SIGCOMM 18 Topic Preview: Congestion Monitoring and Control

Thomas Zinner

Congestion Control

▶ An informal definition of congestion

> **Too many sources sending too much data too fast for network to handle**
>
> **Kurose / Ross**

▶ Impact of congestion on data transmission
- Increased transmission delays due to queueing
- Packet loss due to buffer overflows

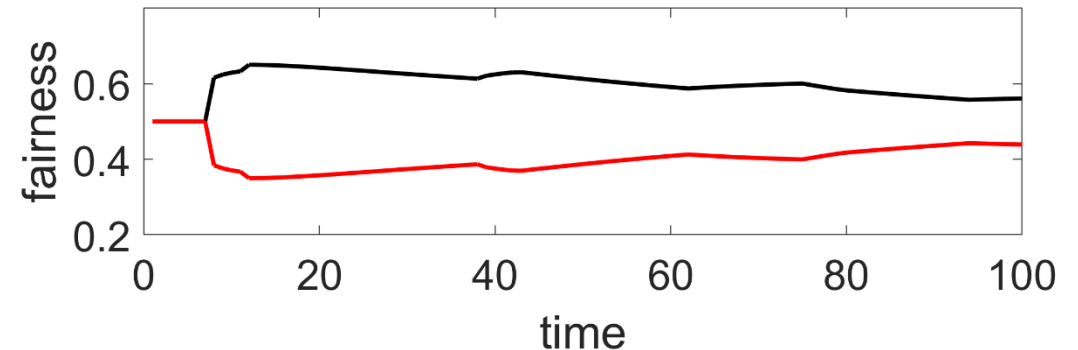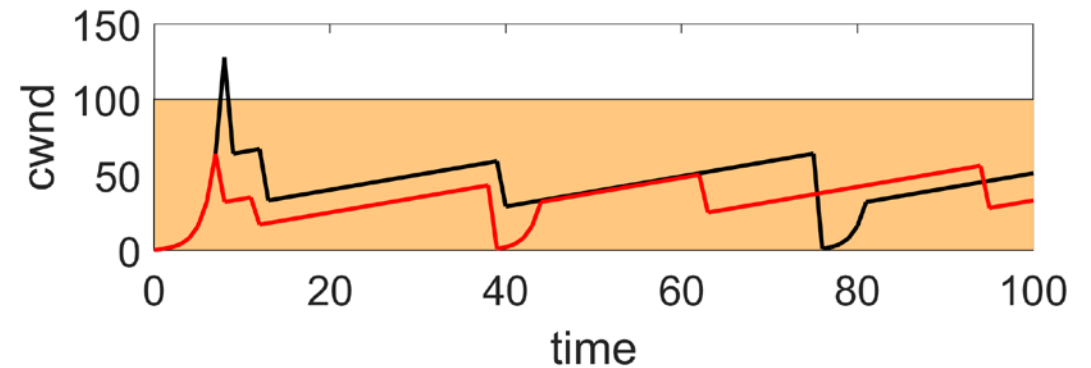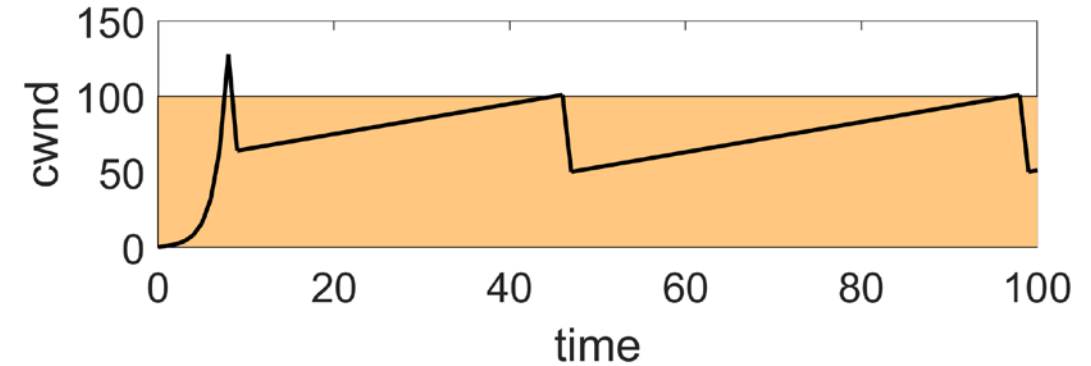▶ Congestion control approaches
- Network-assisted
- End-to-end → TCP

