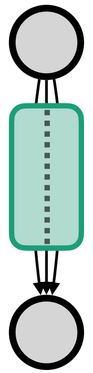
Bundler: A New Middlebox for

Site-to-Site Internet Traffic Control

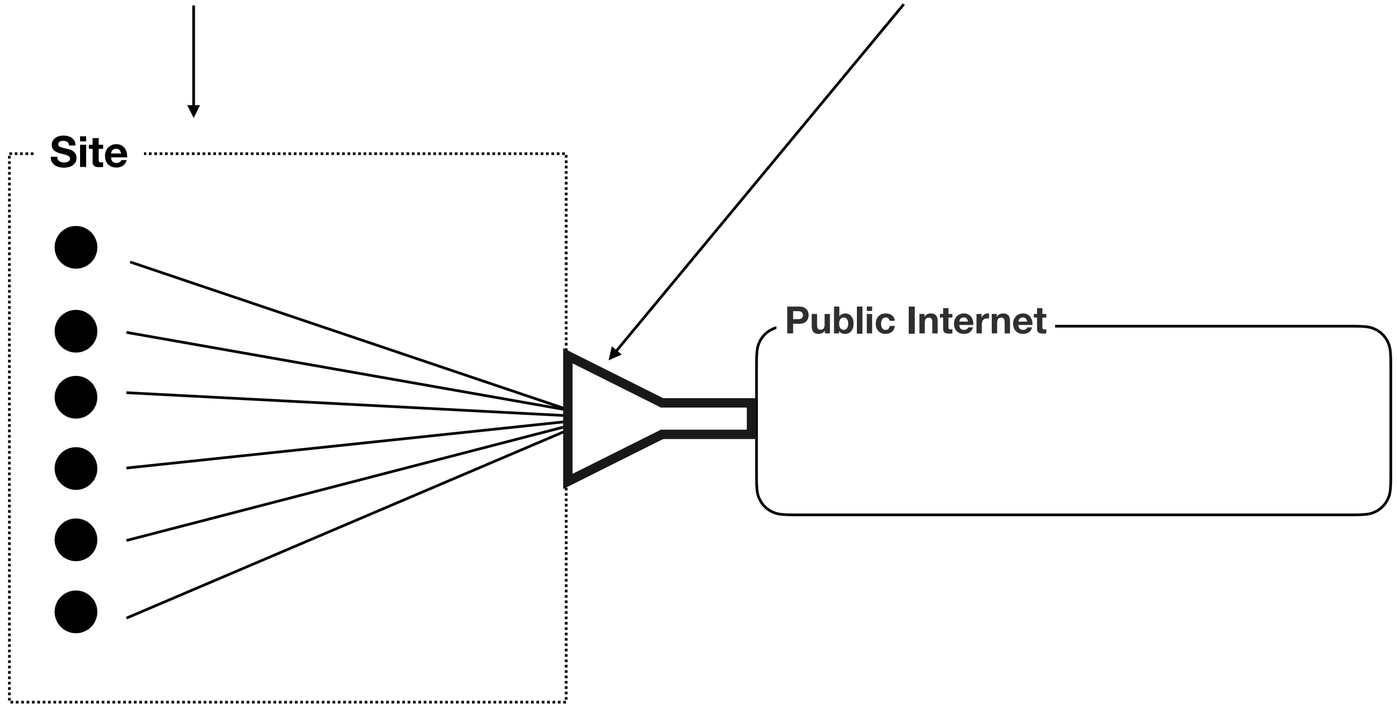


Frank Cangialosi, Akshay Narayan, Prateesh Goyal,
Radhika Mittal, Mohammad Alizadeh, Hari Balakrishnan

MIT CSAIL

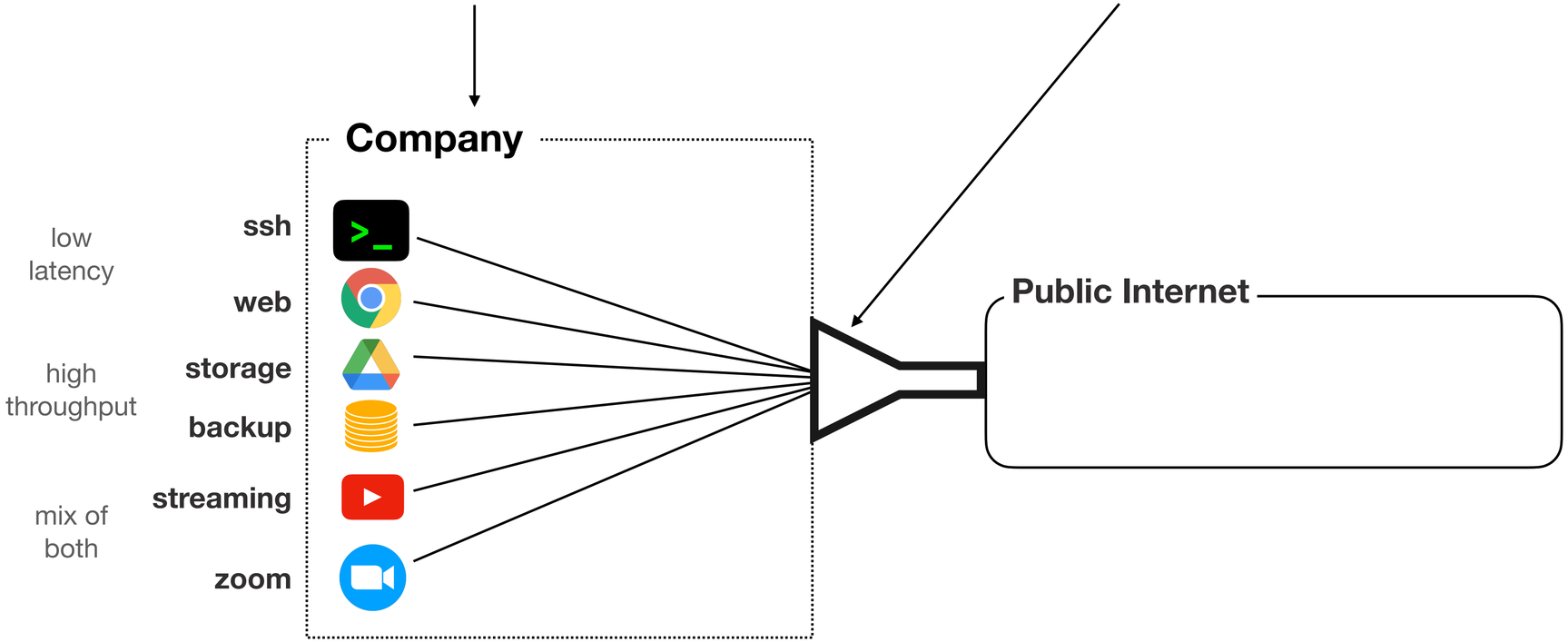
Site-to-Site Internet Traffic Control

“site”: a single physical location with many endpoints sharing an access link to the internet



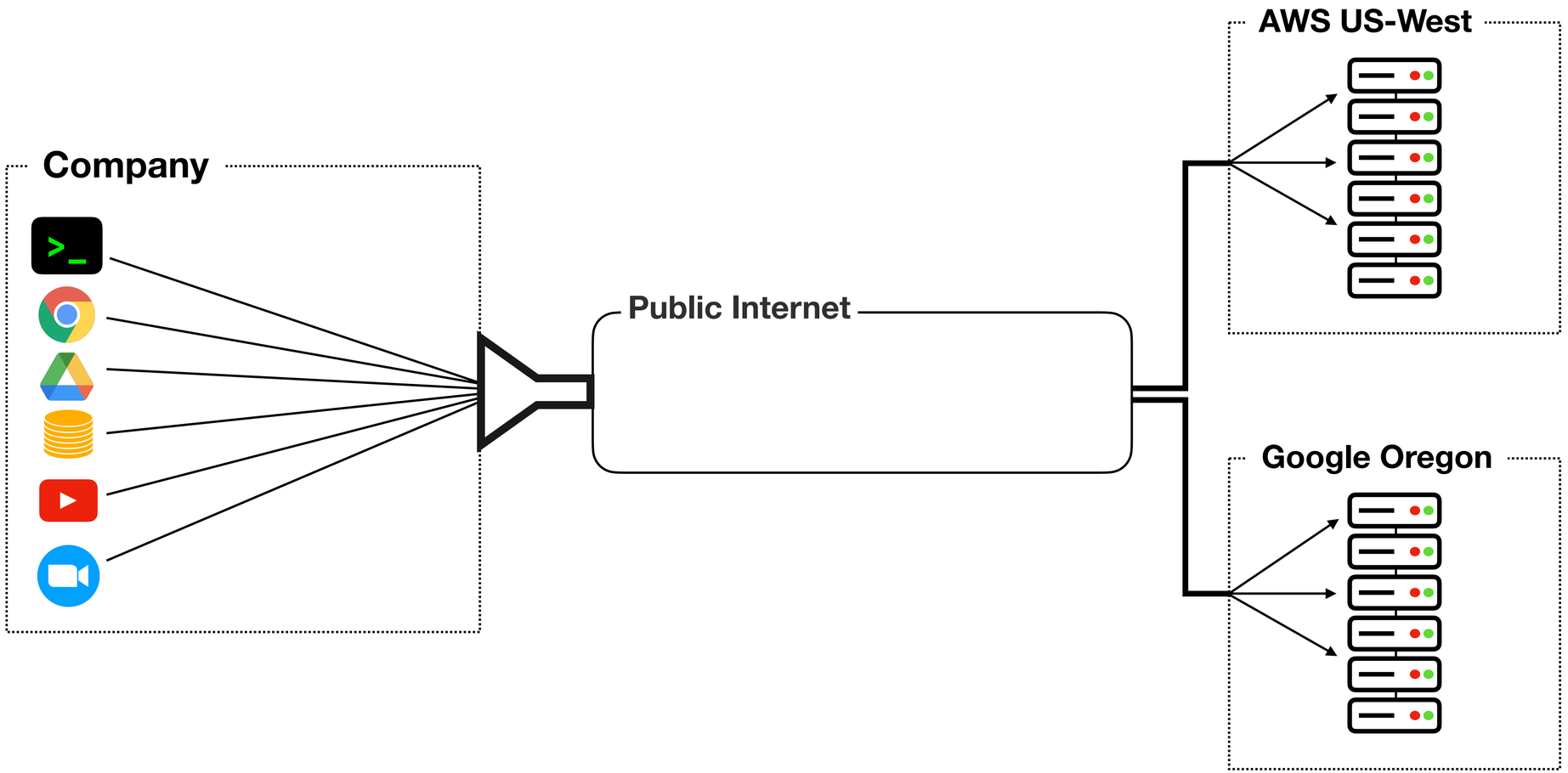
Site-to-Site Internet Traffic Control

“site”: a single physical location with many endpoints sharing an access link to the internet

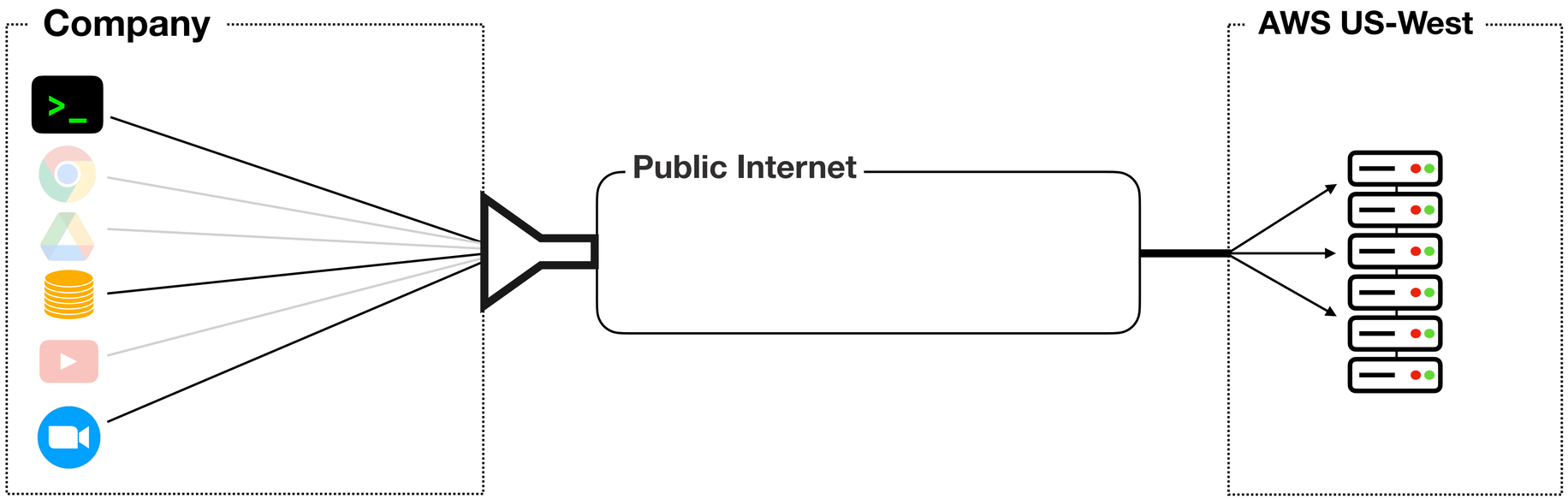


(**Heterogenous** traffic sources and network requirements)

