Reproducing and Solving a Fallback Issue in TCP Prague

Lukas Hagen
Hasso Plattner Institute
lukas.hagen@student.hpi.de

Anna Brunstrom
Karlstad University
anna.brunstrom@kau.se

TCP Prague

- Congestion control for L4S

TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet

TCP Prague

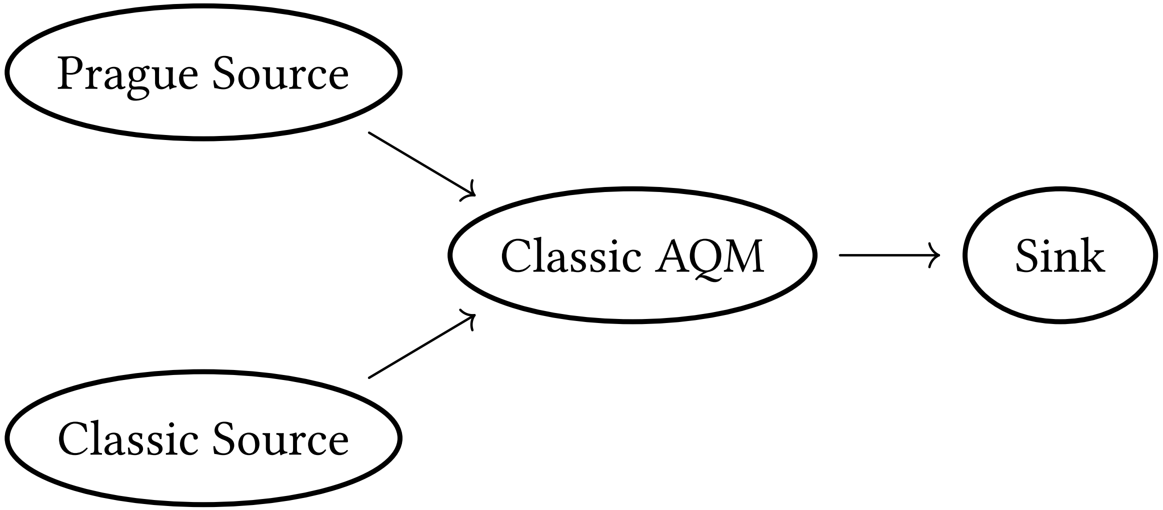
- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)

TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)
- Increases the rate of marked packets, but reduces the response to each packet

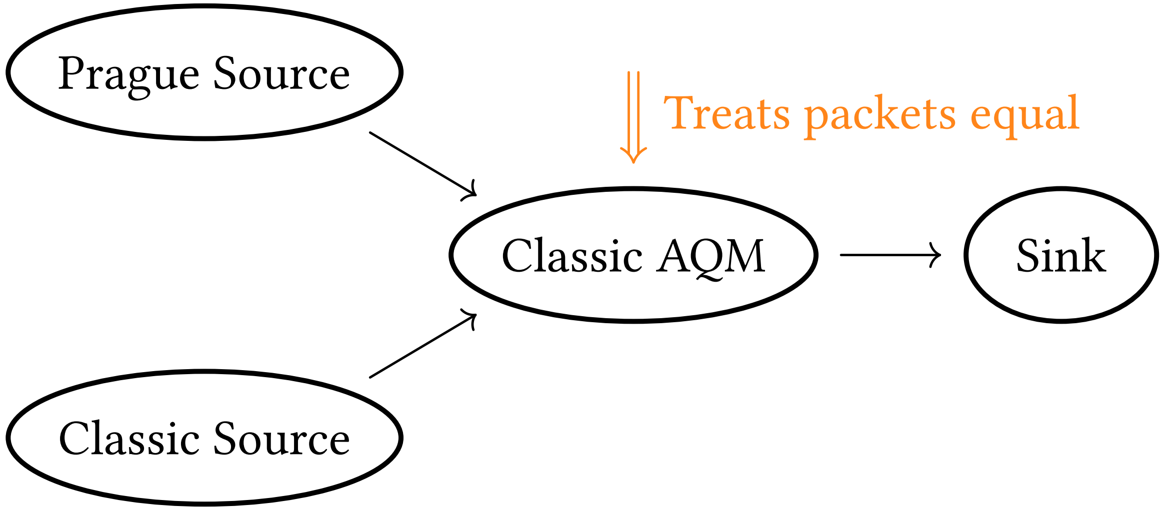
TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)
- Increases the rate of marked packets, but reduces the response to each packet



