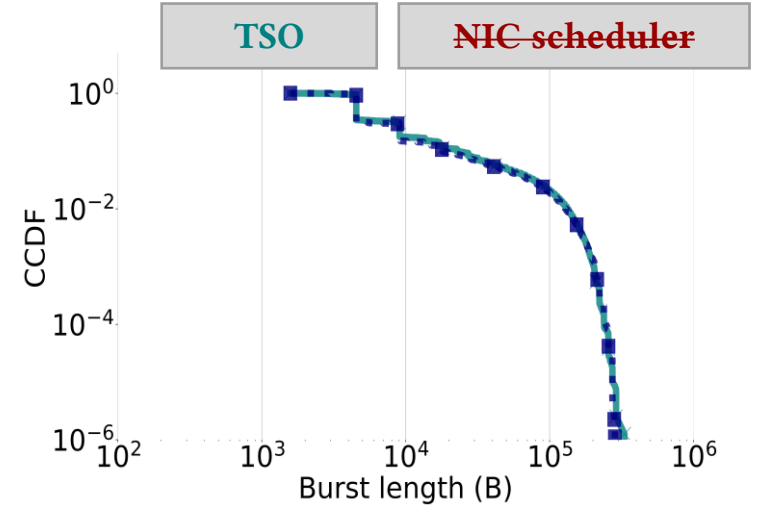
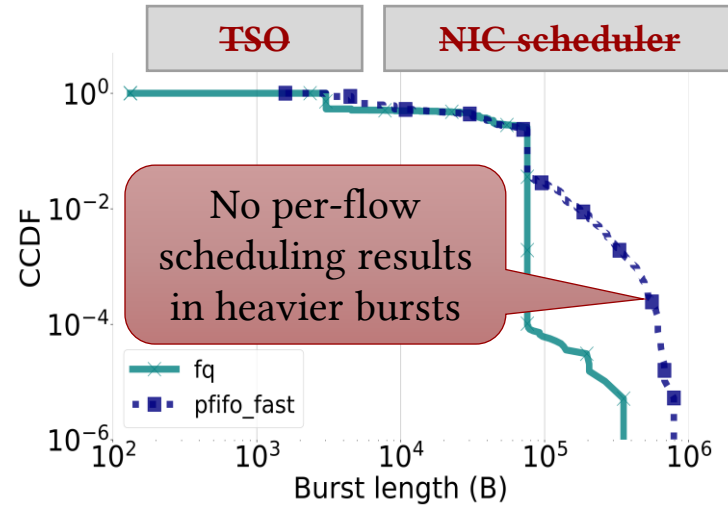
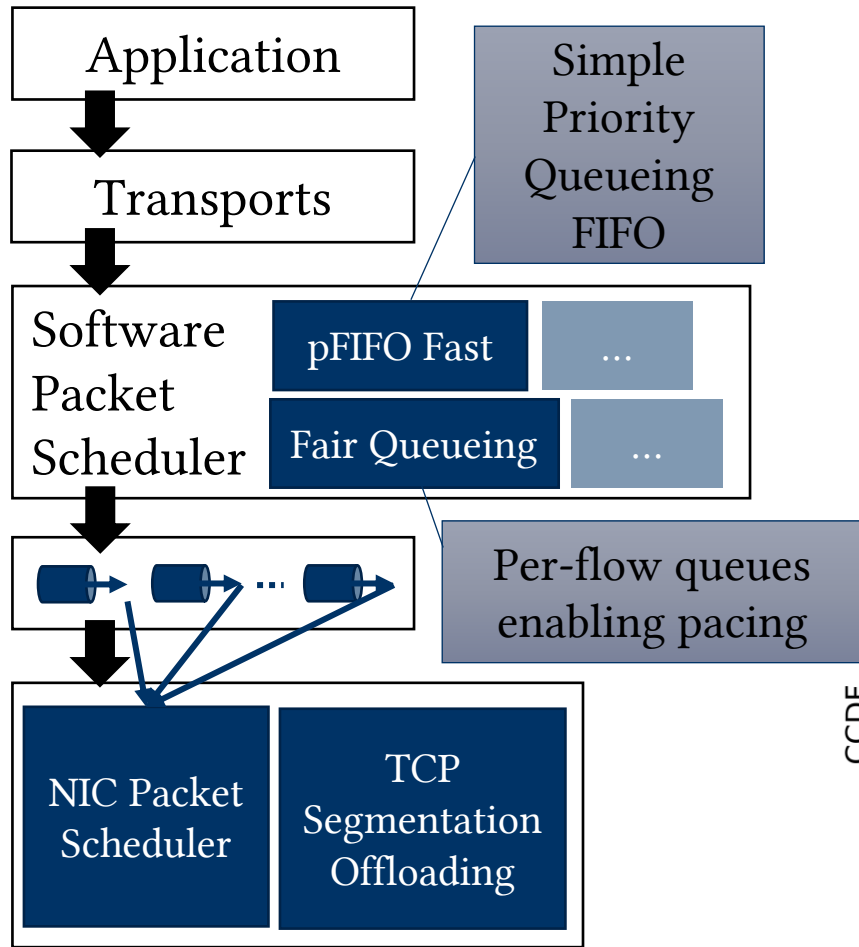


