Swift: Adaptive Video Streaming with Layered Neural Codecs

Mallesham Dasari, Kumara Kahatapitiya,

Samir R. Das, Aruna Balasubramanian, Dimitris Samaras

NSDI 2022

Stony Brook University

Carnegie Mellon University

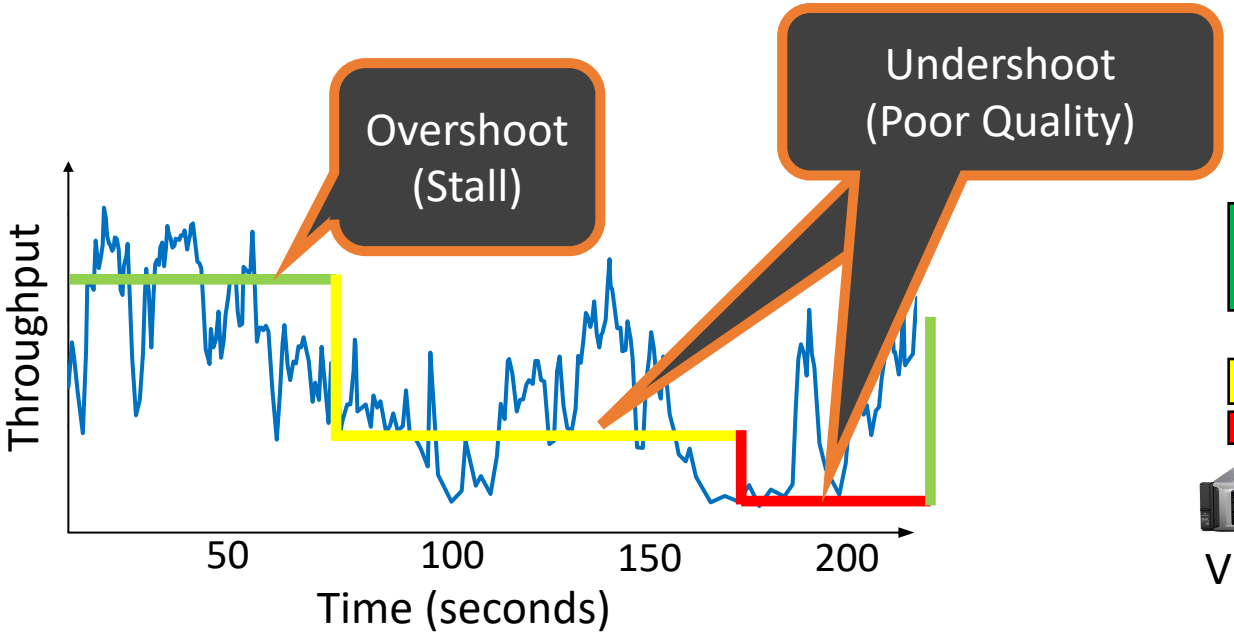
Outline

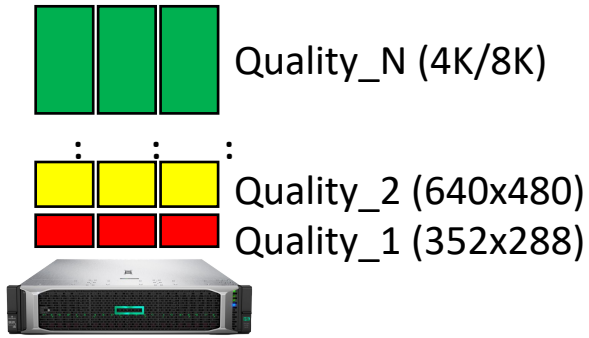
- ❑ Background & Motivation
  - ❑ Adaptive Video Streaming & Today's Practice
  - ❑ Existing Alternatives: Layered Streaming & Challenges
- ❑ Swift
  - ❑ Layered Neural Video Coding
  - ❑ Encoder, Decoder, Streaming Protocol
  - ❑ Evaluation & Comparison with Existing Work
- ❑ Conclusions

Video Streaming Today

Fundamental problems: 1) Limited throughput, 2) Variability

Video Client



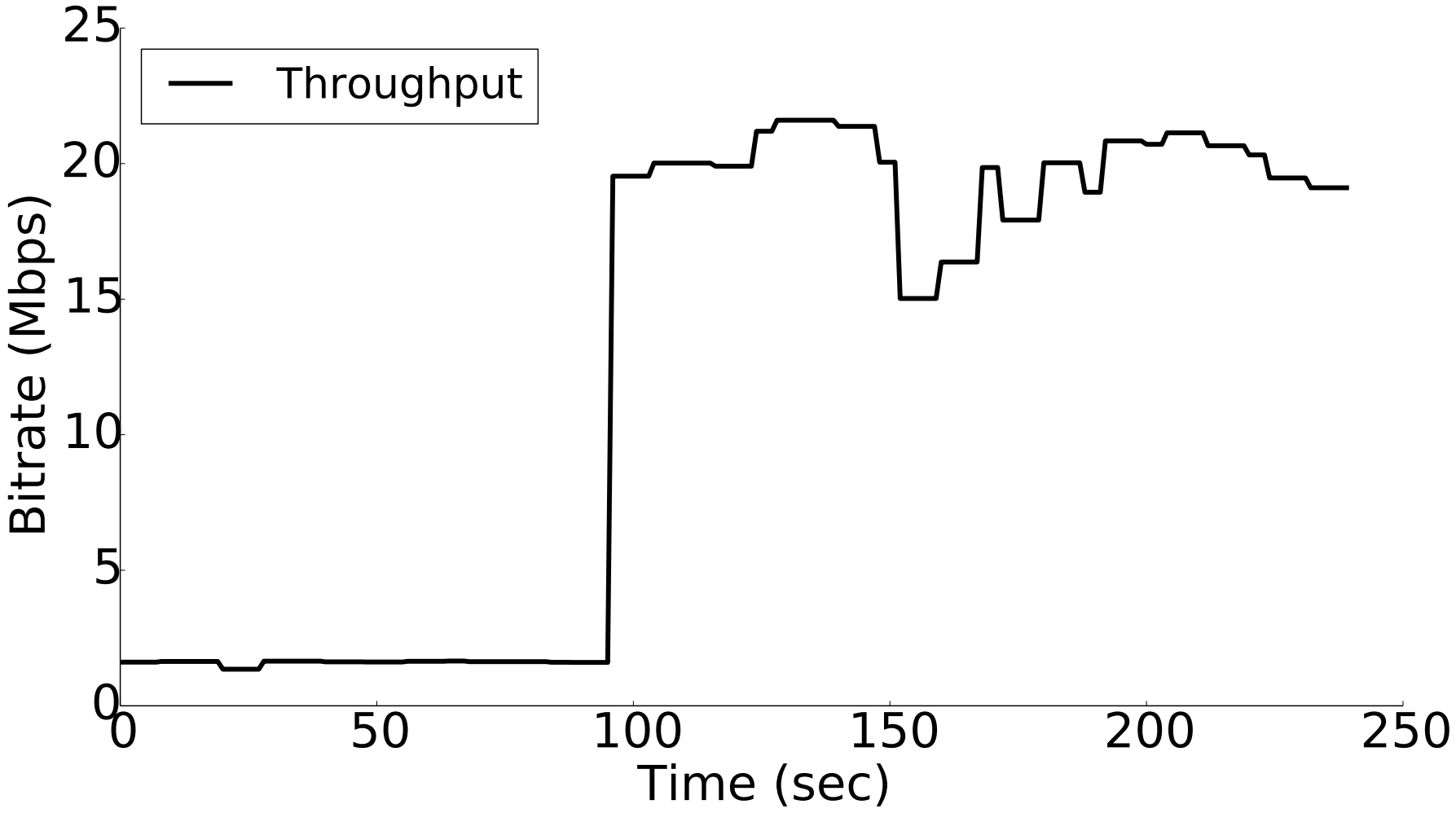


Video Server

Bitrate Selection: A Decade-long Research!

