Practically High Performant Neural Adaptive Video Streaming

Sagar Patel, Junyang Zhang, Nina Narodytska, Sangeetha Abdu Jyothi

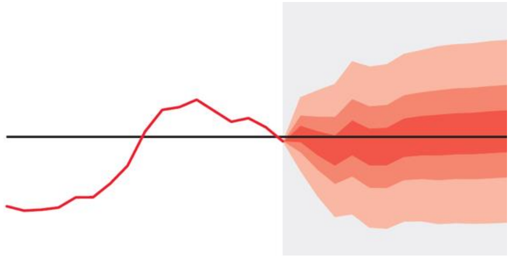
UC Irvine

vmware by Broadcom

Delivering Video over Internet is Challenging

- Video Streaming is a multi-billion dollar business

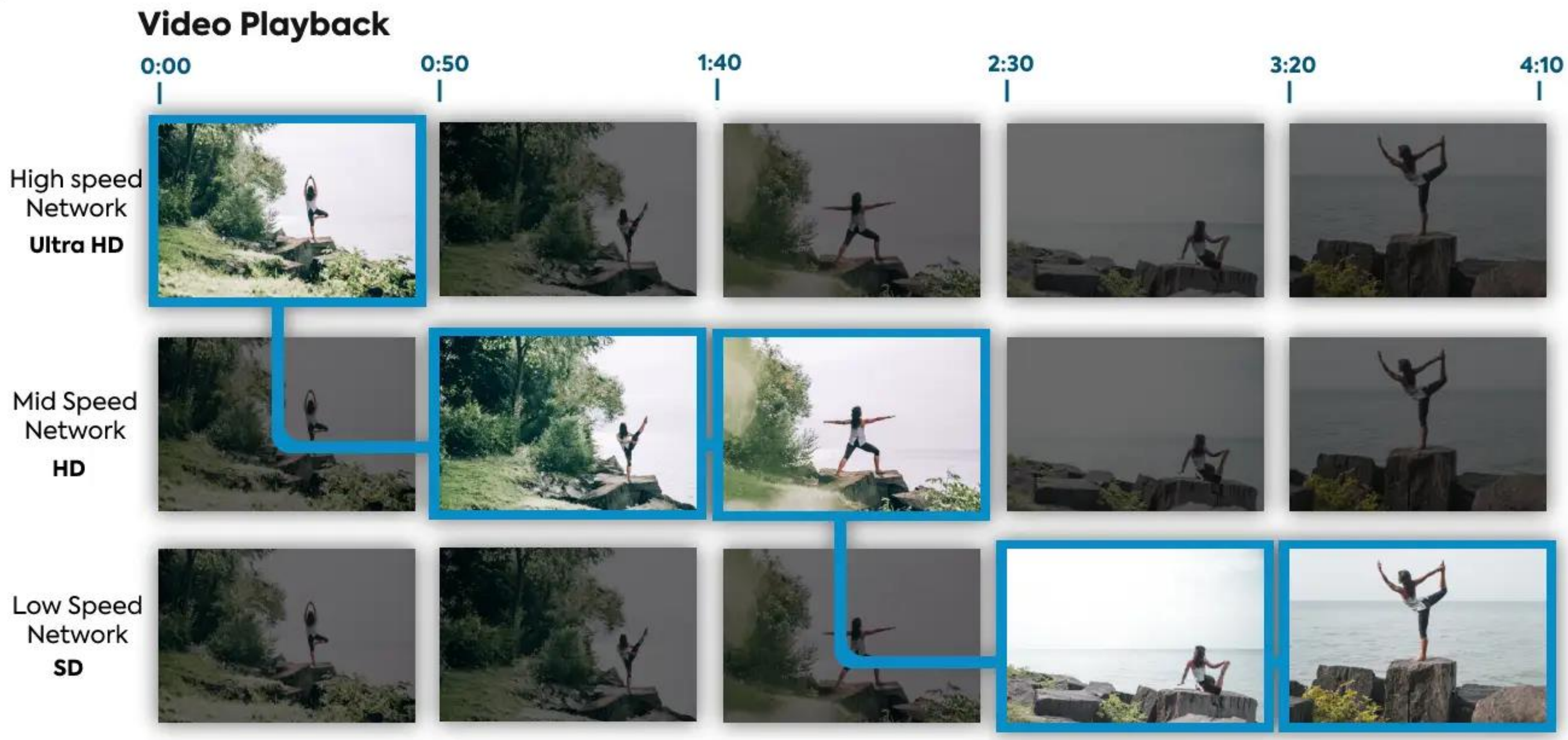
- However, delivery high quality-of-experience is still an open challenge



Network Uncertainty

Growing Video Size

Adaptive Bitrate Streaming



-Teyuto

Supervised Machine Learning

- Standard techniques use machine learning to predict the network and plan ahead

CS2P [SIGCOMM ‘16]

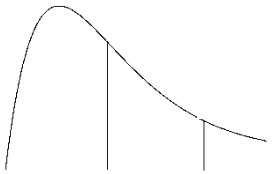
Fugu [NSDI ‘20]

Sensei [NSDI ‘21]

Xatu [SIGMETRICS ‘21]

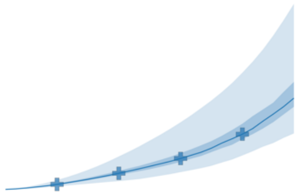
Limitations of Supervised Machine Learning

- However, predicting network behavior is challenging
  - The Internet is heavy-tailed: average behavior is not enough



  - The Internet is partially observable: cross traffic is not visible

  - Prediction Error compounds: Error of early chunks propagates



Can we use Reinforcement Learning?

- Prior RL solutions do not generalize
  - Pensieve [SIGCOMM ‘17]
  - Real-World Pensieve [RL4RealLife ‘19]

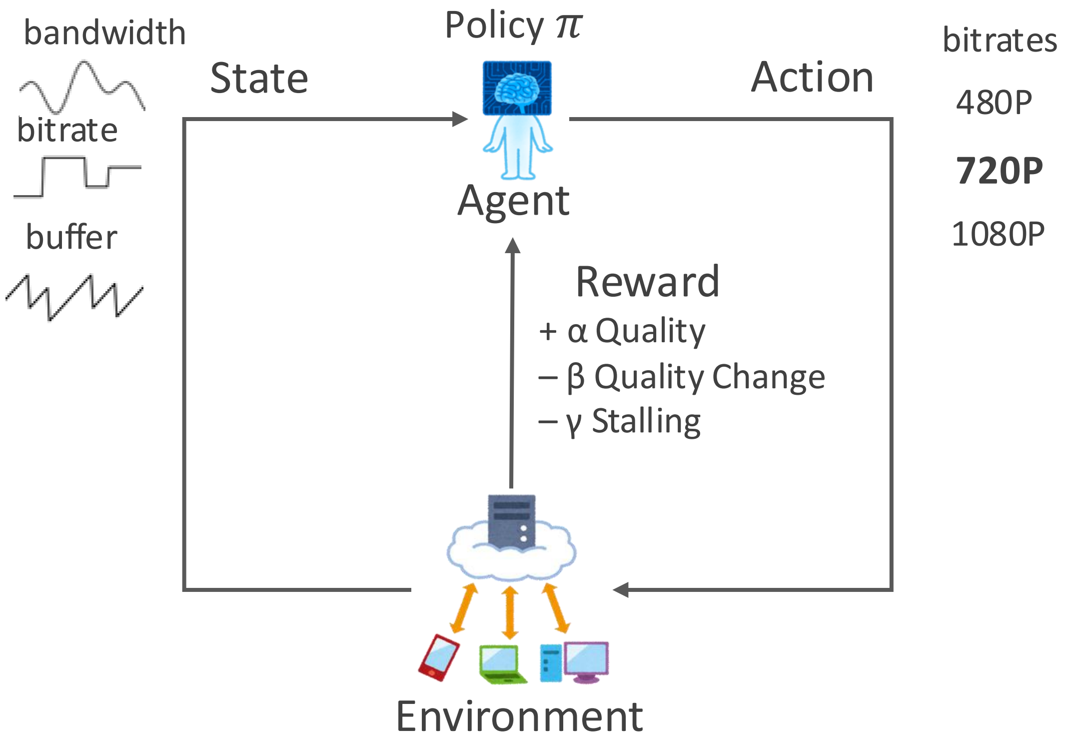
- However, Reinforcement Learning is different
  - Fundamentally different data-driven decision making
  - Does not require predicting the network

- Can we achieve high performance in the real world with reinforcement learning?