TCP-based Congestion Control
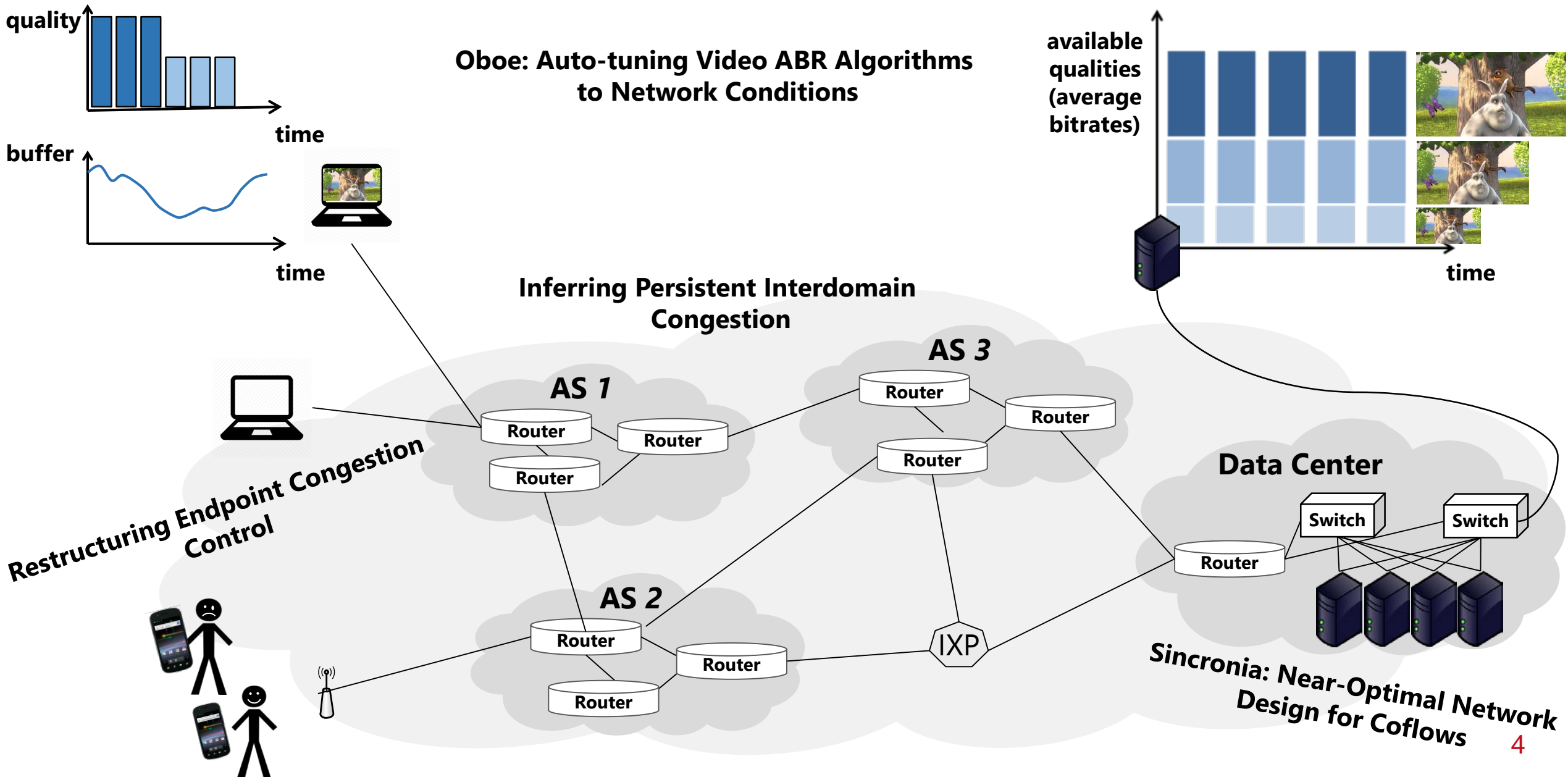
▶ **End-host based congestion control**
  - Network congestion inferred from observed packet loss and delays
  - Transmission rate adjusted w.r.t. congestion level in AIMD fashion

▶ **Interaction of congestion control of multiple clients**

▶ **Fairness: Similar transmission rate per client in case of congestion**

Session Overview

quality

buffer

time

time

Oboe: Auto-tuning Video ABR Algorithms to Network Conditions

available qualities (average bitrates)

time

Inferring Persistent Interdomain Congestion

AS 3

Router

Router

AS 1

Router

Router

Router

Router

Data Center

Restructuring Endpoint Congestion Control

Router

Switch

Switch

Router

AS 2

Router

Router

Router

IXP

Sincronia: Near-Optimal Network Design for Coflows

4

Inferring Persistent Interdomain Congestion

YouTube sucks on French ISP Free, and French regulators want to know why

Robert Andrews & Stacey Higginbotham Jan 2, 2013 - 1:29 PM CDT

https://gigaom.com/2013/01/02/youtube-sucks-on-french-isp-free-french-regulators-want-to-know-why/

verizon✓

06.19.2013 | Policy

Unbalanced Peering and the Whole Story Behind the Verizon/Cogent Dispute

https://www.verizon.com/about/news/unbalanced-peering-and-whole-story-behind-verizoncogent-dispute

HD COMES IN SUPER NOW? —

Time Warner, net neutrality foes cry foul over Netflix Super HD demands

Super HD, 3D content only available to subs whose ISPs play nicely with Netflix.