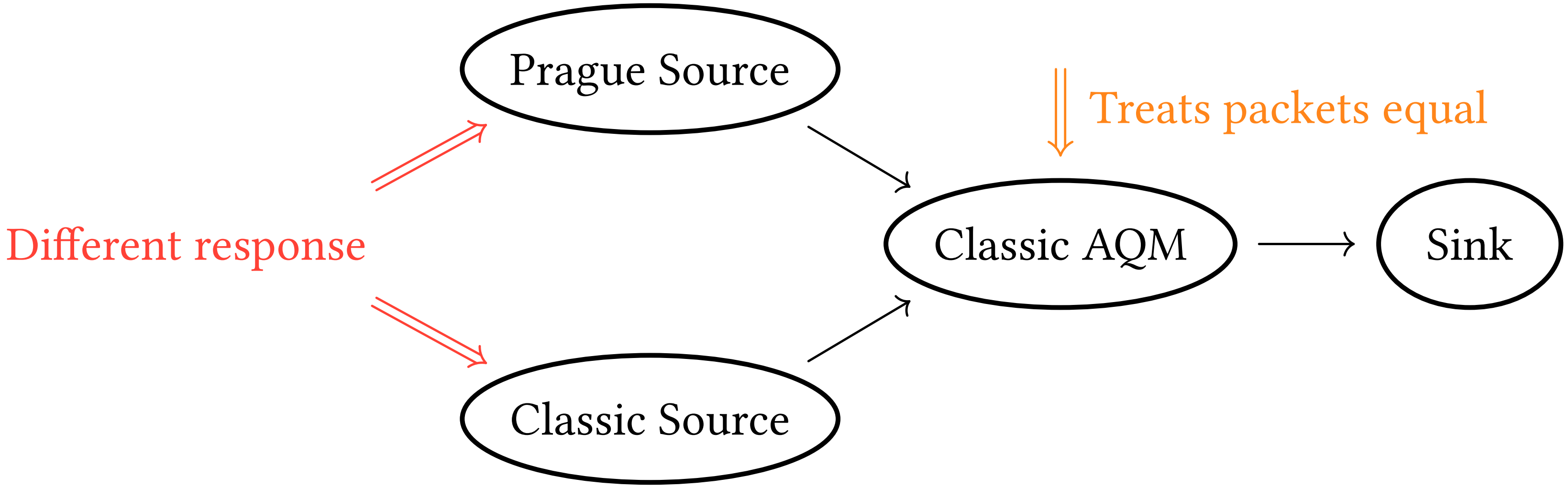
TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)
- Increases the rate of marked packets, but reduces the response to each packet



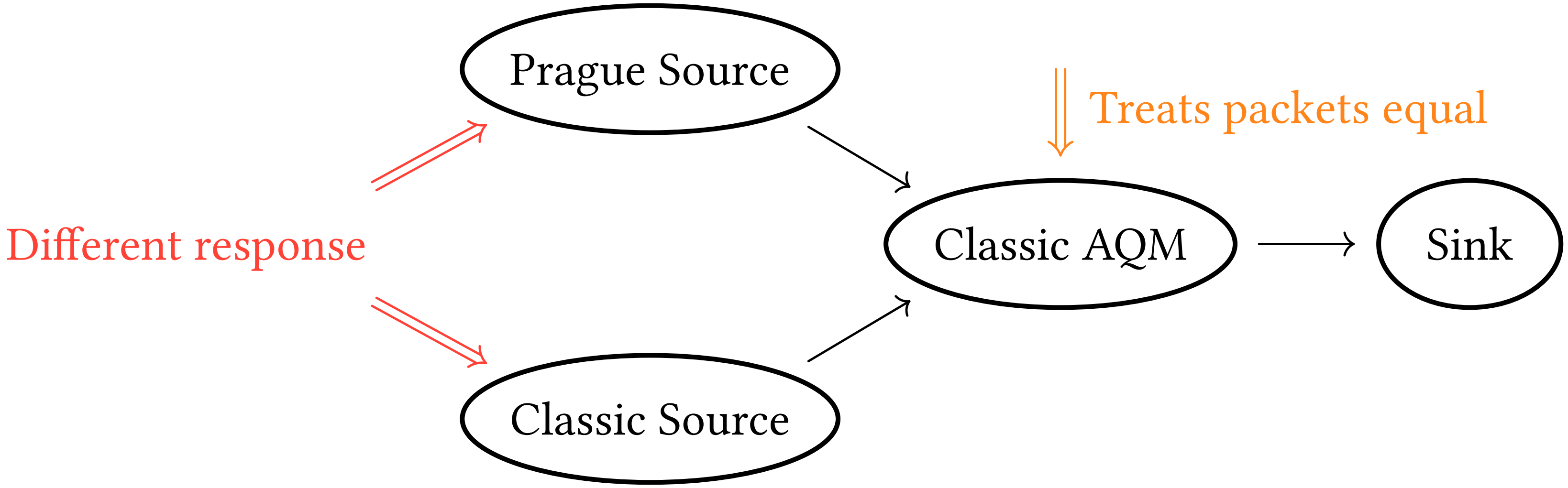
TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)
- Increases the rate of marked packets, but reduces the response to each packet