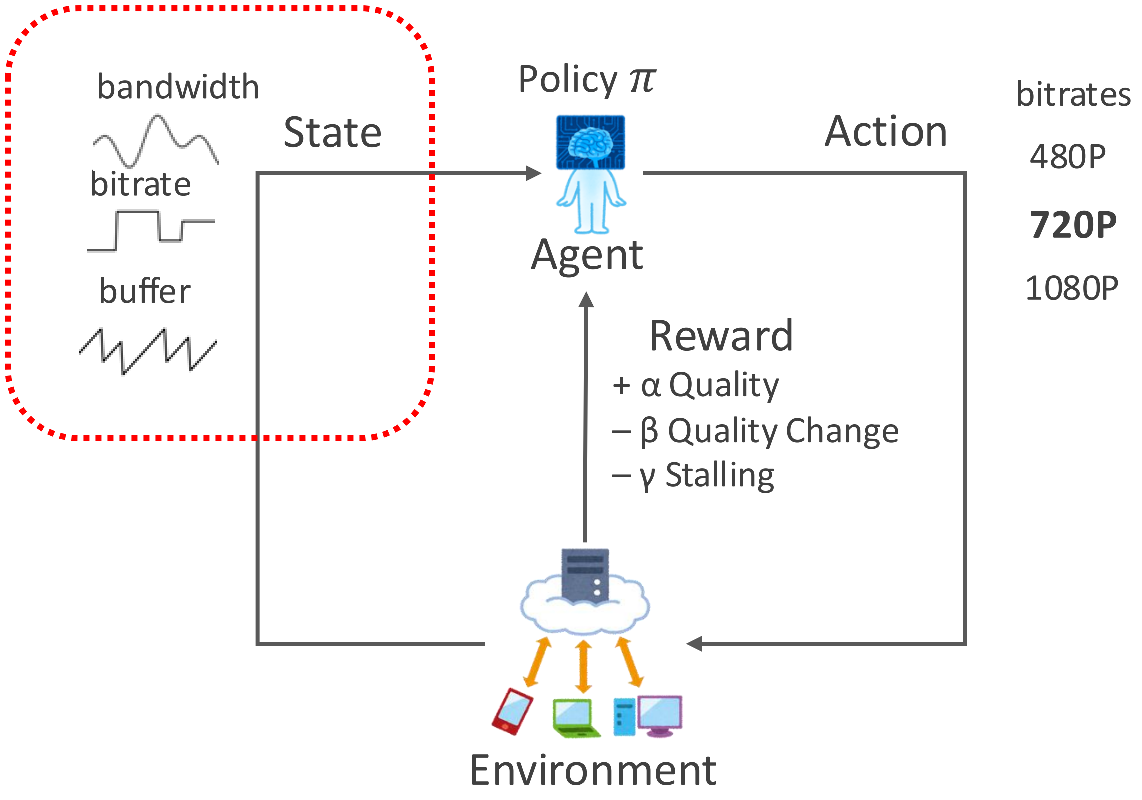
Practically High-Performant ABR: Our solution

- We address these by introducing
  - A new training framework, Plume
  - A new controller architecture, Gelato