https://arstechnica.com/information-technology/2013/01/timewarner-net-neutrality-foes-cry-foul-netflix-requirements-for-super-hd/

BIZ & IT —

Why YouTube buffers: The secret deals that make—and break—online video

When ISPs and video providers fight over money, Internet users suffer.

https://arstechnica.com/information-technology/2013/07/why-youtube-buffers-the-secret-deals-that-make-and-break-online-video/

▶ Large number of peering disputes, however, the causes of congestion and poor performance remain unclear

▶ Identification of the bottleneck important, e.g., for regulators

▶ No lightweight toolset available yet

How to Infer Interdomain Congestion

▶ Geographically distributed vantage points (VPs) in access networks
   → CAIDA's ark measurement infrastructure

▶ Identification of interdomain links via VPs
   → bdrmap tool (based on traceroute, BGP data and accepted assumptions)

▶ Inference of congestion via latency measurements
   → Time-Series Latency Probing (TSLP)
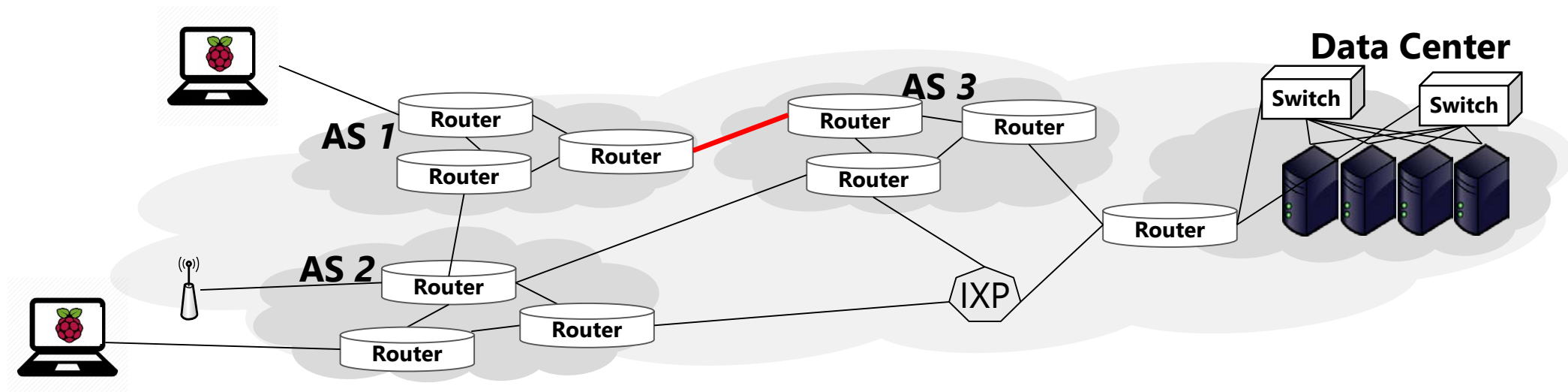
Contribution of the Paper

▶ Design and implementation of a system for congestion measurements of a large number of interdomain links from a set of vantage points

▶ Validation of congestion interferences

▶ Results of a longitudinal measurement study of interdomain congestion in US broadband access networks

▶ Analysis scripts and underlying dataset are made publicly available
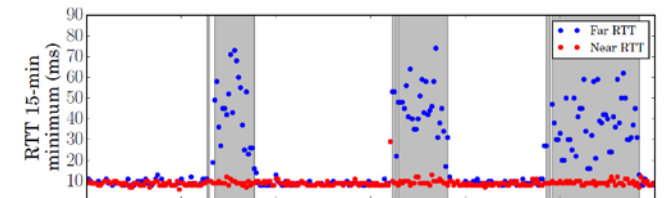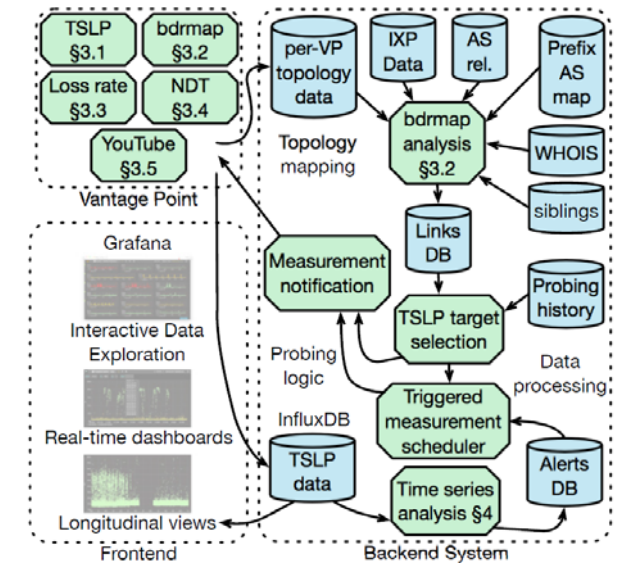