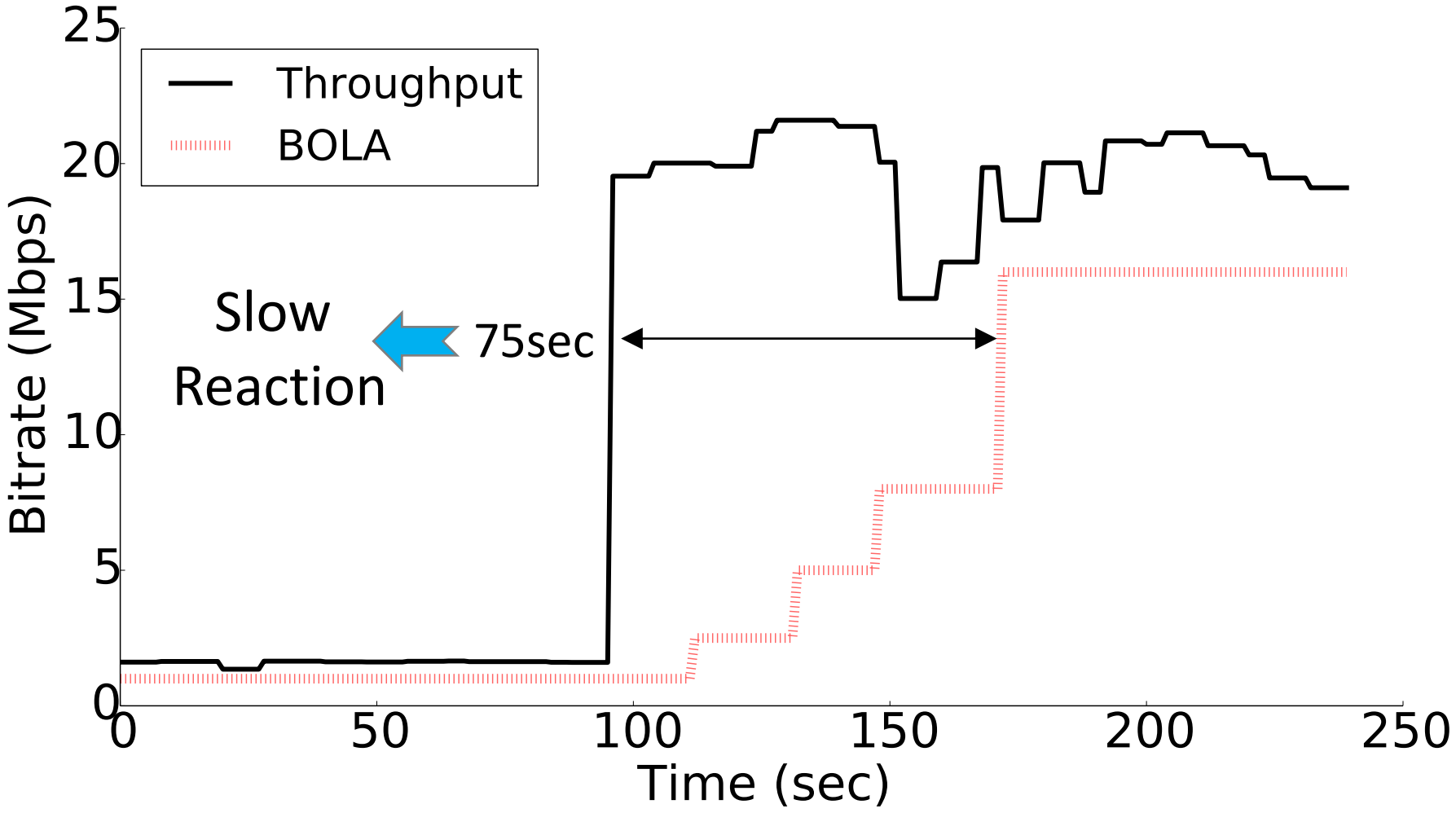
(e.g., Festive [CoNEXT'12], MPC [SIGCOMM'15], BOLA [INFOCOM'16], Pensieve [SIGCOMM'17], Fugu [NSDI'20])

Limitations of Today's Streaming Algorithms

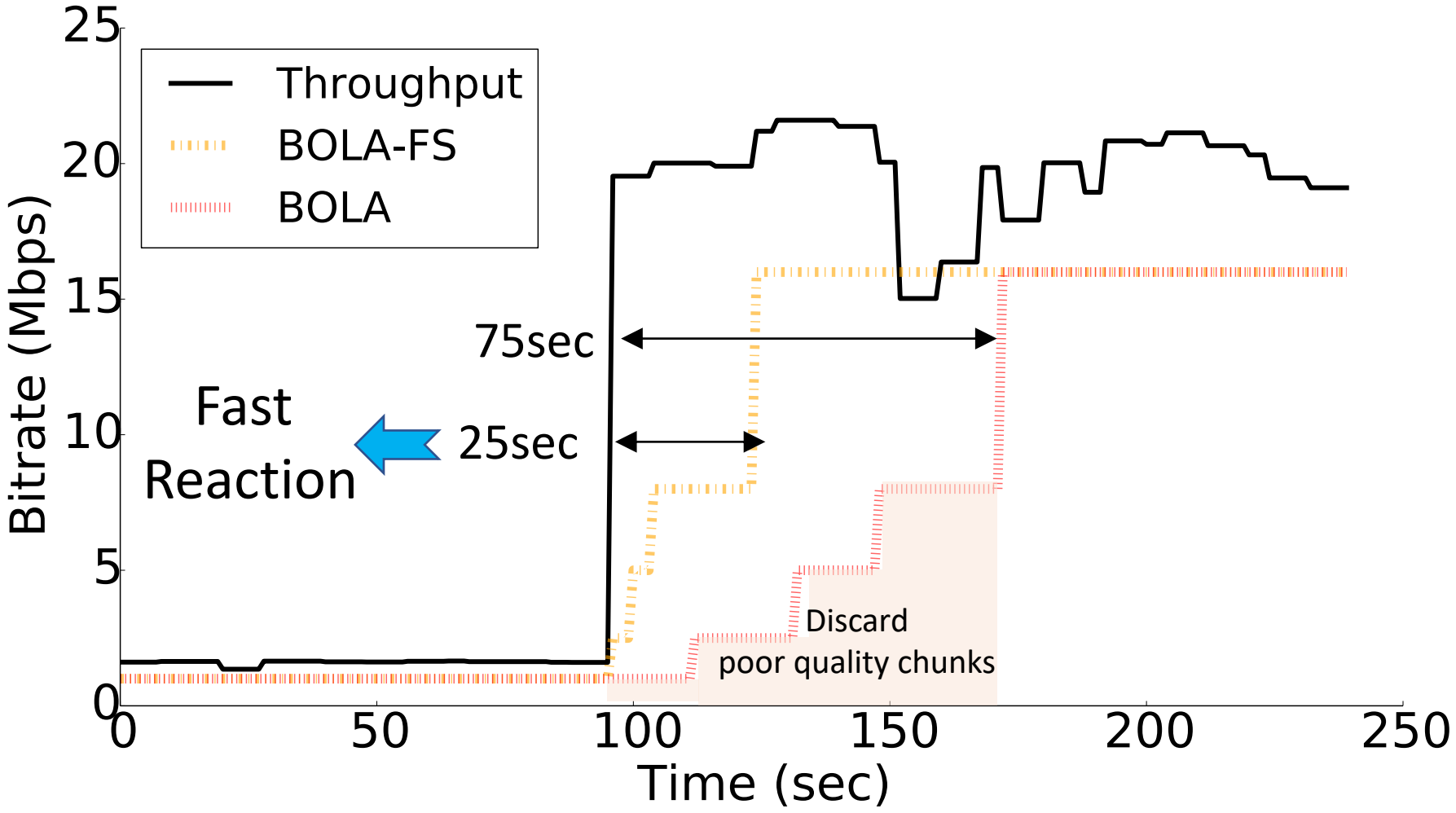


Throughput trace is from ACM/IEEE TON'2020

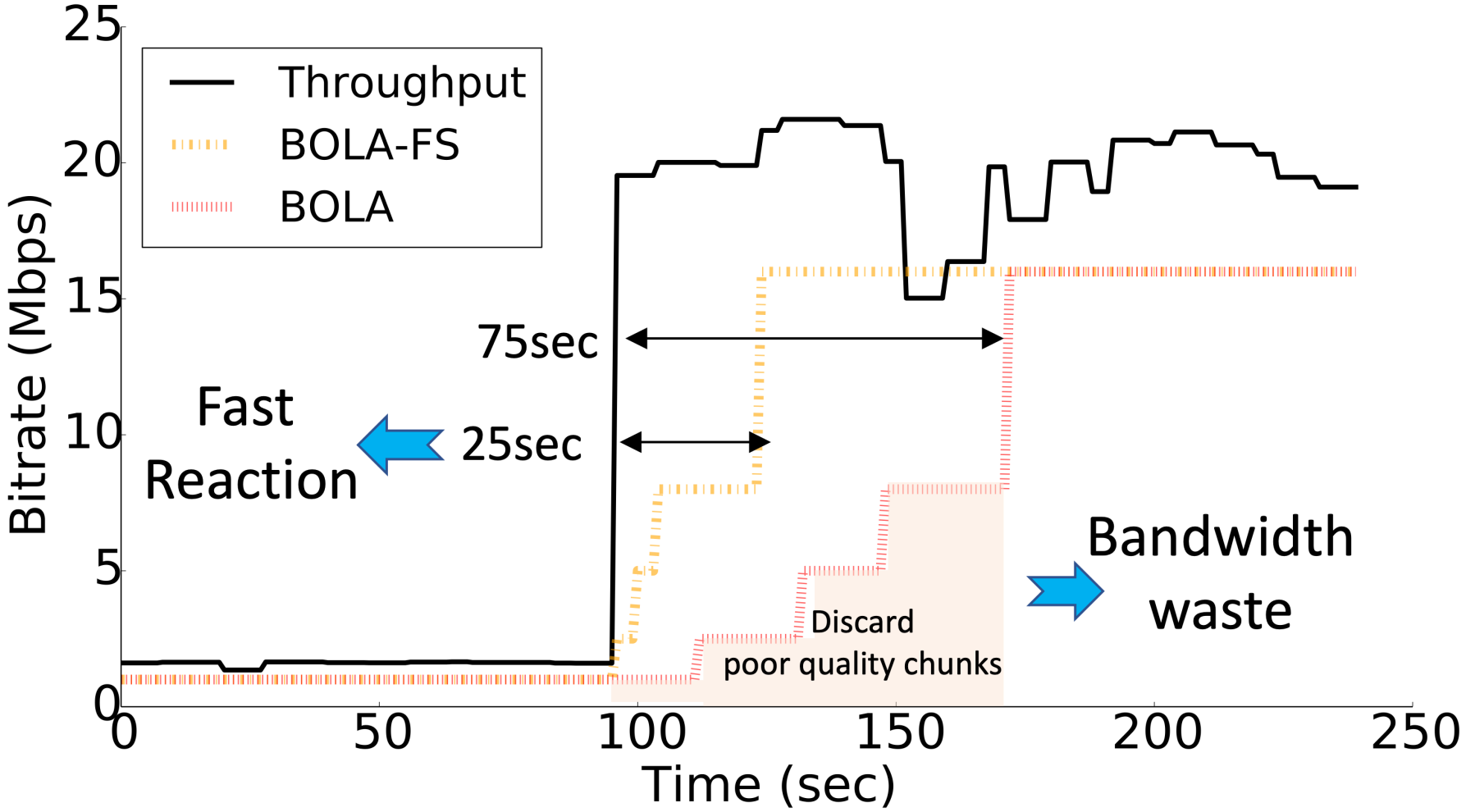
Limitations of Today's Streaming Algorithms