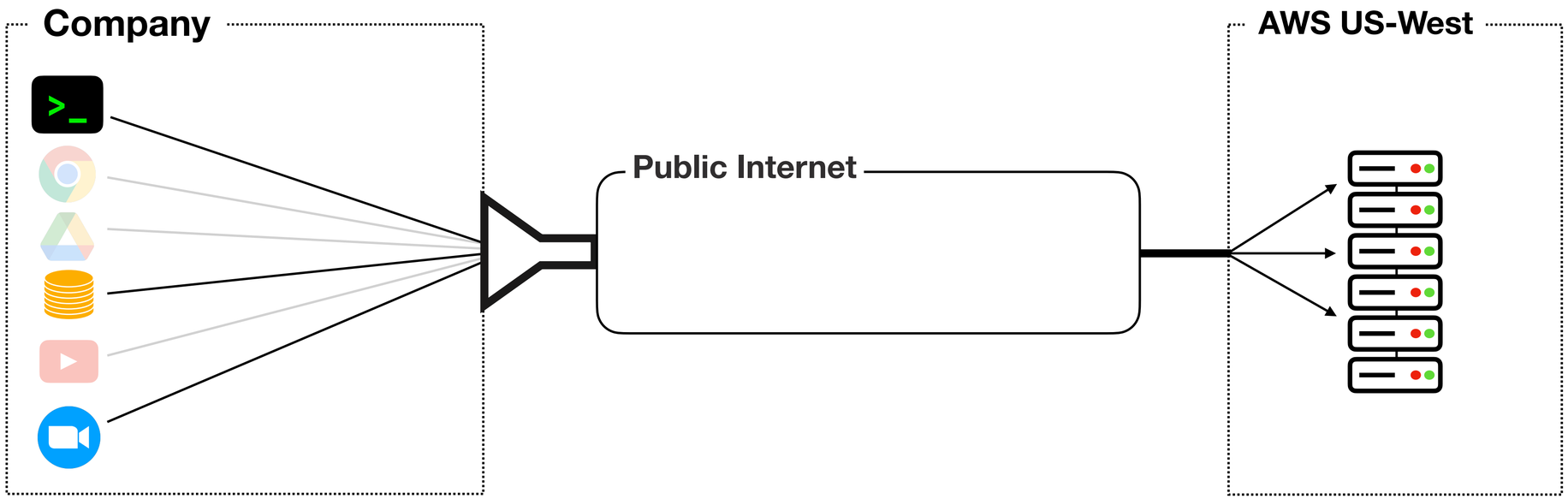
Site-to-Site Internet Traffic Control



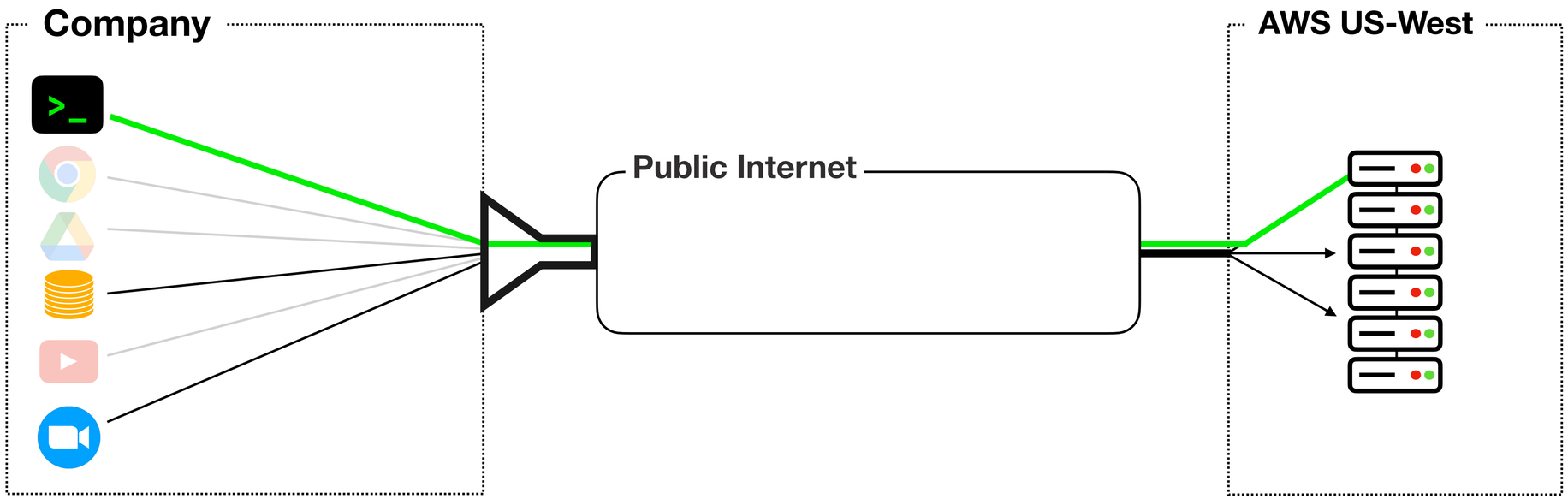
Site-to-Site Internet Traffic Control



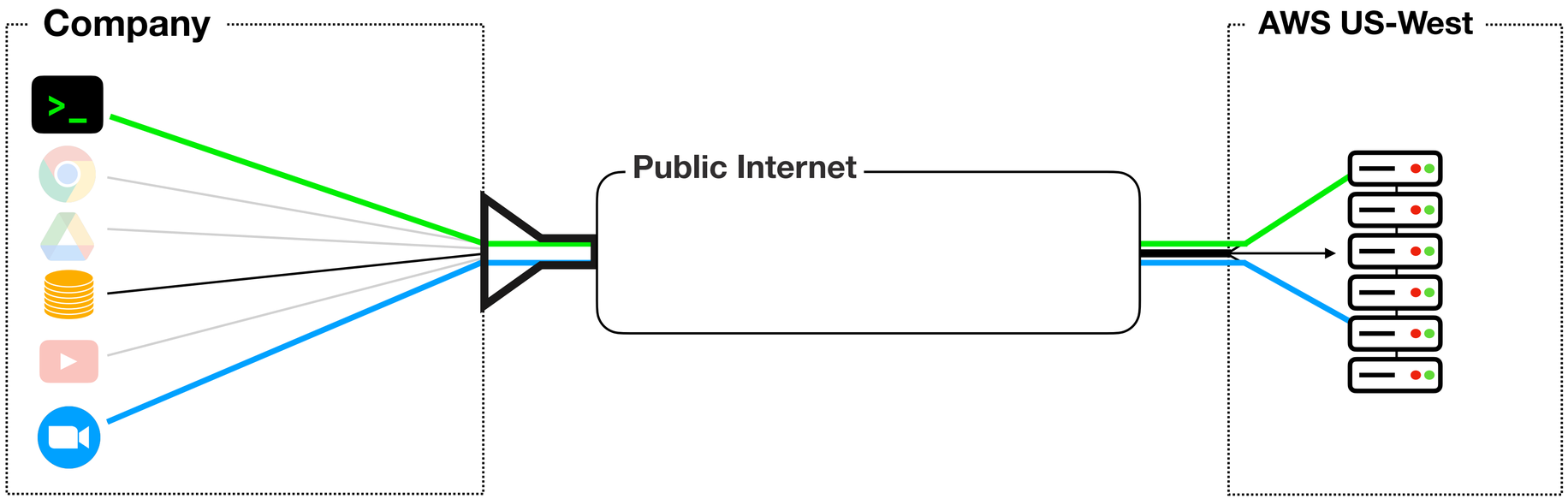
Site-to-Site Internet Traffic Control



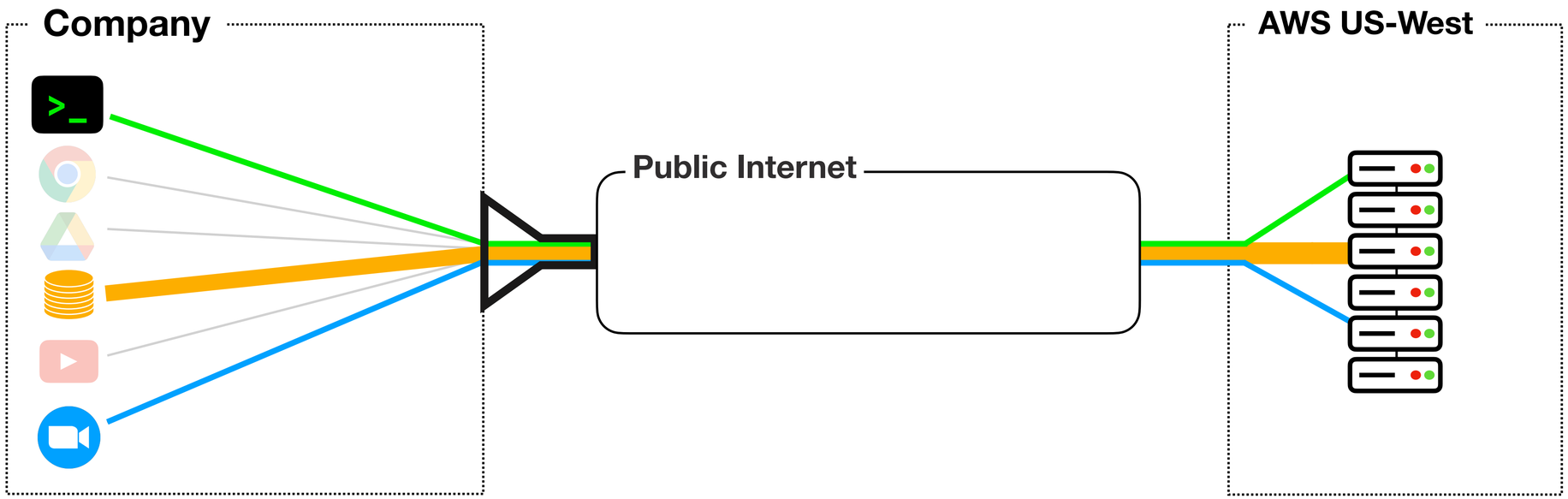
Site-to-Site Internet Traffic Control



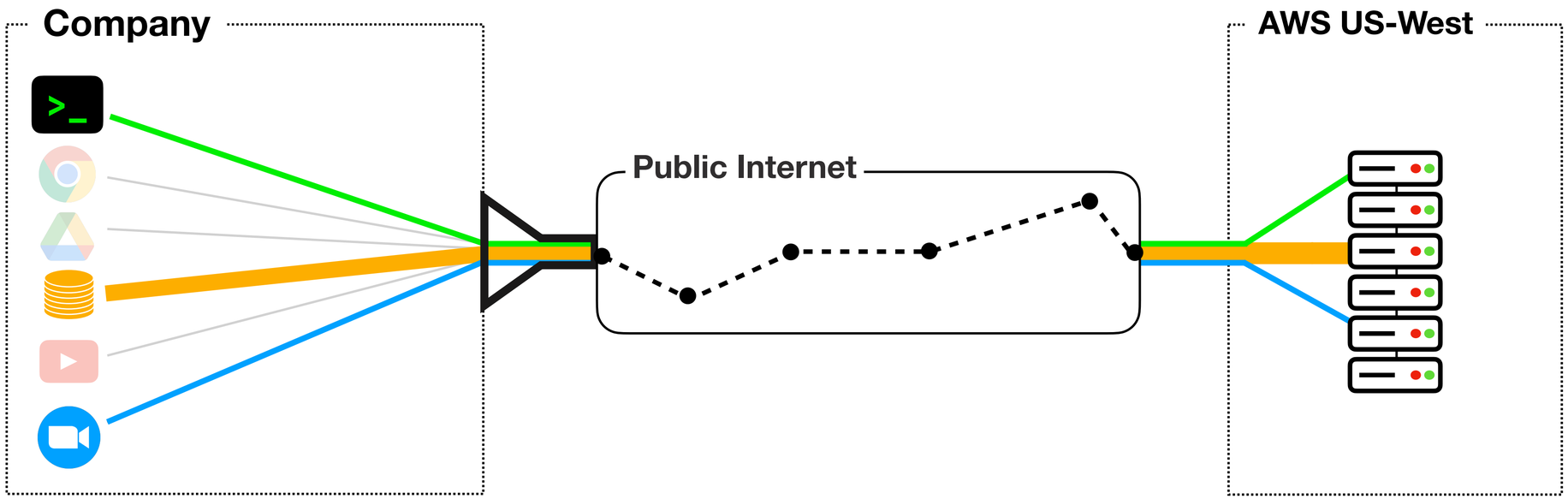
Site-to-Site Internet Traffic Control



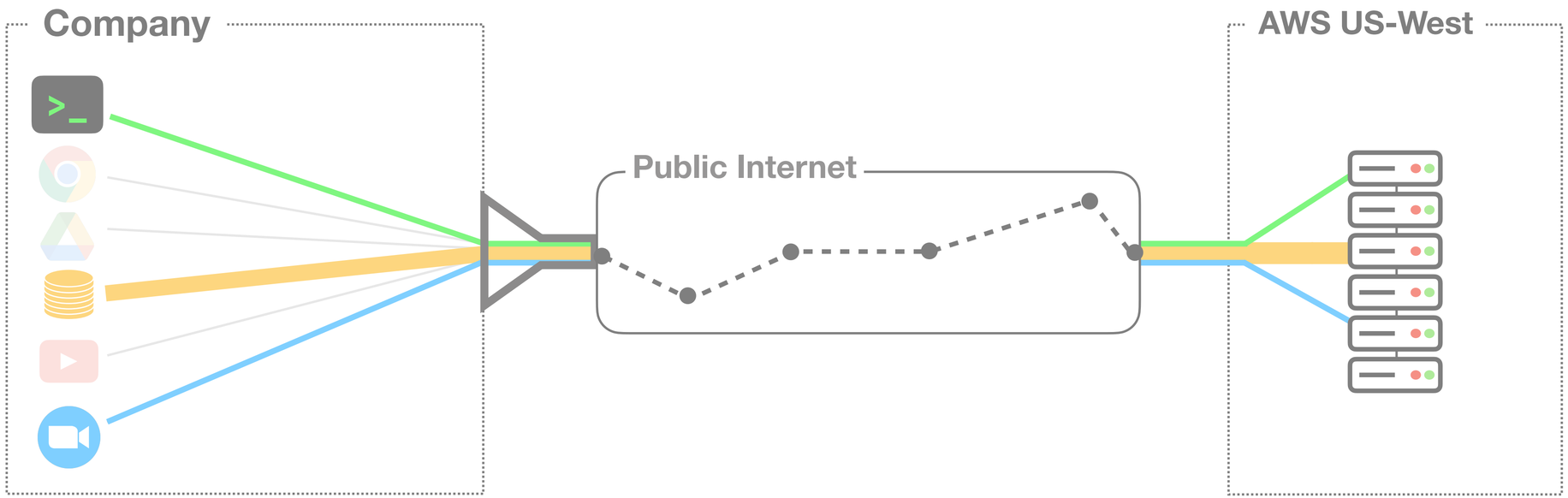
Site-to-Site Internet Traffic Control



Site-to-Site Internet Traffic Control



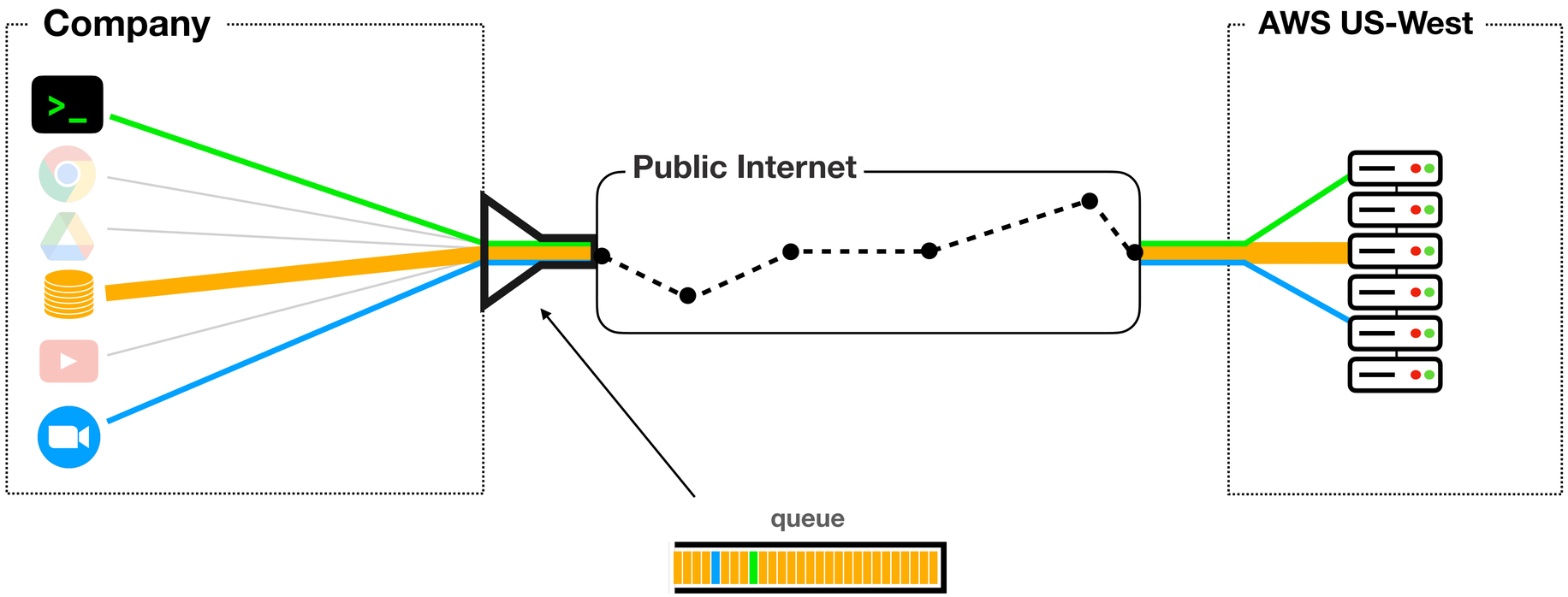
Site-to-Site Internet Traffic Control



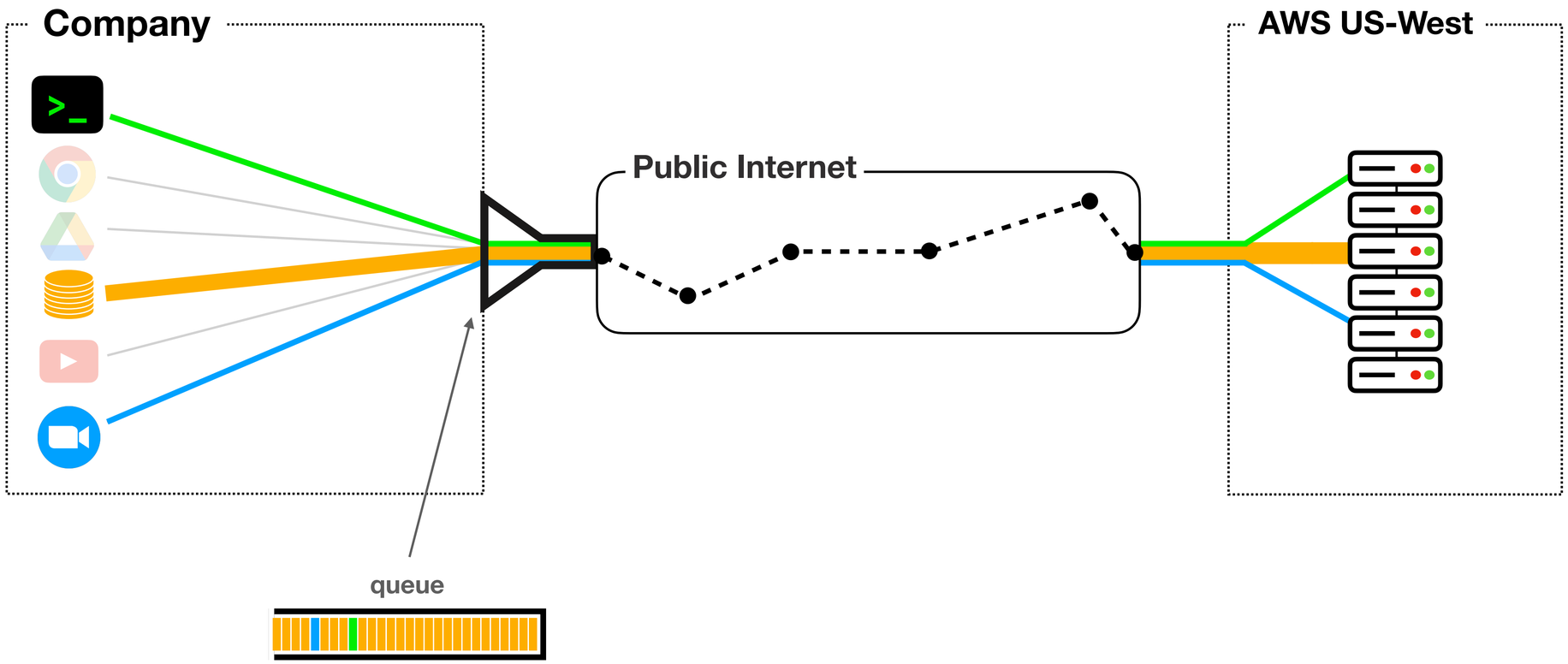
Bottleneck queue

Latency-sensitive apps stuck waiting

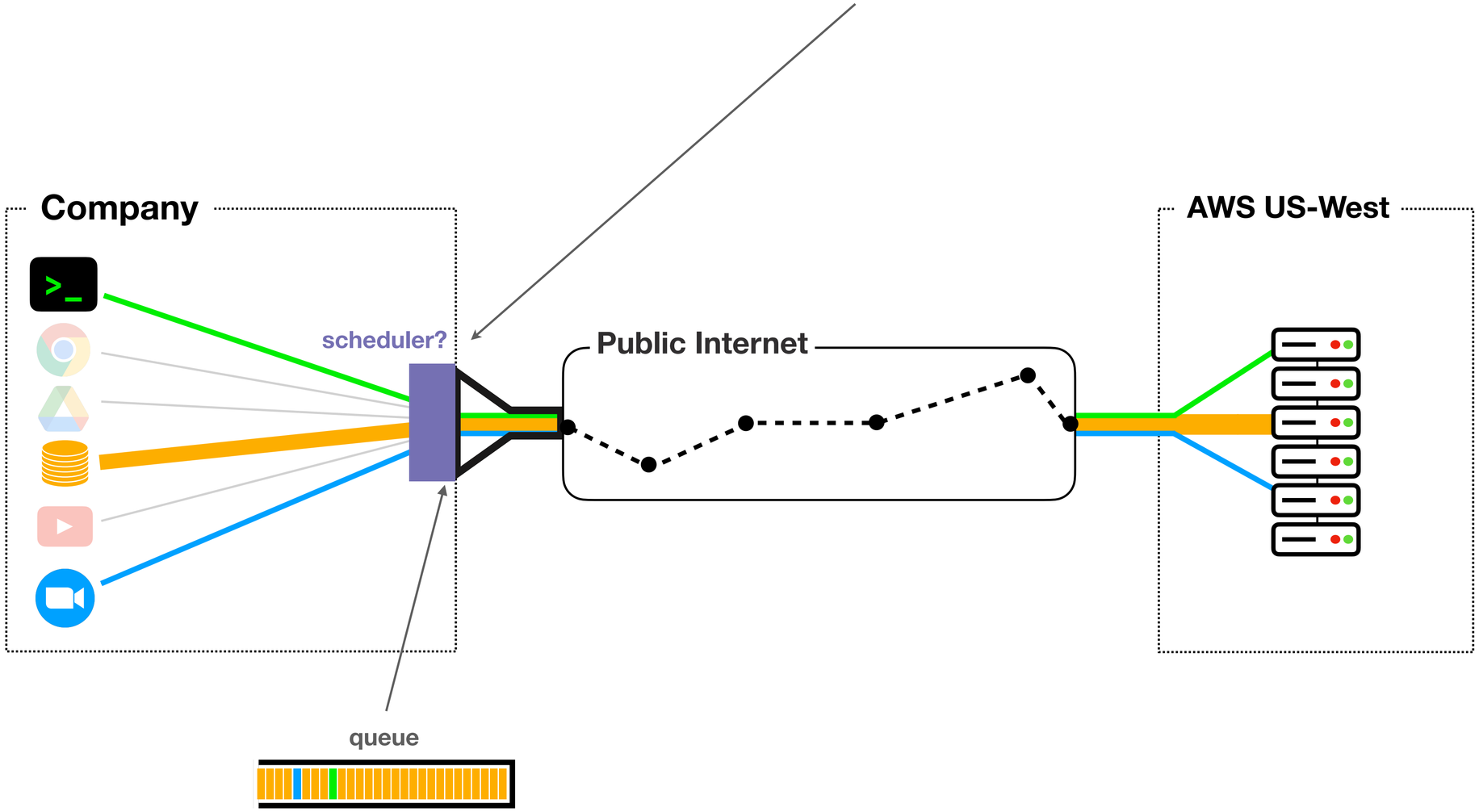
Site-to-Site **Internet Traffic Control**



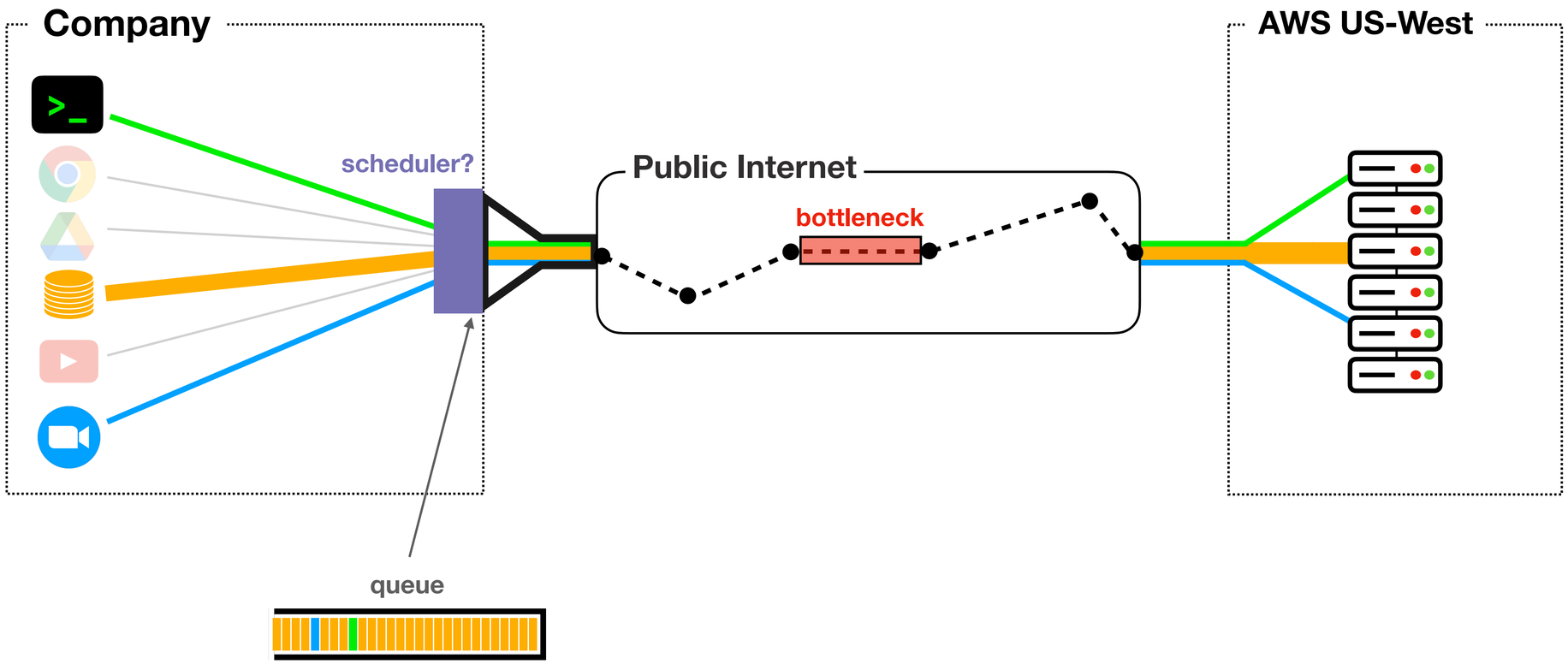
Site-to-Site **Internet Traffic Control**



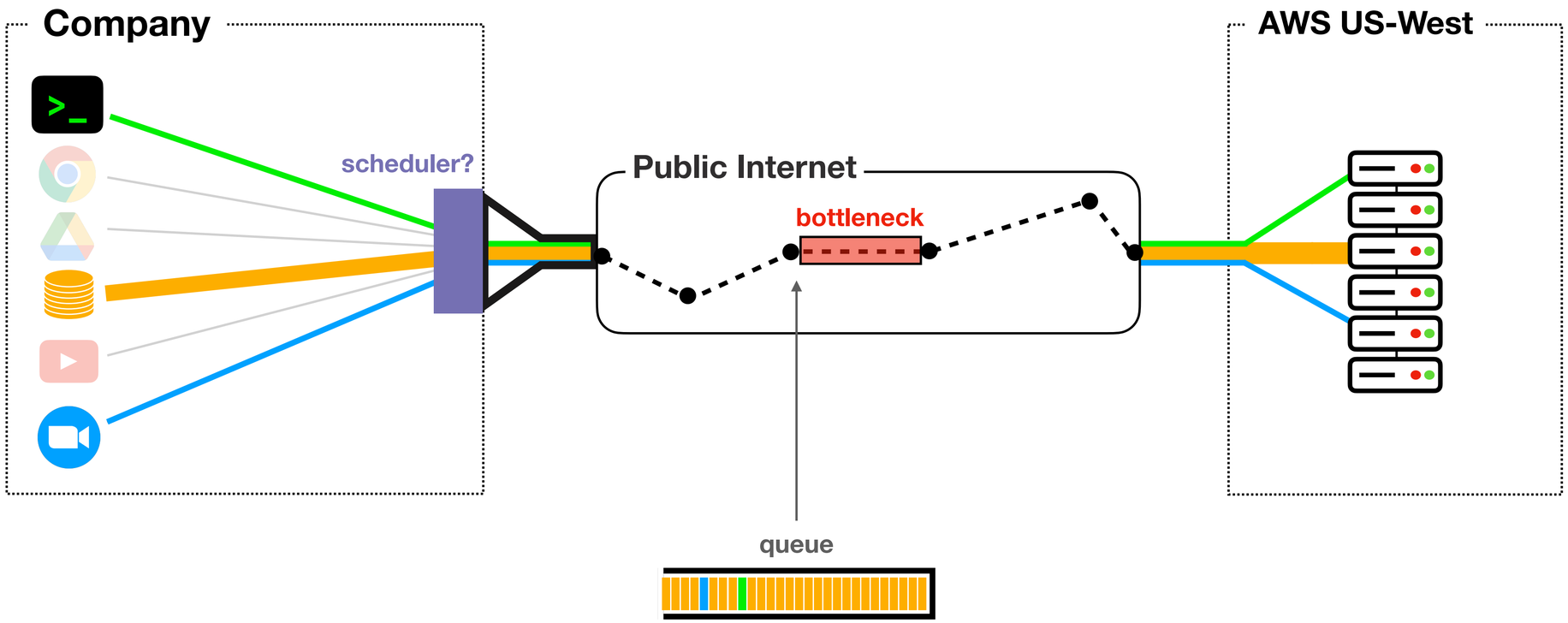
Site-to-Site **Internet Traffic Control**



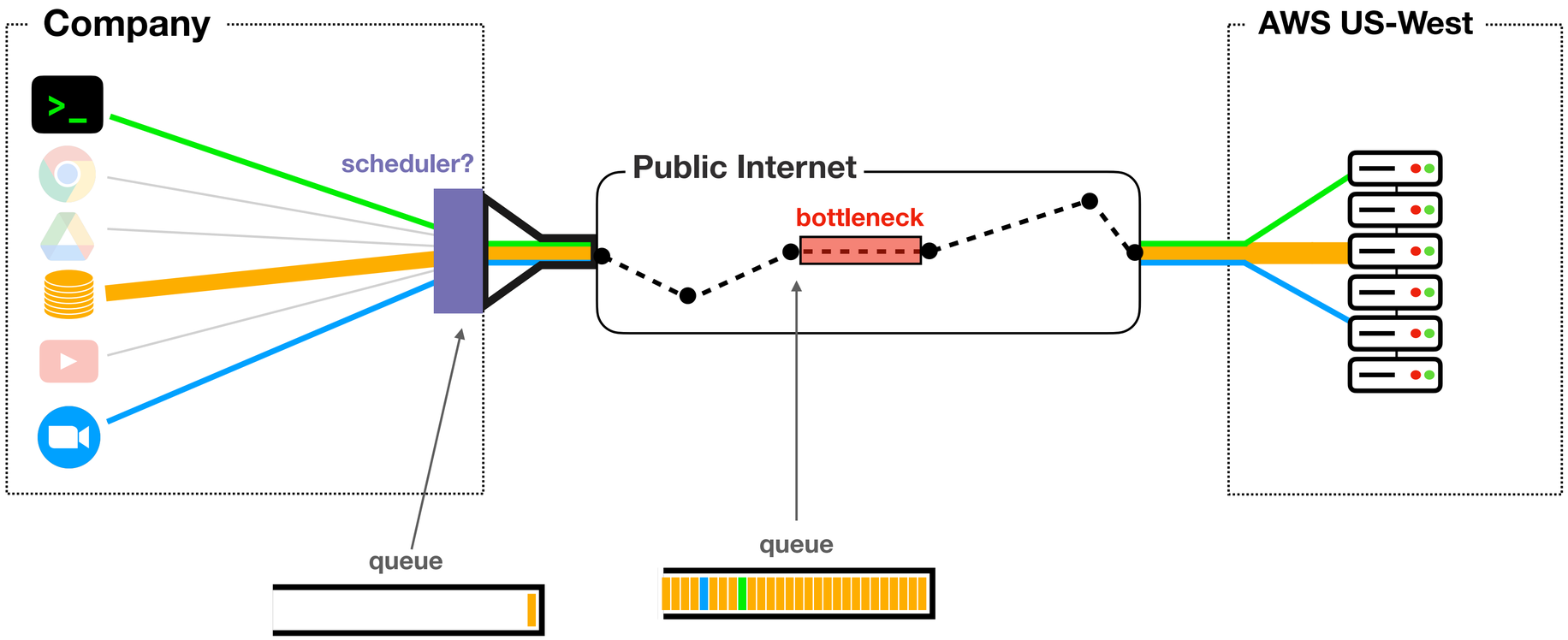
Site-to-Site **Internet Traffic Control**



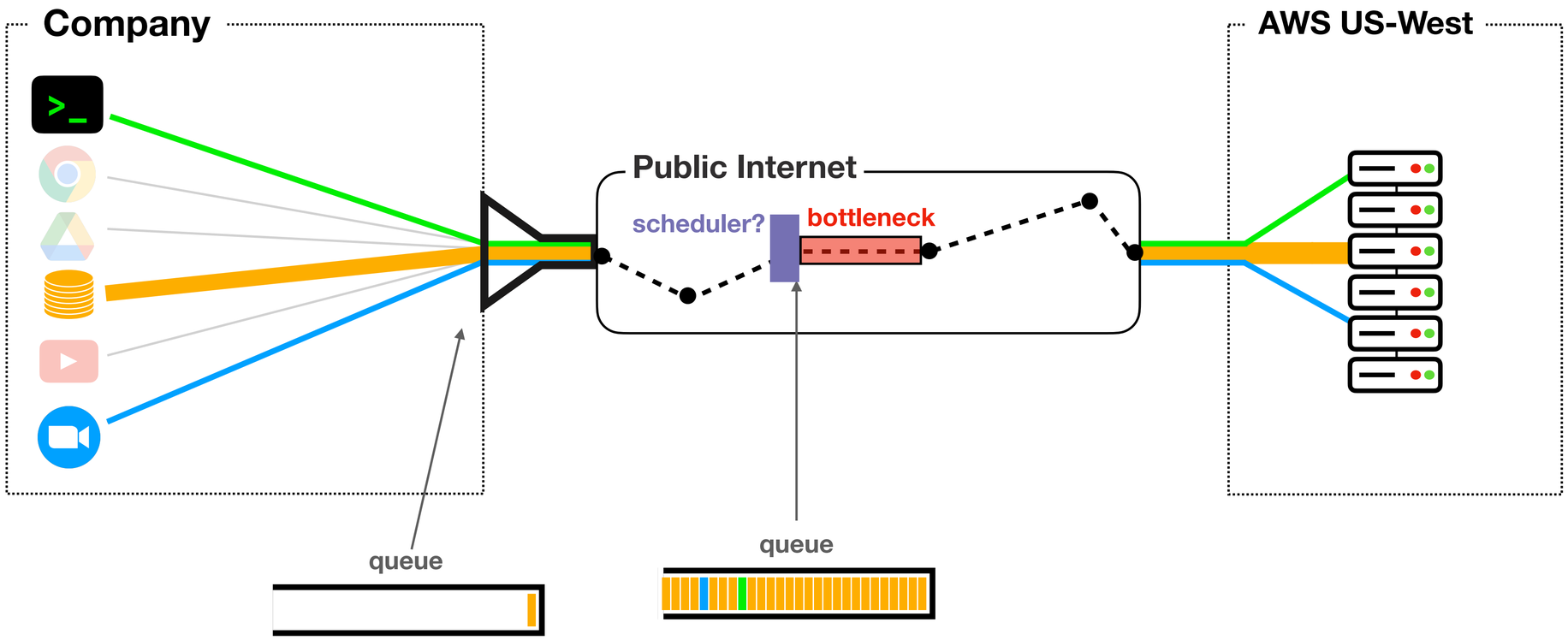
Site-to-Site **Internet Traffic Control**