A. Dhamdhere et al., "Inferring Persistent Interdomain Congestion"

Sincronia: Near-Optimal Network Design for Coflows

▶ Network design typically optimized for latency/throughput for single flows, e.g., typical web applications
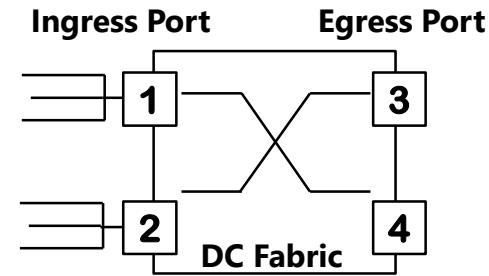
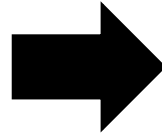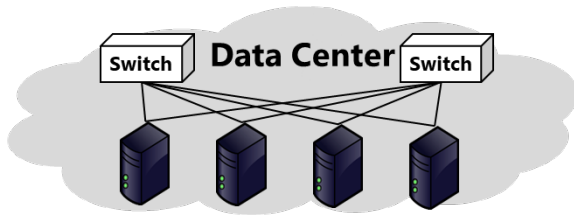- Example: request scheduling instead of fair sharing
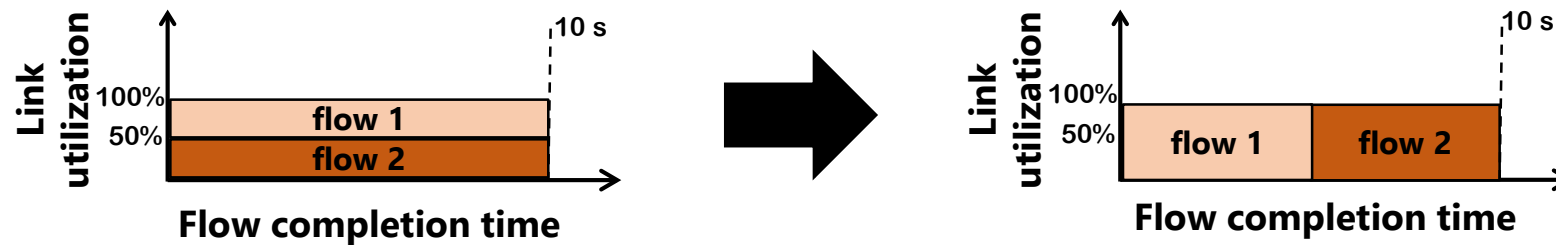
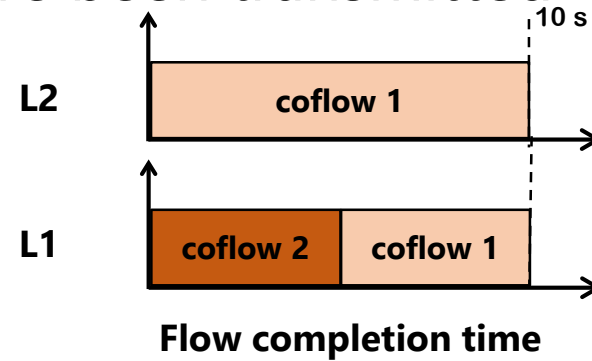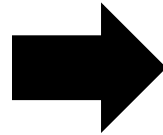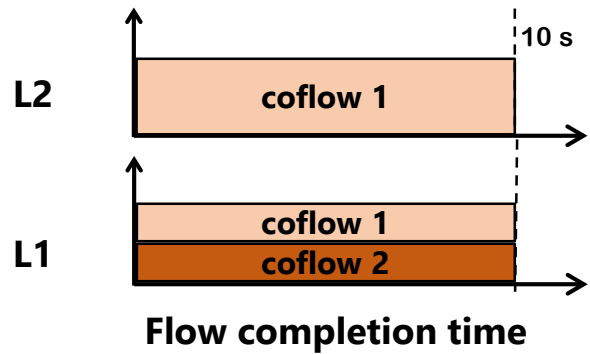▶ Distributed datacenter applications (e.g., MapReduce) require a collection of flows (coflows) to be transmitted
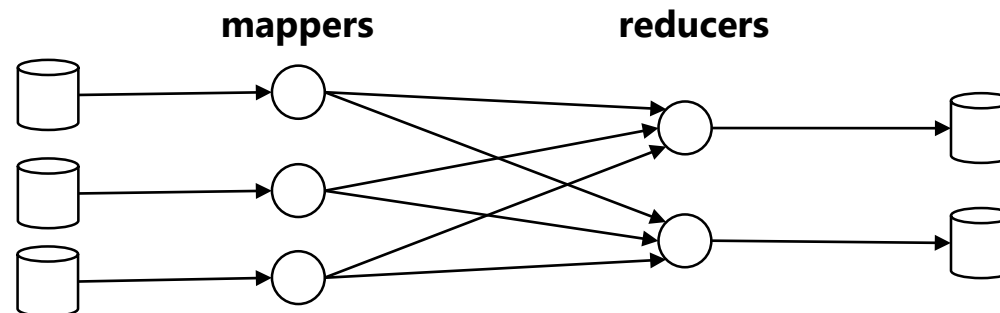
Coflow Scheduling in Data Centers