TCP Prague

- Congestion control for L4S
- Low latency and low loss at full data rate on the Internet
- Leverages Explicit Congestion Notifications (ECN)
- Increases the rate of marked packets, but reduces the response to each packet



- Fairness issue on AQMs that treat classic and L4S similarly

TCP Prague's Fallback

- Detect cases of an RFC 3168 compliant bottleneck AQM

TCP Prague's Fallback

- Detect cases of an RFC 3168 compliant bottleneck AQM

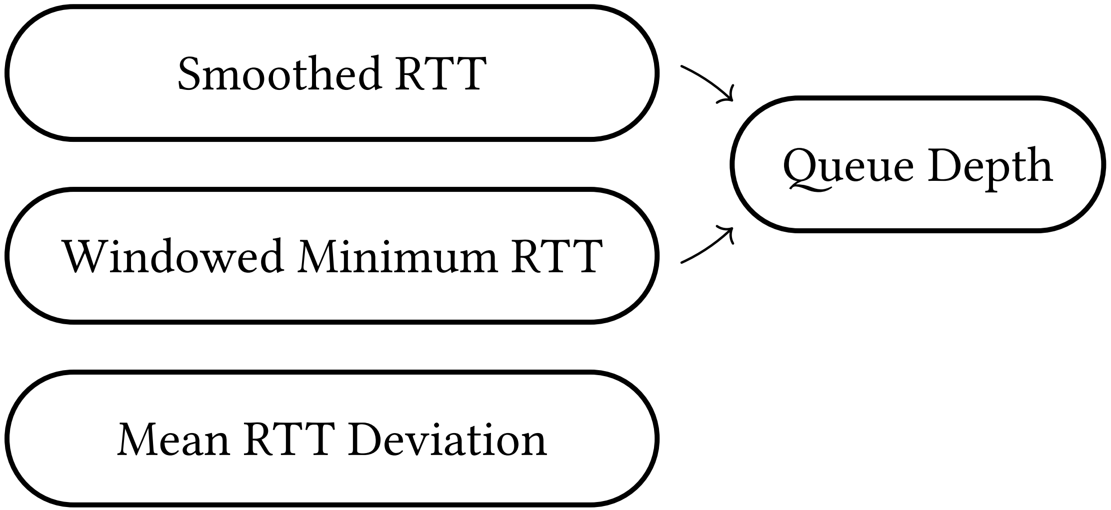
Smoothed RTT

Windowed Minimum RTT

Mean RTT Deviation

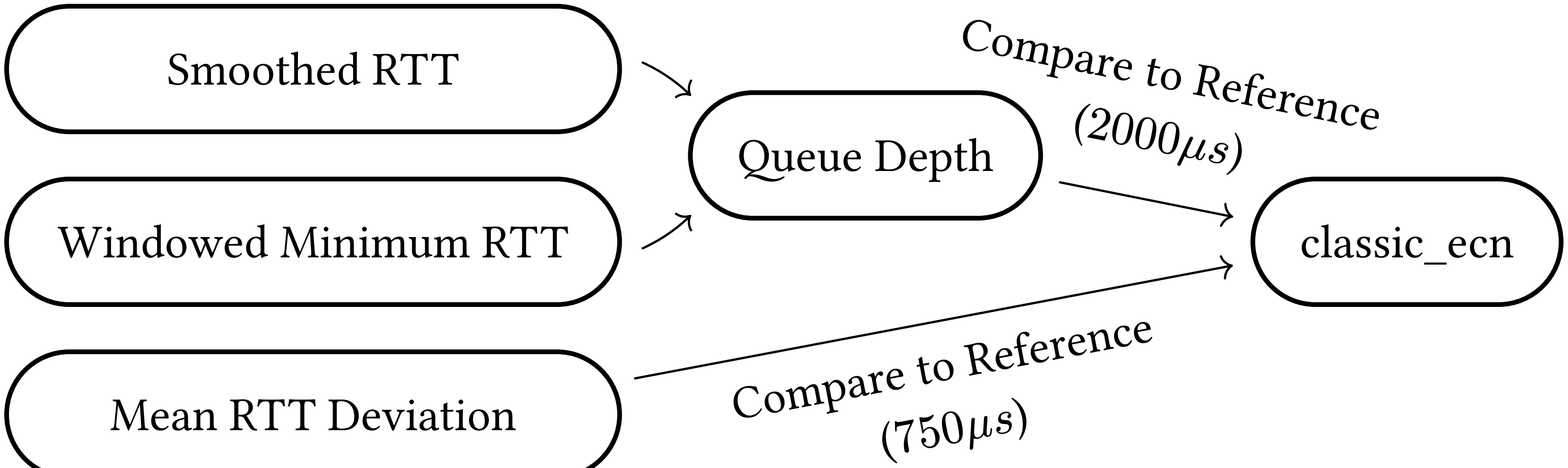
TCP Prague's Fallback

- Detect cases of an RFC 3168 compliant bottleneck AQM



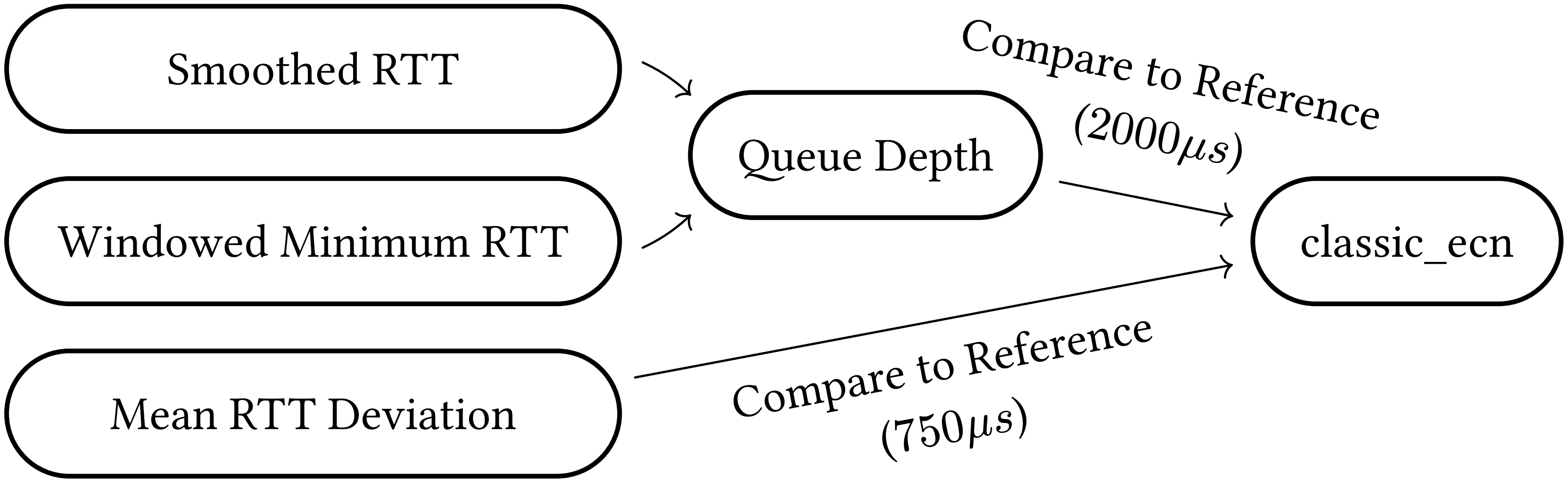
TCP Prague's Fallback

- Detect cases of an RFC 3168 compliant bottleneck AQM



TCP Prague’s Fallback

- Detect cases of an RFC 3168 compliant bottleneck AQM



- Increase response to congestion as classic_ecn increases

Related Work

- Extensively evaluated by its authors [1]

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

- Contradicted by newer research [2]

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

- Contradicted by newer research [2]
- Fixed link rate 100Mbps and RTT 10ms with long running flows

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

- Contradicted by newer research [2]
- Fixed link rate 100Mbps and RTT 10ms with long running flows
- Prague completely starves itself on L4S AQM

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

- Contradicted by newer research [2]
- Fixed link rate 100Mbps and RTT 10ms with long running flows
- Prague completely starves itself on L4S AQM
- Similar issue present in a followup paper [3]

Related Work

- Extensively evaluated by its authors [1]
- Different link rates, RTTs, traffic scenarios
- No misclassification of classic AQMs
- Rarely misclassified L4S AQMs (link rate $\leq$ 12Mbps or RTT $\geq$ 100ms)

- Contradicted by newer research [2]
- Fixed link rate 100Mbps and RTT 10ms with long running flows
- Prague completely starves itself on L4S AQM
- Similar issue present in a followup paper [3]

We investigate this stark difference in the literature

Contributions

We

- reproduce the results of recent literature in Mininet

Contributions

We

- reproduce the results of recent literature in Mininet
- extend their findings to different RTTs and link rates

Contributions

We

- reproduce the results of recent literature in Mininet
- extend their findings to different RTTs and link rates
- do a root cause analysis using eBPF to find a bug in the current fallback implementation

Contributions

We

- reproduce the results of recent literature in Mininet
- extend their findings to different RTTs and link rates
- do a root cause analysis using eBPF to find a bug in the current fallback implementation
- visualize the rate sharing of Prague with the patch