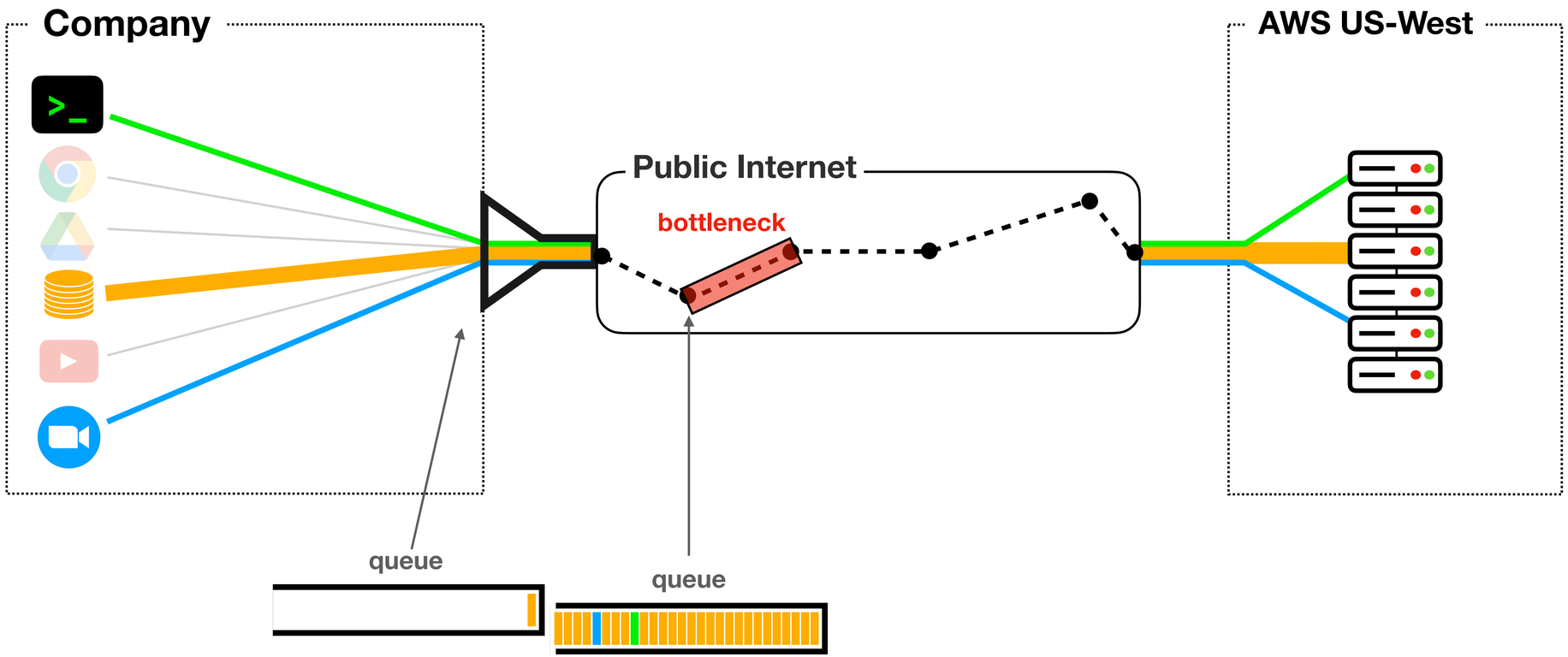
Site-to-Site Internet Traffic Control



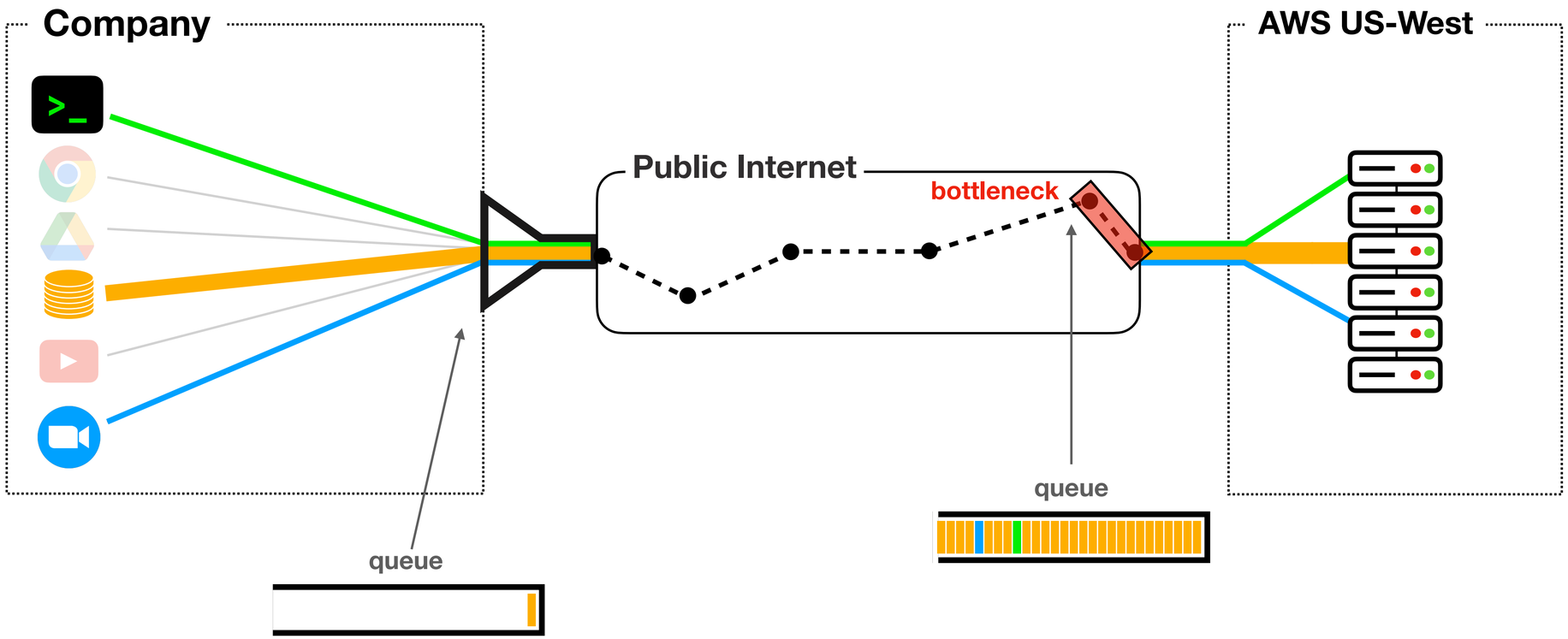
Site-to-Site Internet Traffic Control



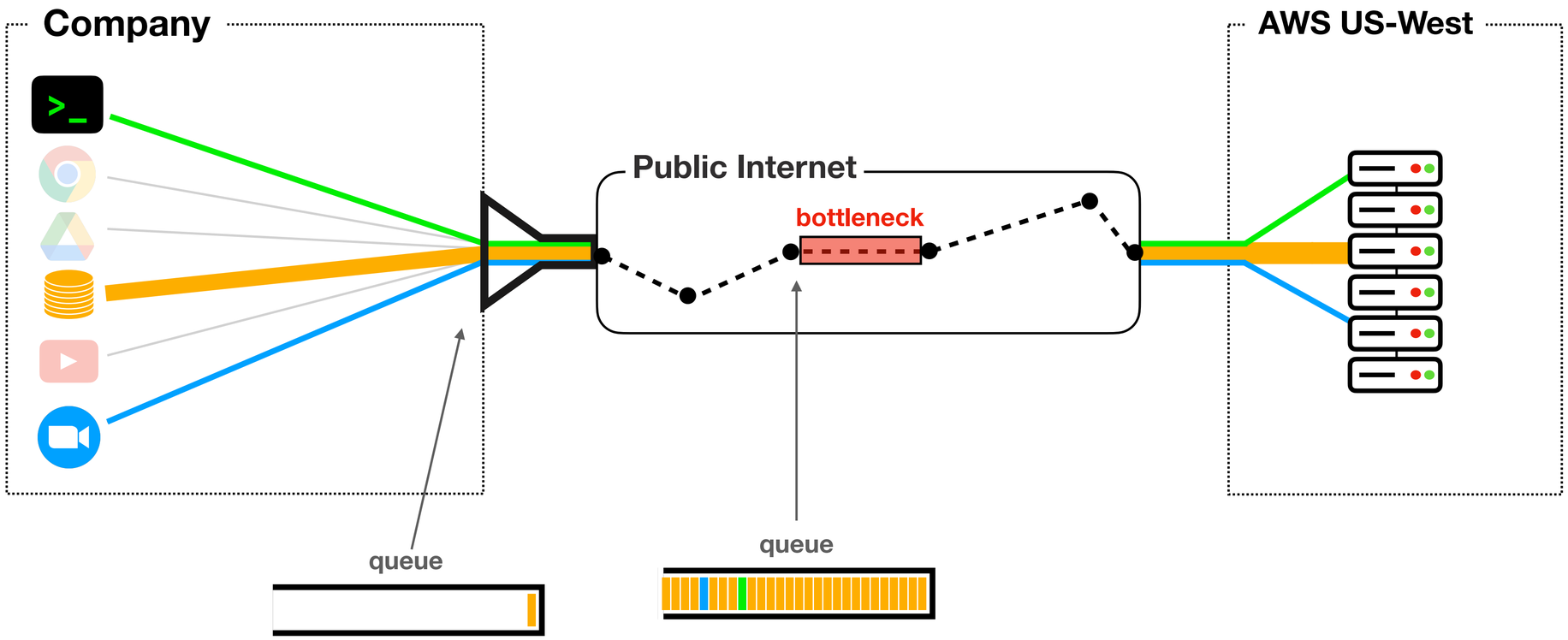
Site-to-Site Internet Traffic Control



Site-to-Site Internet Traffic Control



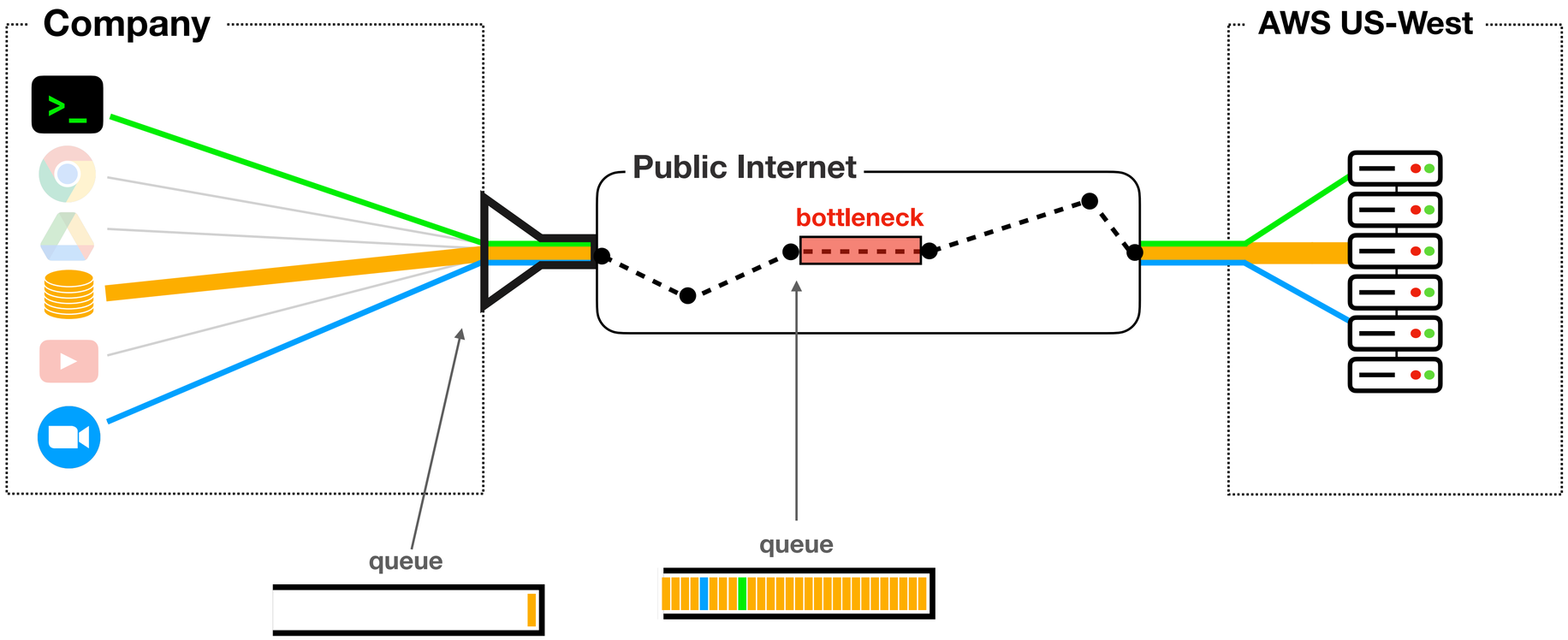
Site-to-Site Internet Traffic Control



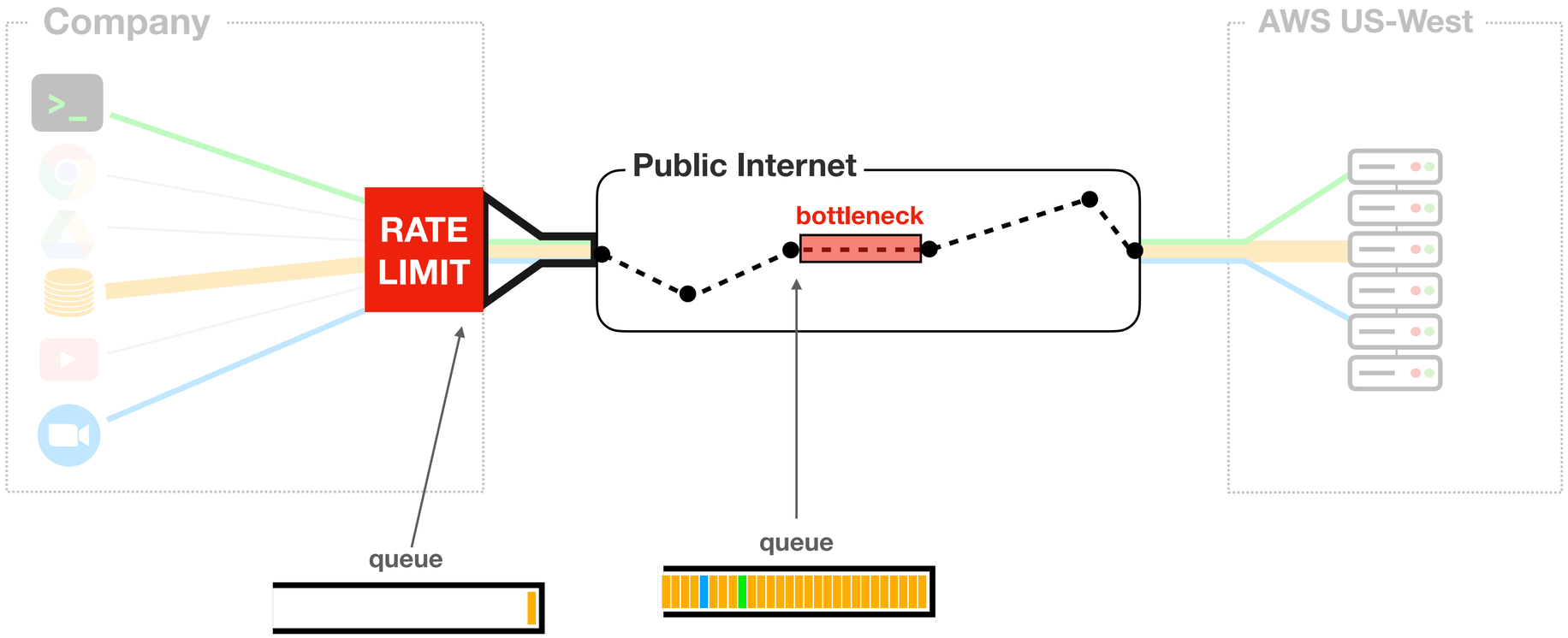
Site-to-Site Internet Traffic Control

Problem: want to enforce scheduling policy for your traffic,
but often don't control the bottleneck where packets queue.