▶ Datacenter network fabric, e.g., fat-tree or leaf-spline, is very complex
→ Abstraction as one big switch with virtual output queues



▶ Job finishes after all flows of a coflow have been transmitted

▶ Drawbacks of current solutions

▪ Approaches utilize per flow rate allocation and heuristics for optimization
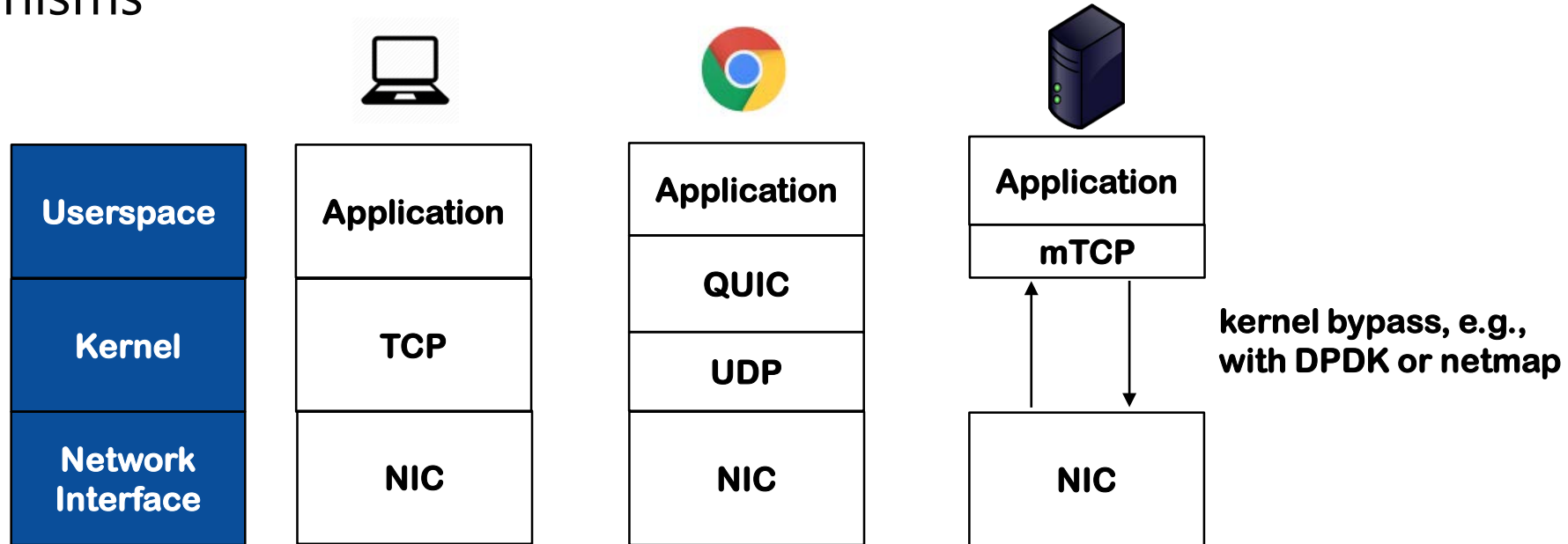→ Missing scalability and optimality bounds

Contribution of the Paper

▶ Presentation of a new datacenter network design that achieves near-optimal average coflow completion time

▶ Sincronia overcomes current drawbacks by
  ▪ Avoiding per-flow rate allocation and rate reallocation after arrivals or departures
  ▪ Utilizing simple prioritization mechanisms
  ▪ Being transport-agnostic

▶ Implementation and evaluation of Sincronia
  ▪ TCP-based with DiffServ for priority scheduling
  ▪ Fat-tree architecture
  ▪ Comparison with state-of-the-art
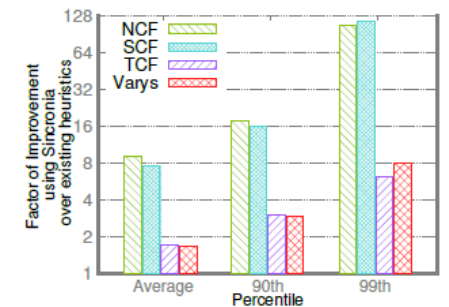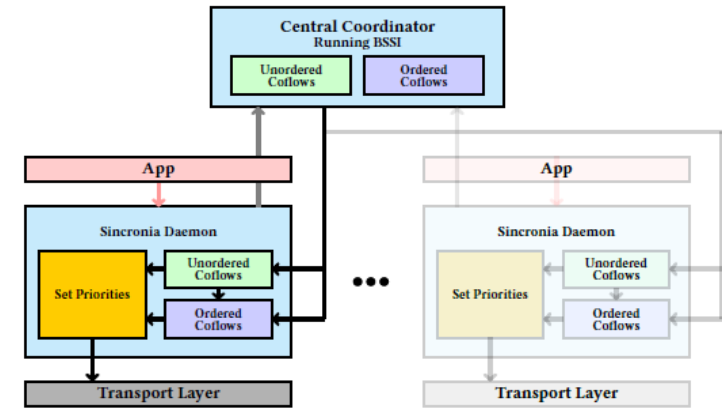
S. Agarwal et al., "Sincronia: Near-Optimal Network Design for Coflows"