Contributions

We

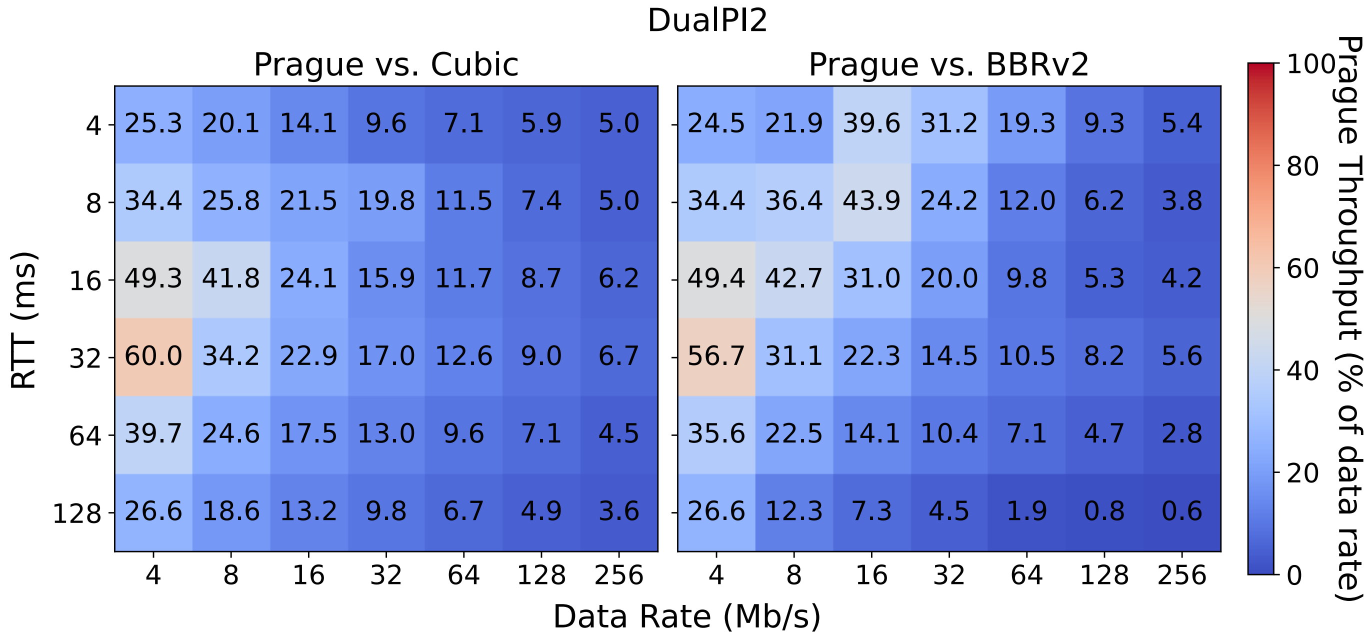
- reproduce the results of recent literature in Mininet
- extend their findings to different RTTs and link rates
- do a root cause analysis using eBPF to find a bug in the current fallback implementation
- visualize the rate sharing of Prague with the patch
- publish our experiment code for reproducibility

Reproducing Previous Results

- Mininet dumbbell topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow

Reproducing Previous Results

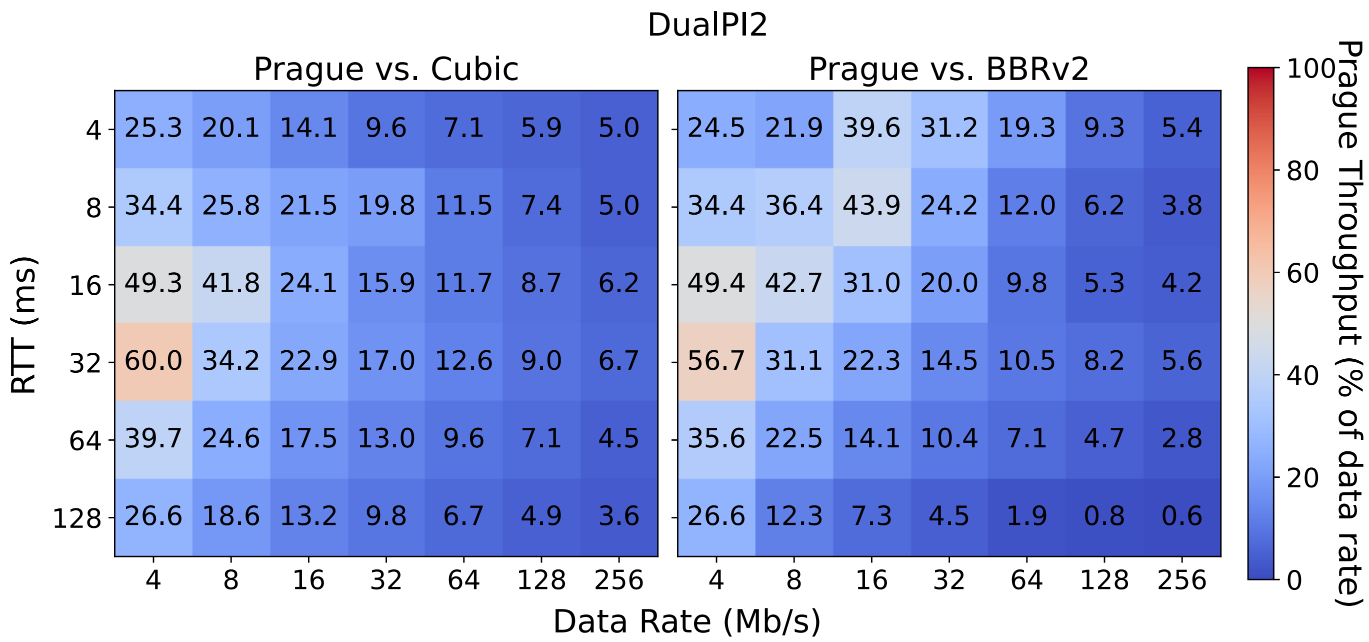
- Mininet dumbbell topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow



RTT (ms) \ Data Rate (Mb/s)	4	8	16	32	64	128	256
Prague vs. Cubic							
4	25.3	20.1	14.1	9.6	7.1	5.9	5.0
8	34.4	25.8	21.5	19.8	11.5	7.4	5.0
16	49.3	41.8	24.1	15.9	11.7	8.7	6.2
32	60.0	34.2	22.9	17.0	12.6	9.0	6.7
64	39.7	24.6	17.5	13.0	9.6	7.1	4.5
128	26.6	18.6	13.2	9.8	6.7	4.9	3.6
Prague vs. BBRv2							
4	24.5	21.9	39.6	31.2	19.3	9.3	5.4
8	34.4	36.4	43.9	24.2	12.0	6.2	3.8
16	49.4	42.7	31.0	20.0	9.8	5.3	4.2
32	56.7	31.1	22.3	14.5	10.5	8.2	5.6
64	35.6	22.5	14.1	10.4	7.1	4.7	2.8
128	26.6	12.3	7.3	4.5	1.9	0.8	0.6

Reproducing Previous Results

- Mininet dumbbell topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow



- Prague starves in *almost* all cases

Root Cause Analysis I

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure the value of classic_ecn at congestion window reduction

Root Cause Analysis I

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure the value of classic_ecn at congestion window reduction

Interval	Count
Pure Scalable	3
Transition	0
Pure Classic	485

Root Cause Analysis I

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure the value of classic_ecn at congestion window reduction

Interval	Count
Pure Scalable	3
Transition	0
Pure Classic	485

- Prague does fall back to classic mode on L4s

Root Cause Analysis II

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure queue depth and RTT deviation at classic_ecn update

Root Cause Analysis II

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure queue depth and RTT deviation at classic_ecn update

	Exceeds Reference Value		Less Than Reference Value	
Estimate	Count	Mean	Count	Mean
RTT Deviation	0		2324	42.7 μs
Queue Depth	3	2436 μs	2321	52.6 μs

Root Cause Analysis II

- Single run Prague vs. Cubic on DualPI2 at 100Mbps and 10ms for 60s
- Measure queue depth and RTT deviation at classic_ecn update

	Exceeds Reference Value		Less Than Reference Value	
Estimate	Count	Mean	Count	Mean
RTT Deviation	0		2324	42.7 μs
Queue Depth	3	2436 μs	2321	52.6 μs

- There is no reason for Prague's fallback to activate

Root Cause Analysis III

- Calculation scaling factor: 2^{24}

Root Cause Analysis III

- Calculation scaling factor: 2^{24}
- Reference value scaling factor: 2^{20}

```
#define V0_LG   (10014683ULL >> V)  /* reference queue V of ~750us */
#define D0_LG   (11498458ULL >> D)  /* reference queue D of ~2ms */
```

Root Cause Analysis III

- Calculation scaling factor: 2^{24}
- Reference value scaling factor: 2^{20}

```
#define V0_LG   (10014683ULL >> V)  /* reference queue V of ~750us */
#define D0_LG   (11498458ULL >> D)  /* reference queue D of ~2ms */
```

- Reference queue depth of $1.6\mu s$ and RTT deviation $1.5\mu s$

Root Cause Analysis III

- Calculation scaling factor: 2^{24}
- Reference value scaling factor: 2^{20}

```
#define V0_LG   (10014683ULL >> V)  /* reference queue V of ~750us */
#define D0_LG   (11498458ULL >> D)  /* reference queue D of ~2ms */
```