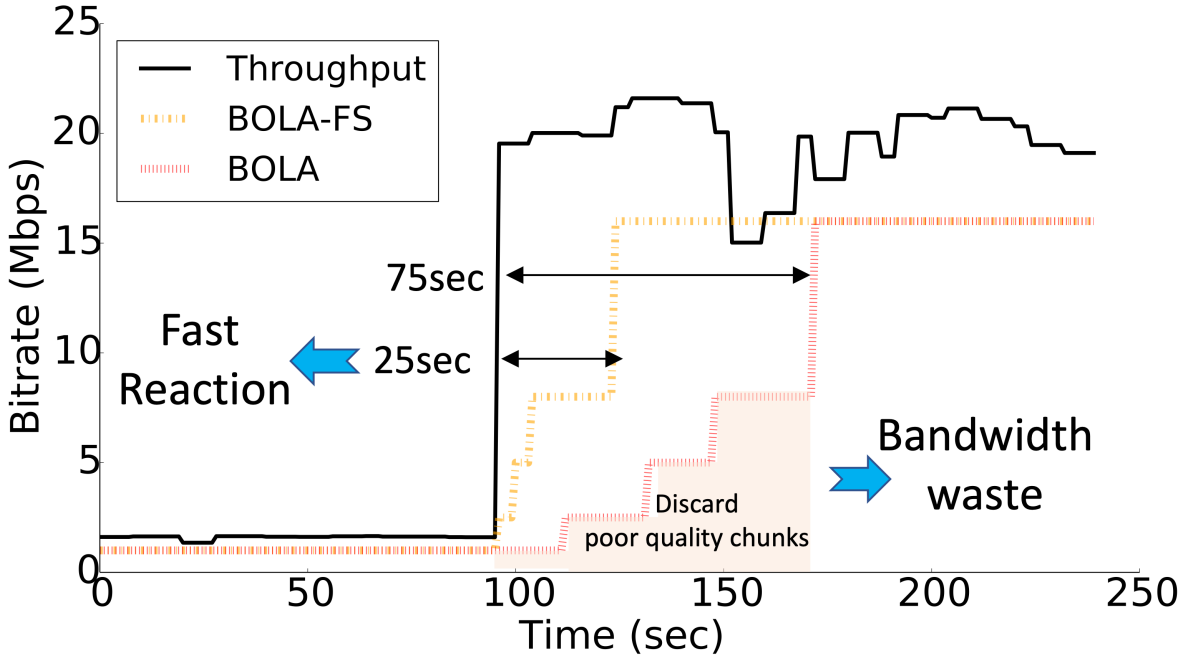
Limitations of Today's Streaming Algorithms



Limitations of Today's Streaming Algorithms



Limitations of Today's Streaming Algorithms

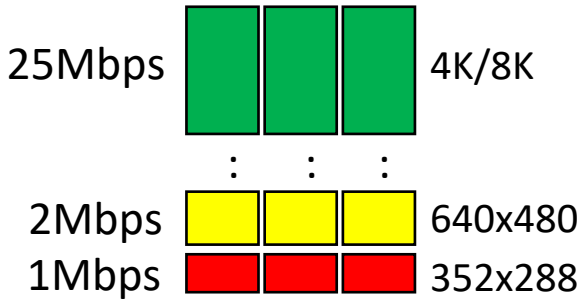


- ❑ Bandwidth Efficient
- ❑ Slow Reaction -> Poor QoE
  - ❑ BOLA, Pensieve, MPC

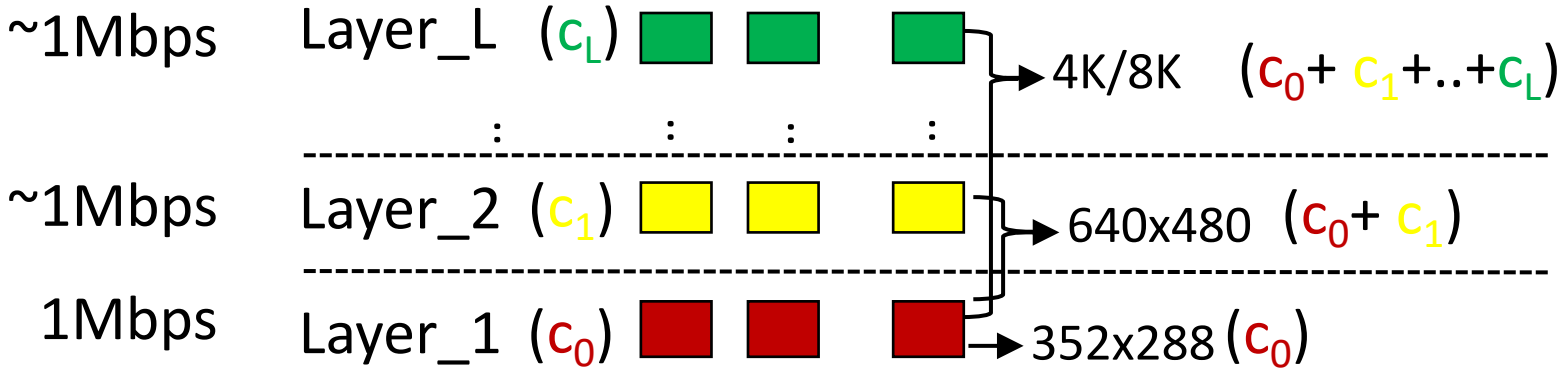
- ❑ Bandwidth Inefficient
- ❑ Fast Reaction -> Better QoE
  - ❑ BOLA-FS

Existing Alternative: Layered Video Streaming

- ❑ Layered Coding
- ❑ A well-suited technique for streaming under variable network conditions