Restructuring Endpoint Congestion Control

▶ Increasing number of different datapaths using congestion control (CC) mechanisms

Userspace	Application	Application	Application
			mTCP
Kernel	TCP	QUIC	
		UDP	
Network Interface	NIC	NIC	NIC

kernel bypass, e.g., with DPDK or netmap

▶ Shortcomings

- New CC mechanisms difficult to implement for specific datapaths
- Limited code reusability due to tight interweaving with the datapath
- Hard to pass congestion information between flows

Off-Datapath Congestion Control Plane

▶ Datapath encapsulates congestion signals and periodically provides this data to an off-datapath module
→ Congestion control algorithms can run outside the datapath

▶ Control relevant transmission parameters such as window size, pacing rate and transmission pattern via an interface for each datapath
→ Data transmitted according to congestion control policy

▶ Benefits of an off-datapath congestion control plane

- Write-one, run-anywhere

- Higher pace of development

- New capabilities

Contribution of the Paper

▶ Event-driven language to specify congestion control algorithms

- Specification of congestion control using combinations of events and conditions
- Simple computations performed directly in the datapath
- Complex logic deferred to a user-space component

▶ Specification of datapath responsibilities

- Congestion signals a datapath should maintain
- Simple framework to execute directives from a CCP program

▶ Evaluation of the fidelity of CCP

- Comparison with in-kernel implementations
- Investigation under a variety of link conditions
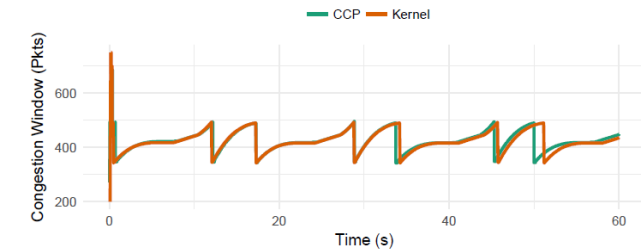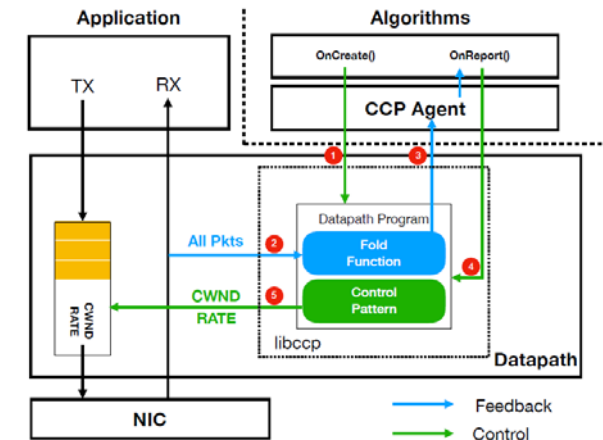
**A. Narayan et al., "Restructuring Endpoint Congestion Control"**