Hurst = 0.6  
cubic traffic is self-similar

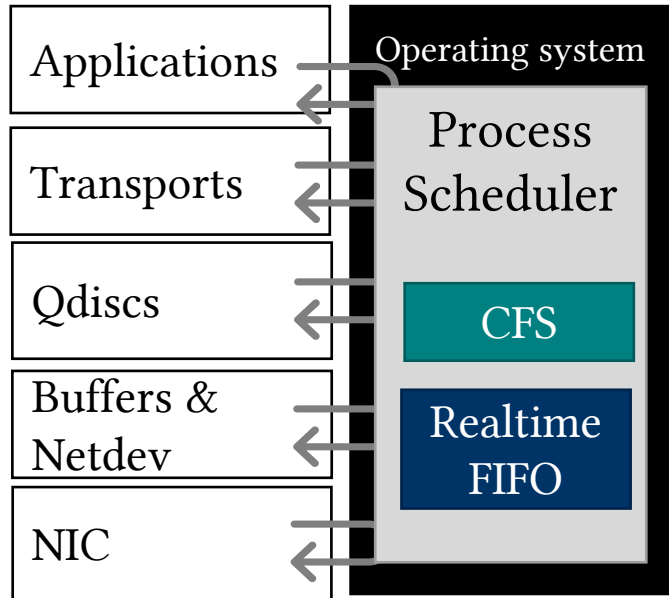
Hurst  $\leq$  0.5  
BBR and DCTCP traffic are not self-similar