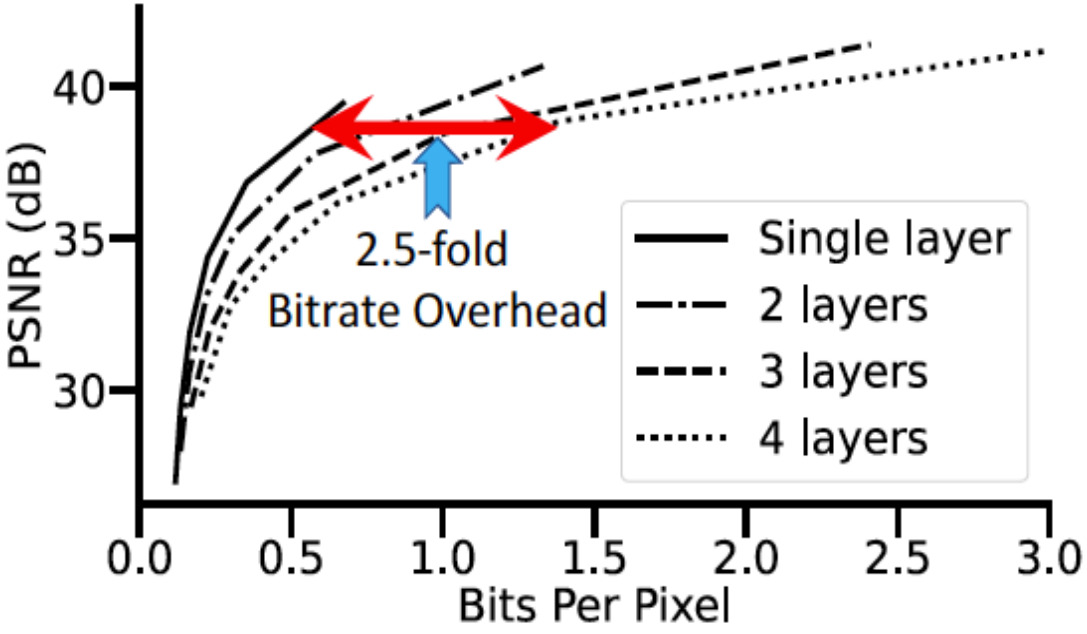
Classical/Regular Coding



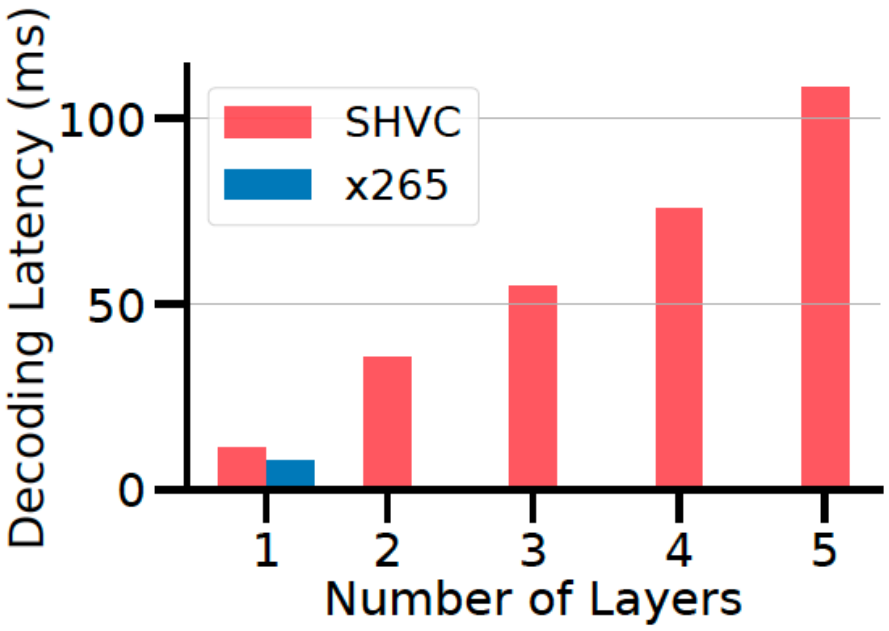
Existing Alternative: Layered Video Streaming

Challenges of using traditional layered coding in practice

Cross Layer Compression Overhead



High Coding Latency

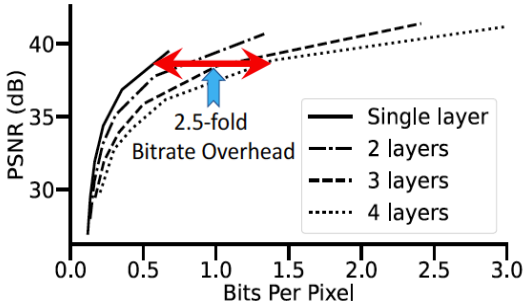


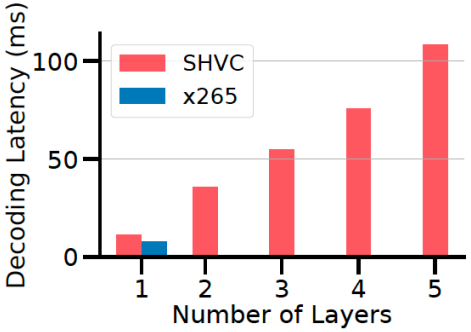
Existing Alternative: Layered Video Streaming

Challenges of using traditional layered coding in practice

Cross Layer Compression Overhead

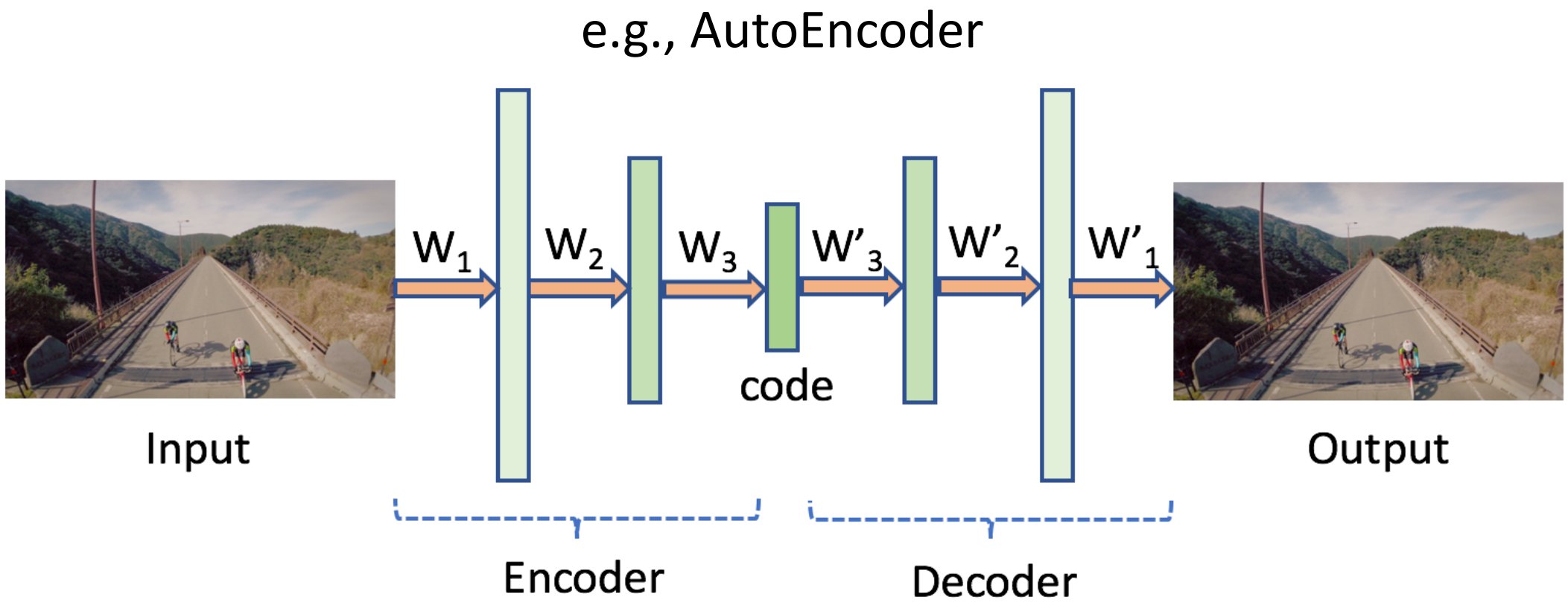
High Coding Latency





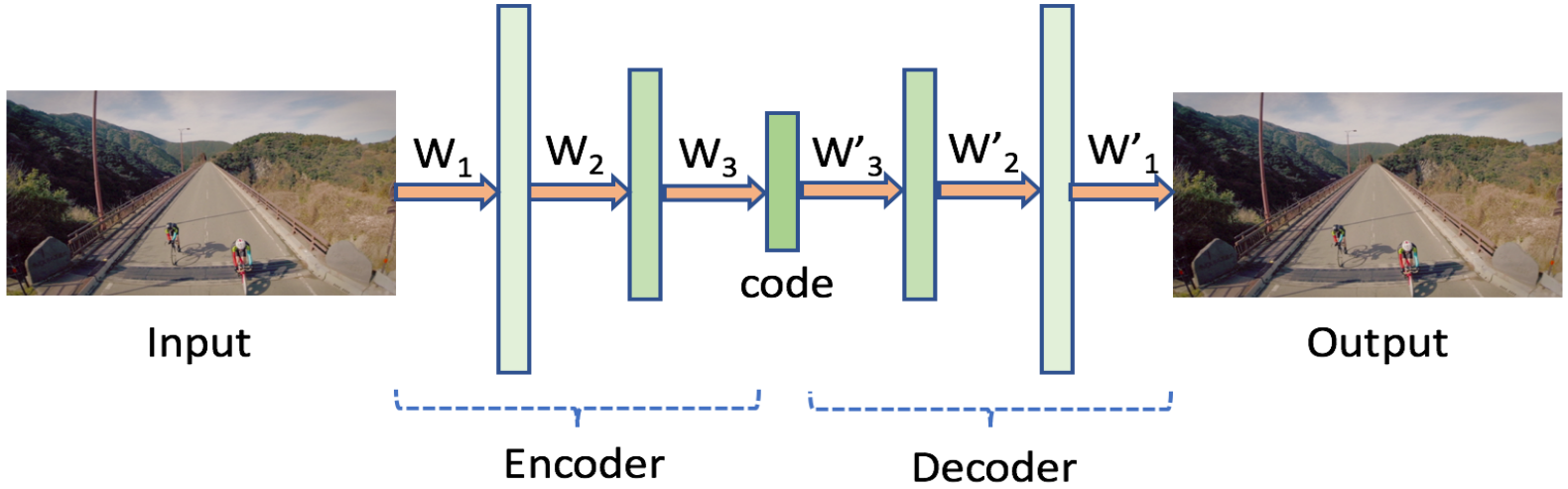
Remains an open problem & Not used in practice!