Oboe: Auto-tuning Video ABR Algorithms to Network Conditions

▶ Adaptive bitrate (ABR) algorithms for video streaming

- Client requests small chunks via HTTP
- Client selects the quality of next segment to download selected based on the client's current state (e.g., video buffer, video quality, size of the next segment, estimated download time,...)
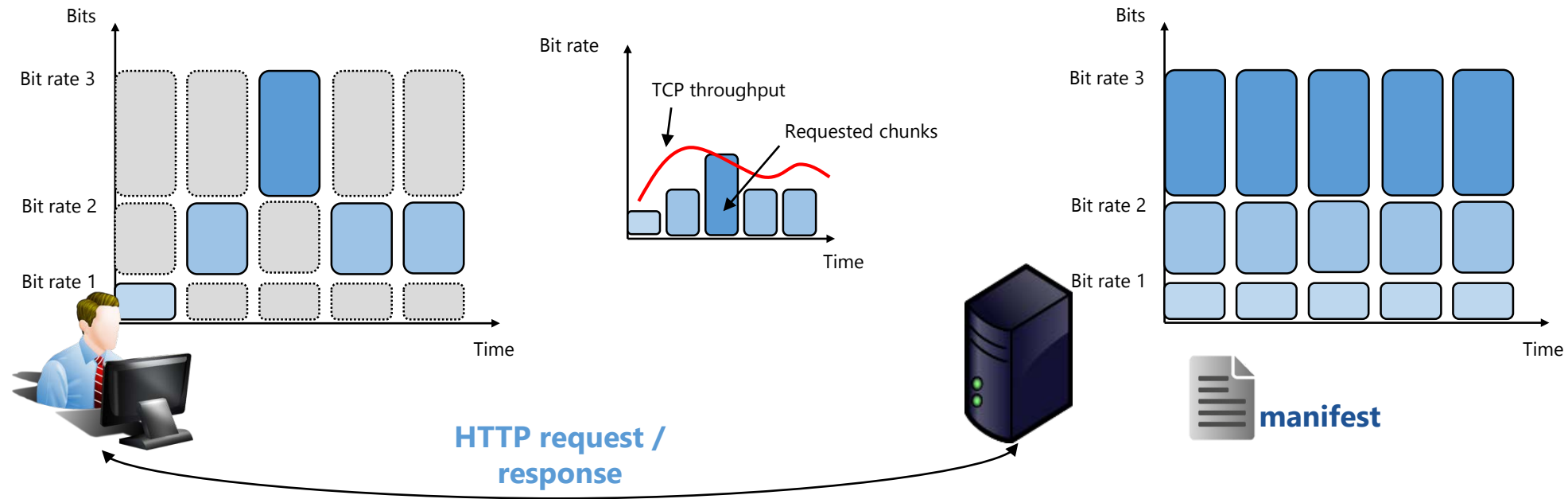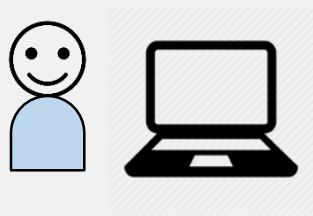
QoE Impact Fators of Adaptive Video Streaming
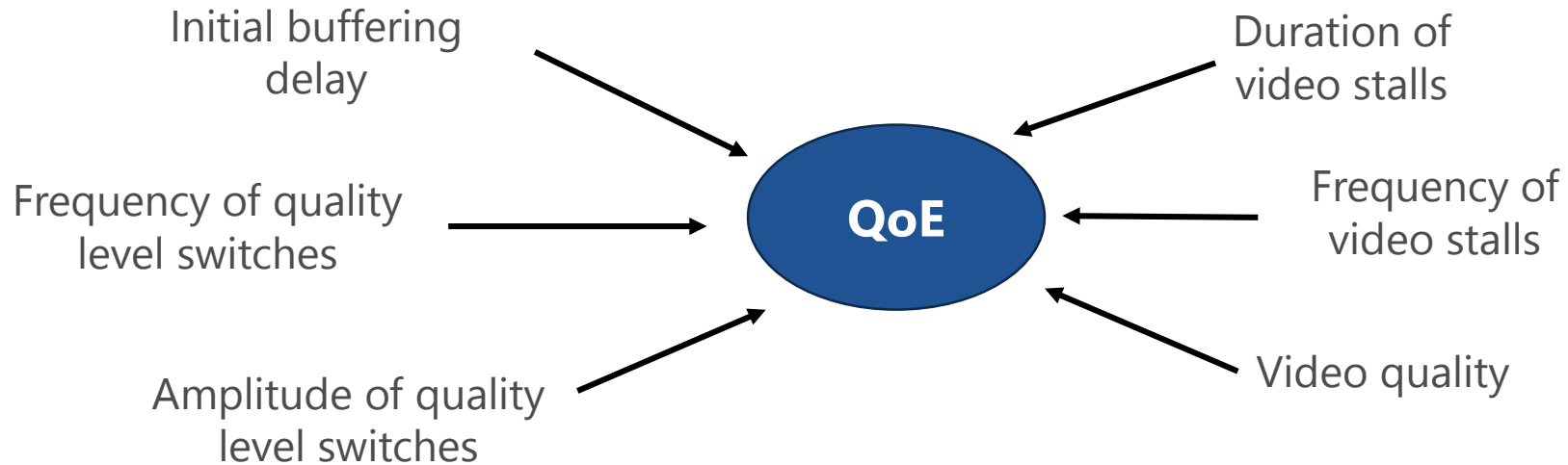
▶ Influence of impact factors on QoE

- Video stalls and amplitude of quality switches typically have significant impact
- Impact of video quality depends on specific context
- Frequency of quality switches and initial buffering delay have minor impact

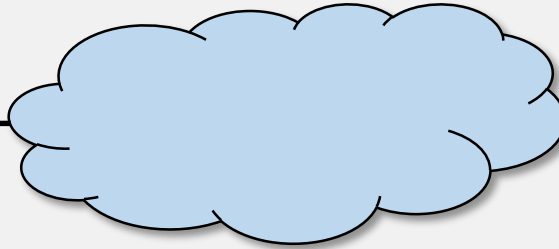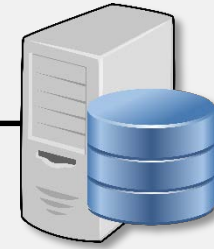# QoE Influence Factors on Adaptive Video Streaming

DASH Client

- Adaptive bitrate (ABR) algorithms
- Player parameters (e.g. Video buffer size)

Network

- Type of access network
- QoS parameters
- Variation of QoS parameters

Server

- Segment Duration
- Number of quality representations
- Bitrate of quality representations
- Video encoding

Adaptive Bitrate Algorithms

▶ Many different adaptive bitrate algorithms have been proposed

- Buffer- and throughput-based approaches vs. purely buffer-based approaches
- Using control theory, queueing theory, machine learning...
- ABR algorithms typically have parameters which can be used to tune the algorithm

▶ ABR examples

- RobusMPC: Throughput prediction and buffer occupancy. Reduces predicted throughput by discount factor 1+d
- Bola: Buffer occupancy. Parameter γ used to control rebuffering
- HYB: Throughput prediction and buffer occupancy. Parameter β used to tune the aggressiveness

▶ ABR algorithms perform well on average, but some users experience bad performance.