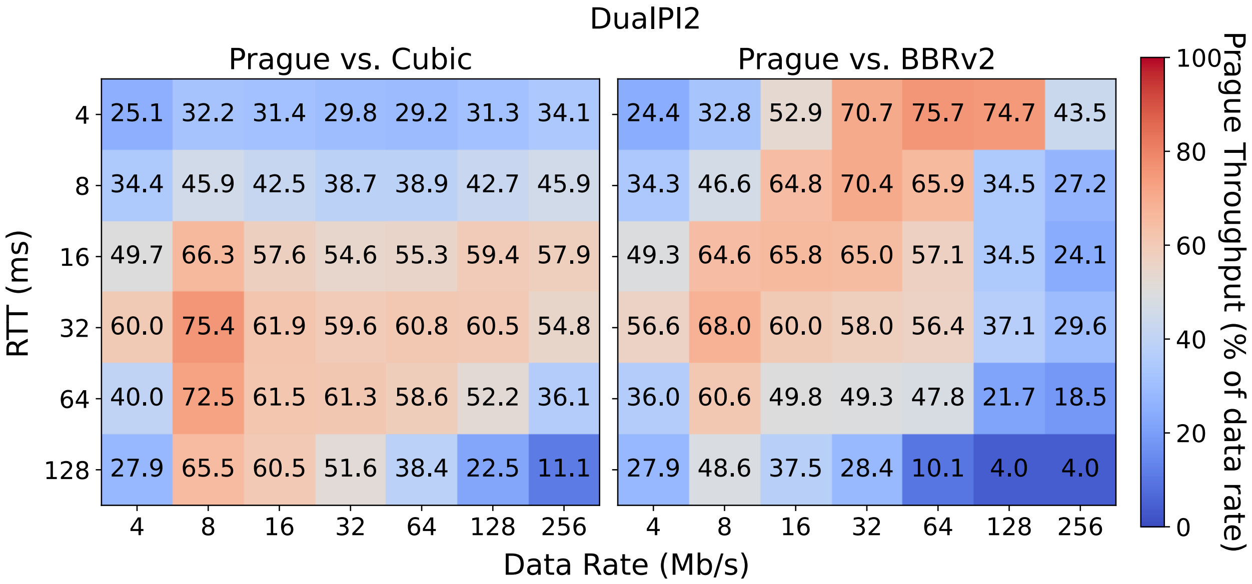
- Reference queue depth of $1.6\mu s$ and RTT deviation $1.5\mu s$
- Not achievable in any realistic setting $\Rightarrow$ Prague's fallback always activates

Evaluation of the Corrected Fallback: DualPI2

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague’s throughput against a single classic flow

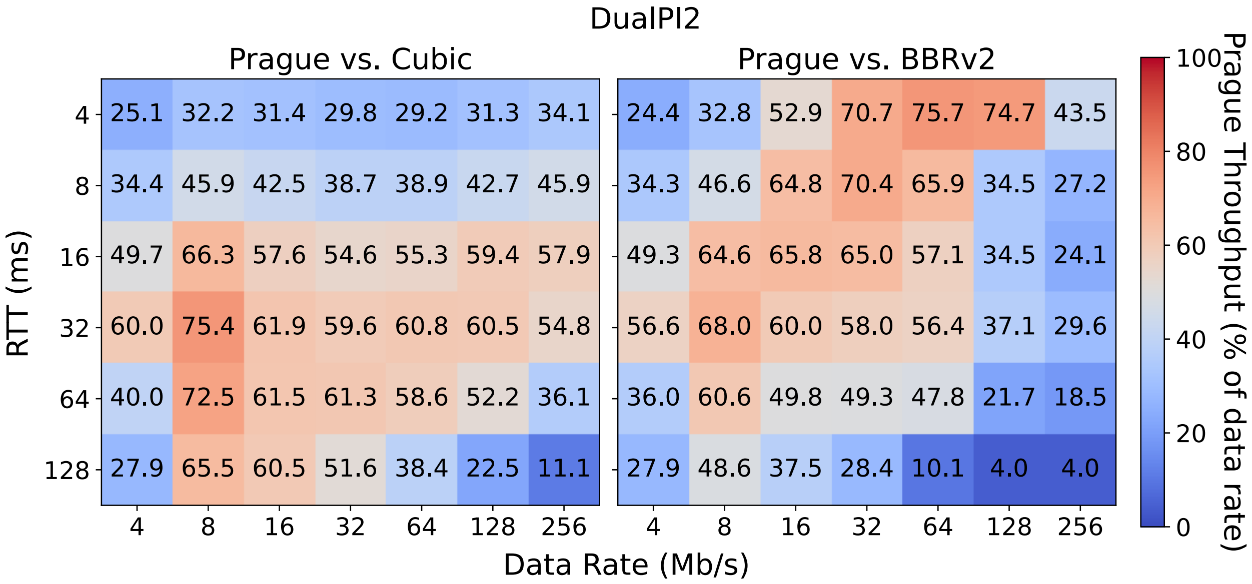
Evaluation of the Corrected Fallback: DualPI2