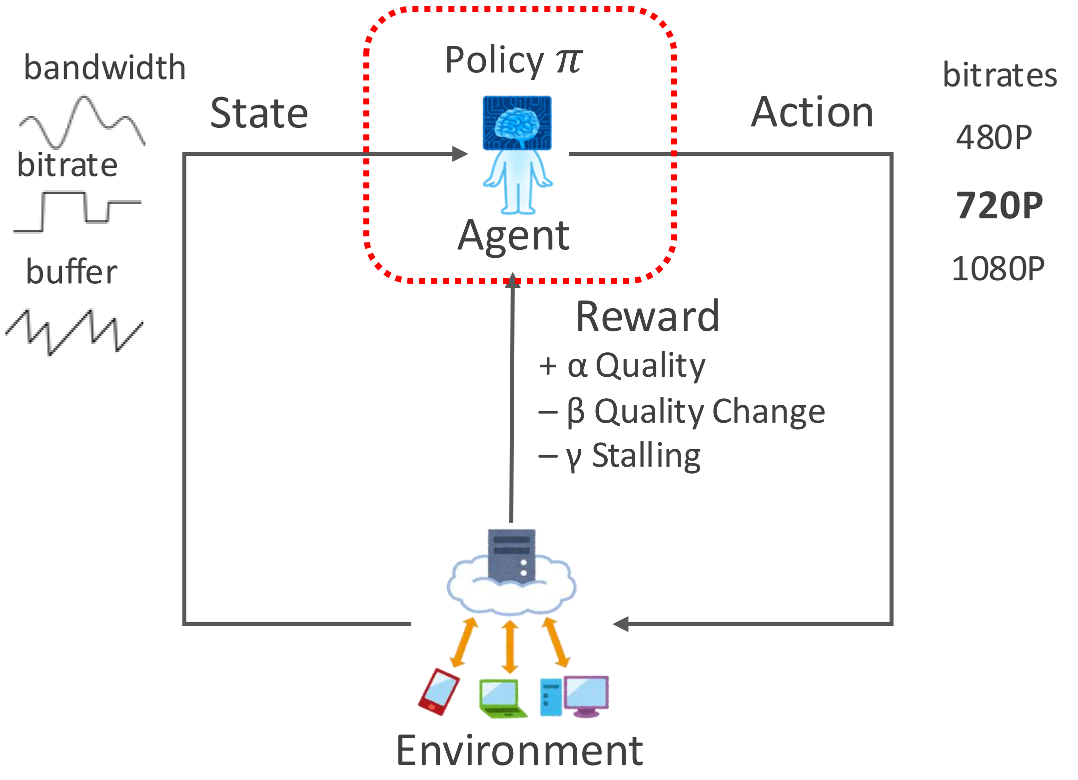
Reinforcement Learning: ABR Perspective



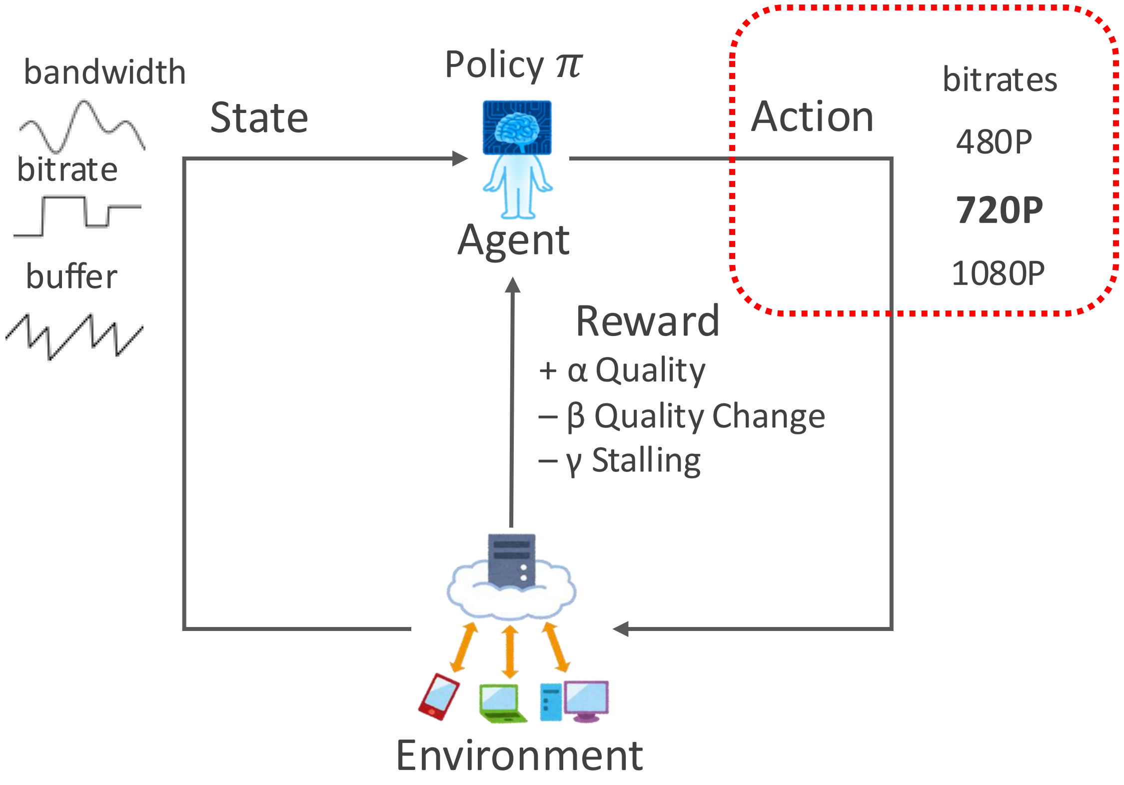
Reinforcement Learning: ABR Perspective



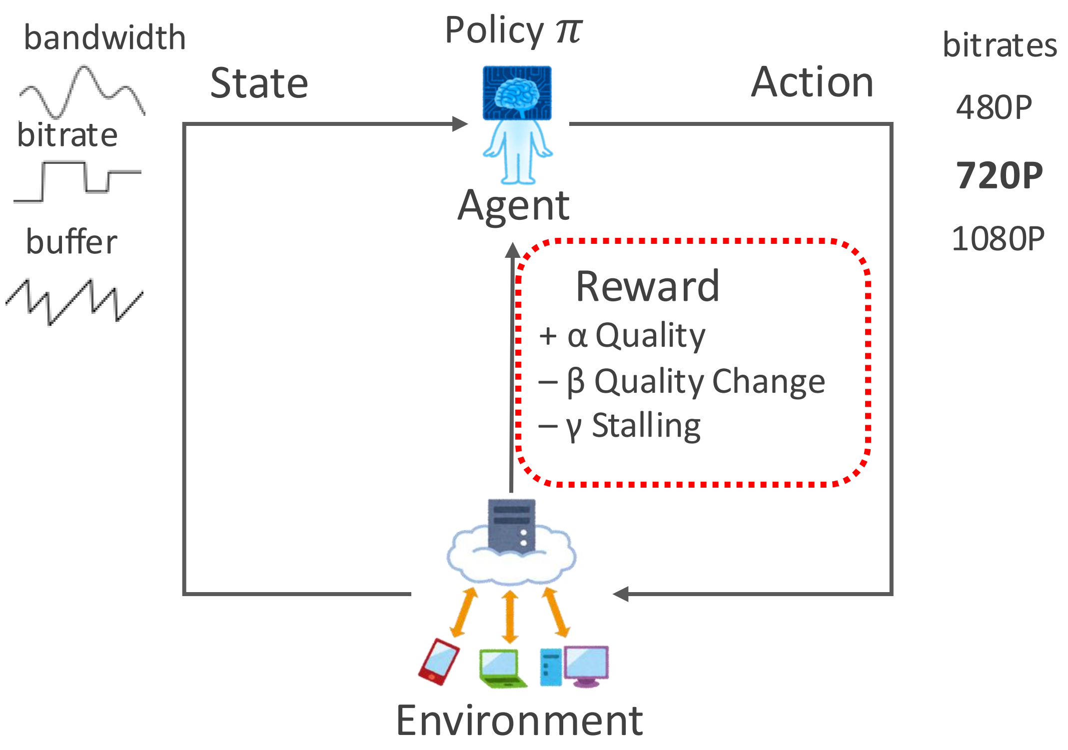
Reinforcement Learning: ABR Perspective



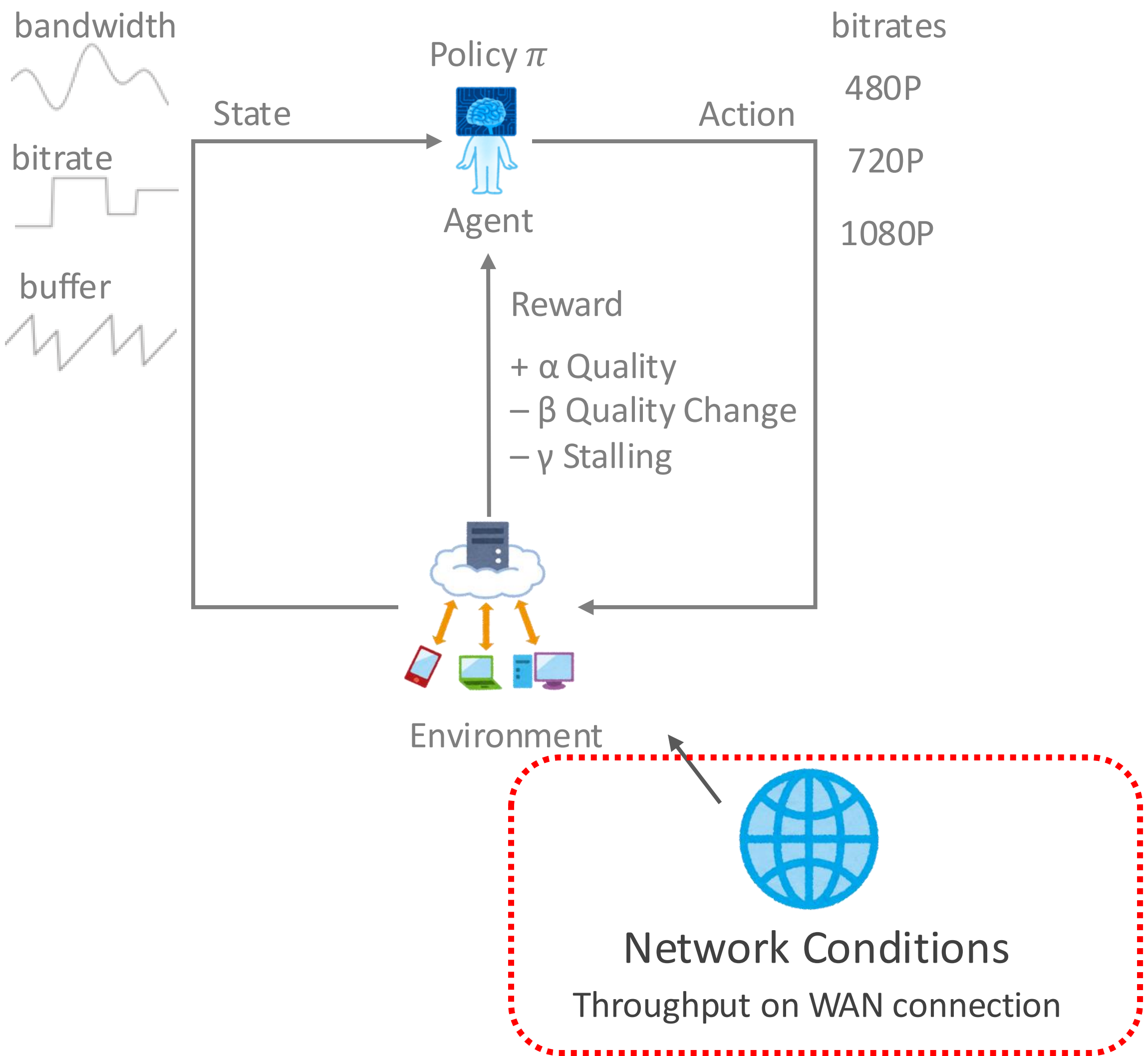
Reinforcement Learning: ABR Perspective



Reinforcement Learning: ABR Perspective

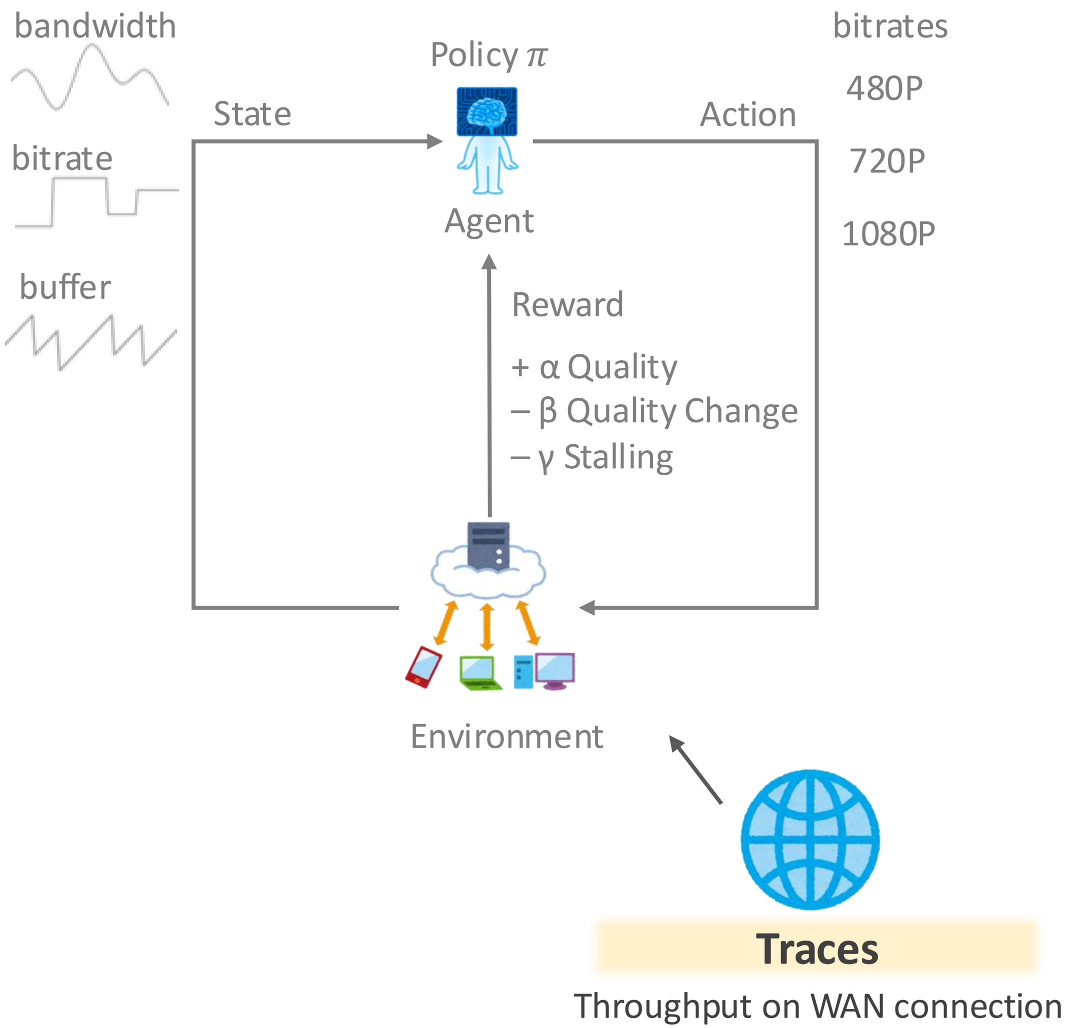


Network Conditions in the Environment



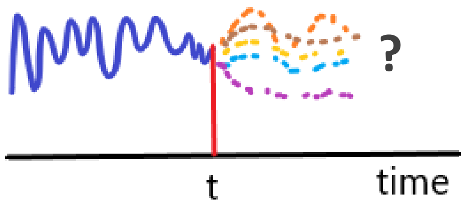
- Environment depends on external network conditions
  - “Inputs”

Network Conditions (cont.)



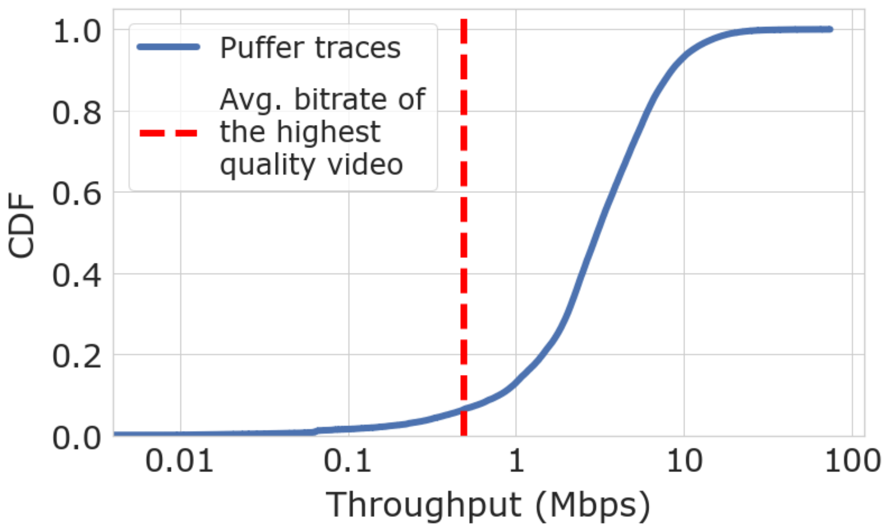