Contribution of the Paper

▶ Improve ABR algorithms to better perform across a range of network conditions by automatically tuning ABR behavior

  - Modelling of variable network states as piecewise-stationary sequence

  - Pre-computing the best parameter configuration for a specific ABR algorithm in an offline manner

▶ System implementation and evaluation

  - Comparison with state-of-the-art ABR algorithms

  - Further comparison with Pensive (Sigcomm 2017)
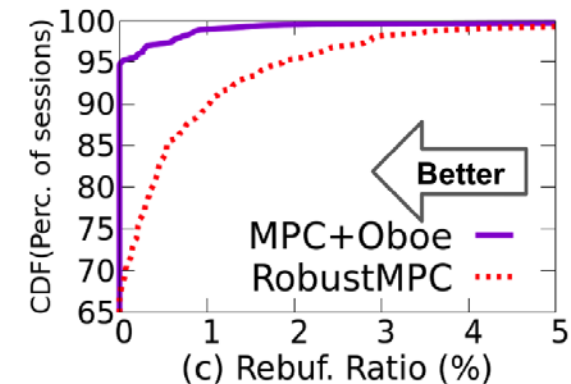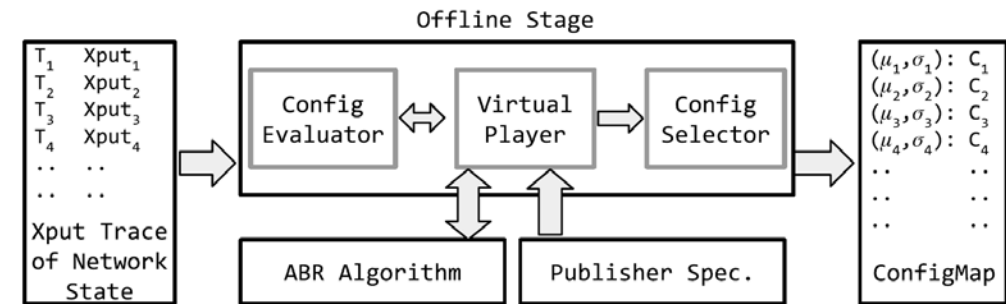
Z. Akhtar et al., "Oboe: Auto-tuning Video ABR Algorithms to Network Conditions"

11:00 am - 12:40 pm Main-Conference Session 1: Congestion Monitoring and Control

Session Chair: Brad Karp *(UCL, UK)*
Location: Vigadó, 2nd-Floor Ceremonial Hall

11:00 am - 11:25 am	**Inferring Persistent Interdomain Congestion** Amogh Dhamdhere *(CAIDA, USA)*, David D. Clark *(MIT, USA)*, Alexander Gamero-Garrido *(CAIDA, USA)*, Matthew Luckie *(Waikato, New Zealand)*, Ricky K. P. Mok, Gautam Akiwate, Kabir Gogia *(CAIDA, USA)*, Vaibhav Bajpai *(TU Munich, Germany)*, Alex C. Snoeren *(UCSD, USA)*, kc claffy *(CAIDA, USA)*
11:25 am - 11:50 am	**Sincronia: Near-Optimal Network Design for Coflows** Saksham Agarwal, Shijin Rajakrishnan *(Cornell, USA)*, Akshay Narayan *(MIT, USA)*, Rachit Agarwal, David Shmoys *(Cornell, USA)*, Amin Vahdat *(Google, USA)*
11:50 am - 12:15 pm	**Restructuring Endpoint Congestion Control** Akshay Narayan, Frank Cangialosi, Deepti Raghavan, Prateesh Goyal, Srinivas Narayan *(MIT, USA)*, Radhika Mittal *(Berkeley, USA)*, Mohammad Alizadeh, Hari Balakrishnan *(MIT, USA)*
12:15 pm - 12:40 pm	**Oboe: Auto-tuning Video ABR Algorithms to Network Conditions** Zahaib Akhtar *(USC, USA)*, Yun Seong Nam *(Purdue, USA)*, Ramesh Govindan *(USC, USA)*, Sanjay Rao *(Purdue, USA)*, Jessica Chen *(UWindsor, Canada)*, Ethan Katz Bassett *(Columbia, USA)*, Bruno Martins Ribeiro *(Purdue, USA)*, Jibin Zhan, Hui Zhang *(Conviva, USA)*