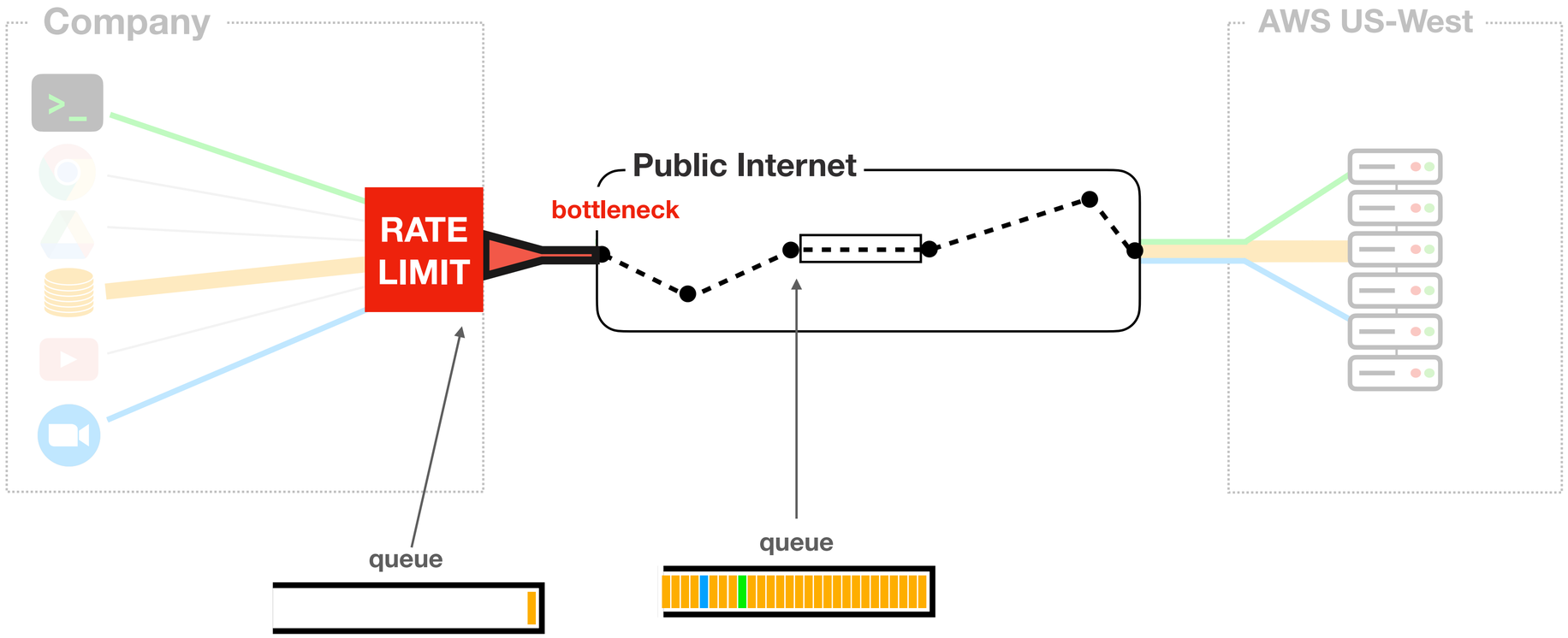
Key Idea: “shifting” the queues



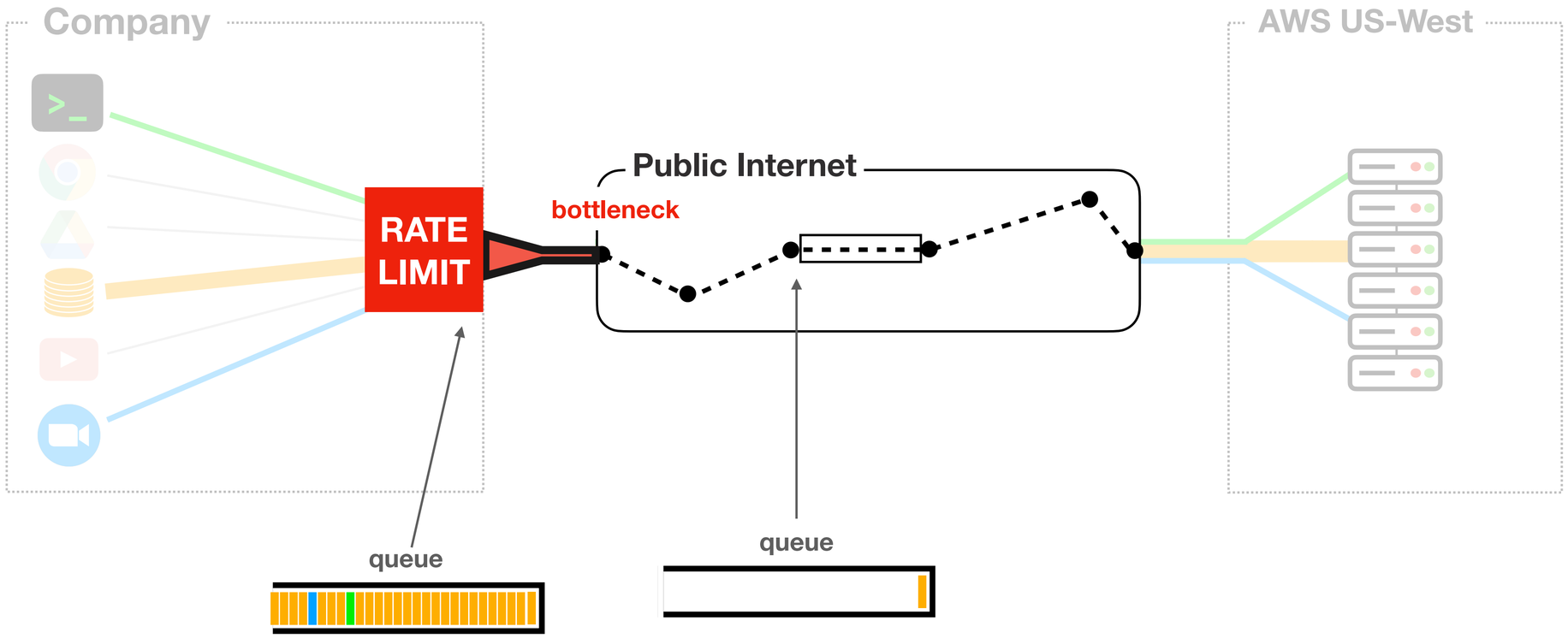
Key Idea: “shifting” the queues



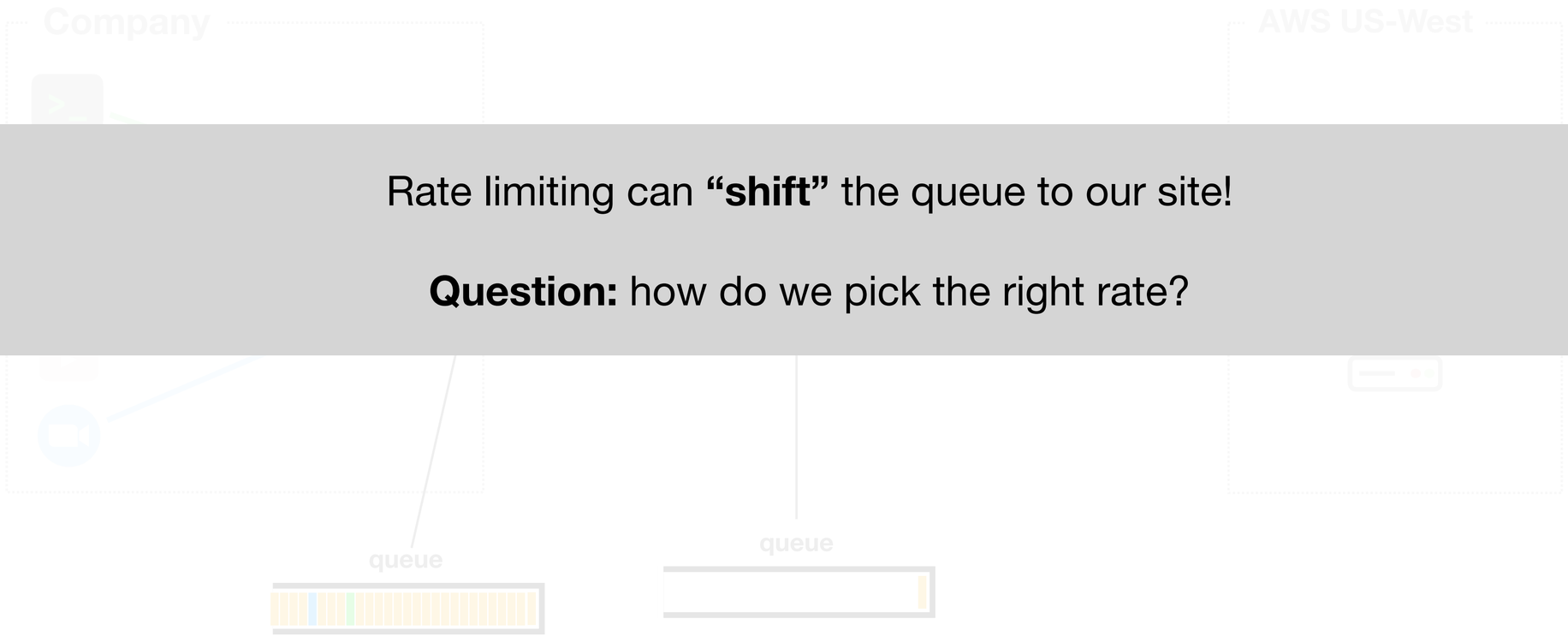
Key Idea: “shifting” the queues



Key Idea: “shifting” the queues



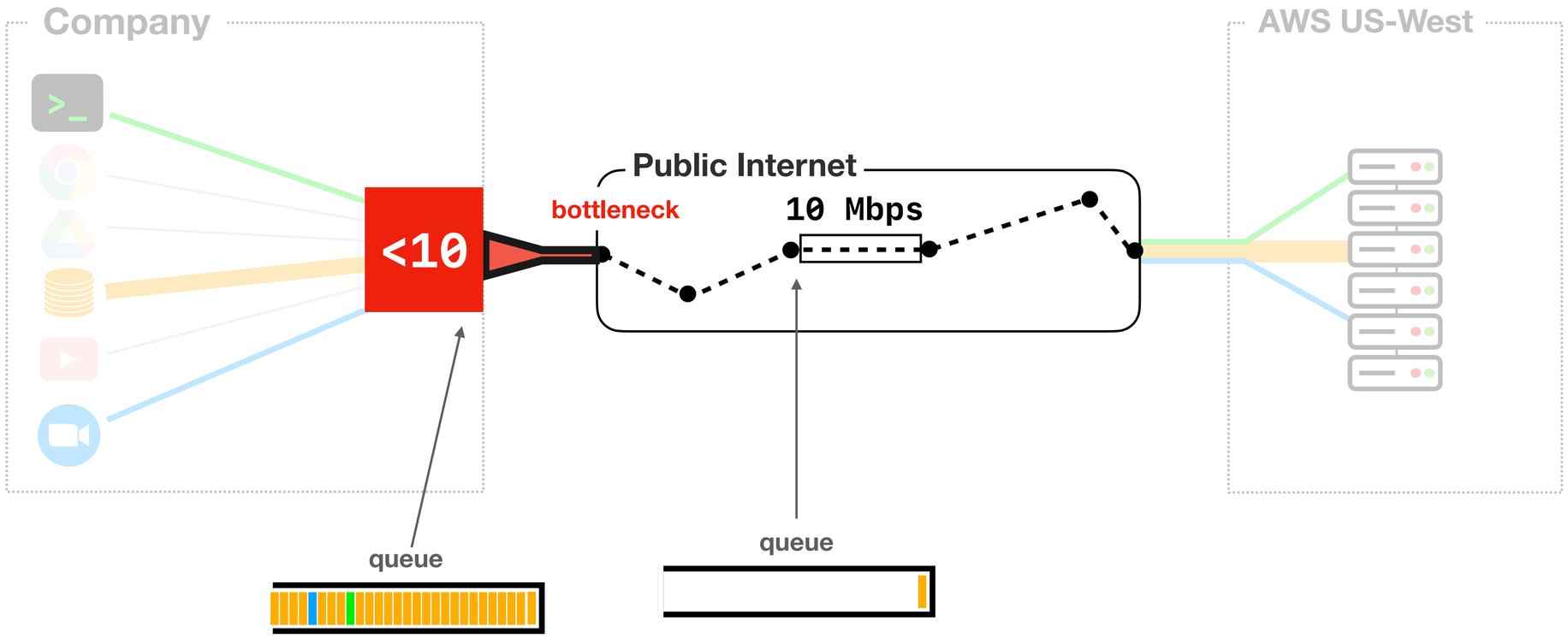
Key Idea: “shifting” the queues



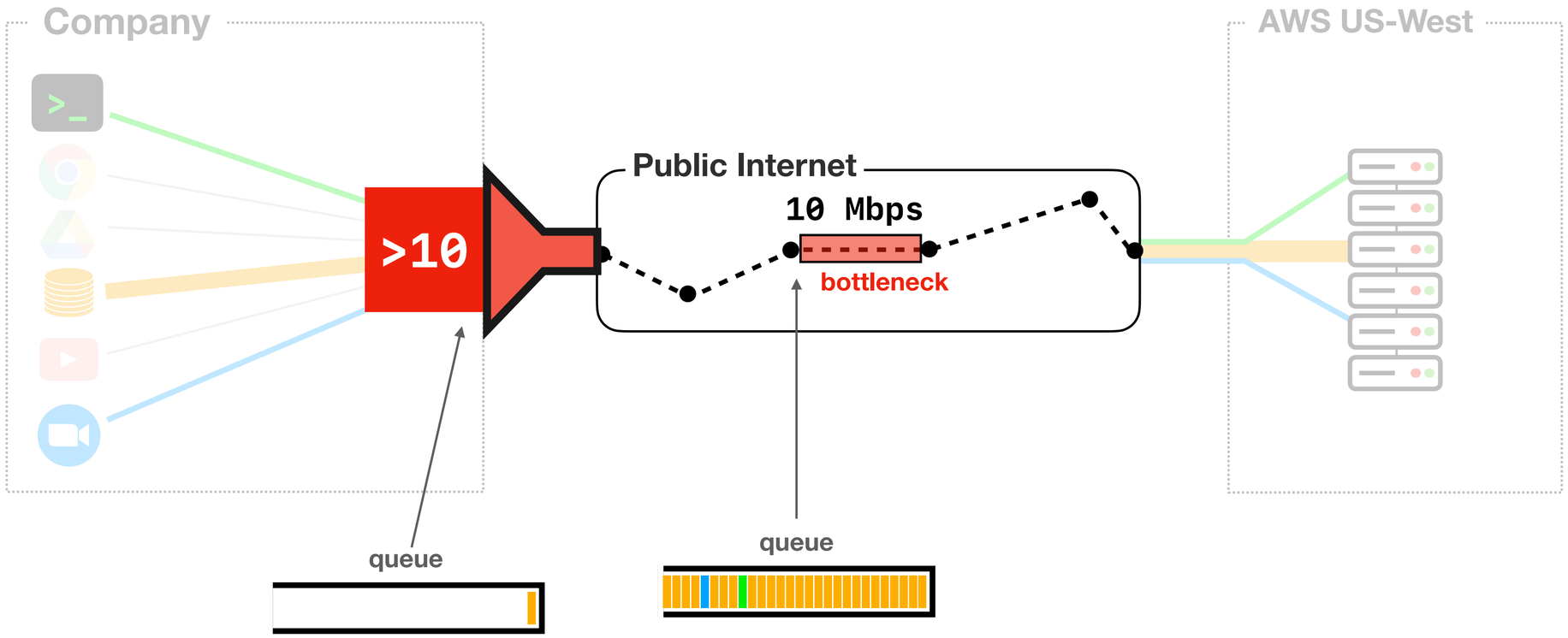
Rate limiting can **“shift”** the queue to our site!

Question: how do we pick the right rate?