Swift: Layered Coding Using Neural Codecs



Swift: Layered Coding Using Neural Codecs

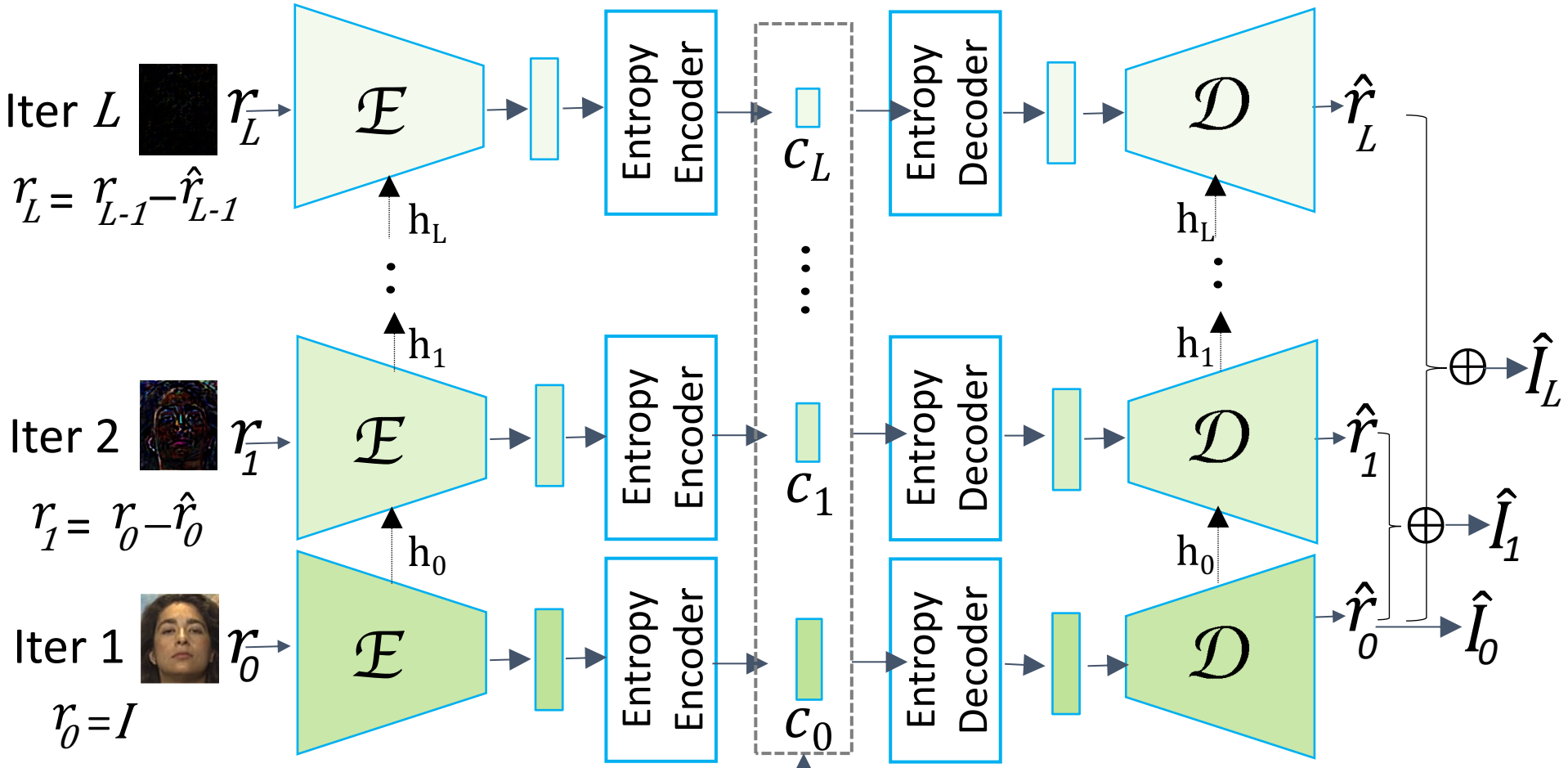
e.g., AutoEncoder



Swift's Components: 1) Encoder, 2) Decoder, 3) Streamer

1) Layered Neural Video Encoder

Residual codes (Todericci et al [CVPR'16], VCII [ECCV'18], HLVC [CVPR'19])



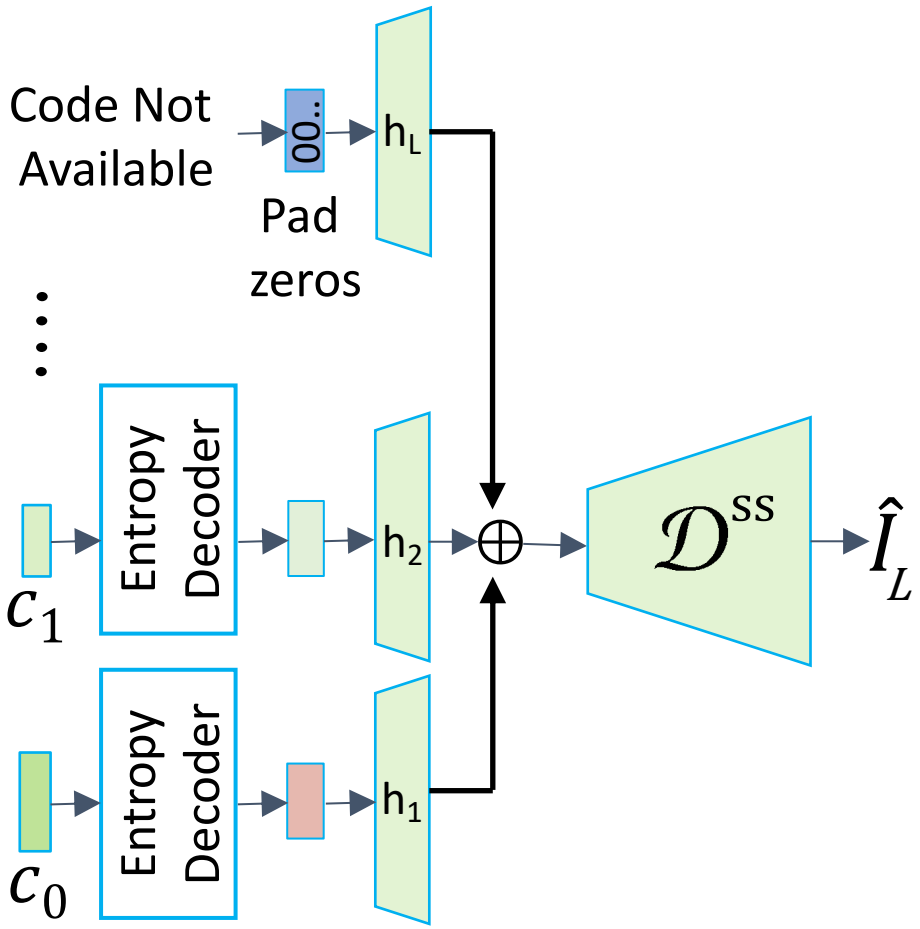
Layered codes transmitted over the Internet