- Replayed using a dataset of traces during training

- Future is unknown



The Problem: Traces are skewed

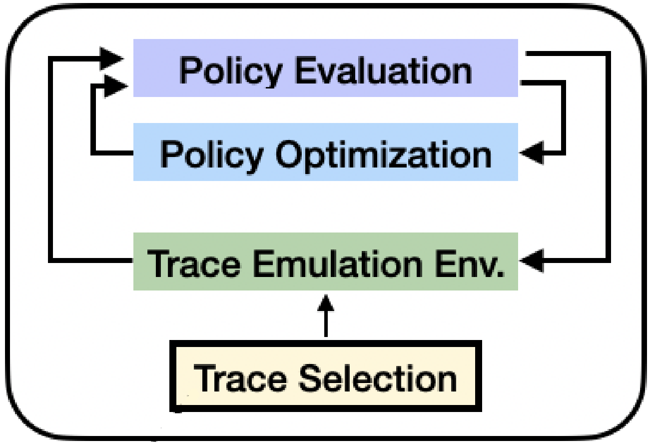
- Real World Network is skewed
  - 93% of YouTube streams *never* stall [SIGCOMM ‘17]
  - Even during lockdown’s demand, only 8% of Facebook video streams are “bad sessions” [IMC ‘20]



Distribution of Throughput observed on the livestreaming platform Puffer over 2 months

The Problem: Impact of Trace Skew

- Traces sit outside of the RL loop and indirectly control the training
- With skewed traces, controller training
  - Can overfit
  - Be inefficient
  - Have noisy or divergent updates