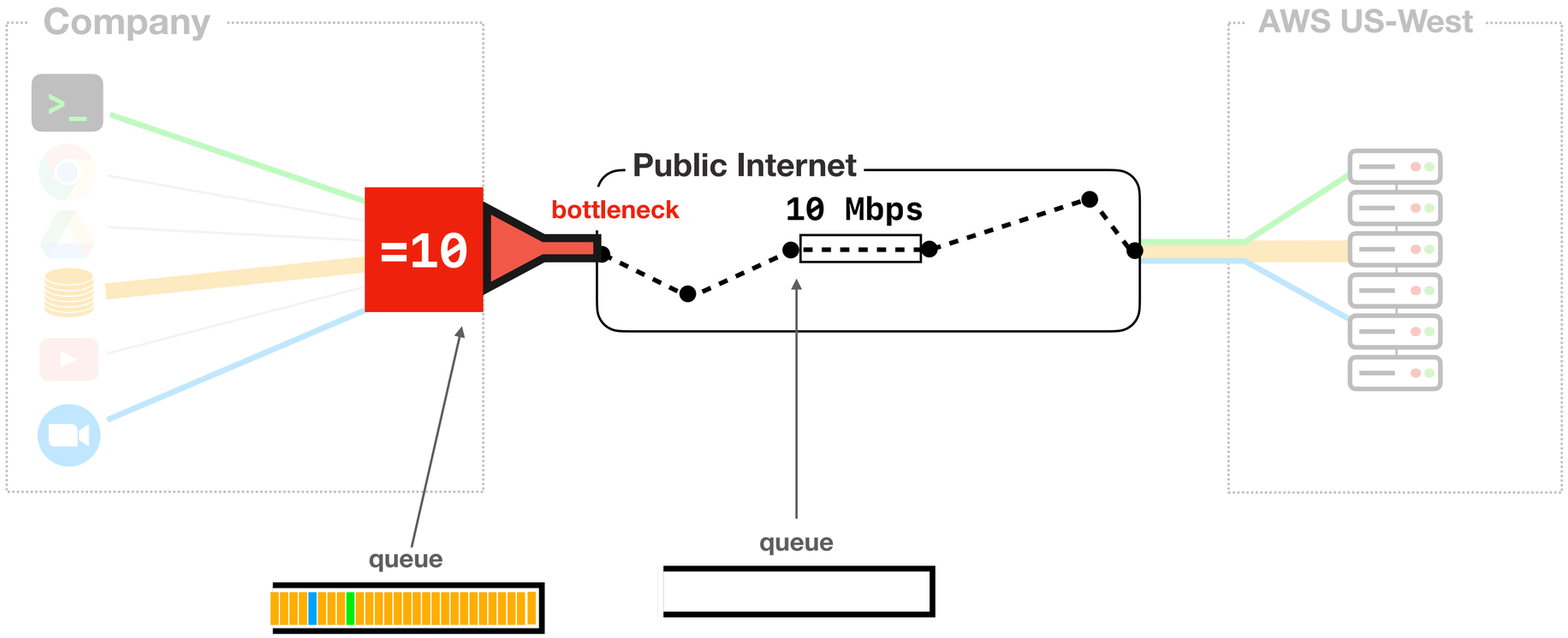
Key Idea: “shifting” the queues



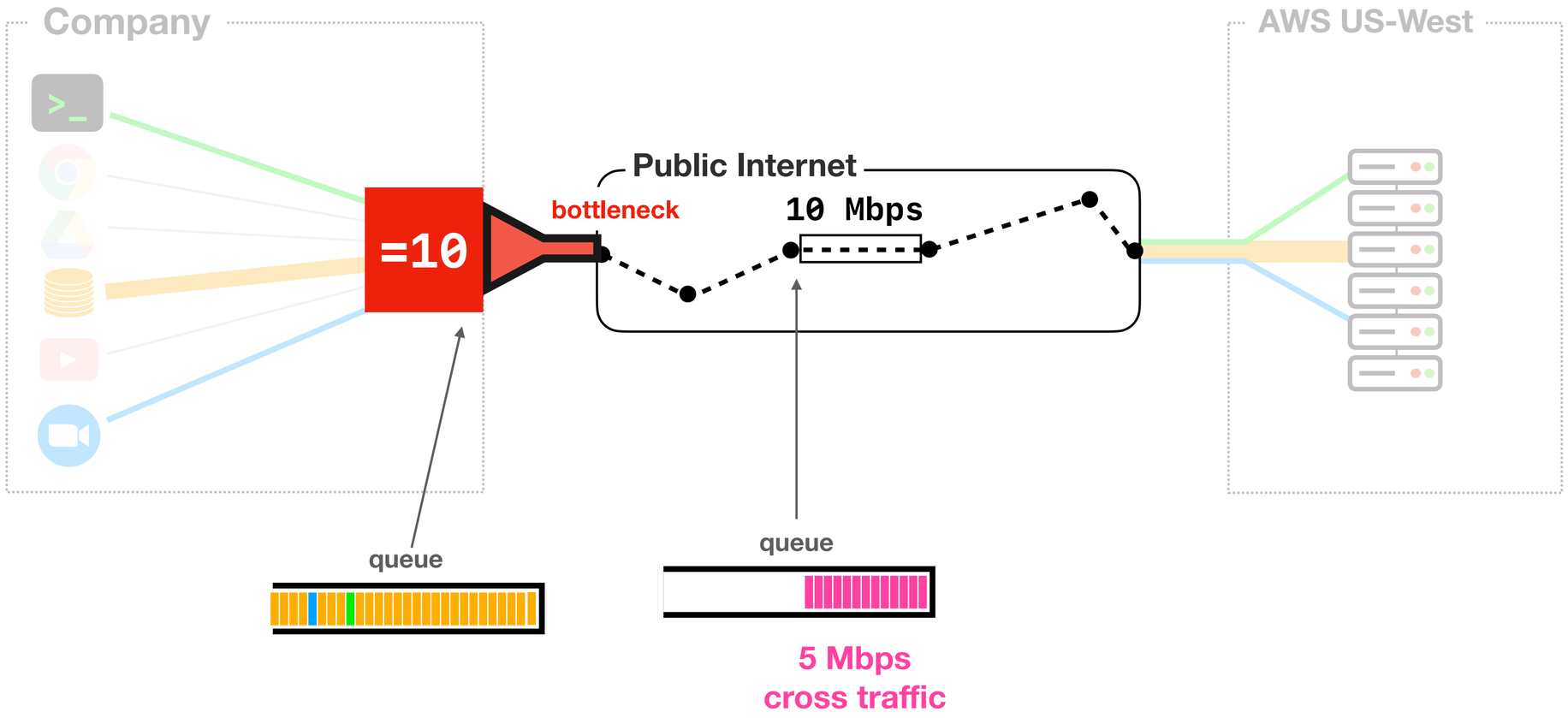
Key Idea: “shifting” the queues



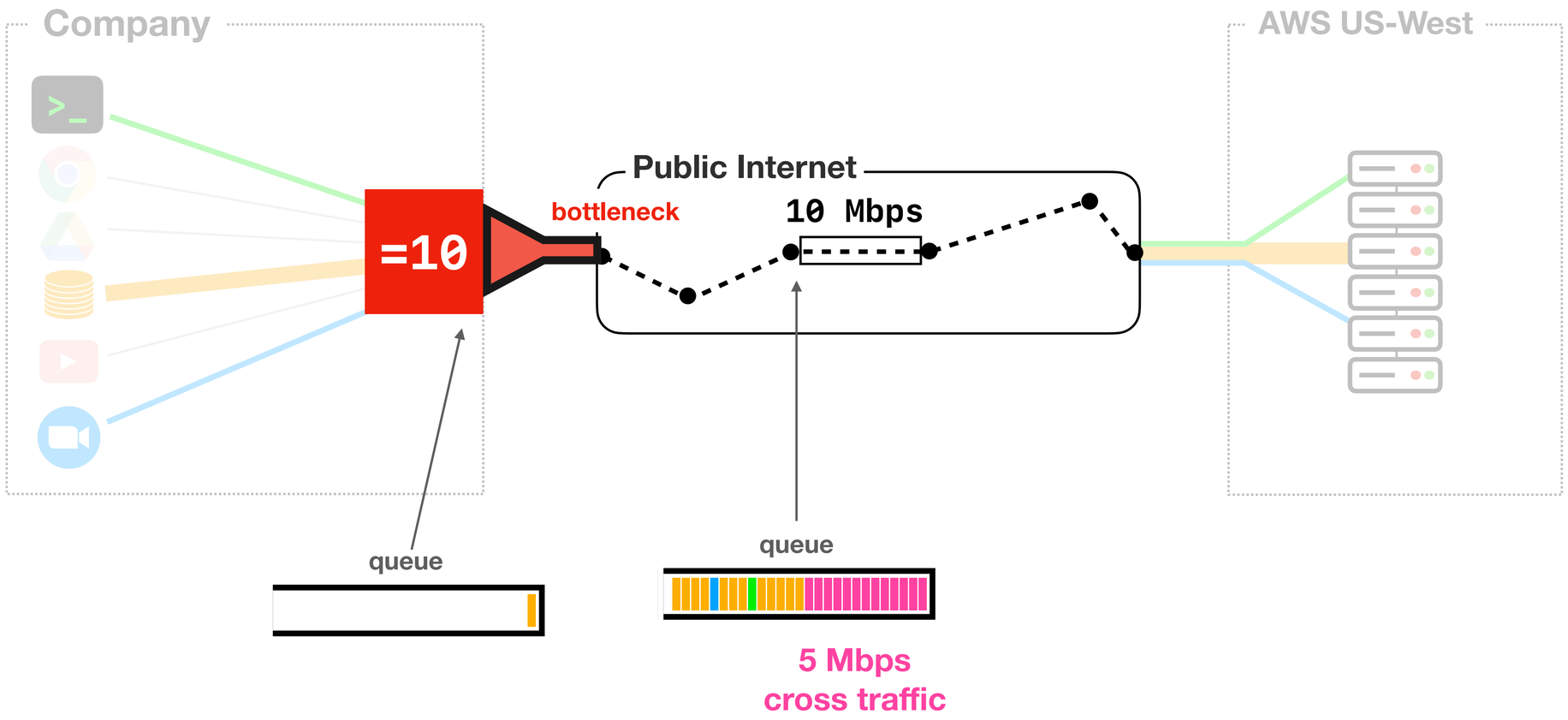
Key Idea: “shifting” the queues



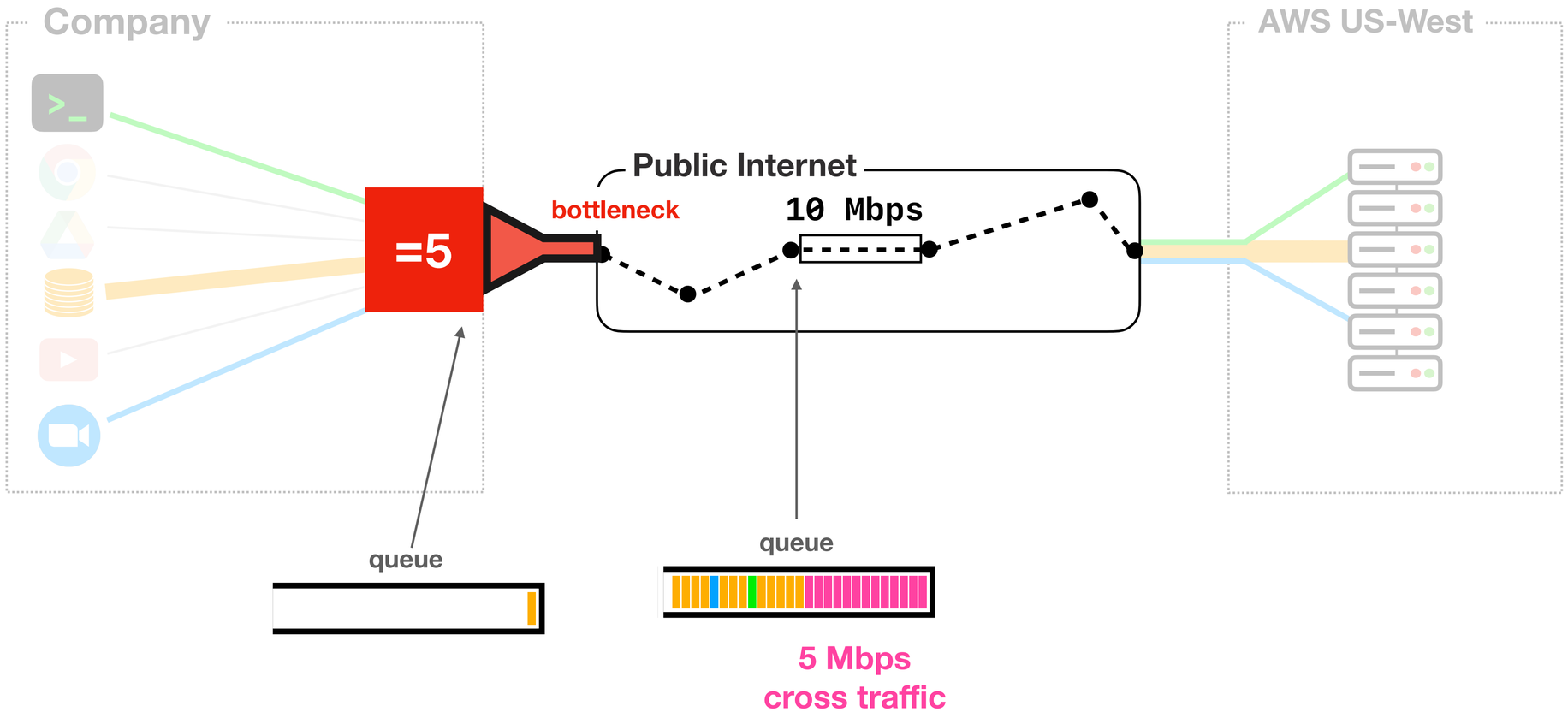
Key Idea: “shifting” the queues



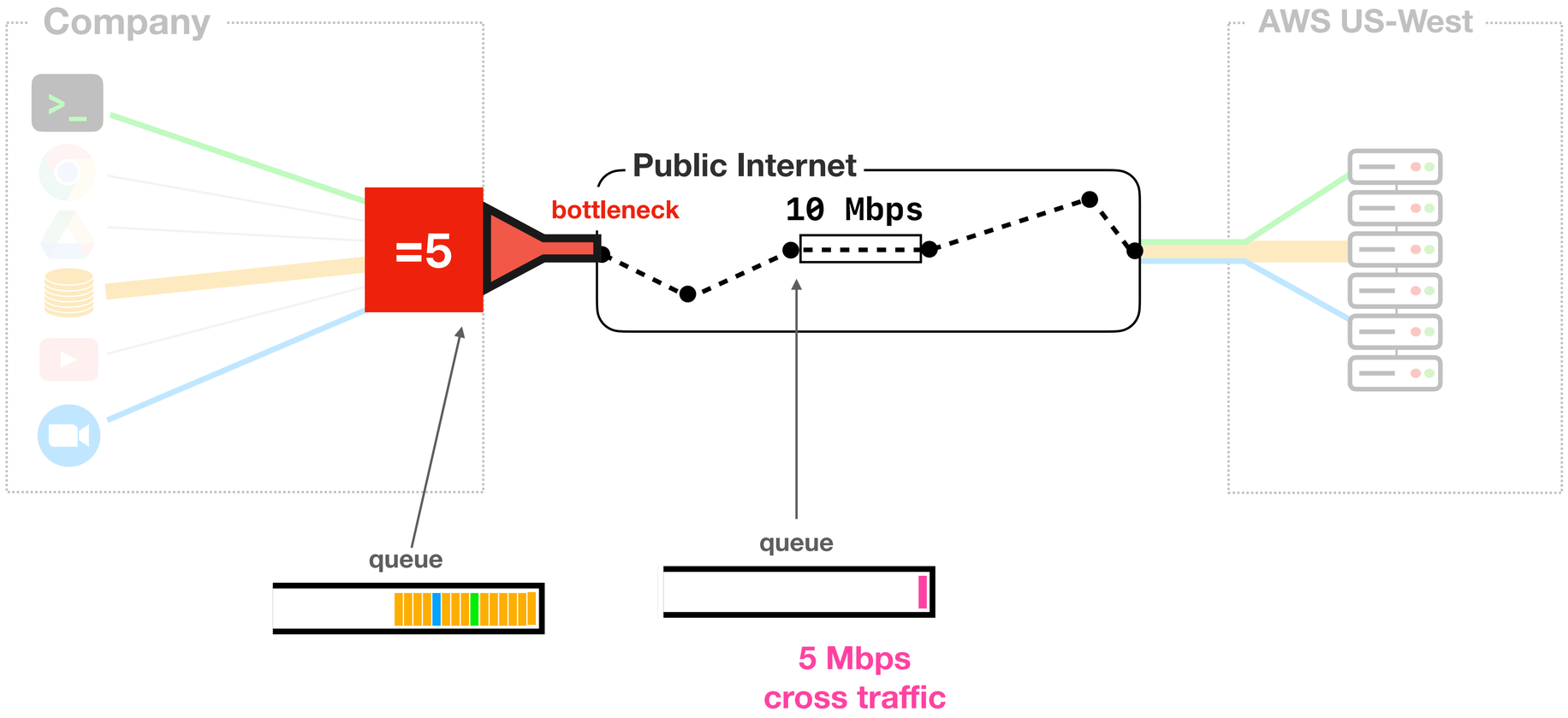
Key Idea: “shifting” the queues



Key Idea: “shifting” the queues

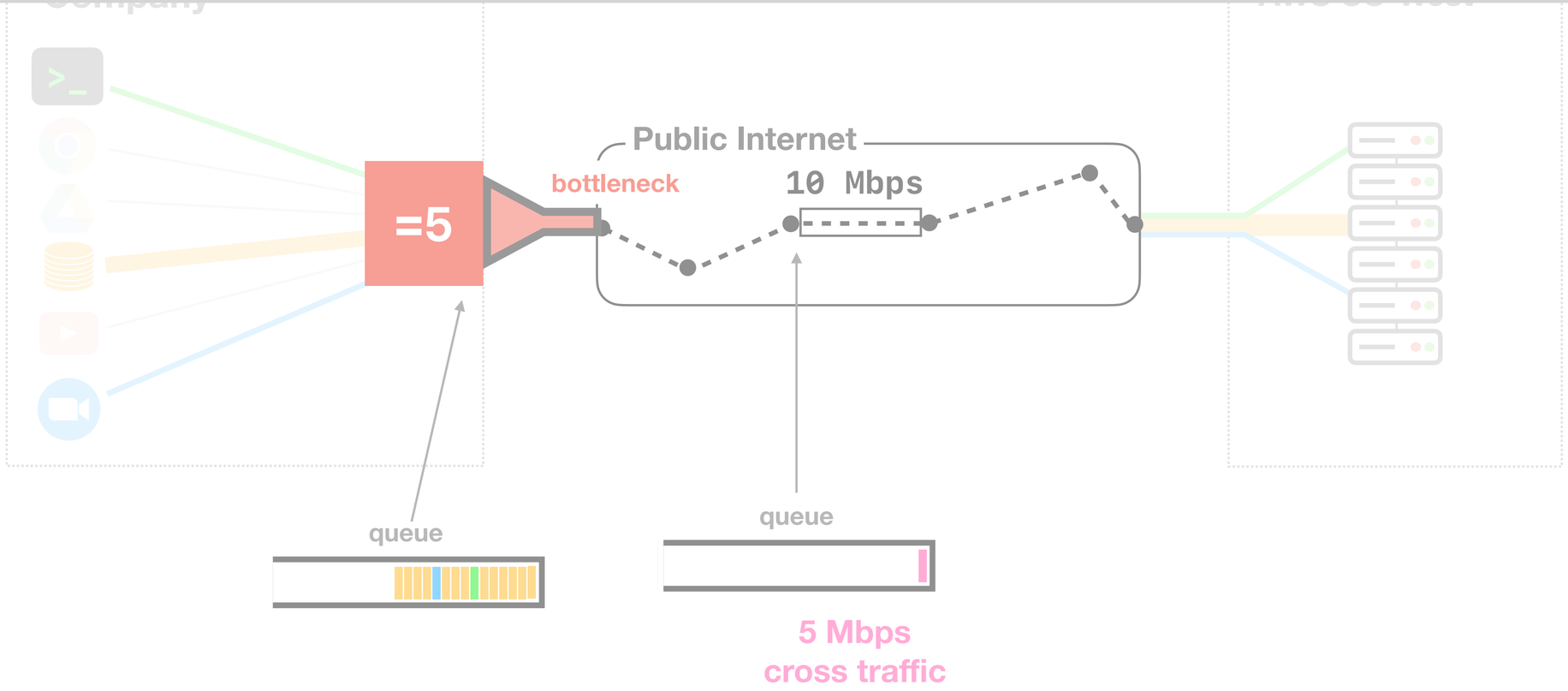


Key Idea: “shifting” the queues

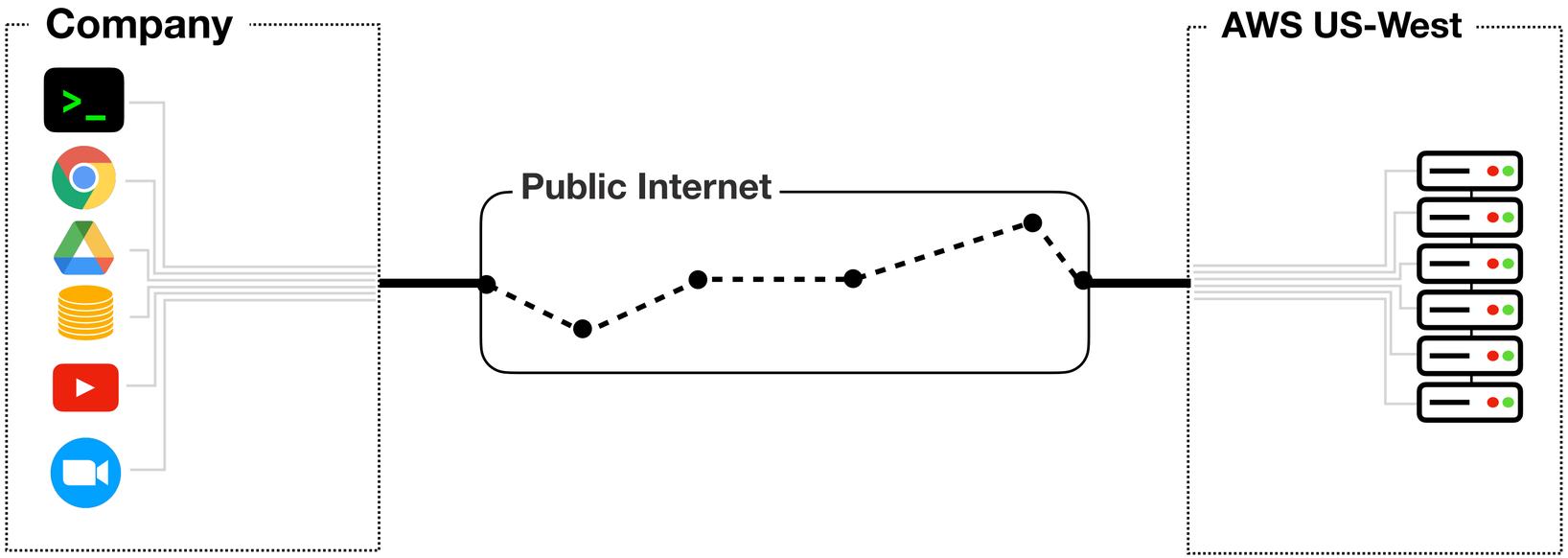


Key Idea: “shifting” the queues

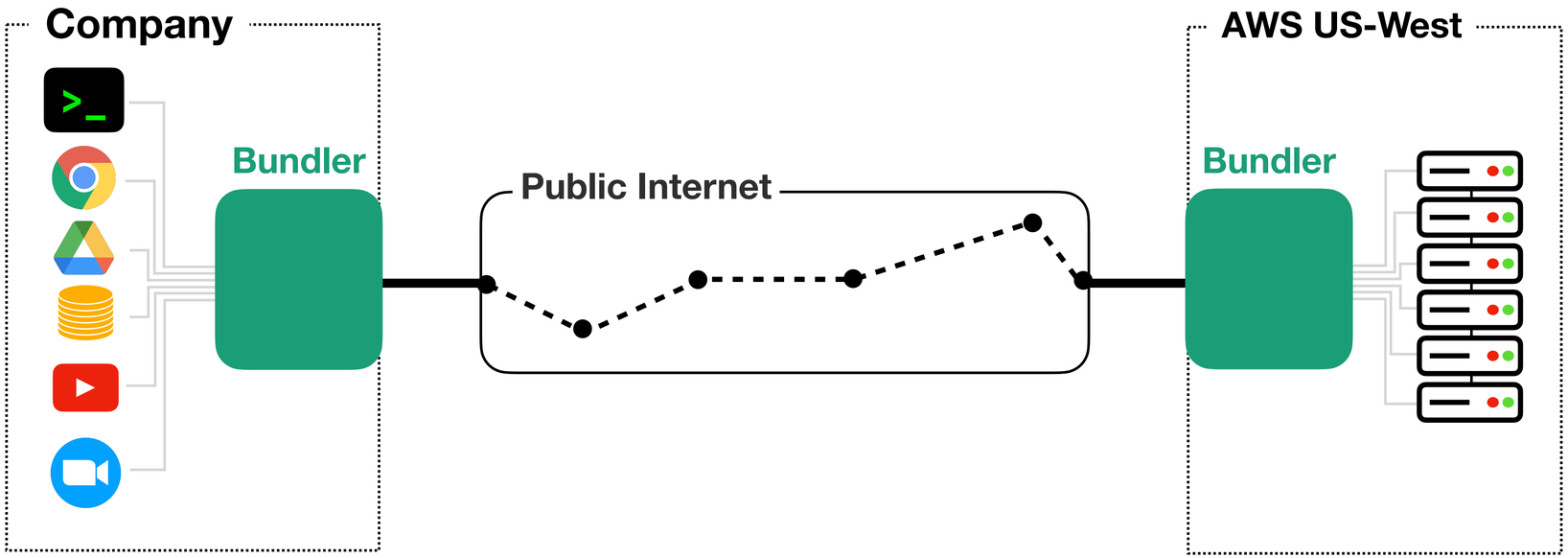
Congestion Control algorithms
aim to calculate exactly the rate we need!



Bundler's Architecture

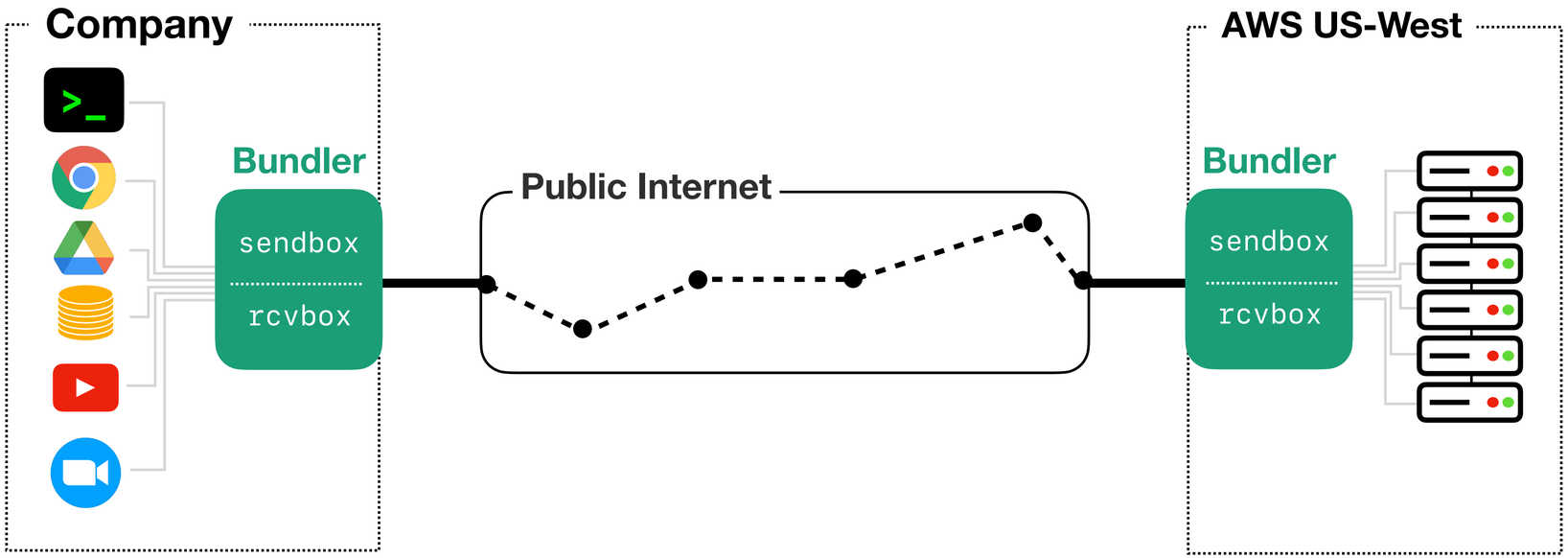


Bundler's Architecture



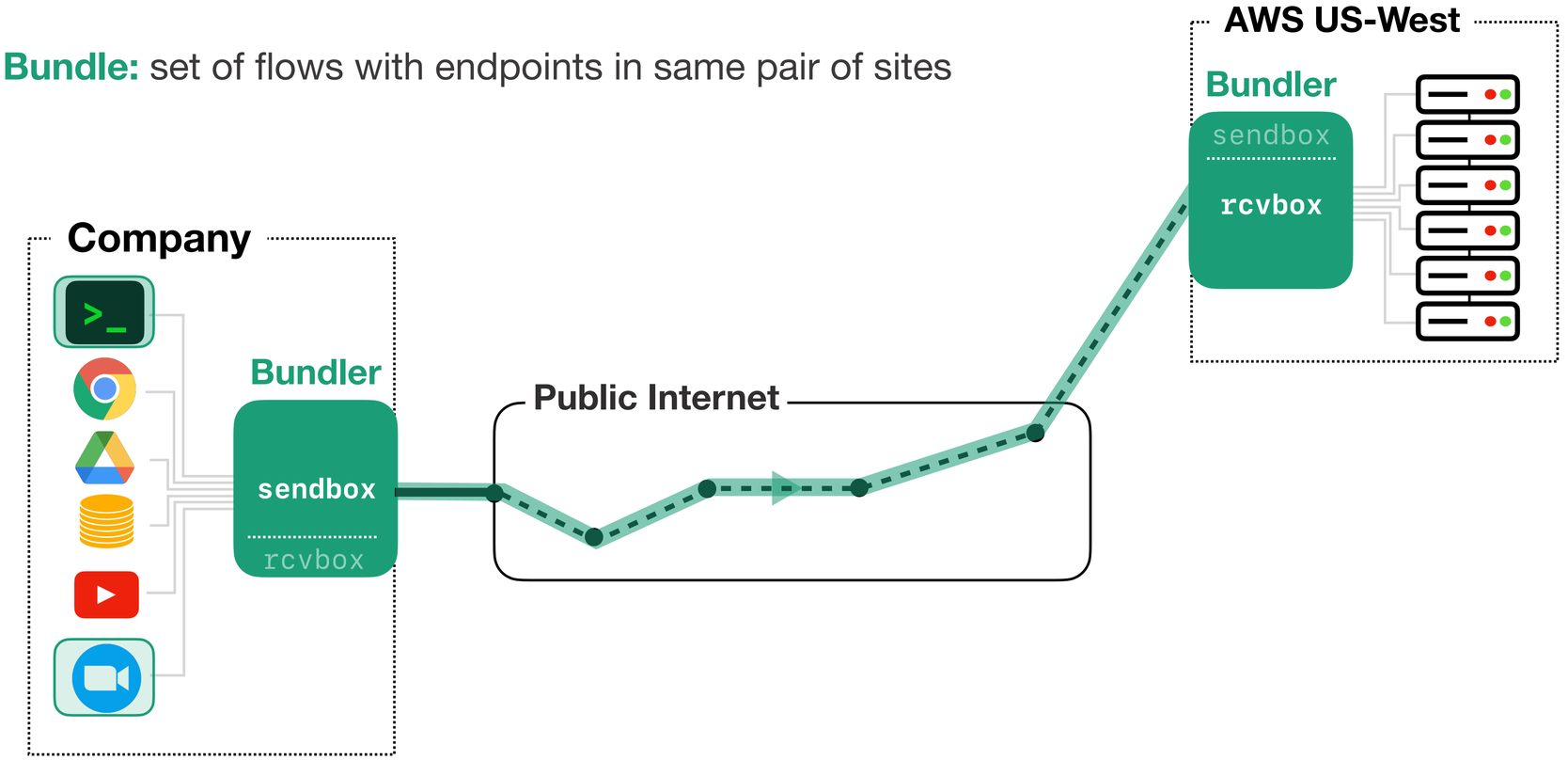
Bundler's Architecture

Bundle: set of flows with endpoints in same pair of sites



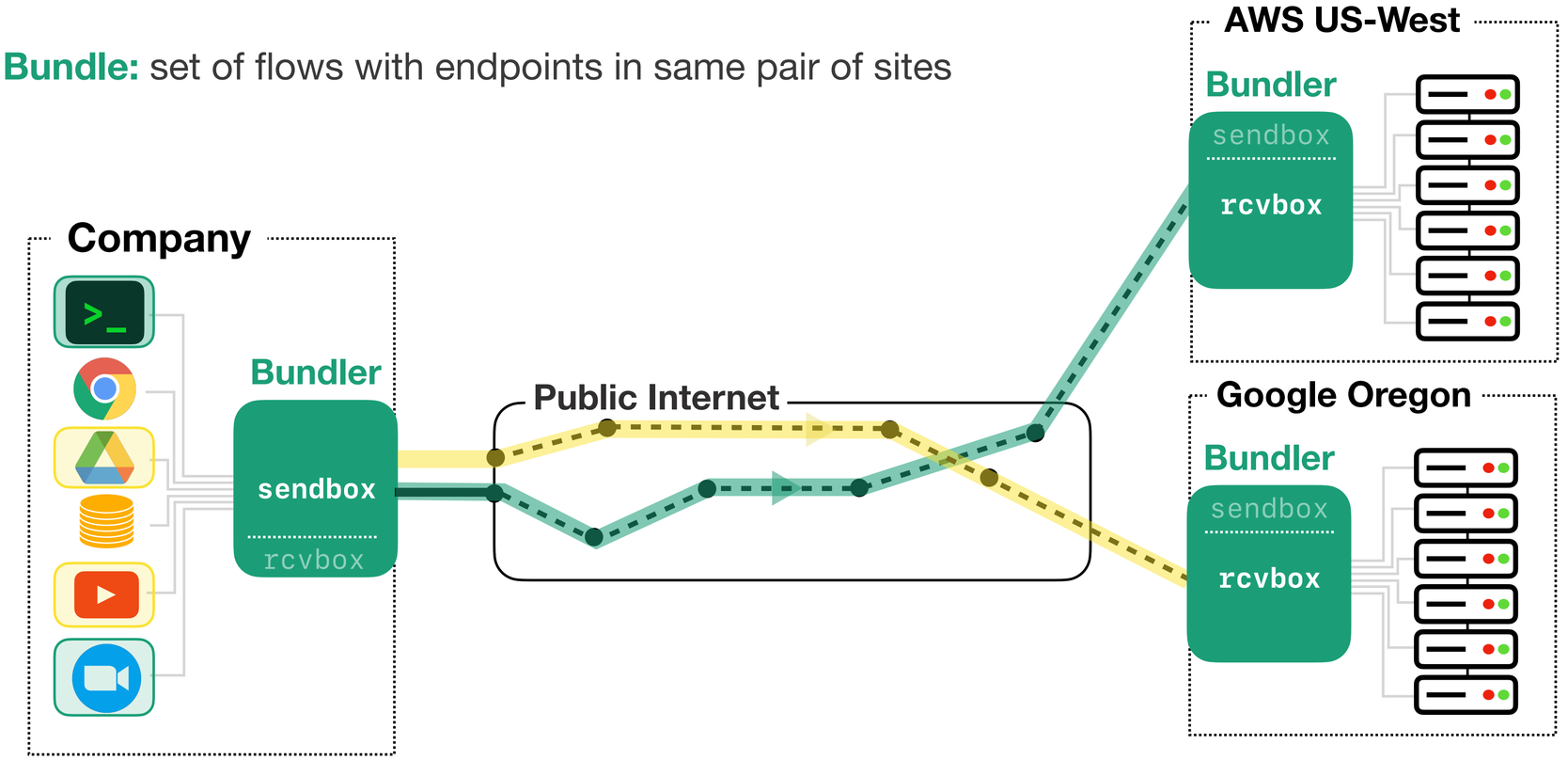
Bundler's Architecture

Bundle: set of flows with endpoints in same pair of sites



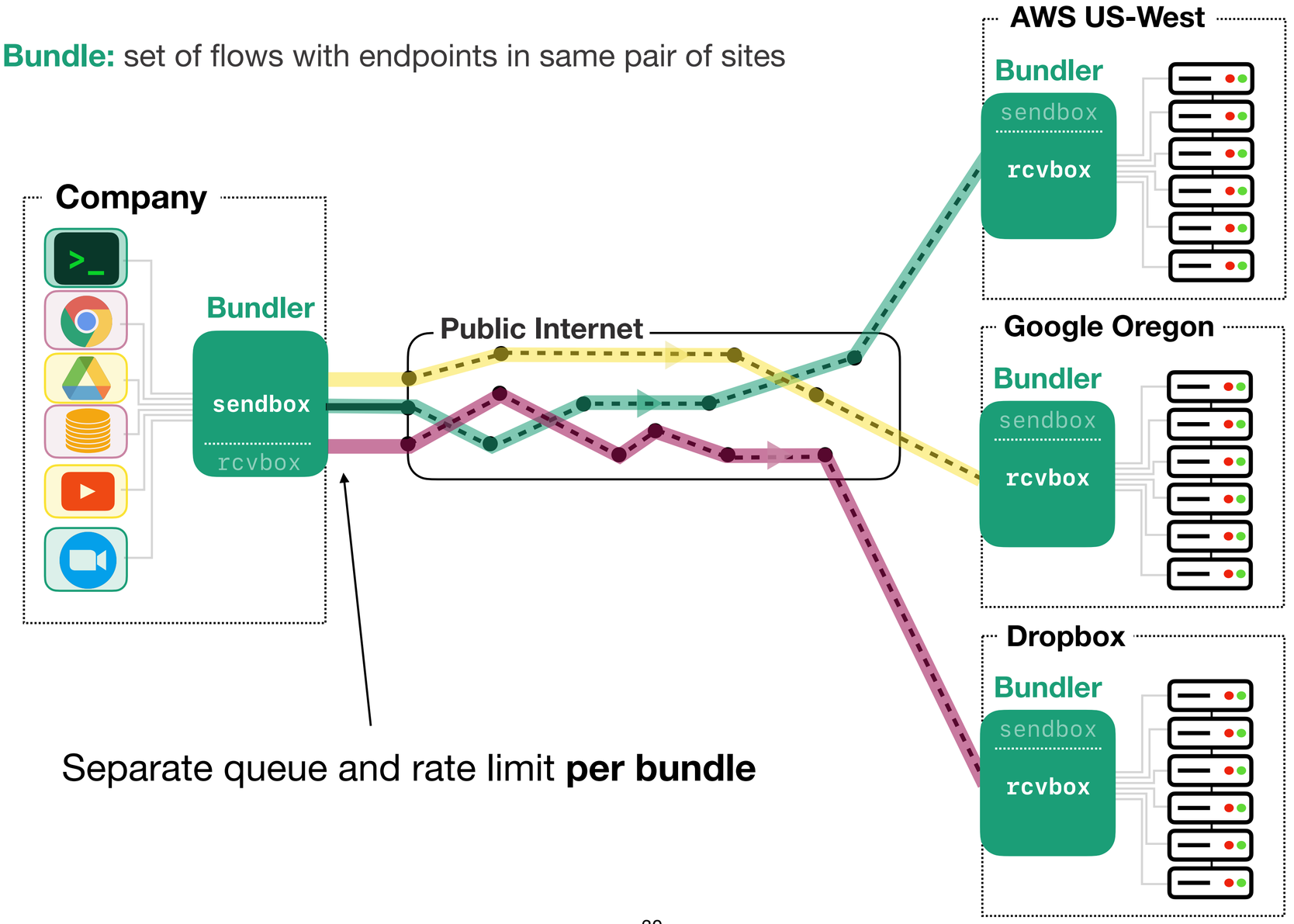
Bundler's Architecture

Bundle: set of flows with endpoints in same pair of sites



Bundler's Architecture

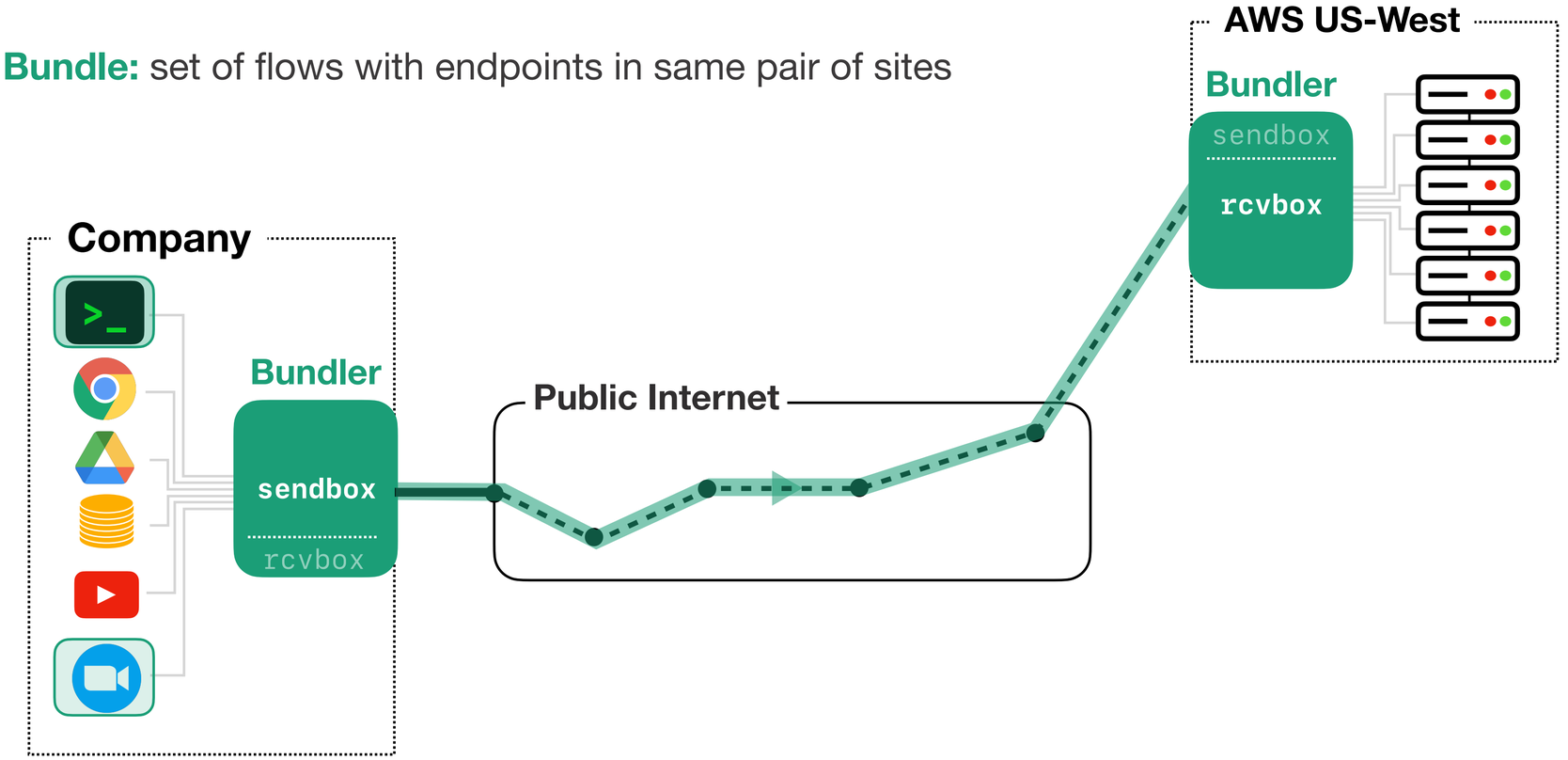
Bundle: set of flows with endpoints in same pair of sites