2) Layered Neural Video Decoder

Challenge1. Iterative Decoding is Slow

- ❑Need real-time decoding for playback

Single-shot Decoder

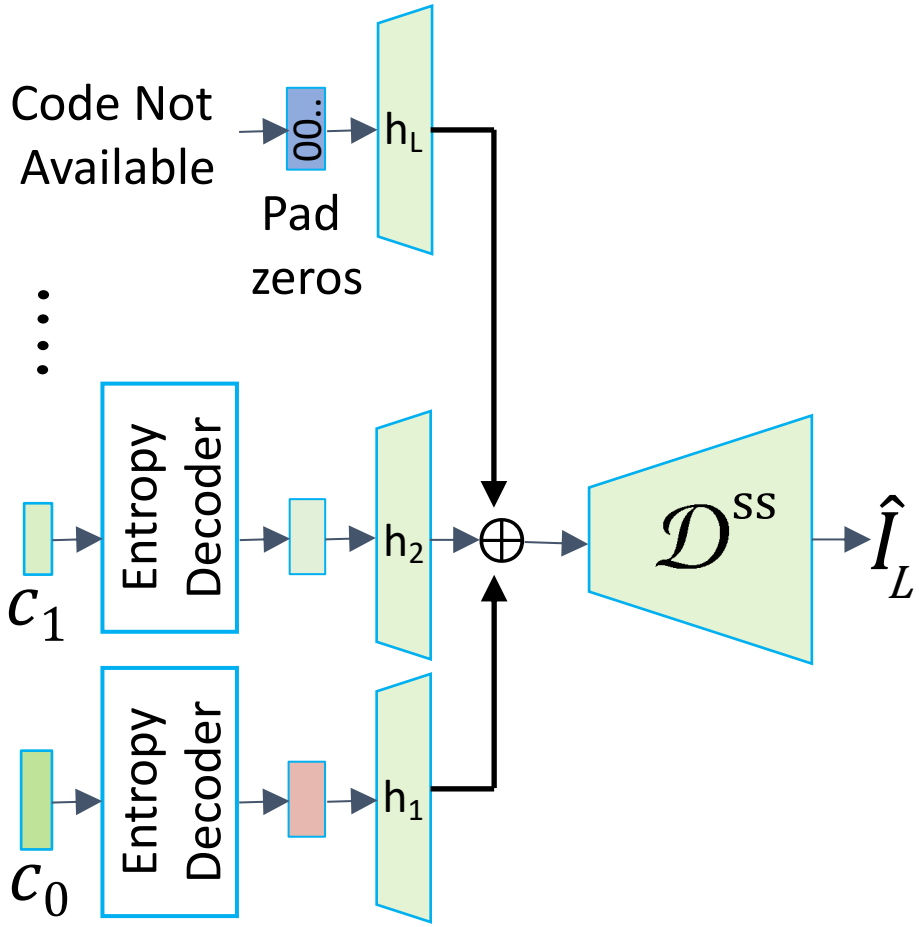


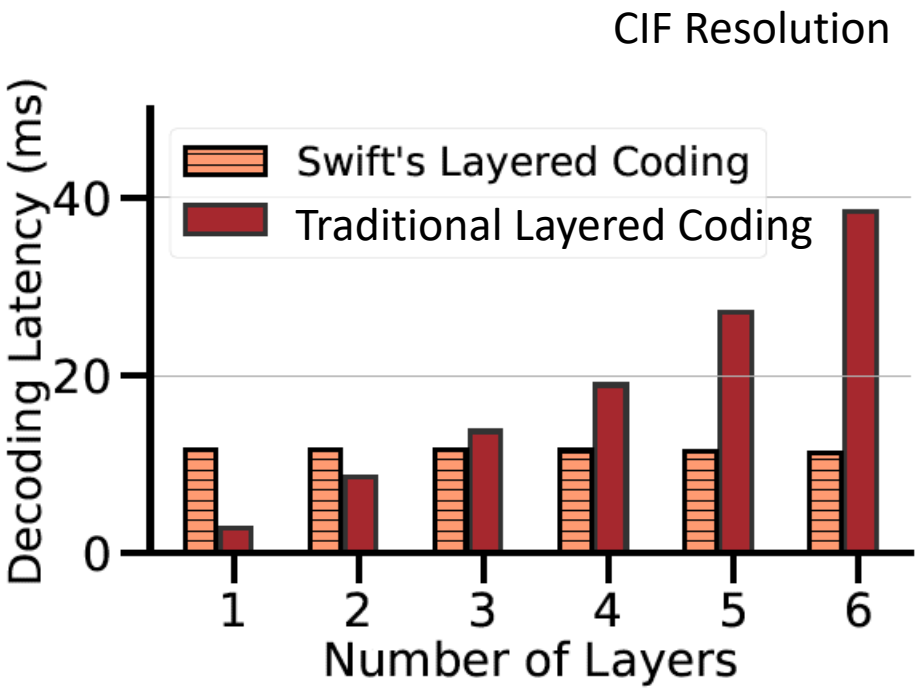
2) Layered Neural Video Decoder

Challenge1. Iterative Decoding is Slow

- ❑Need real-time decoding for playback

Single-shot Decoder





Key Takeaway

Latency is now independent of layers

2) Layered Neural Video Decoder

Challenge2. Variable Compute Capacity

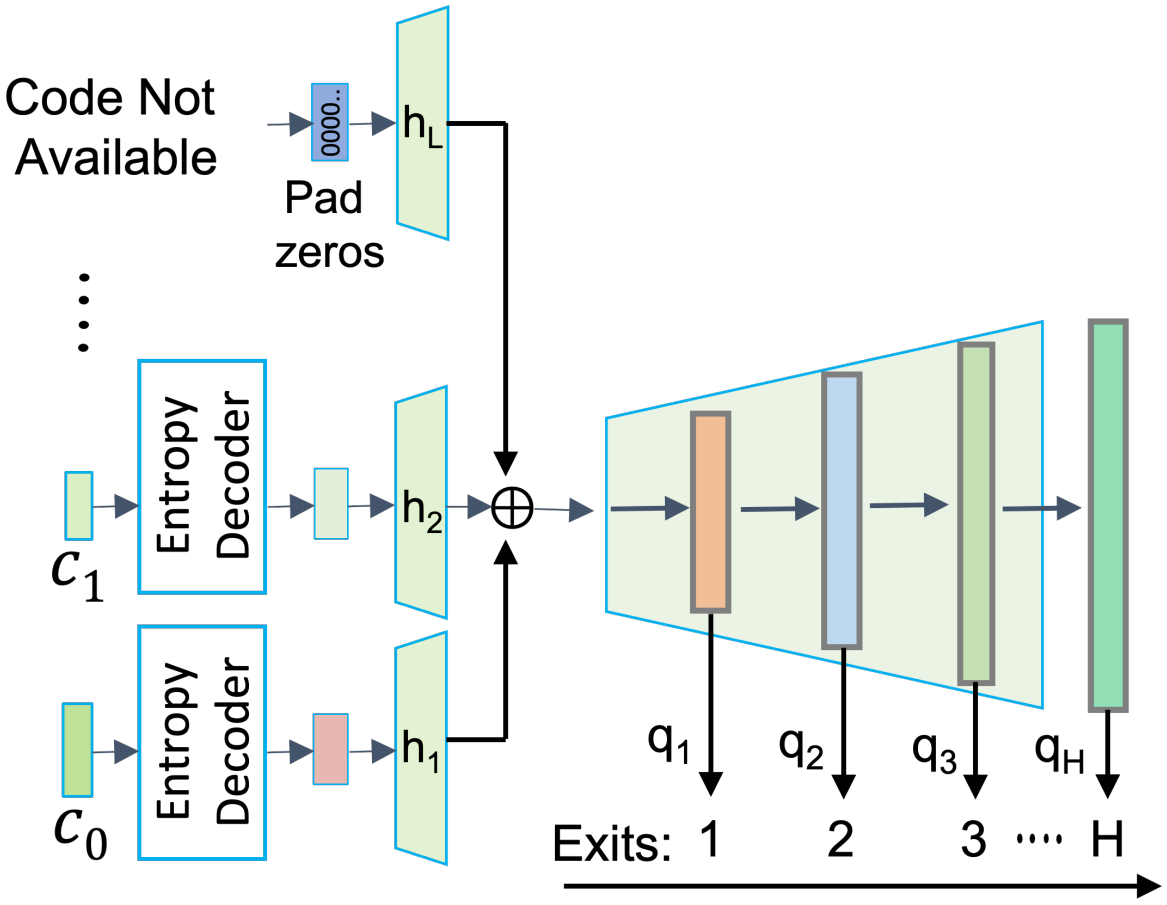
- ❑Software-driven
- ❑Need to scale well with other applications

2) Layered Neural Video Decoder

Challenge2. Variable Compute Capacity

- ❑Software-driven
- ❑Need to scale well with other applications

Multiple Exits

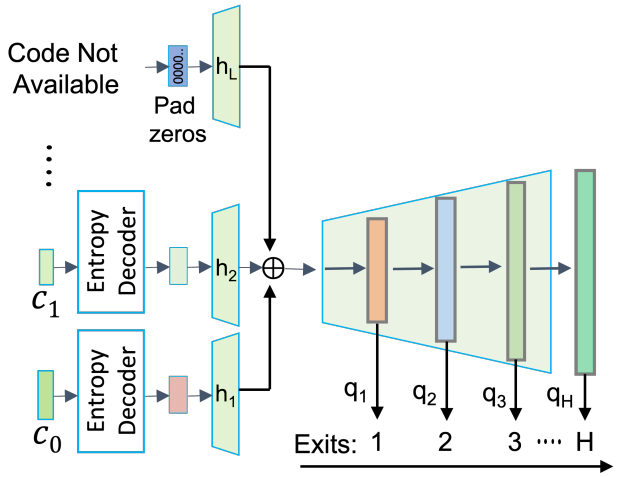


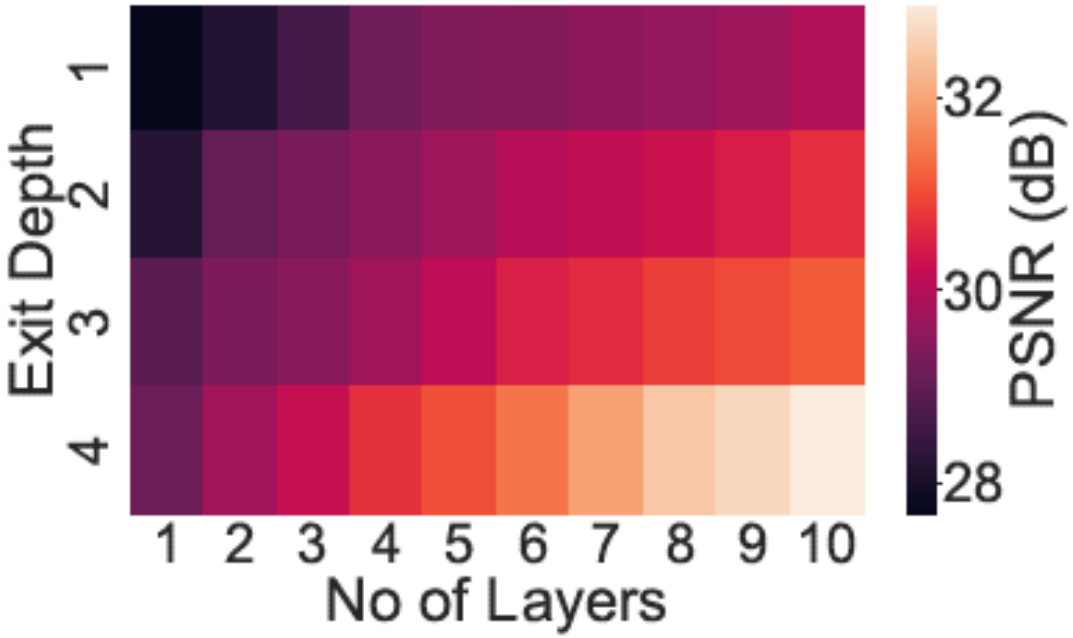
2) Layered Neural Video Decoder

Challenge2. Variable Compute Capacity

- ❑Software-driven
- ❑Need to scale well with other applications

Multiple Exits





Video quality is a function of variable compute and network capacity

3) Streaming Protocol/Algorithm (ABR)