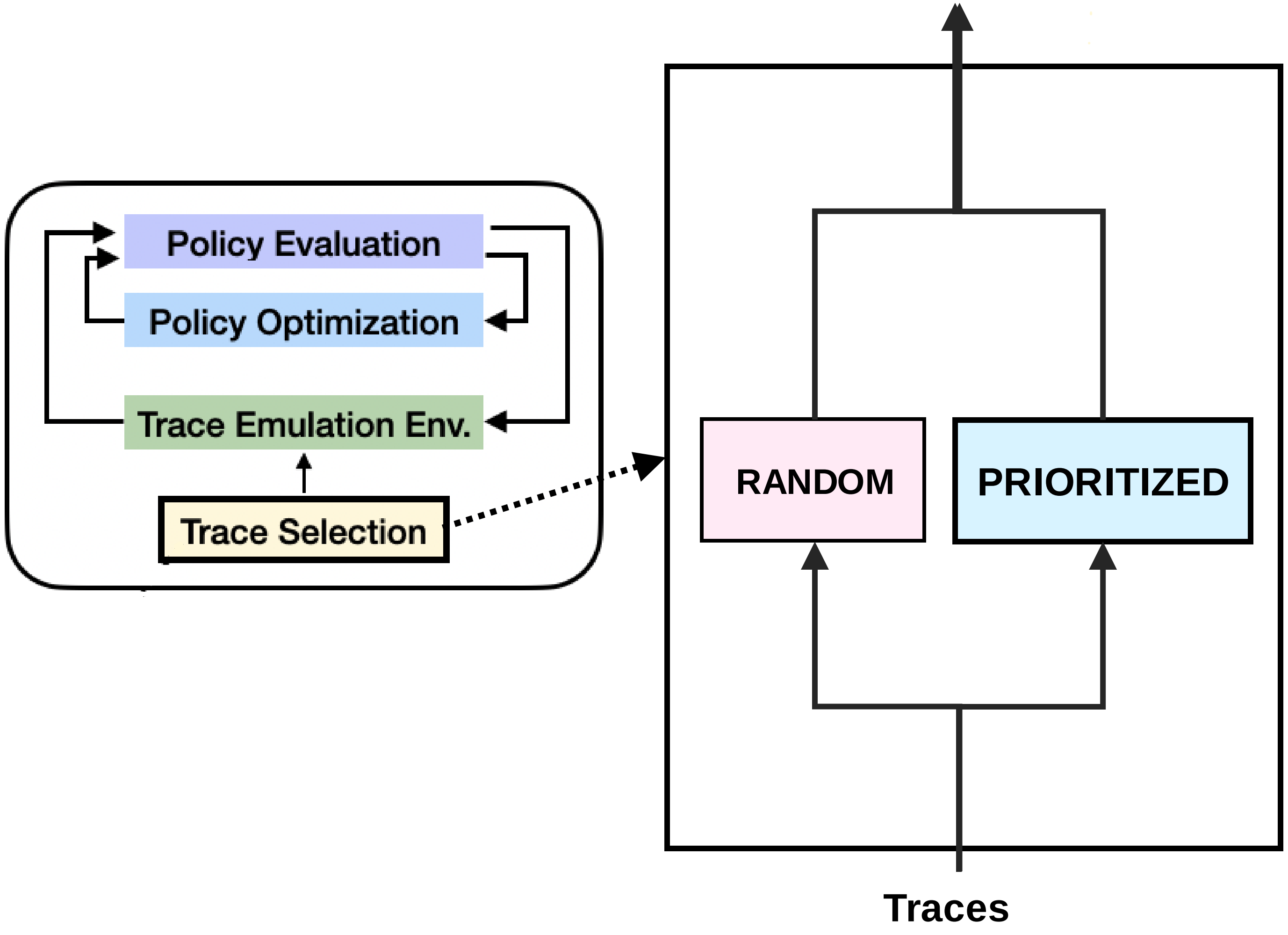


Our Solution

Key Idea

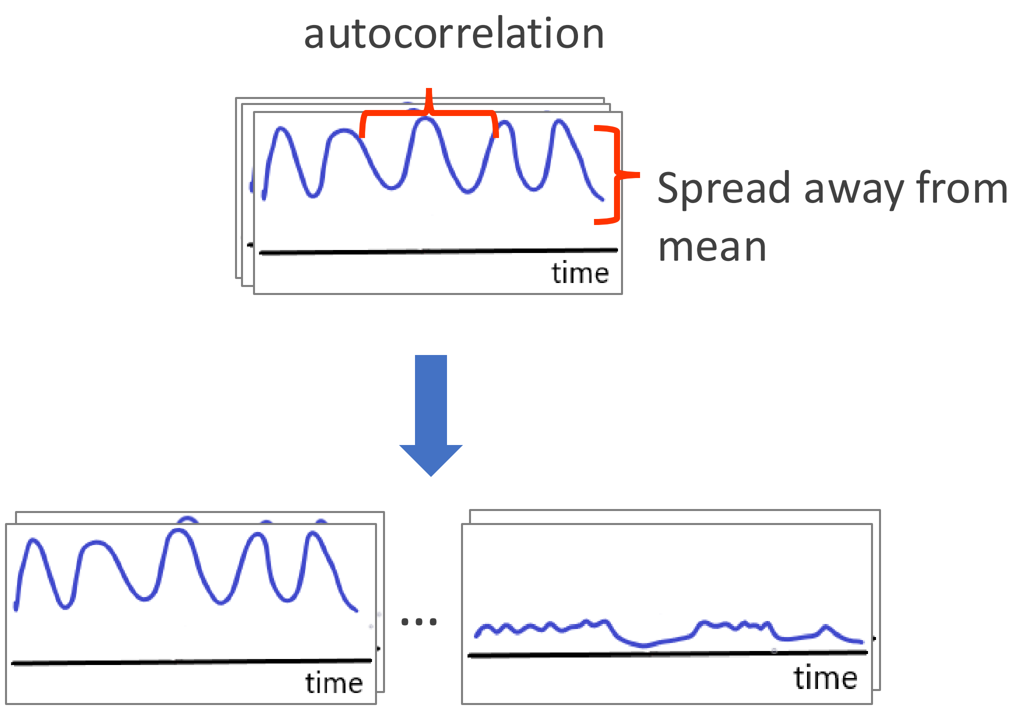
- Balance the distribution with **reward-to-go**: difference between optimal and current
- Reward-to-go = $Reward^{\pi^*} - Reward^{\pi}$

Prioritization

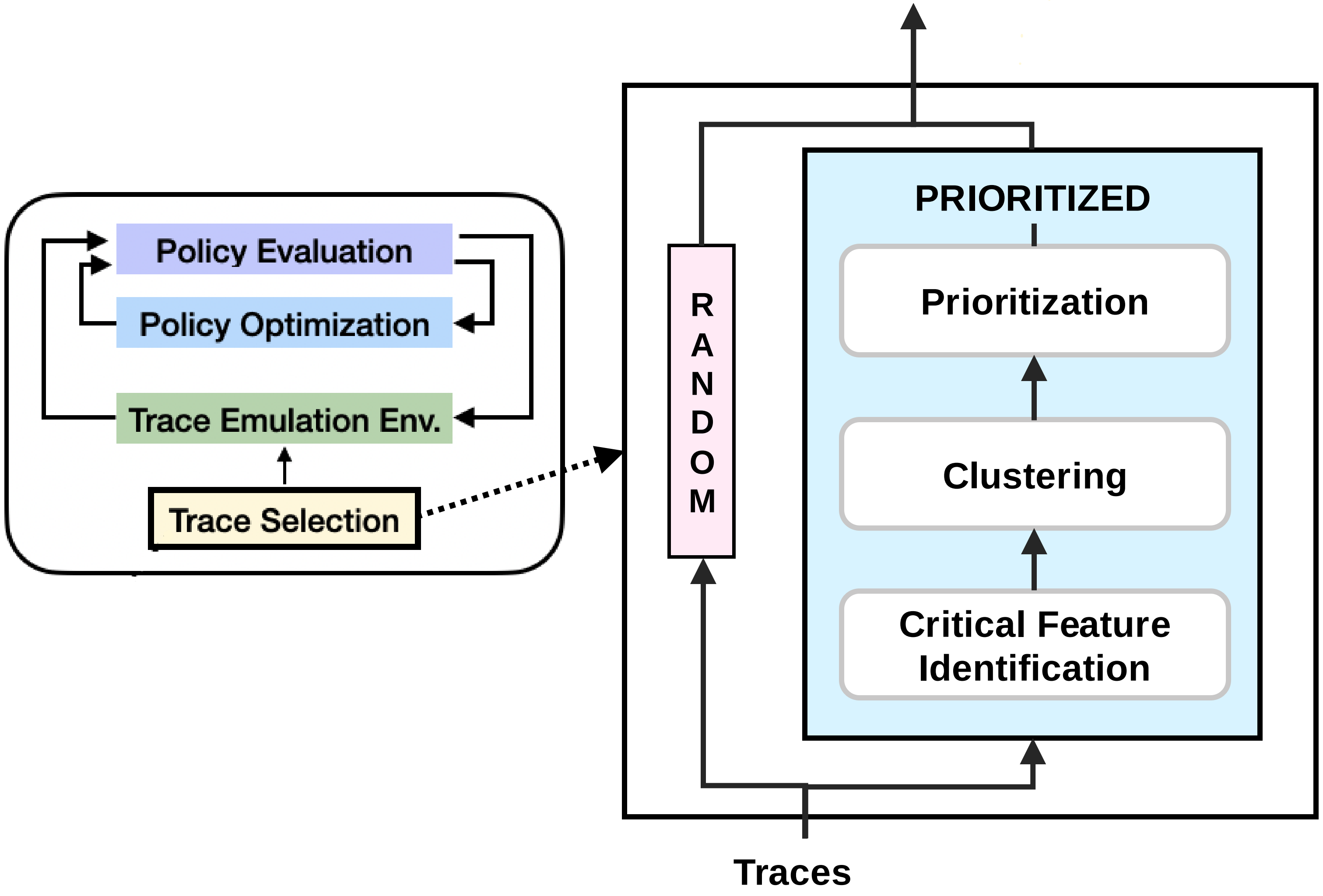


Scaling to ABR

- Real-world datasets are large
  - Can require 10^7 interactions
- Intuition: Traces with similar features have similar behavior
  1. Identify critical features of the traces
  2. Cluster the traces by their features
  3. Weight them the same