Separate queue and rate limit **per bundle**

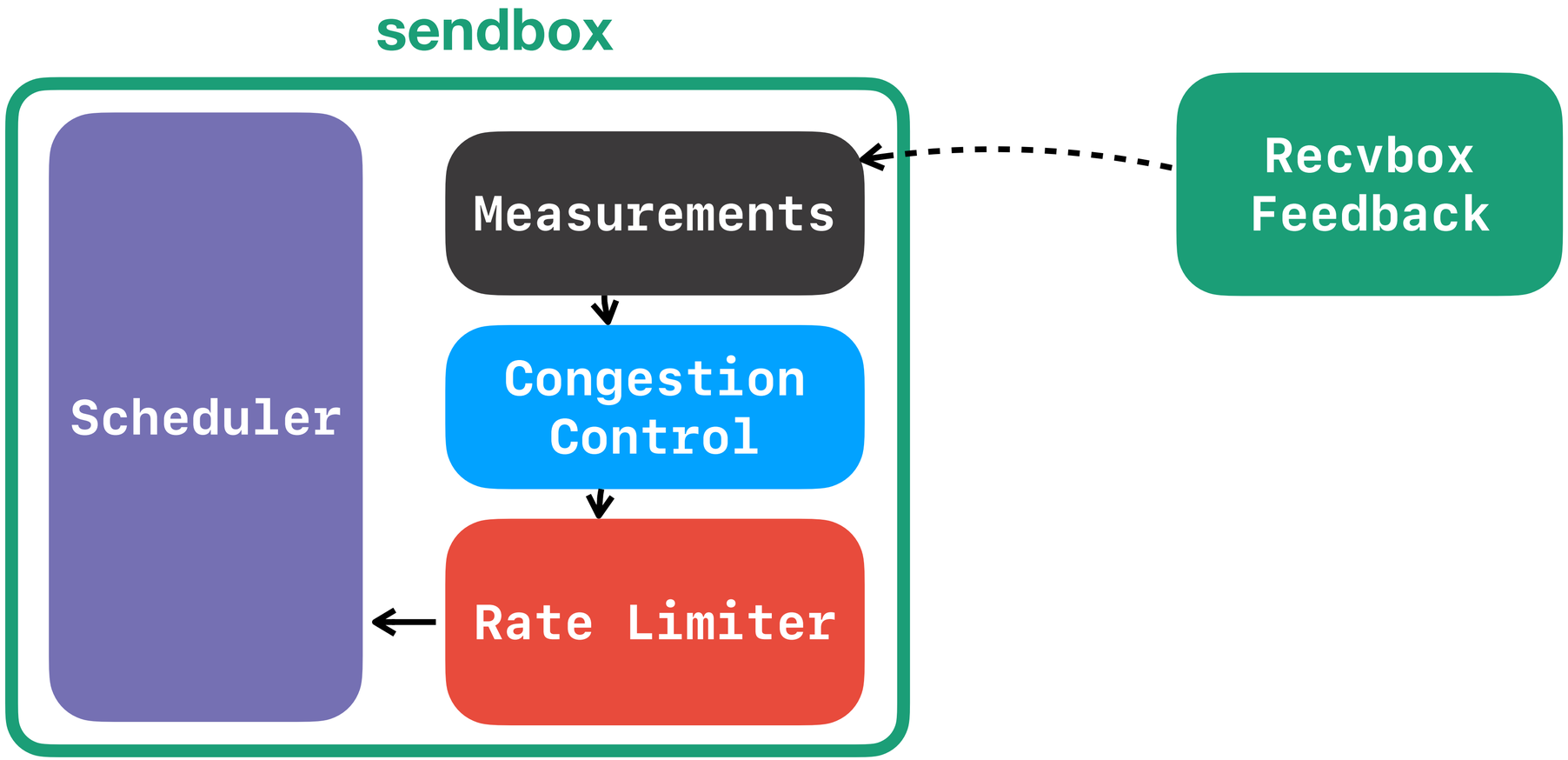
Bundler's Architecture

Bundle: set of flows with endpoints in same pair of sites



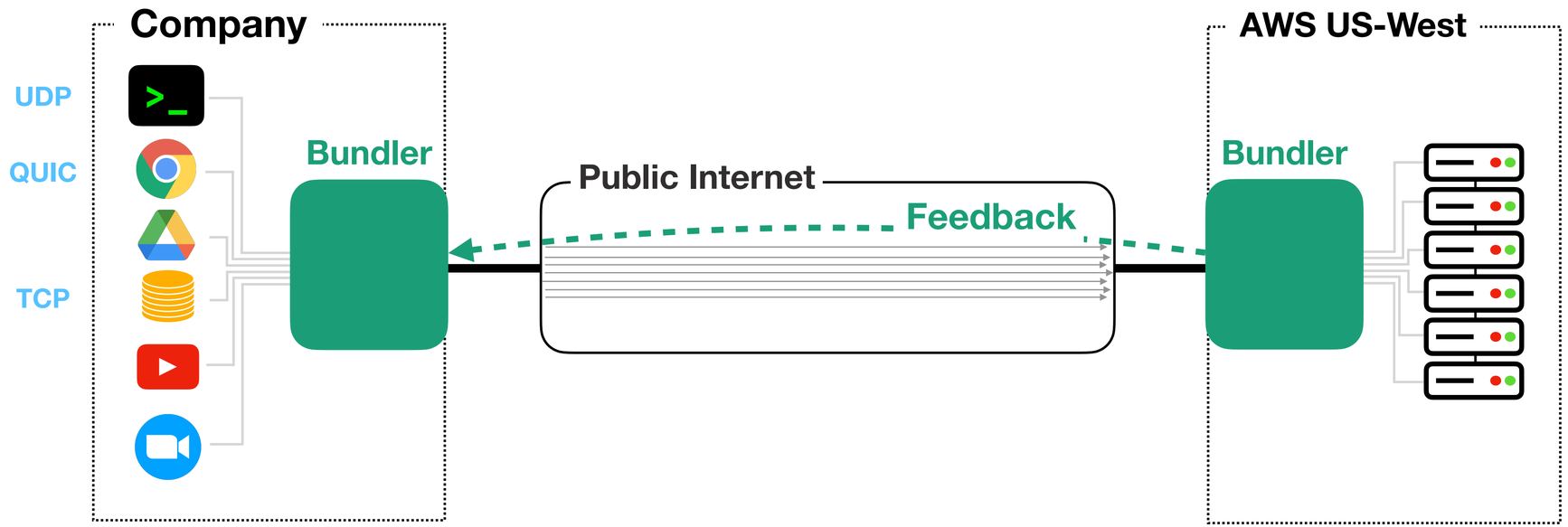
Bundler’s Architecture

Bundler's Architecture



Transparent Measurement Scheme

- Leave connections **intact**
  - Don’t modify packets
  - Don’t disrupt end-to-end connections
- Out-of-band feedback **per RTT**
- Sample the same packets at both boxes **without communication**



Transparent Measurement Scheme

- Leave connections **intact**
  - Don’t modify packets
  - Don’t disrupt end-to-end connections
- Out-of-band feedback **per RTT**
- Sample the same packets at both boxes **without communication**

Compared to alternatives (e.g., TCP proxy)…

Transparent Measurement Scheme

- Leave connections **intact**
  - Don’t modify packets
  - Don’t disrupt end-to-end connections
- Out-of-band feedback **per RTT**
- Sample the same packets at both boxes **without communication**

Compared to alternatives (e.g., TCP proxy)…

Low overhead and complexity

Transparent Measurement Scheme

- Leave connections **intact**
  - Don’t modify packets
  - Don’t disrupt end-to-end connections
- Out-of-band feedback **per RTT**
- Sample the same packets at both boxes **without communication**

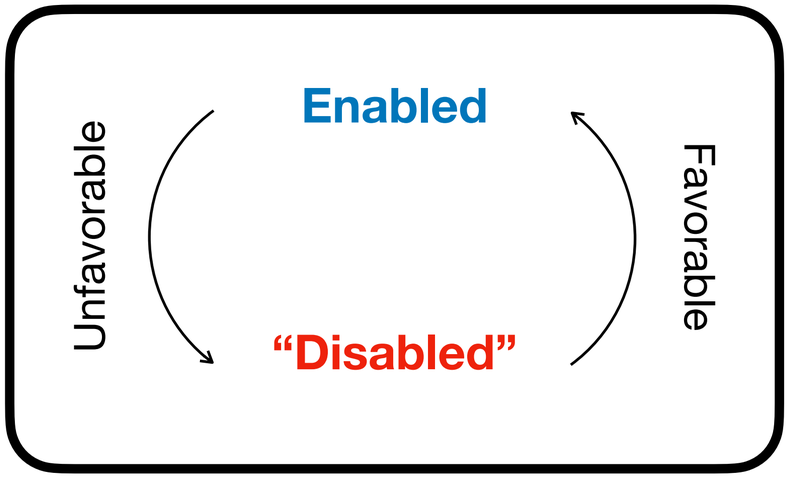
Compared to alternatives (e.g., TCP proxy)…

Low overhead and complexity

Simple datapath

Handling Unfavorable Conditions

1. Flows in a bundle **don't share** the same bottleneck
2. Bundle competing with long-lasting **buffer-filling** cross traffic



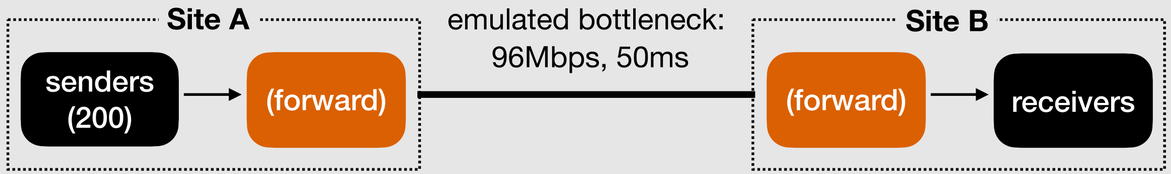
normal throughput
+ scheduling benefits

normal throughput
(but no scheduling)

But… in our experience,
unfavorable conditions are rare.

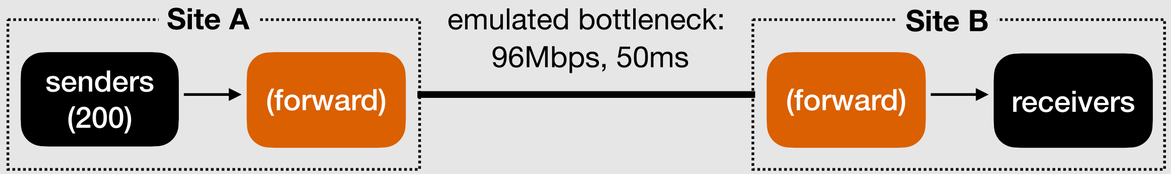
Bundler (using FQ) improves FCTs

Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution



Bundler (using FQ) improves FCTs

Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution



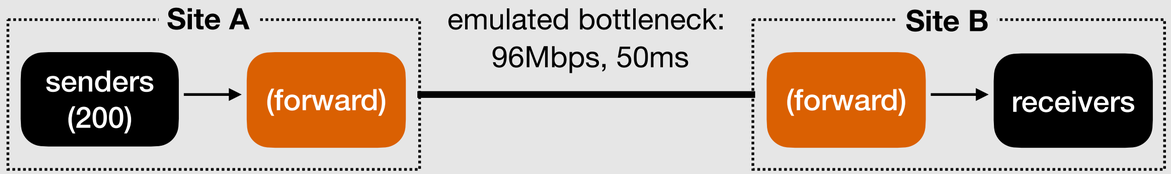
Small
(<10KB, 97.6%)

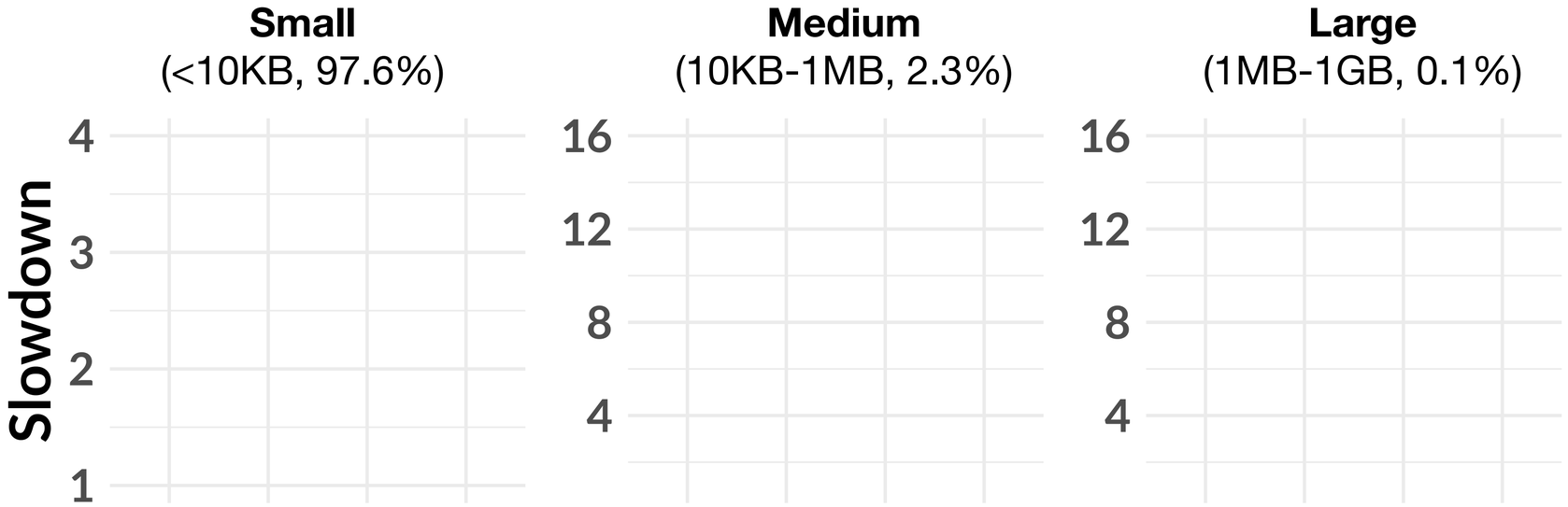
Medium
(10KB-1MB, 2.3%)

Large
(1MB-1GB, 0.1%)

Bundler (using FQ) improves FCTs

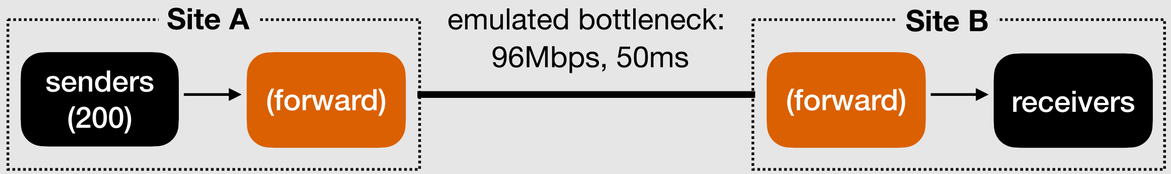
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

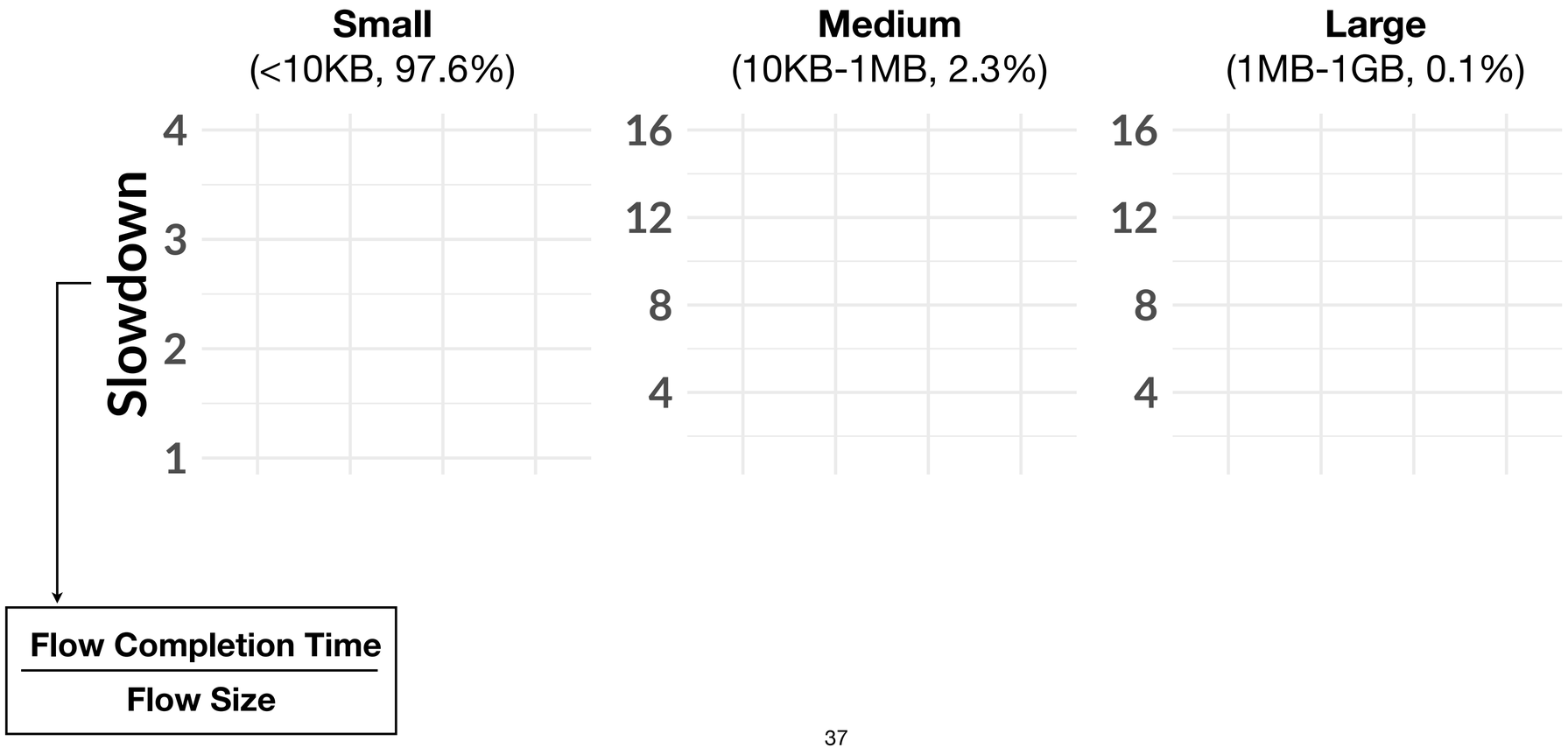




Bundler (using FQ) improves FCTs

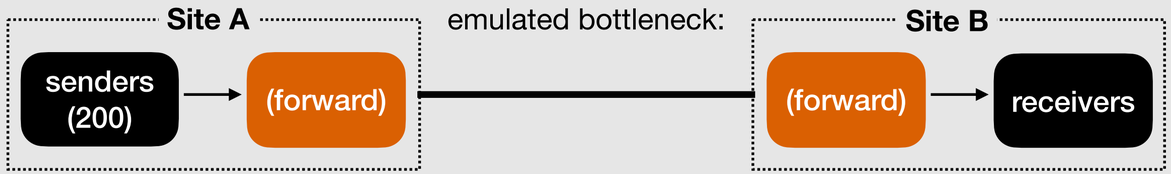
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

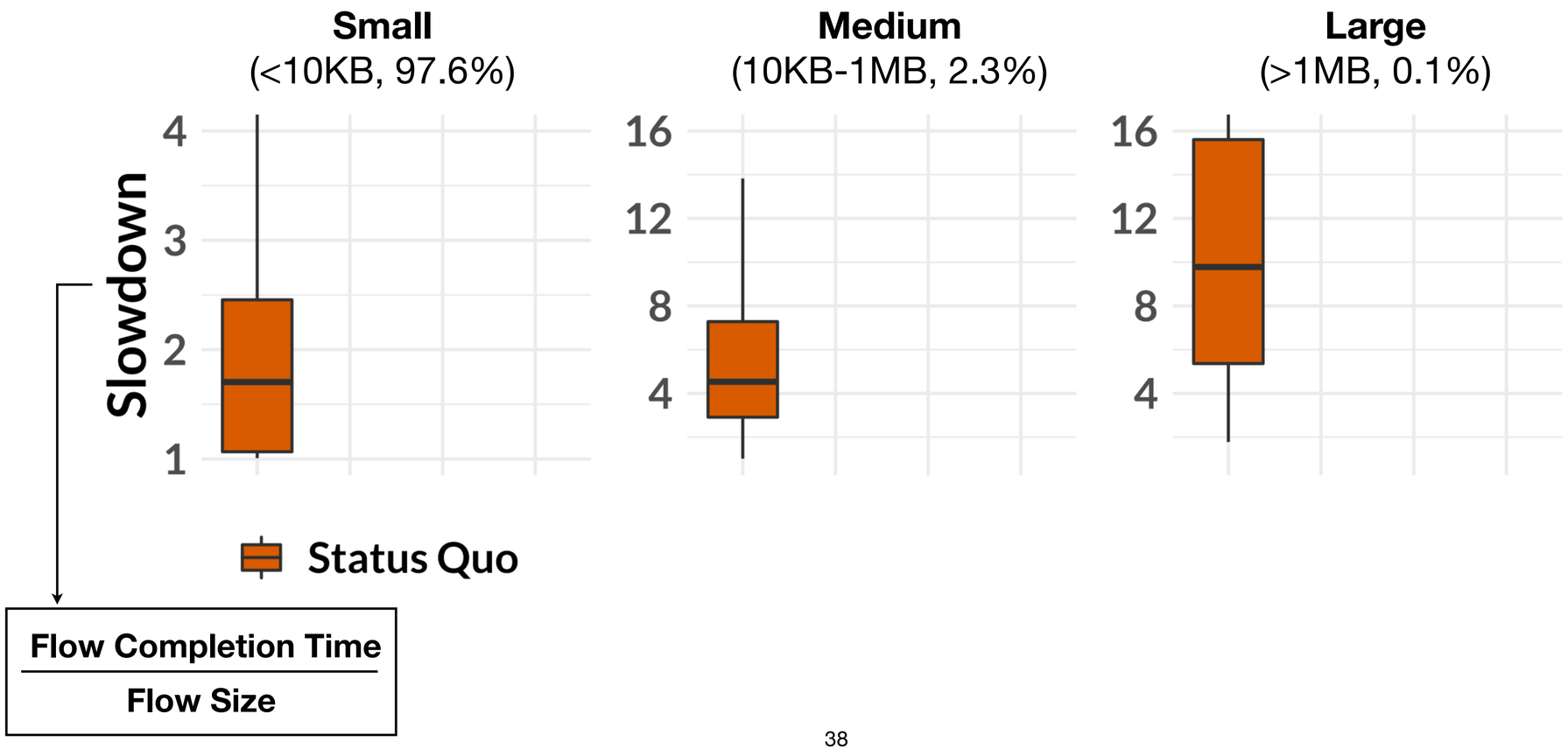




Bundler (using FQ) improves FCTs

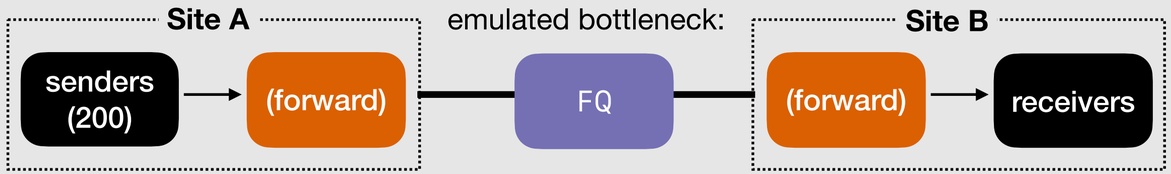
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