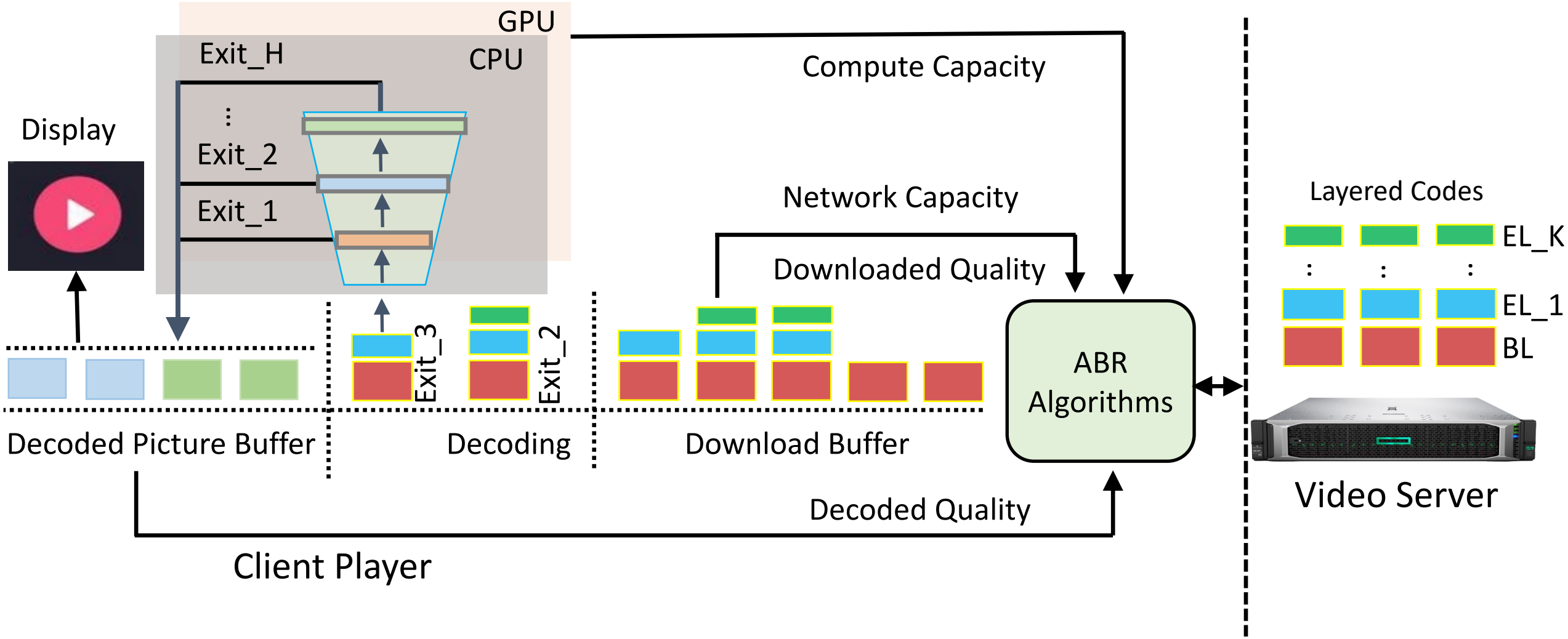
- ❑Question
  - ❑When and How to Enhance Buffered Low-Quality Segment?
- ❑Objective
  - ❑Maximize QoE e.g., F(Quality, Stalls, Smoothness)

Network history
Buffer occupancy
Download qualities
Compute history
Decoded qualities
Playback times
Quality matrix for next segment

→ ABR →

Enhance or fresh download?
How many layers?

Putting Everything Together: Swift



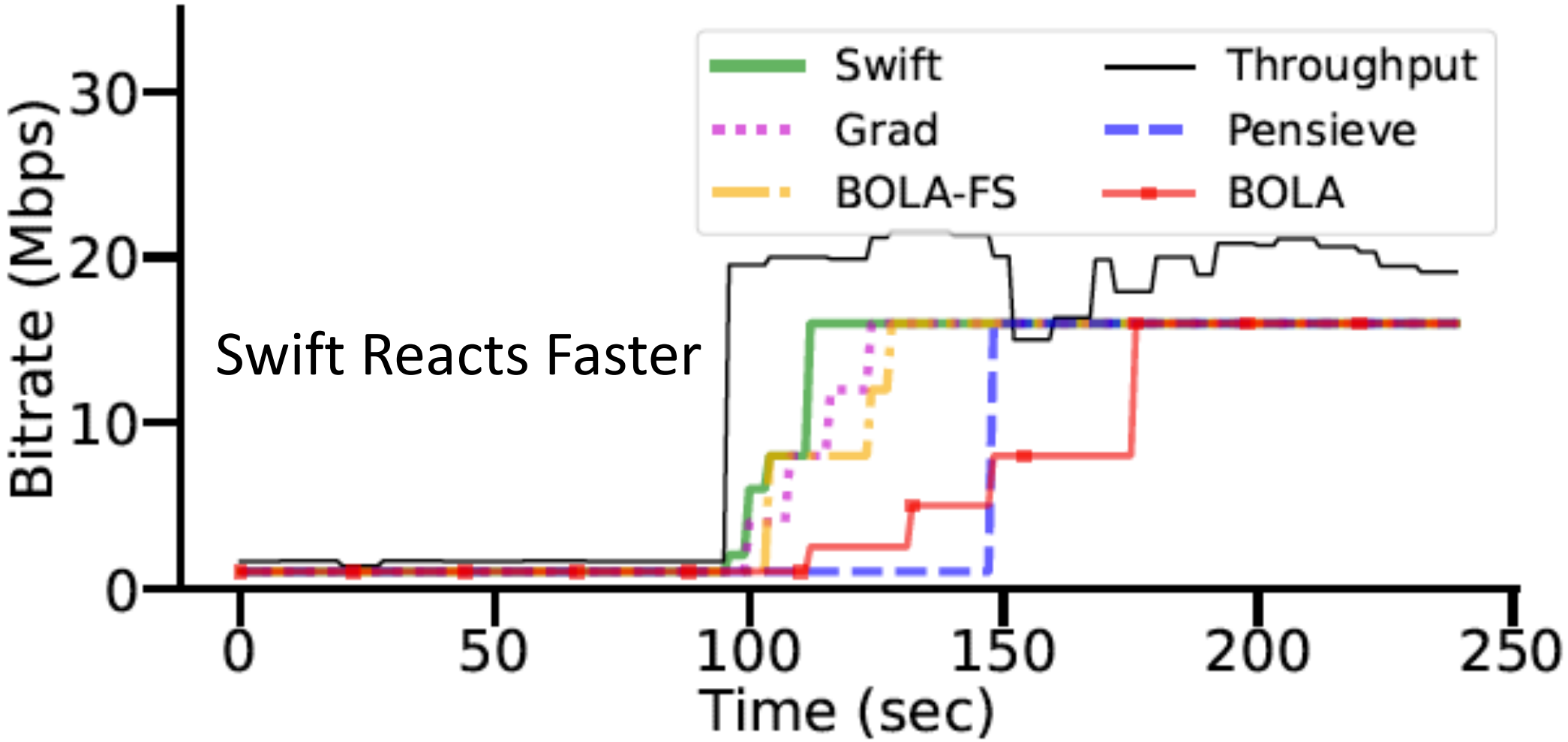
Swift's Performance Comparison

- Compression Baselines
  - H.264
  - SHVC

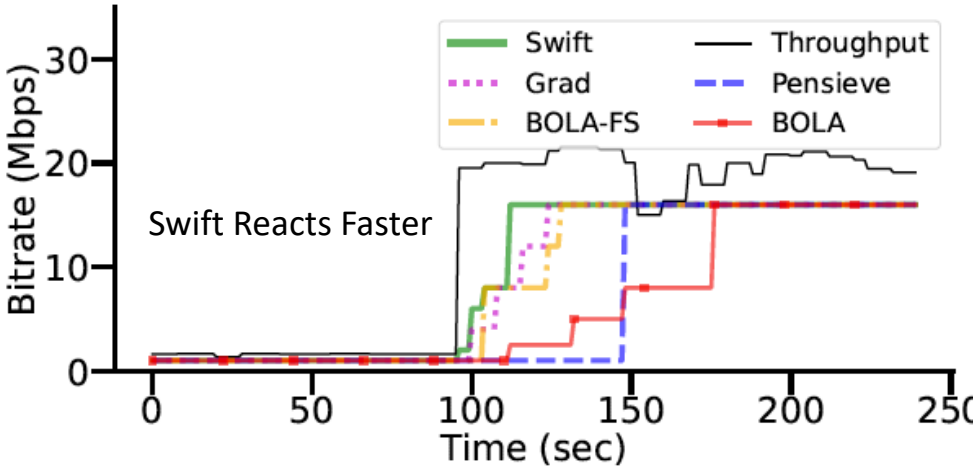
- Streaming Baselines
  - BOLA [INFOCOM'16]
  - BOLA-FS [MMSys'18]
  - Pensieve [SIGCOMM'17]
  - Grad [MM'20]