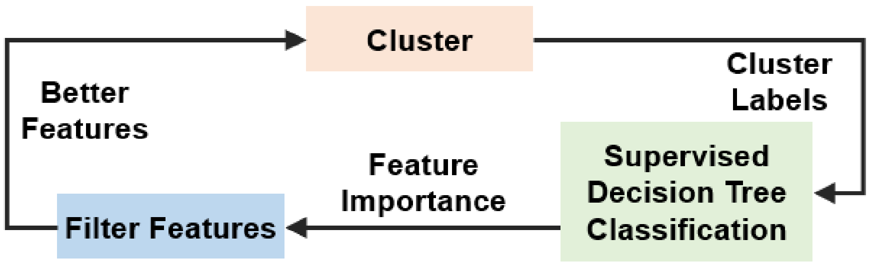


Prioritization



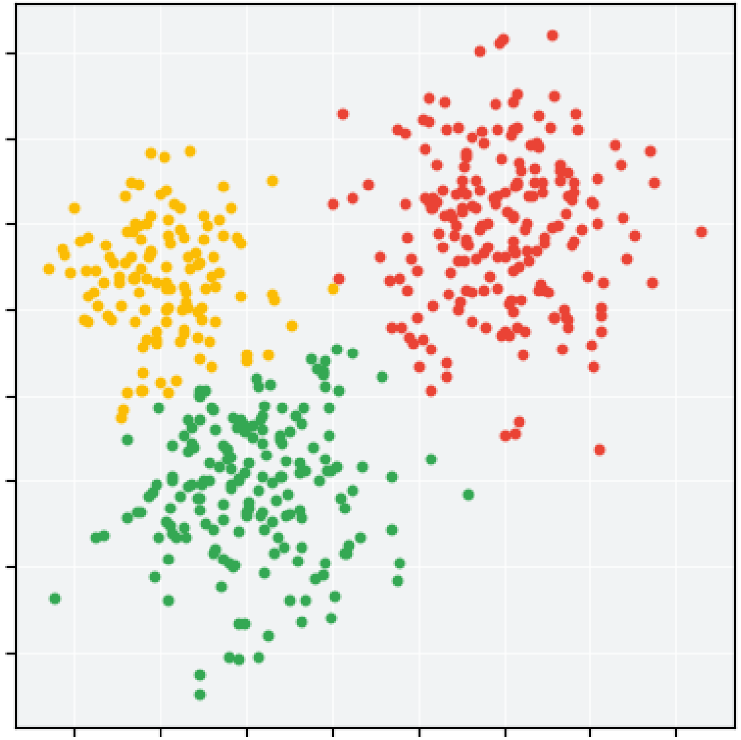
Critical Feature Identification

- Critical feature identification
  - Begin with a large set of standard timeseries features
  - Eliminate irrelevant features
    1. Group traces by features
    2. Train decision tree to predict groups
    3. Filter by most important features



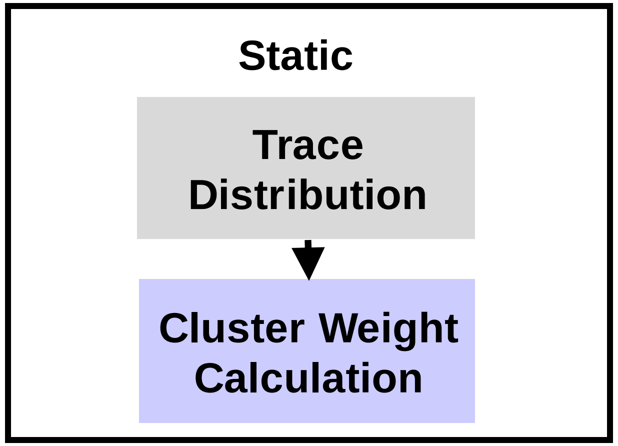
Clustering

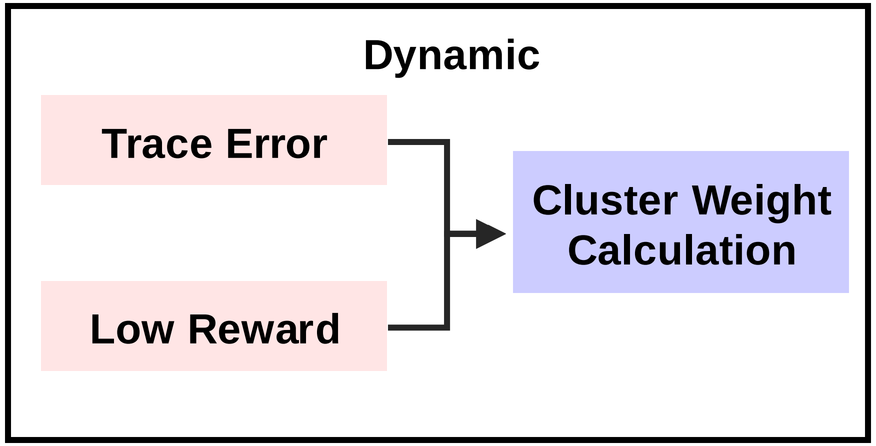
- Gaussian Mixture Model clustering
- Balance spread across features
- Automatically search for optimal number of clusters



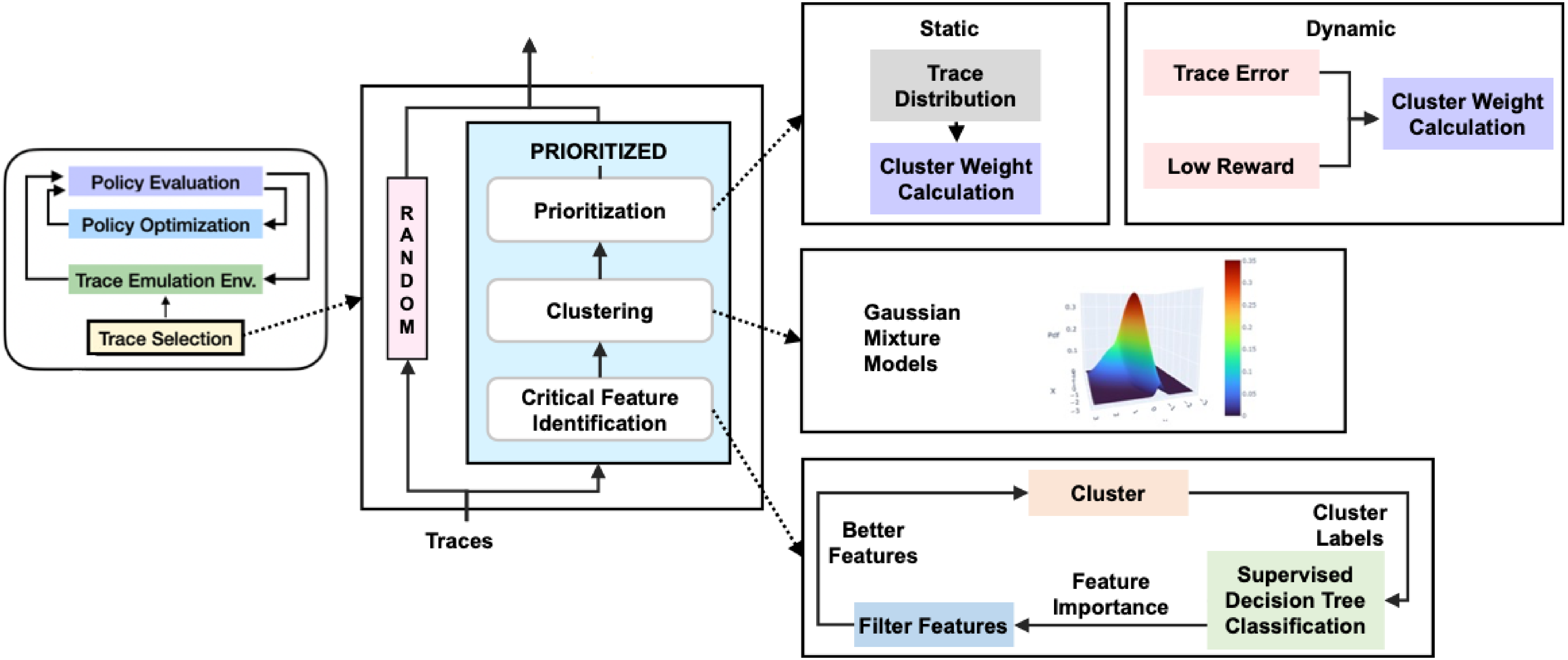
Prioritization

- $Reward^{\pi^*}$ cannot be computed
- We introduce two approximation techniques
  - Static
    - Balance traces to uniform distribution
  - Dynamic
    - Empirically approximate