**Buffer sizes** can be tuned based on self-similarity of congestion control variants

# Is software effective in shaping the traffic?



# Even process scheduling matters!



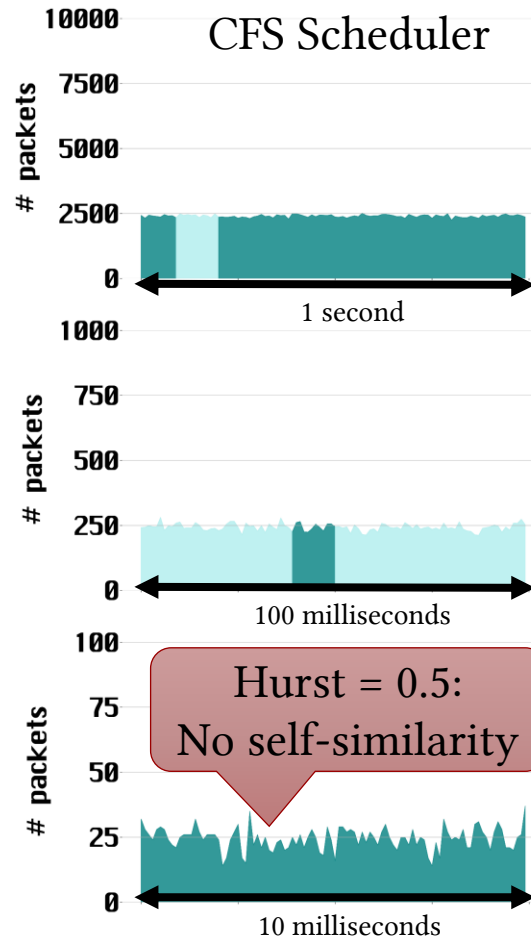
## CFS

- Coarse-grained timeslicing
- Fairness

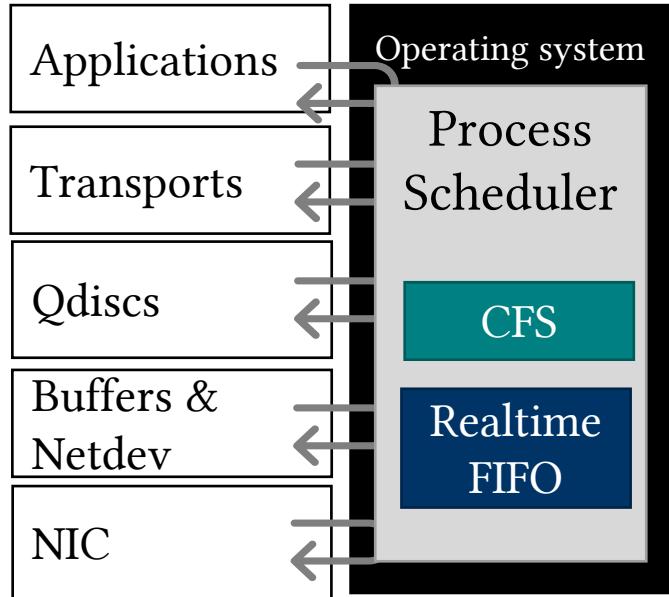
## Realtime FIFO

- Strict prioritization

Not-contended workload