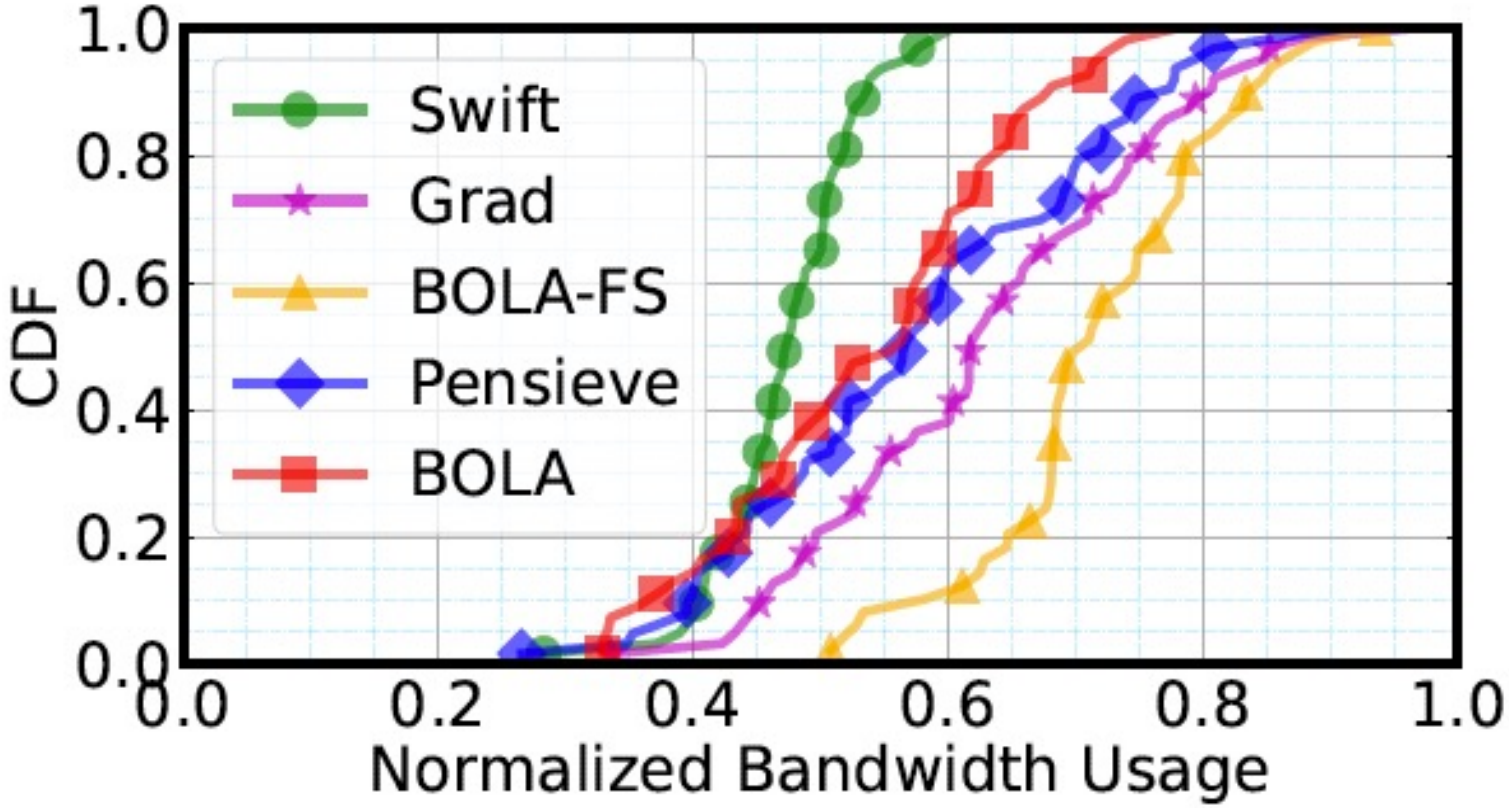
- Swift Evaluation
  - QoE (Quality, Stalls, Temporal Smoothness)
  - Bandwidth Utilization
  - Reaction time
  - Compression Efficiency & Latency

Swift's Performance

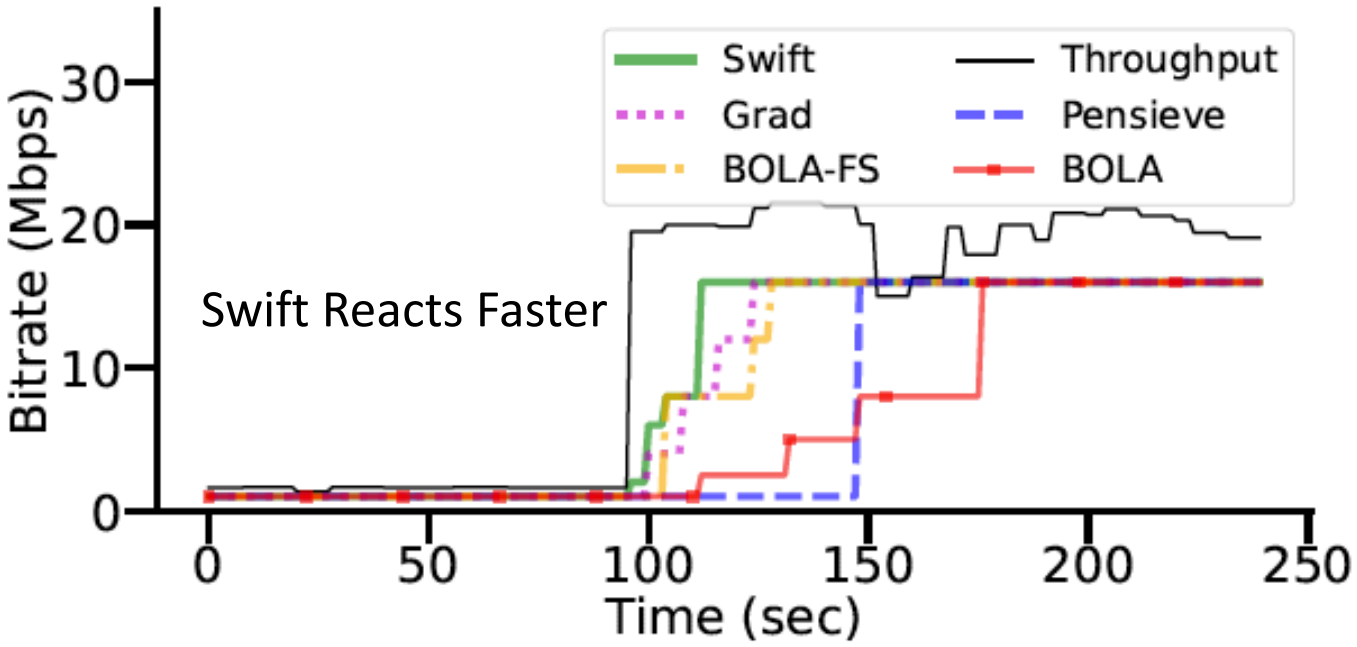


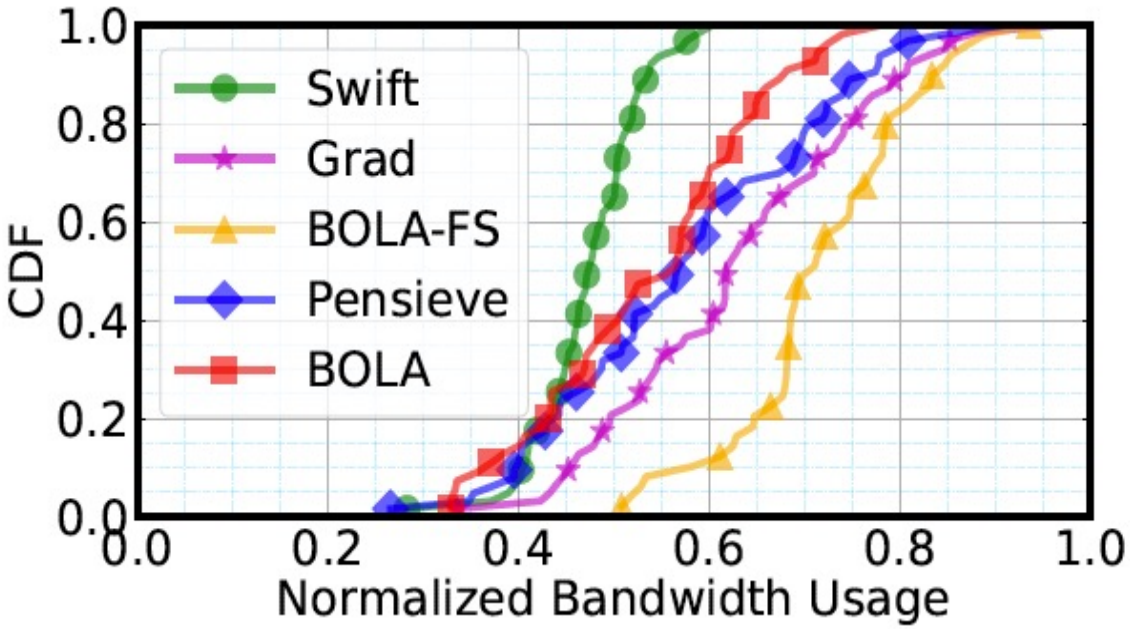
Swift's Performance





Swift's Performance





Swift's faster reaction improves QoE by 45% while saving the bandwidth by 18%

Summary

- Layered coding provides fine-grained rate adaptation but traditionally it is nontrivial to devise an algorithm
- Neural video codecs are becoming mainstream and can efficiently realize layered coding on GPU-equipped devices
- We present Swift, an adaptive video streaming system using layered neural coding, that includes an encoder, a decoder, and a streaming protocol

Thank you!

Questions? malleshd@andrew.cmu.edu