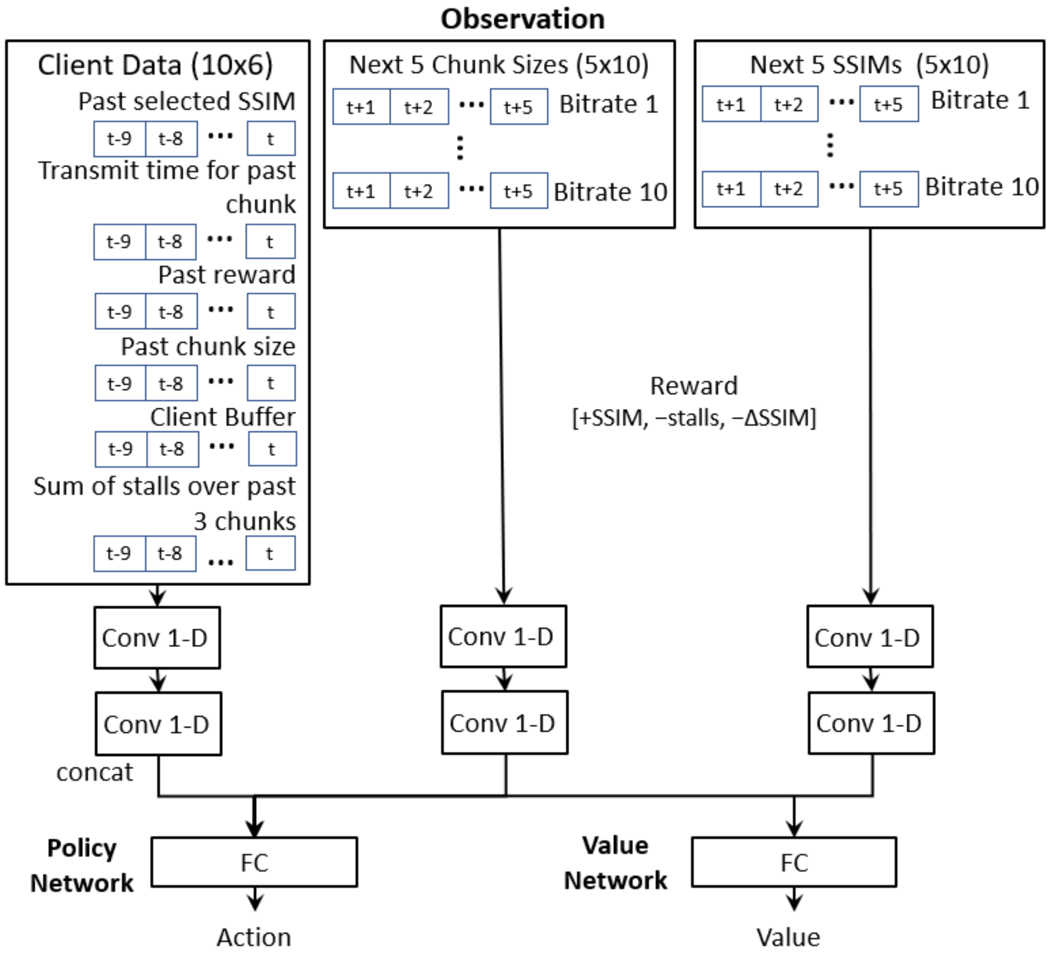
Plume: All together



Practically High-Performant ABR: Our solution

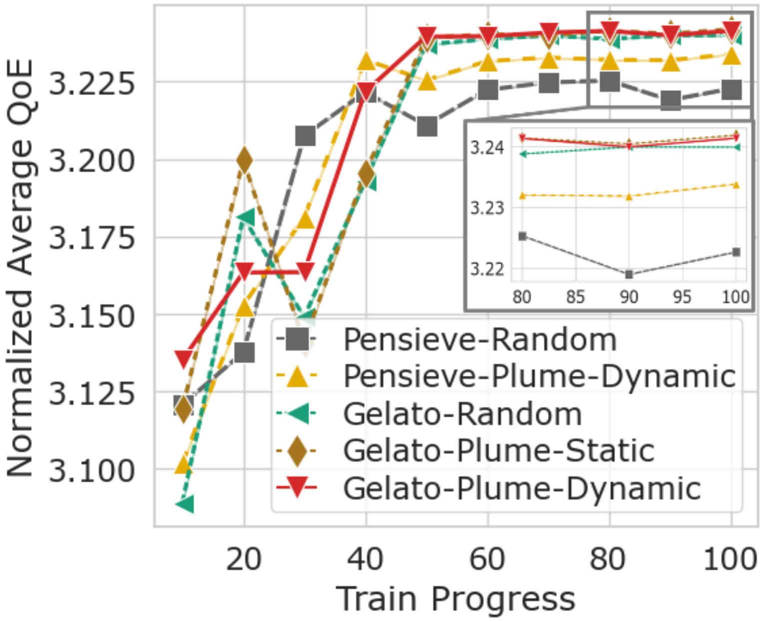
- We address these by introducing
  - A new training framework, Plume
  - A new controller architecture, Gelato

Practical ABR with Gelato

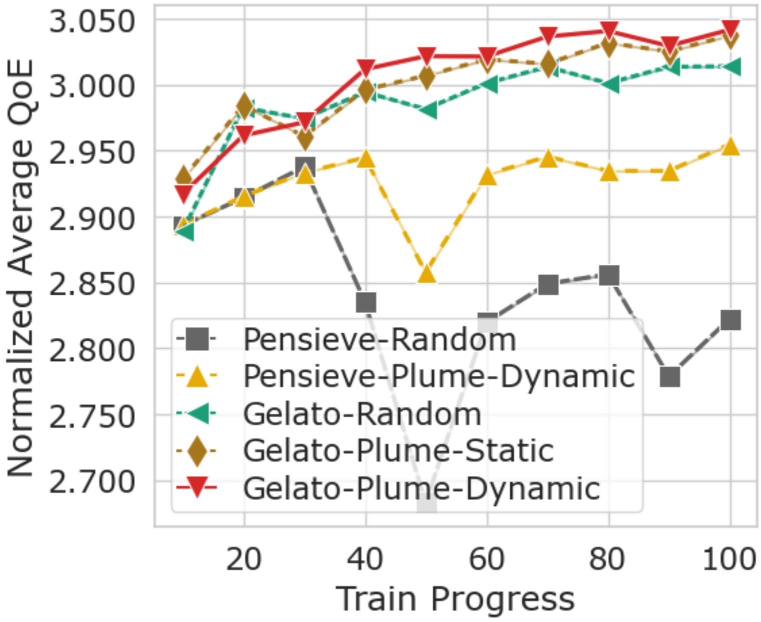


Results

Simulation Evaluation



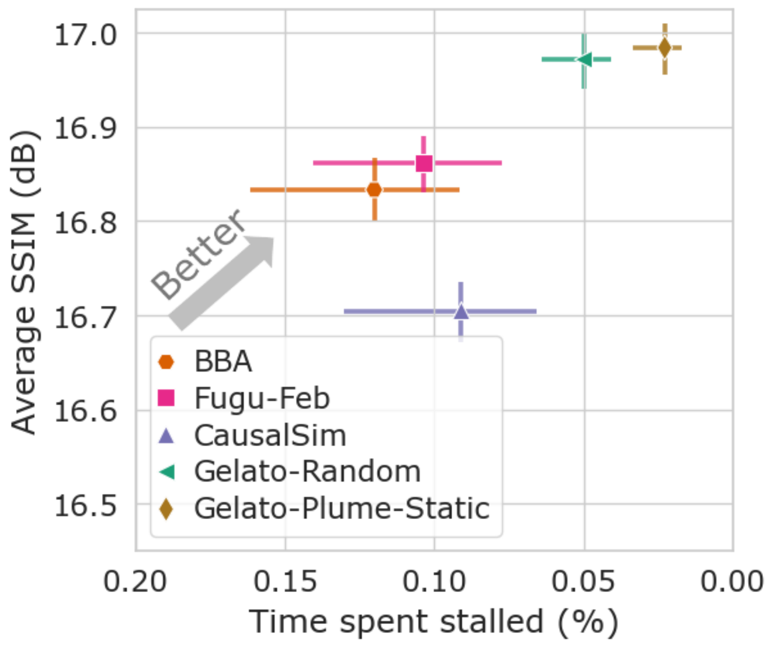
All Traces



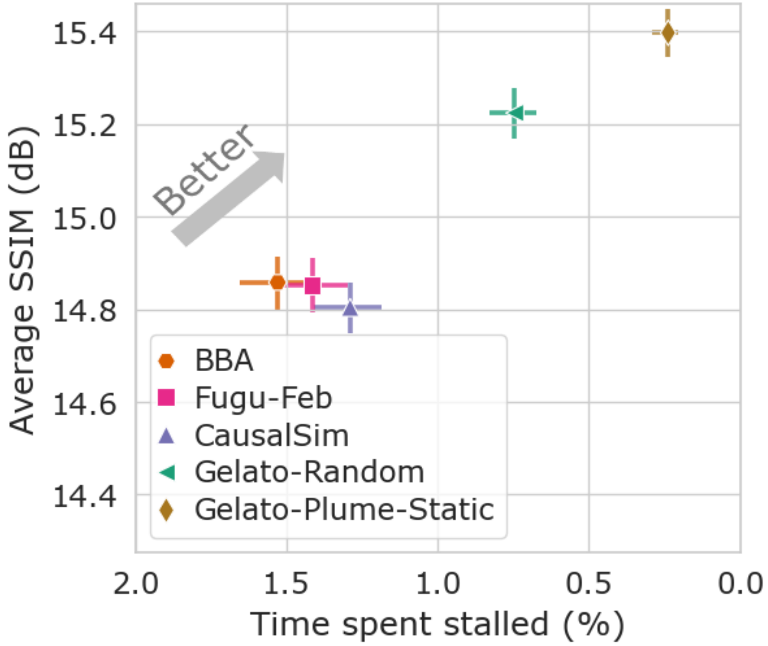
Slow Traces

Plume improves overall performance of our controller Gelato and Pensieve [SIGCOMM ’17]