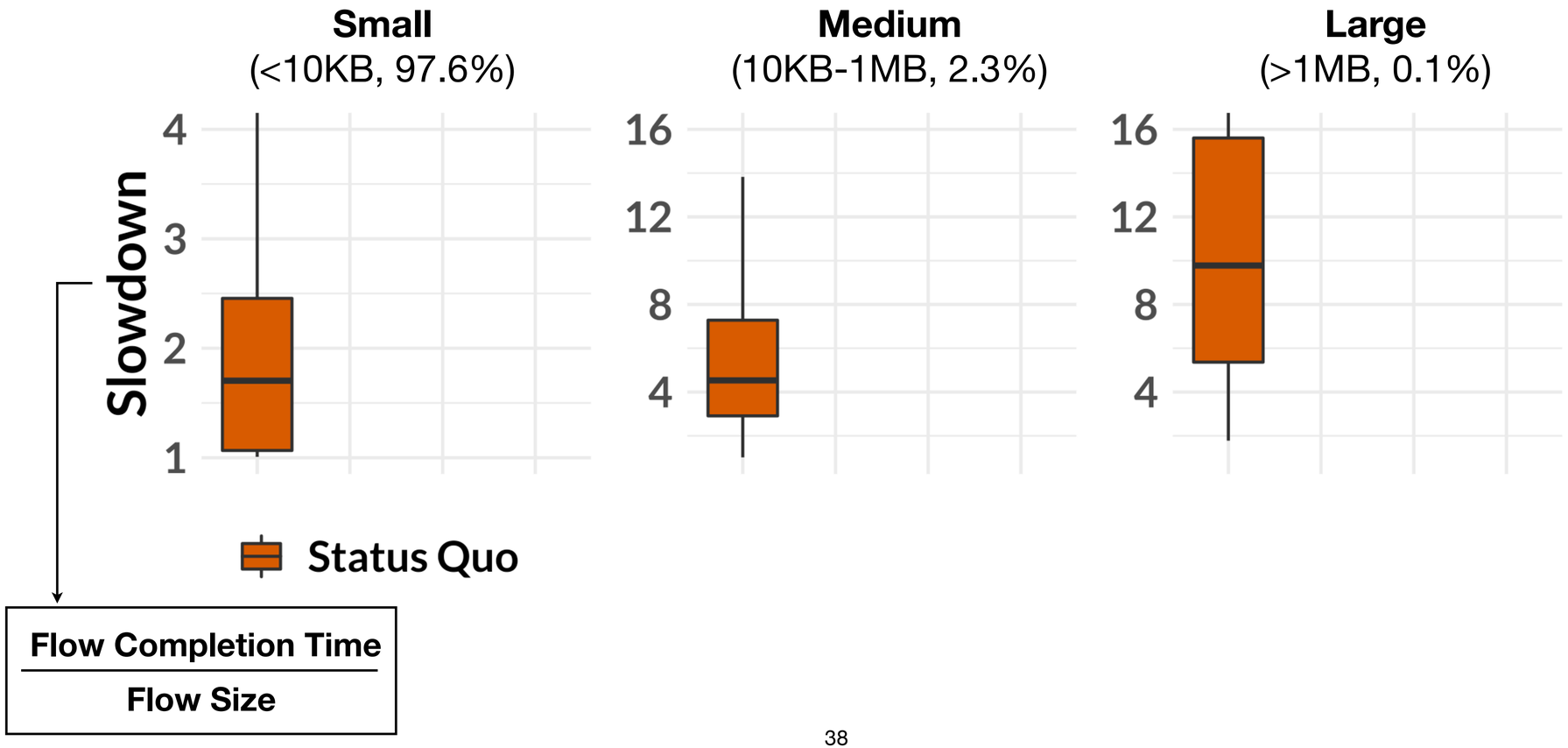




Bundler (using FQ) improves FCTs

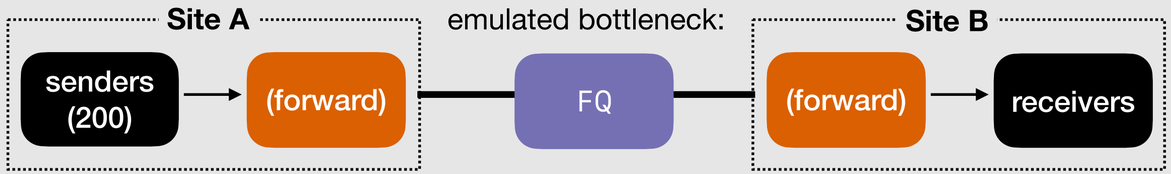
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

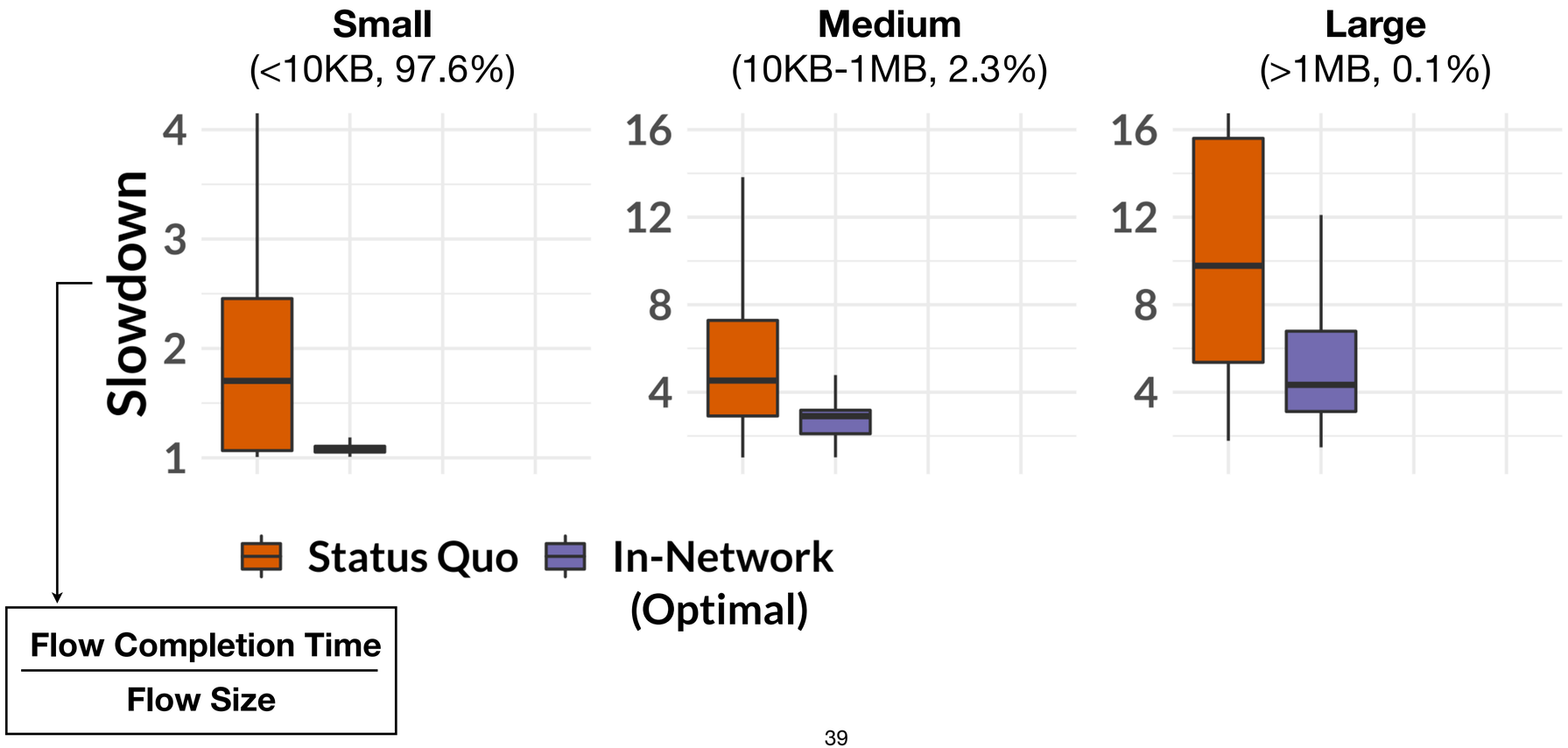




Bundler (using FQ) improves FCTs

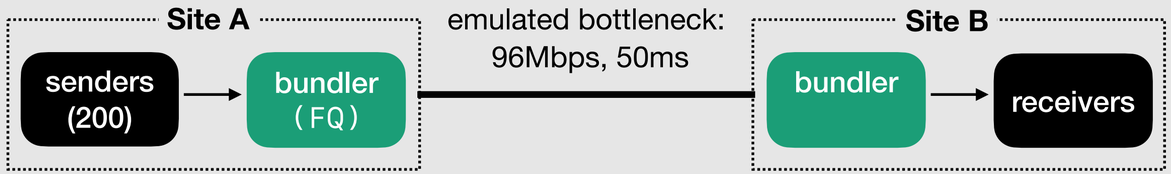
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

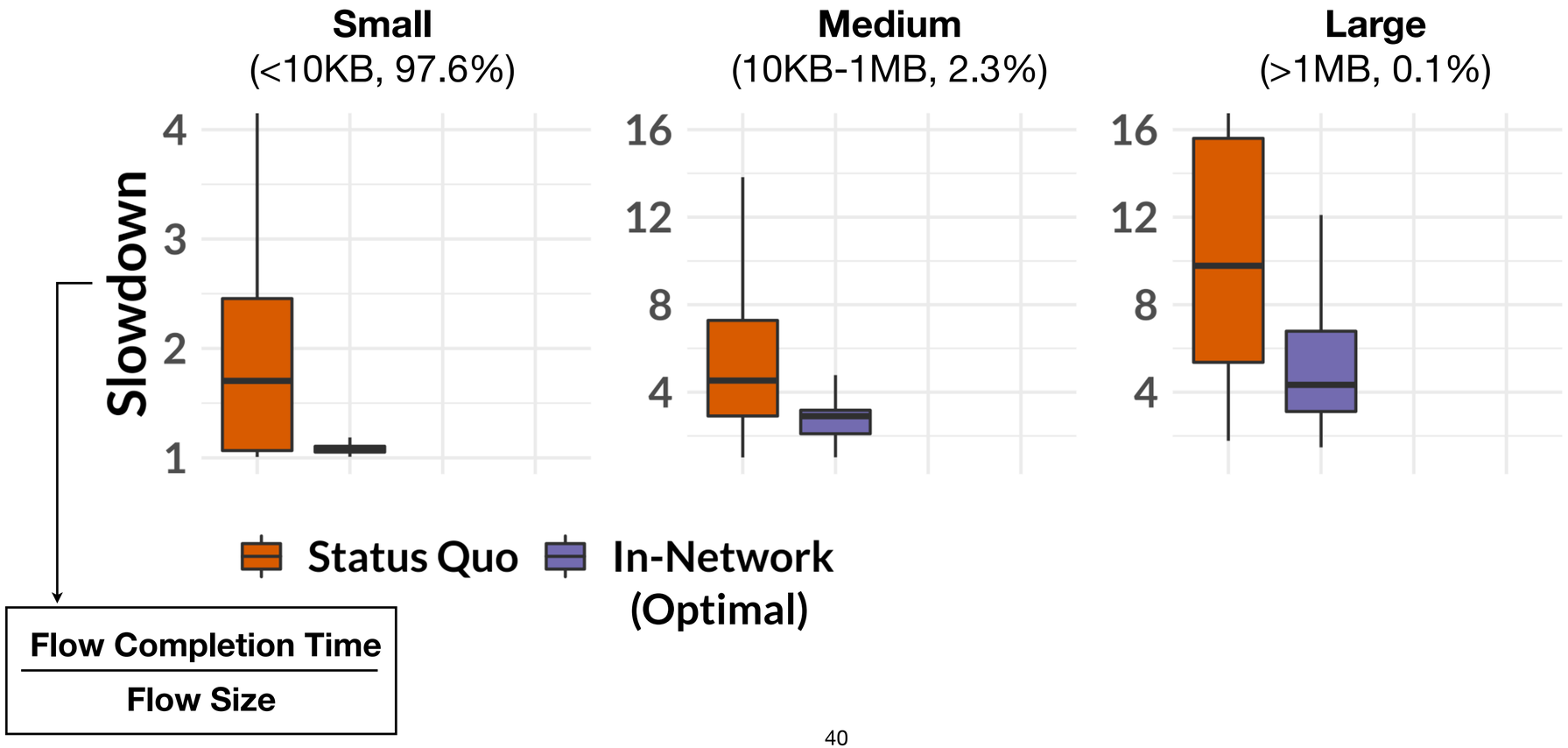




Bundler (using FQ) improves FCTs

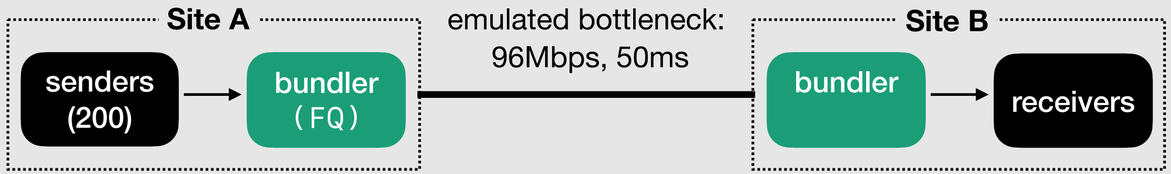
Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution

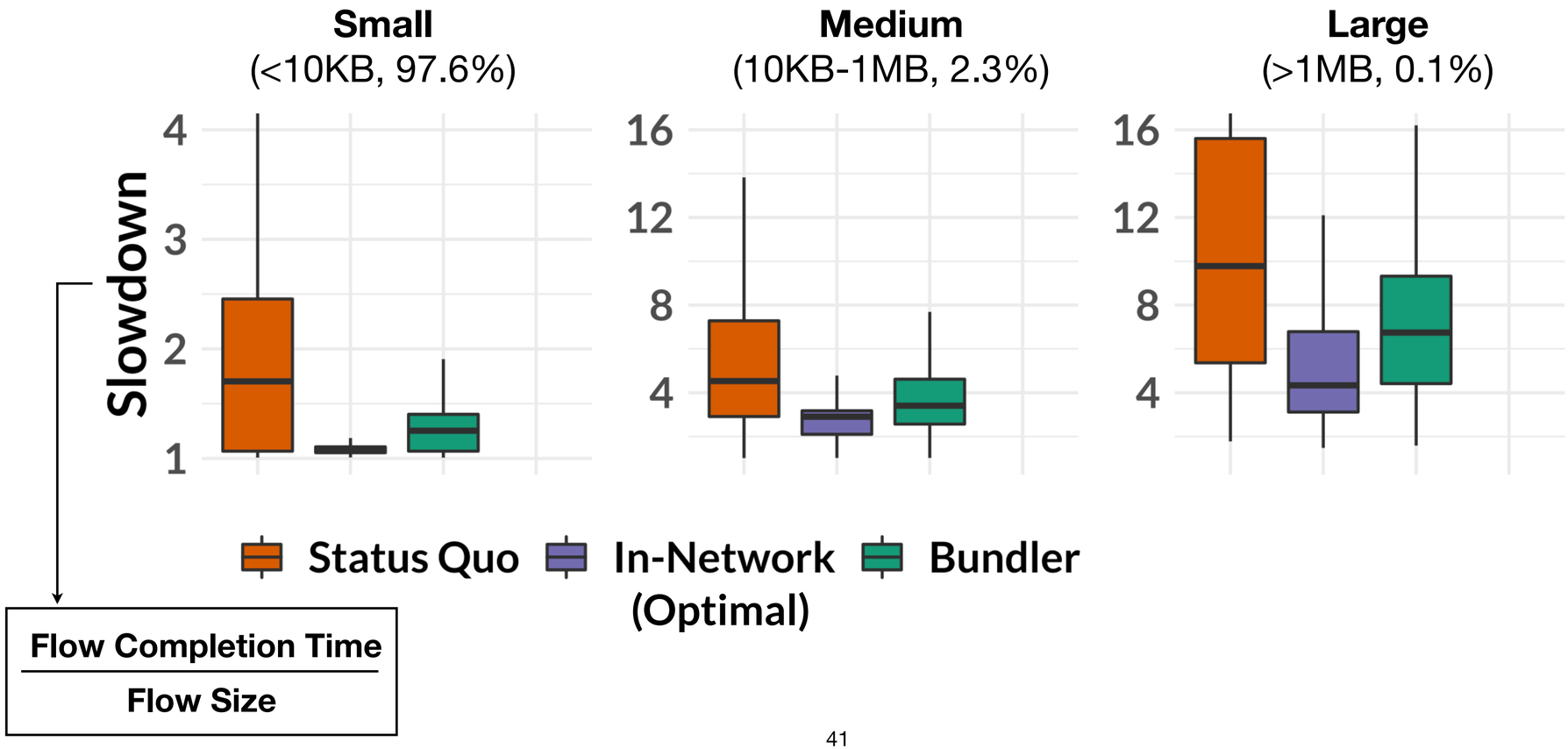




Bundler (using FQ) improves FCTs

Traffic: 1 million TCP cubic flows, sizes sampled from CAIDA internet backbone distribution





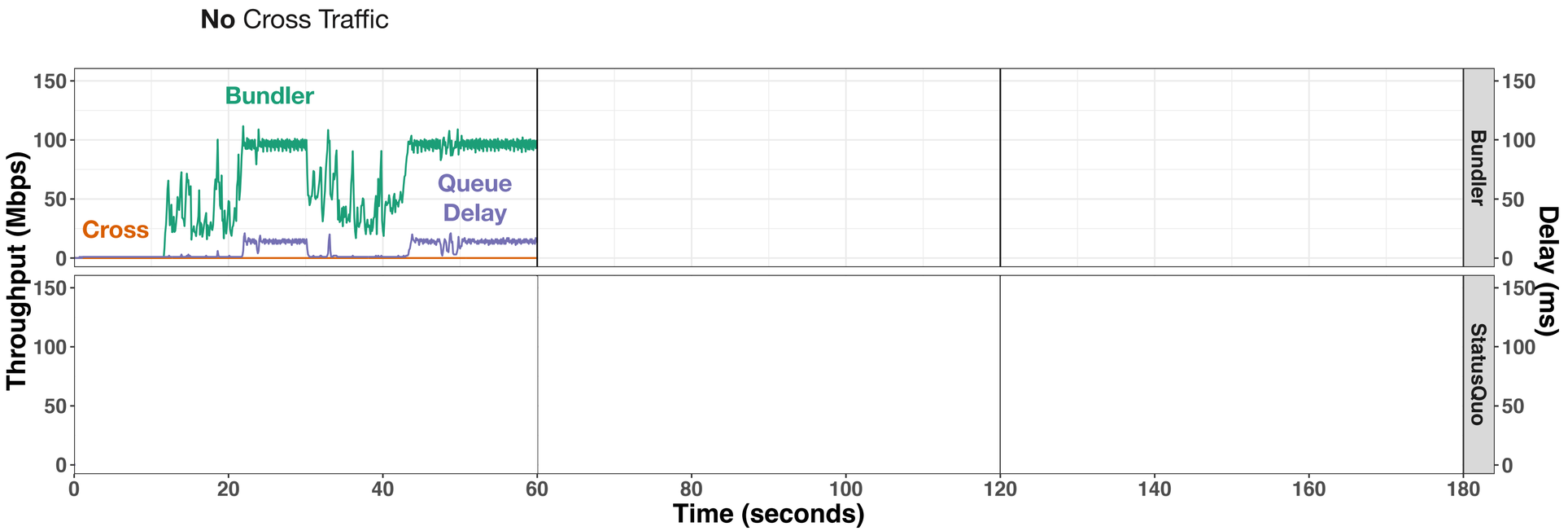
Summary

Bundler is a new middlebox that enables scheduling regardless of where congestion occurs in the network

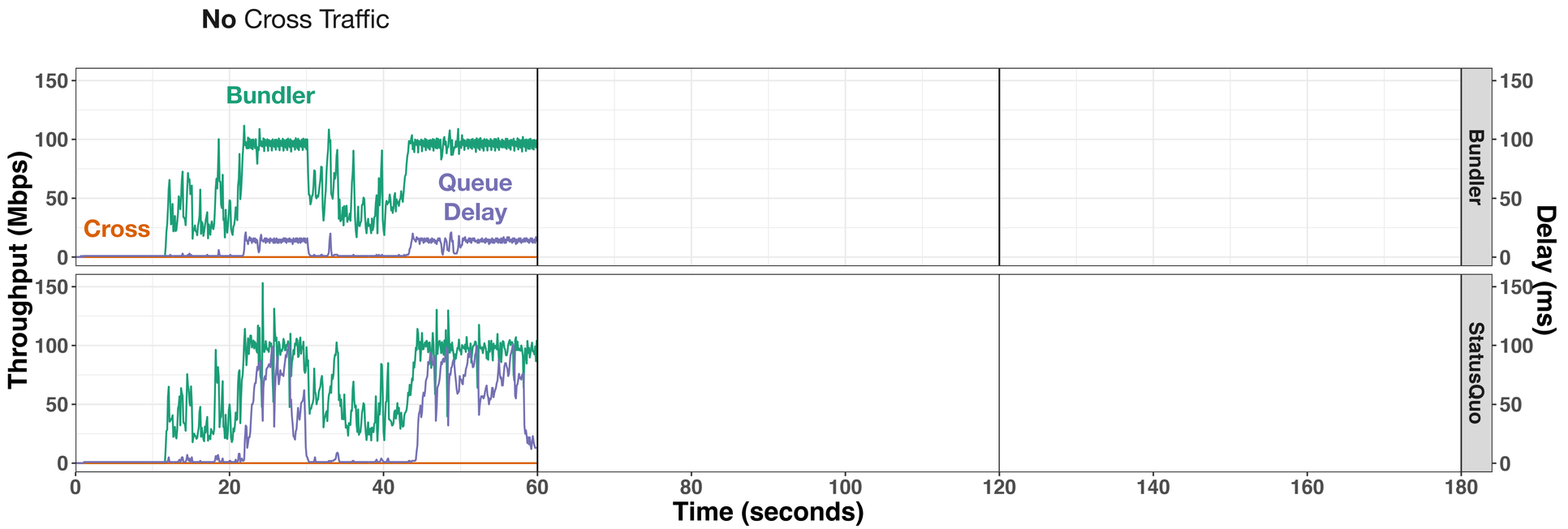
Source code and evaluation scripts available at:

github.com/bundler-project

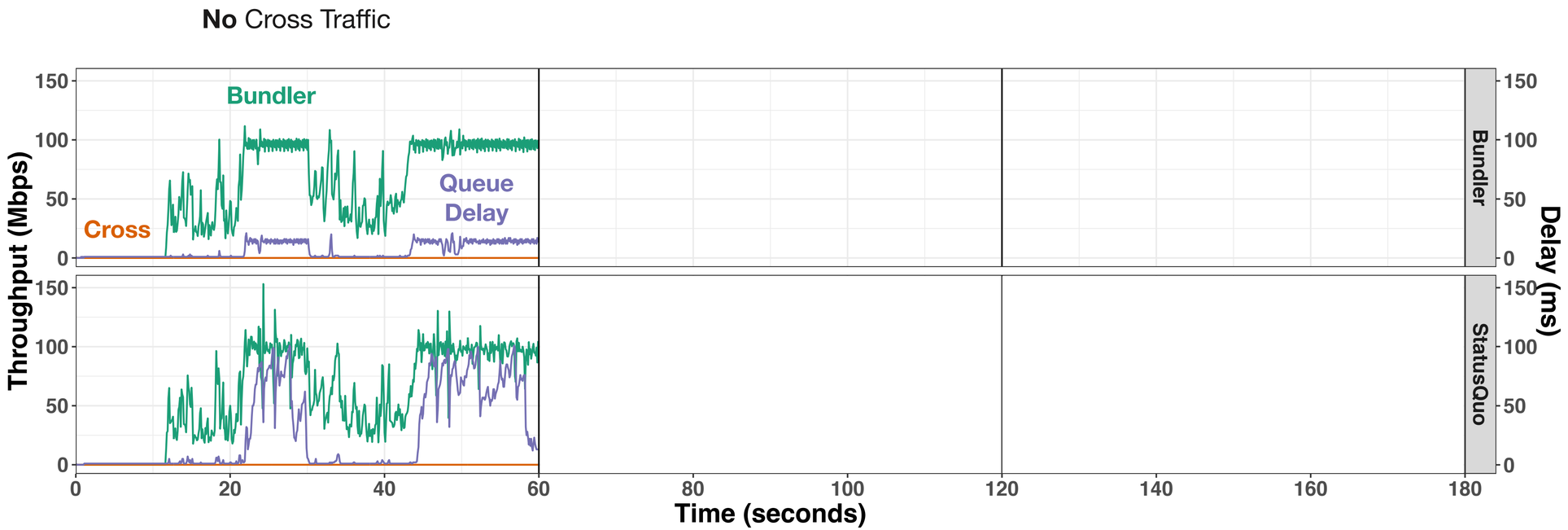
Bundler dynamically adjusts to cross-traffic