# Even process scheduling matters!

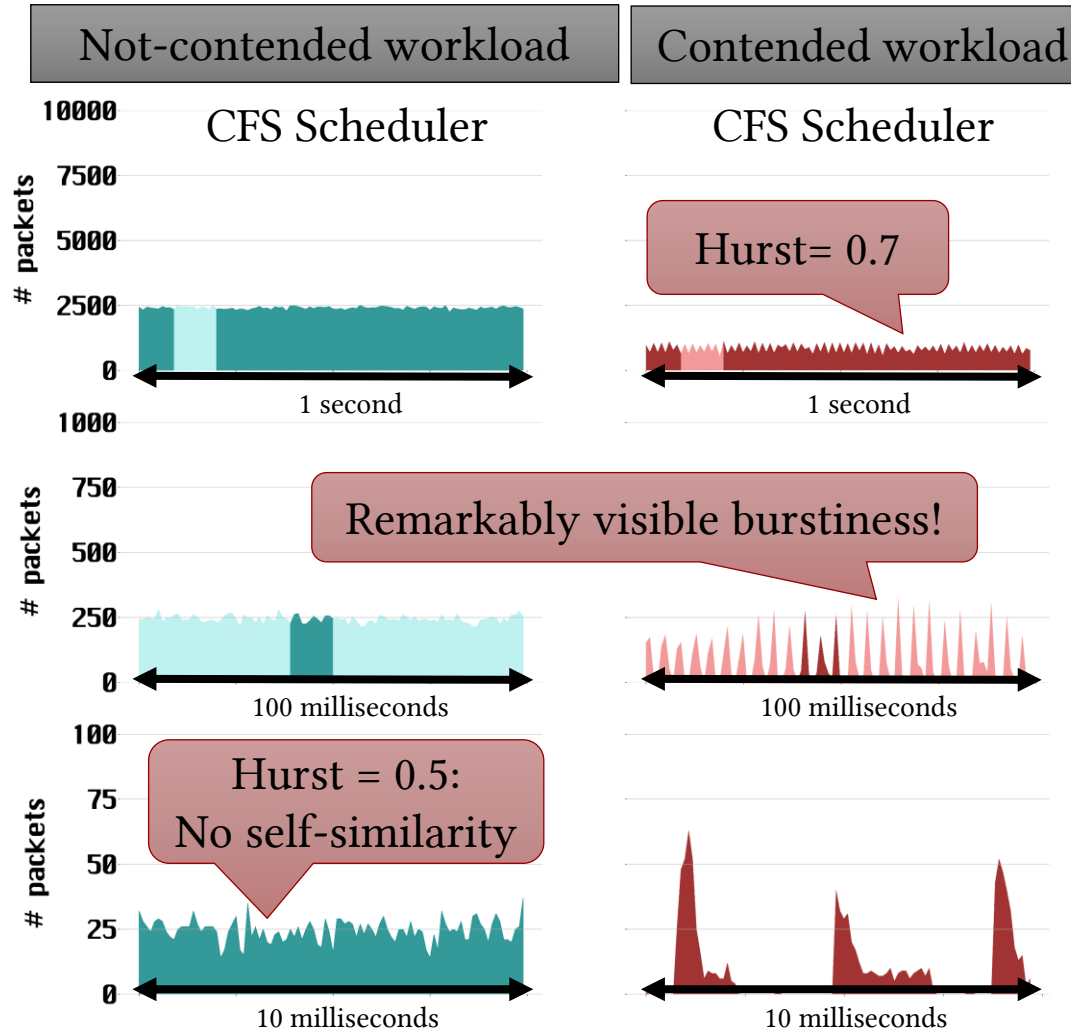


## CFS

- Coarse-grained timeslicing
- Fairness

## Realtime FIFO

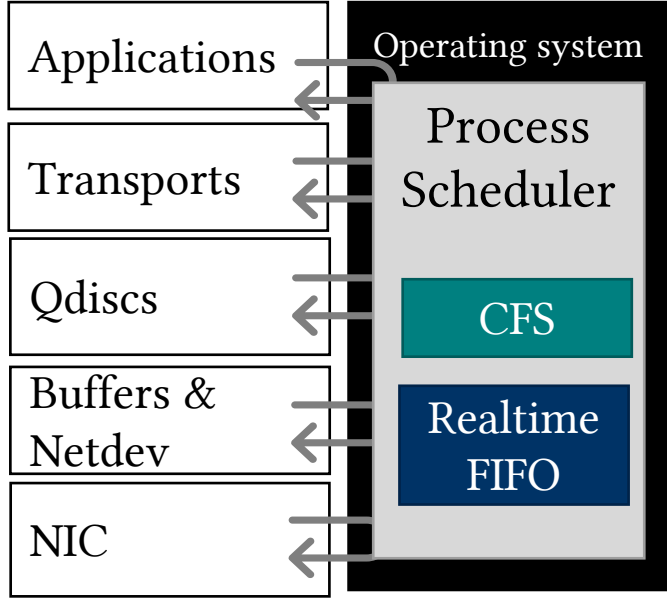
- Strict prioritization





# Even process scheduling matters!

Designing **burst-aware process schedulers**: A tradeoff between **latency**, support for **collocation**, and **burstiness**!