Streaming Live TV across the Internet



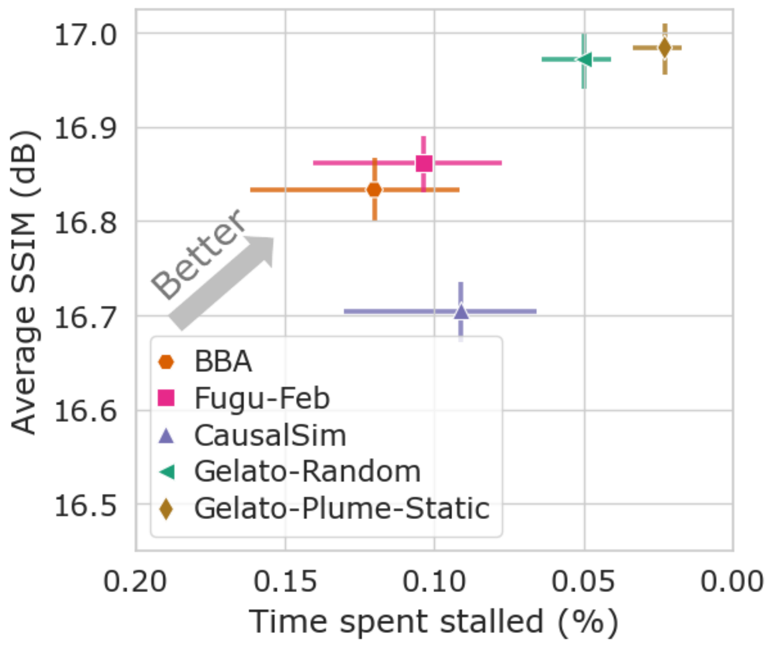
All Traces



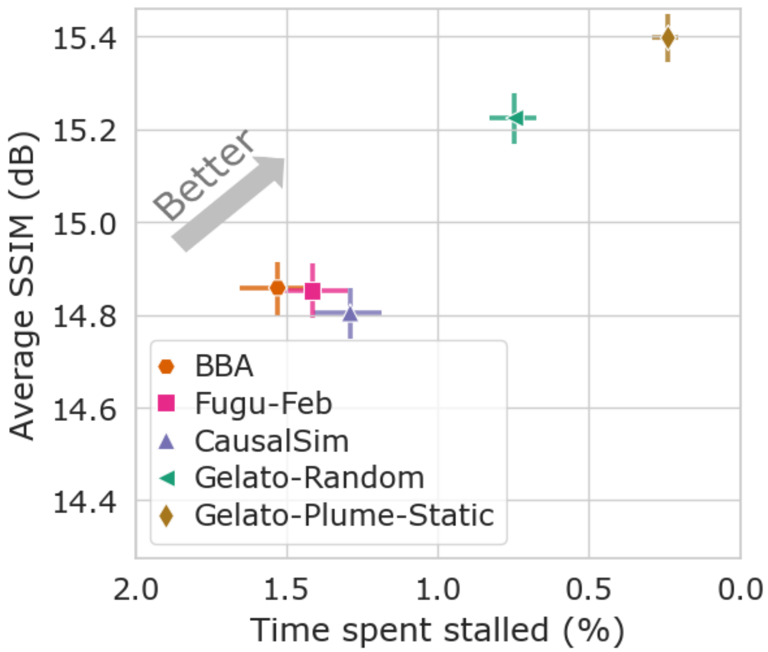
Slow Traces

Gelato-Plume achieves state-of-the-art performance on Puffer [NSDI '20], streaming **58.9** stream-years of live TV to **250k+ users** over 8 months

Streaming Live TV across the Internet



All Traces



Slow Traces

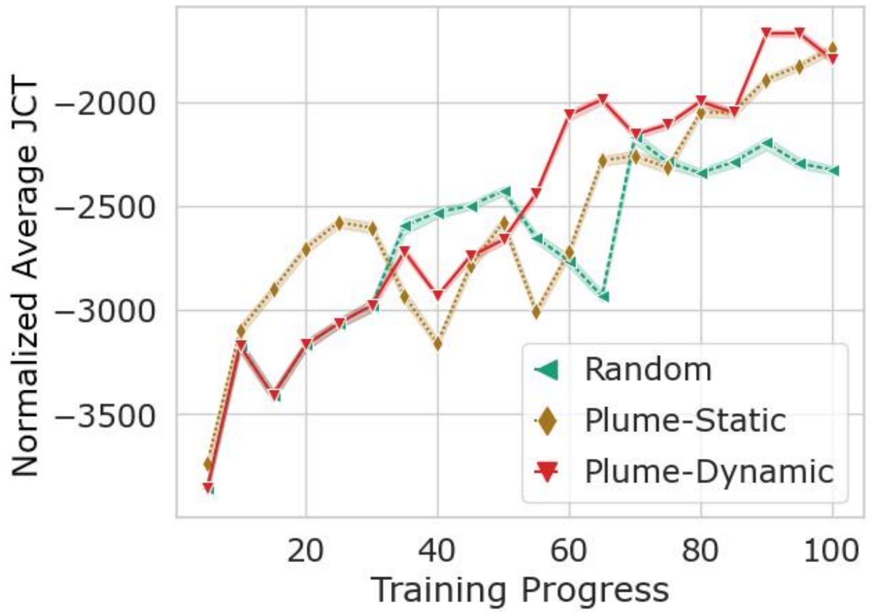
1. 75% stall reduction
2. Statistically significant quality improvement
3. No retraining

Generalization

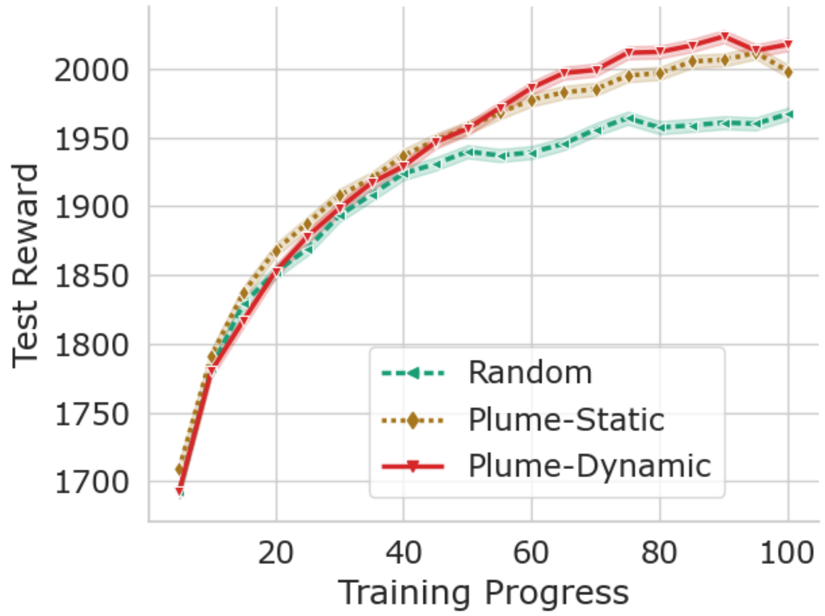
Evaluating Plume's Generalizability

- Conditions in ABR cover a small set of all networking applications
- We benchmark Plume further
  - Across networking Applications
    - Load Balancing
    - Congestion Control
  - Across Trace Distributions
    - Varying distribution of throughput

Benchmarking Plume across applications



Load Balancing



Congestion Control

Plume generalizes across networking applications

Summary

- Deep Reinforcement Learning is a key technique to bring practical ML to adaptive video streaming
- Skewed trace distribution make RL training difficult
- We systematically balance traces with Plume
- Gelato with Plume achieves state-of-the-art performance in the real world, streaming live TV to internet users in the wild
  - First controller on Puffer to achieve both video quality and stalling improvement
  - Up to 75% reduction in stalling

Contact
sagar.patel@uci.edu

Code
github.com/sagar-pa/plume