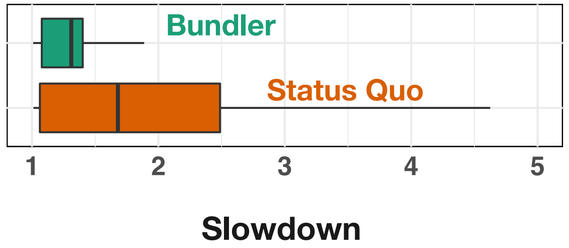


Bundler dynamically adjusts to cross-traffic

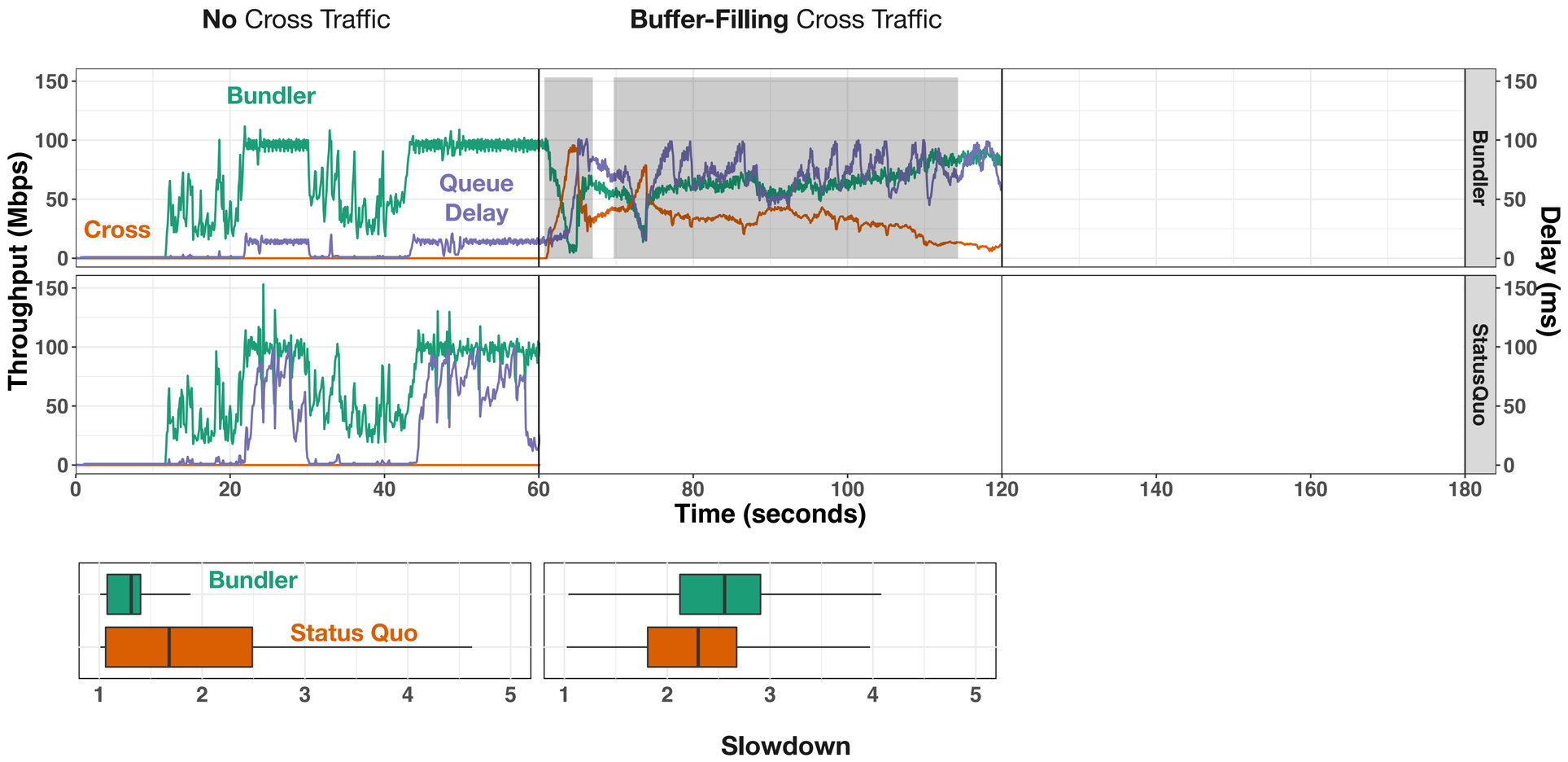


Bundler dynamically adjusts to cross-traffic

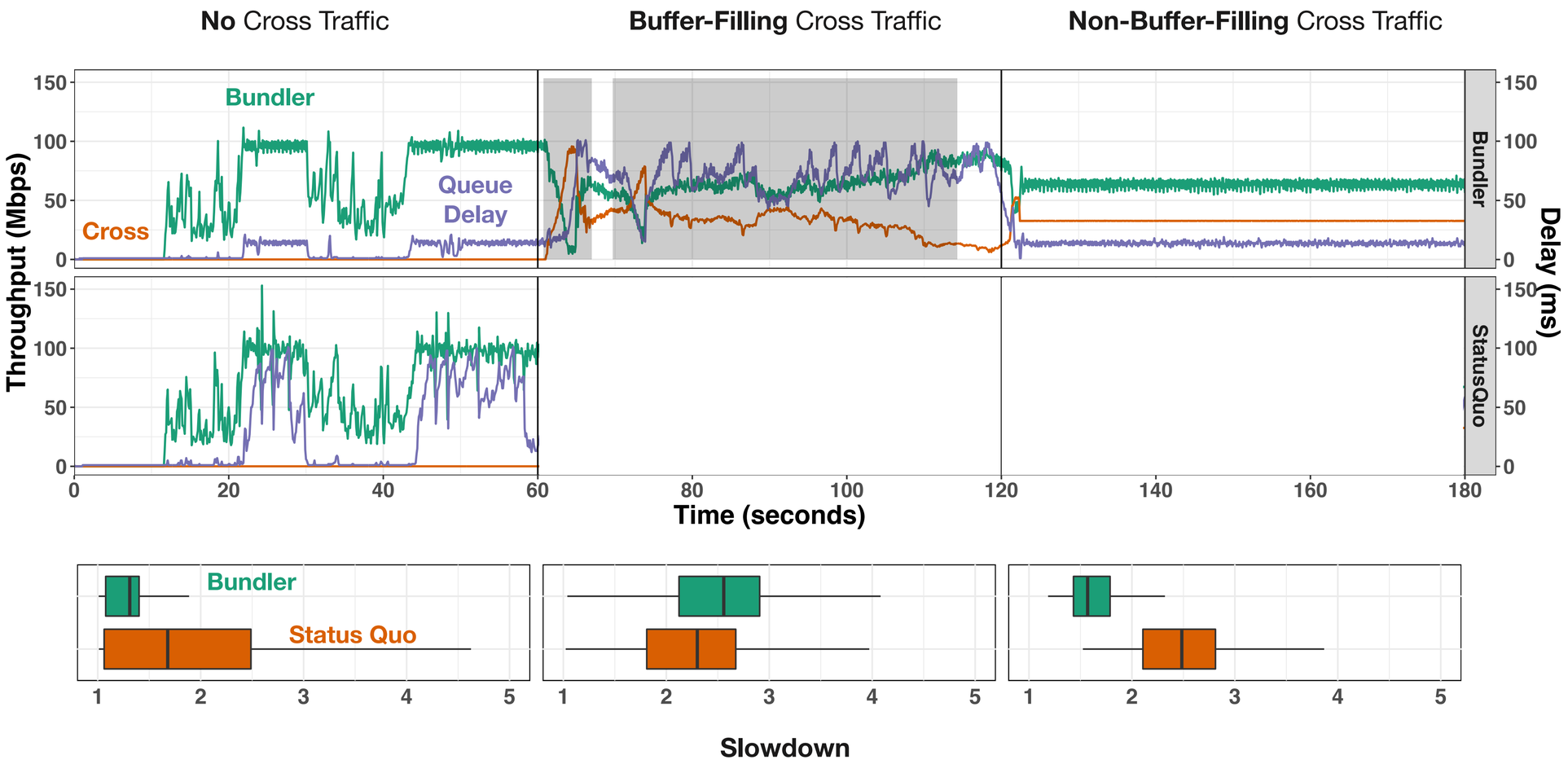




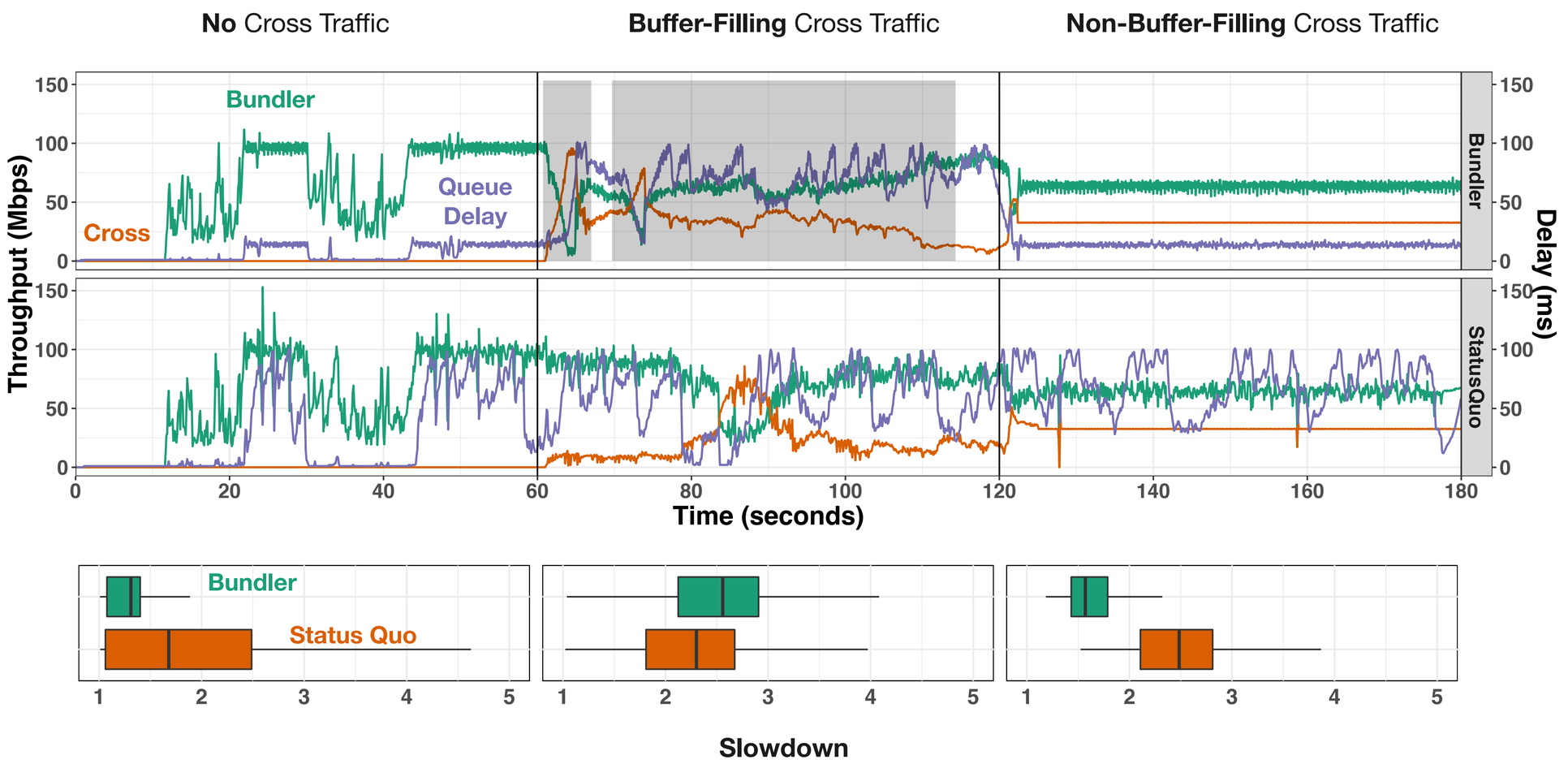
Bundler dynamically adjusts to cross-traffic



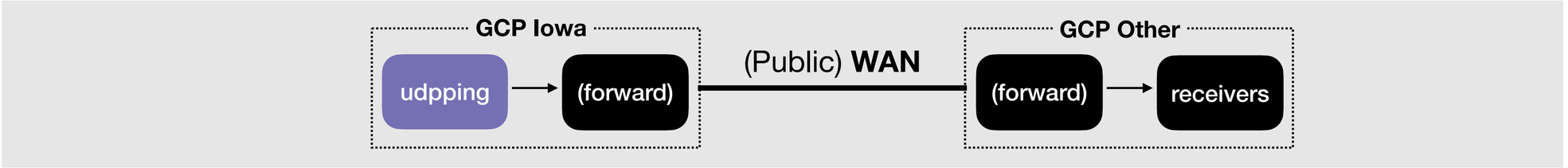
Bundler dynamically adjusts to cross-traffic

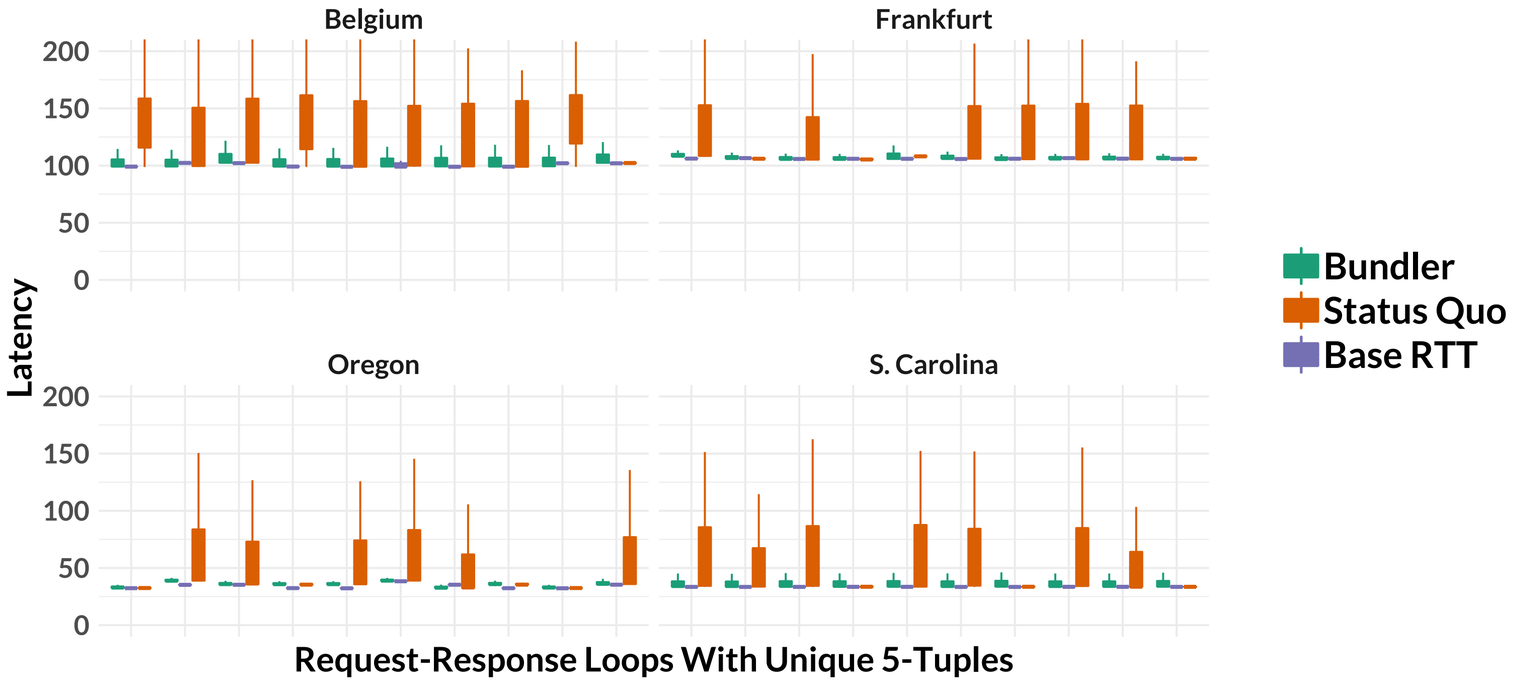


Bundler dynamically adjusts to cross-traffic

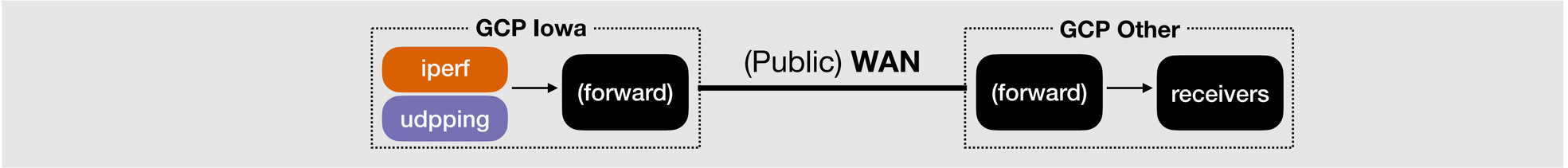


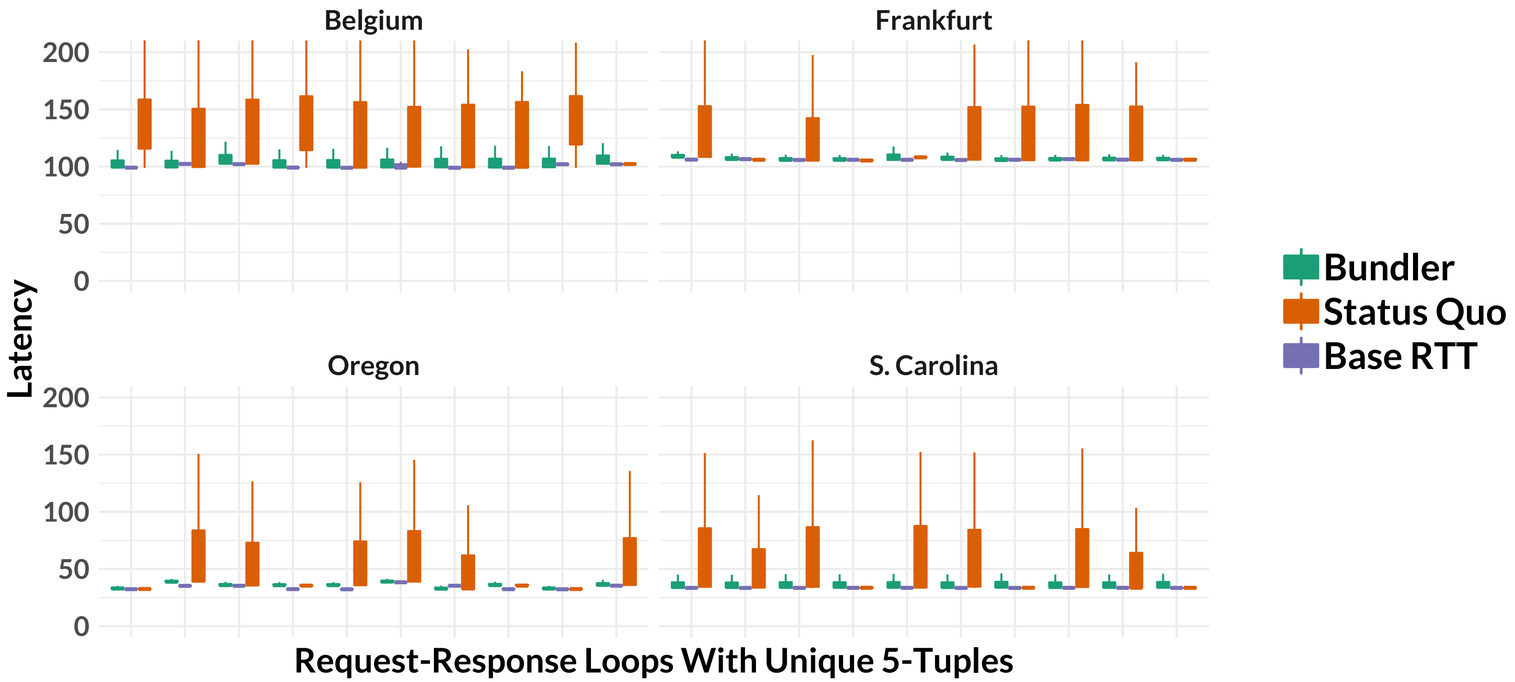
Bundler can shift queues on real internet paths



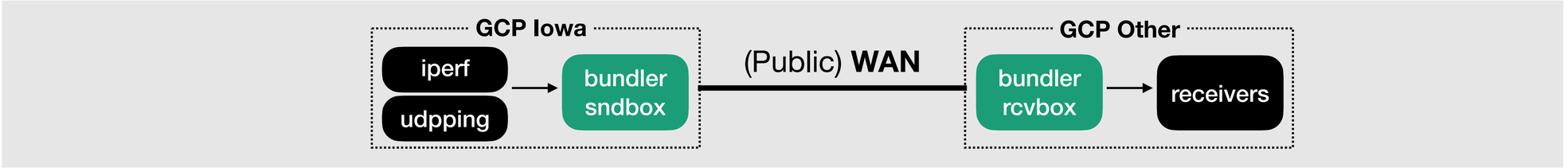


Bundler can shift queues on real internet paths





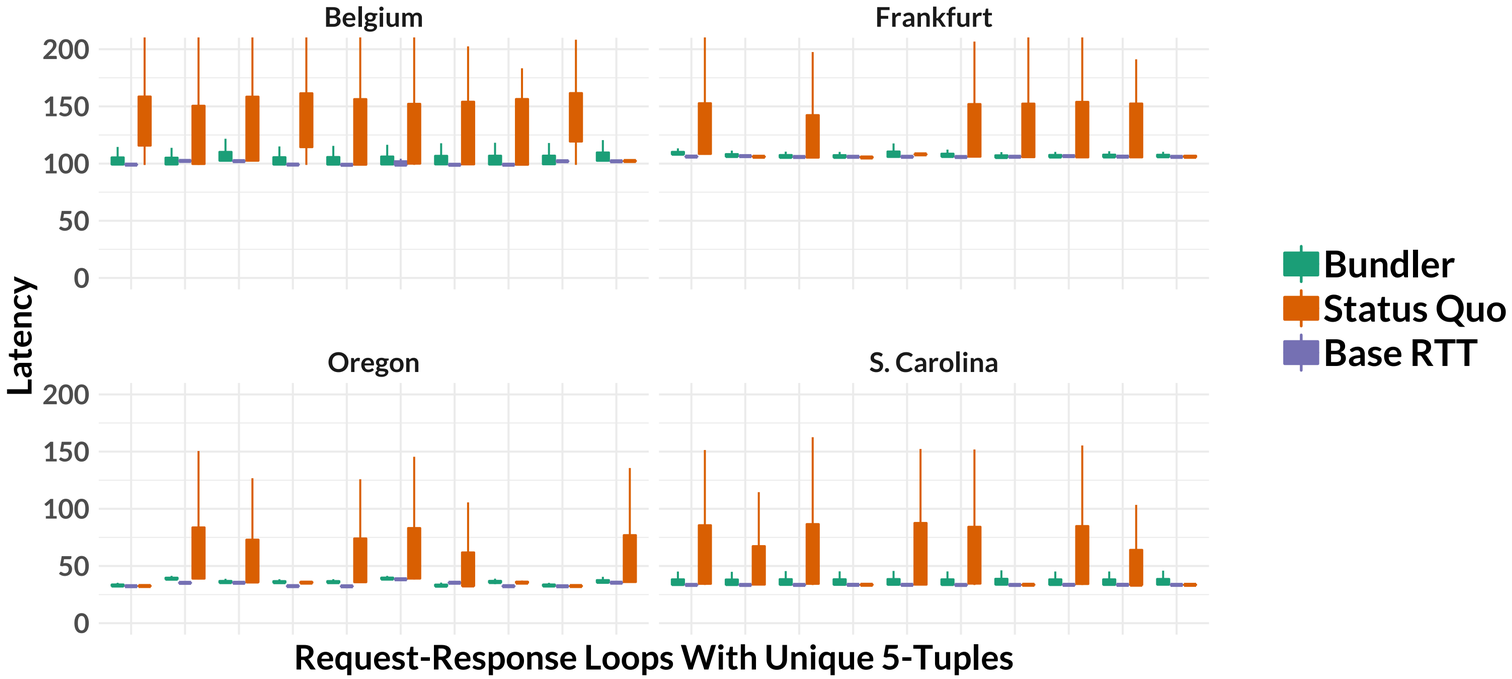
Bundler can shift queues on real internet paths



With Bundler:

Throughput stays the same as status quo,

But latency is reduced back to base RTT



Bundler improves rate stability