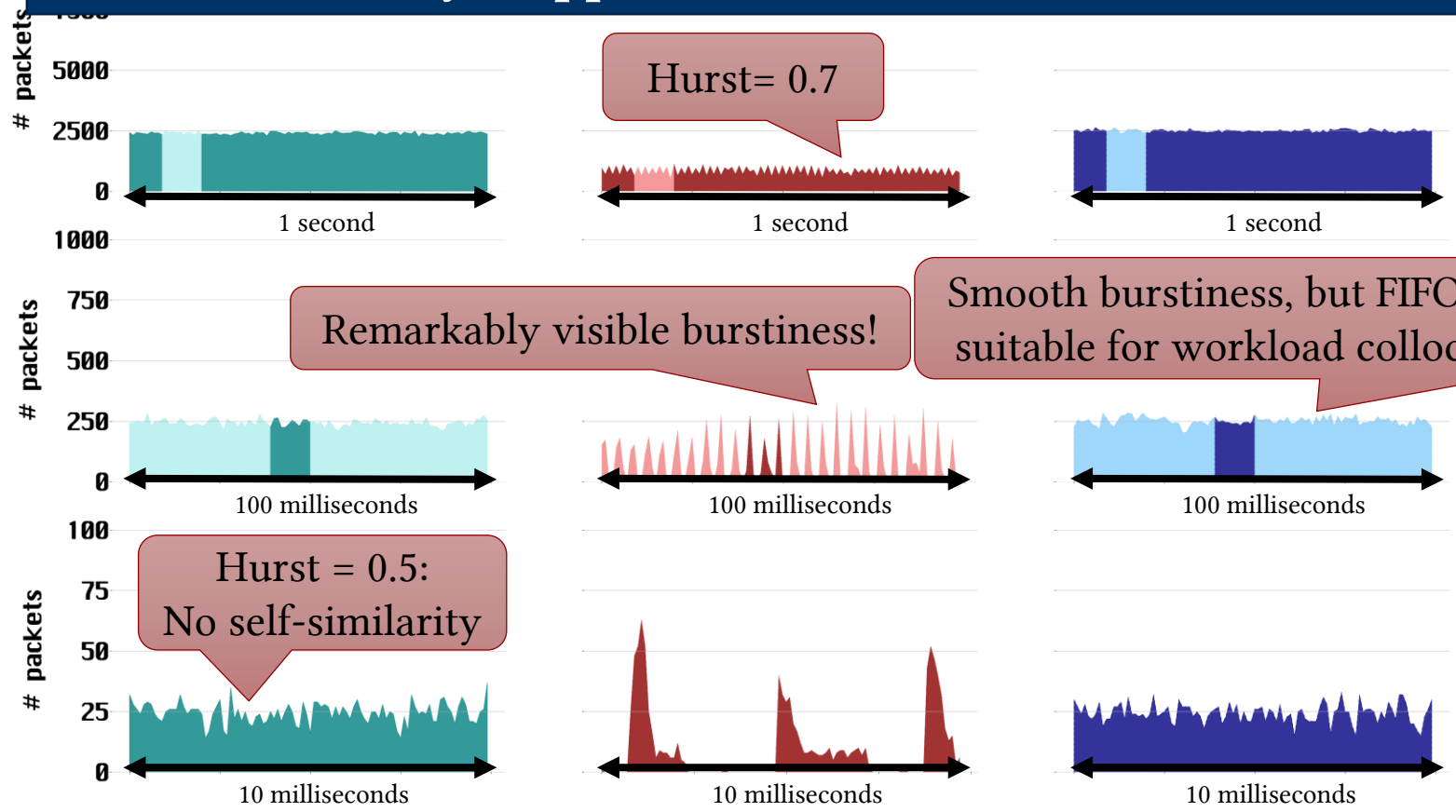


**CFS**

- Coarse-grained timeslicing
- Fairness

**Realtime FIFO**

- Strict prioritization



# Implications for network design

- Valinor measures the burstiness of individual network stack components.
- Lower layers of the stack **compromise** software shaping.
- Existing burst countermeasures in the software **are not effective!**
- Pacing and shaping must be **pushed down** the stack.
- Network stack layers must be **co-designed** with **burstiness** in mind.
- Visit <https://hopnets.github.io/valinor> for Valinor artifacts.