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow



Evaluation of the Corrected Fallback: DualPI2

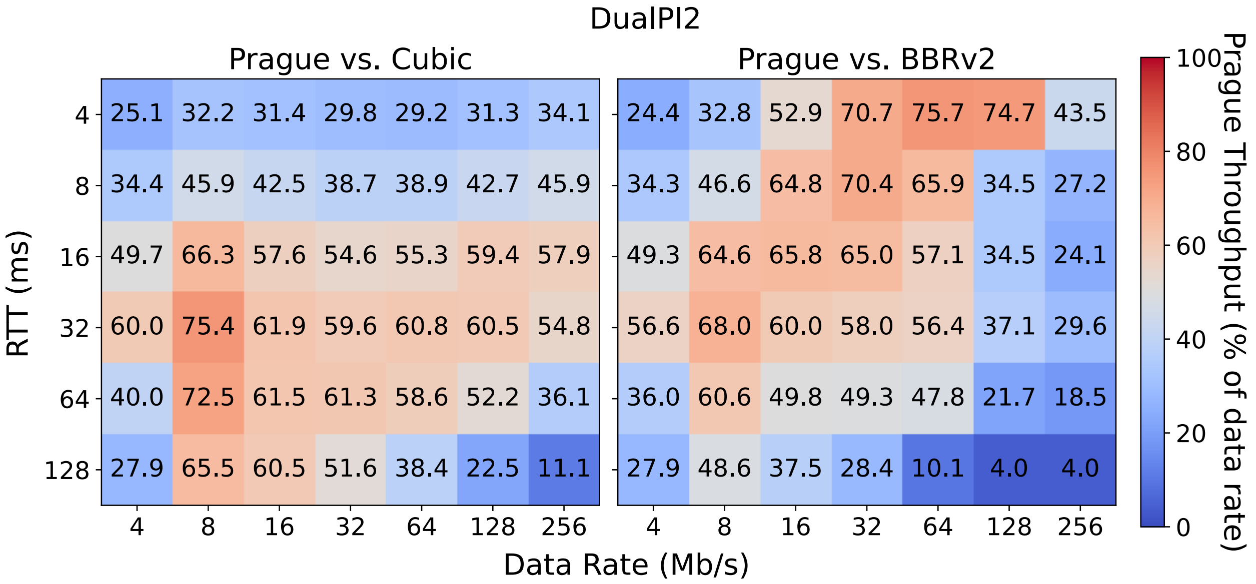
- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague’s throughput against a single classic flow



- Prague shares the bottleneck fairly in *most* cases

Evaluation of the Corrected Fallback: DualPI2

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow



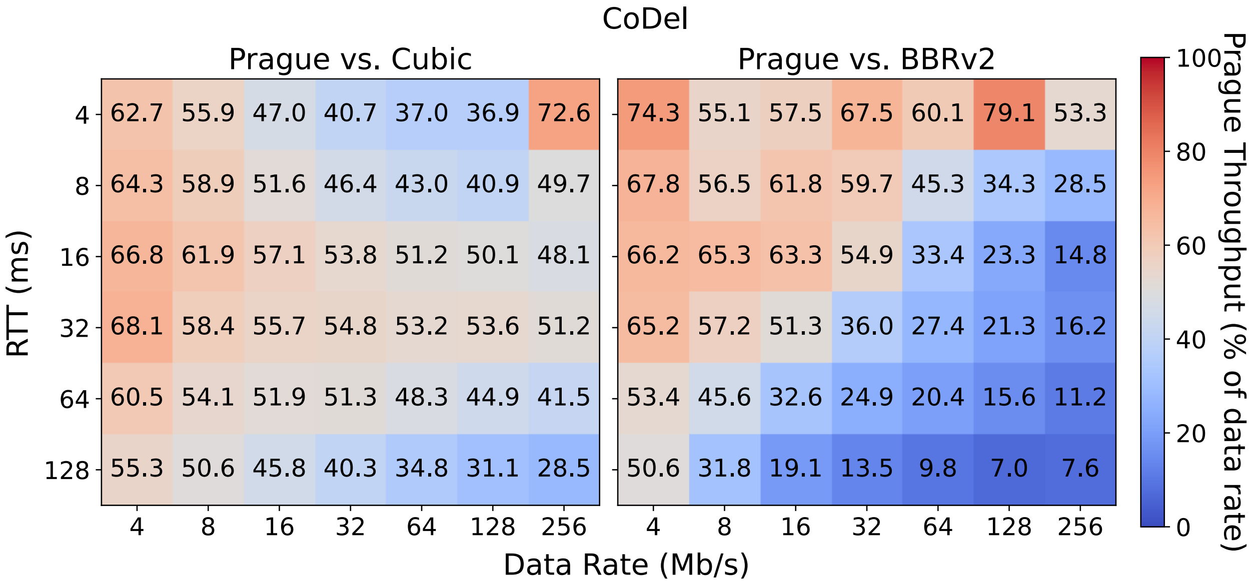
- Prague shares the bottleneck fairly in *most* cases
- Prague's fallback only activated at $\leq$ 4Mbps, similar to the original authors results [1]

Evaluation of the Corrected Fallback: CoDel

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague’s throughput against a single classic flow on a classic AQM

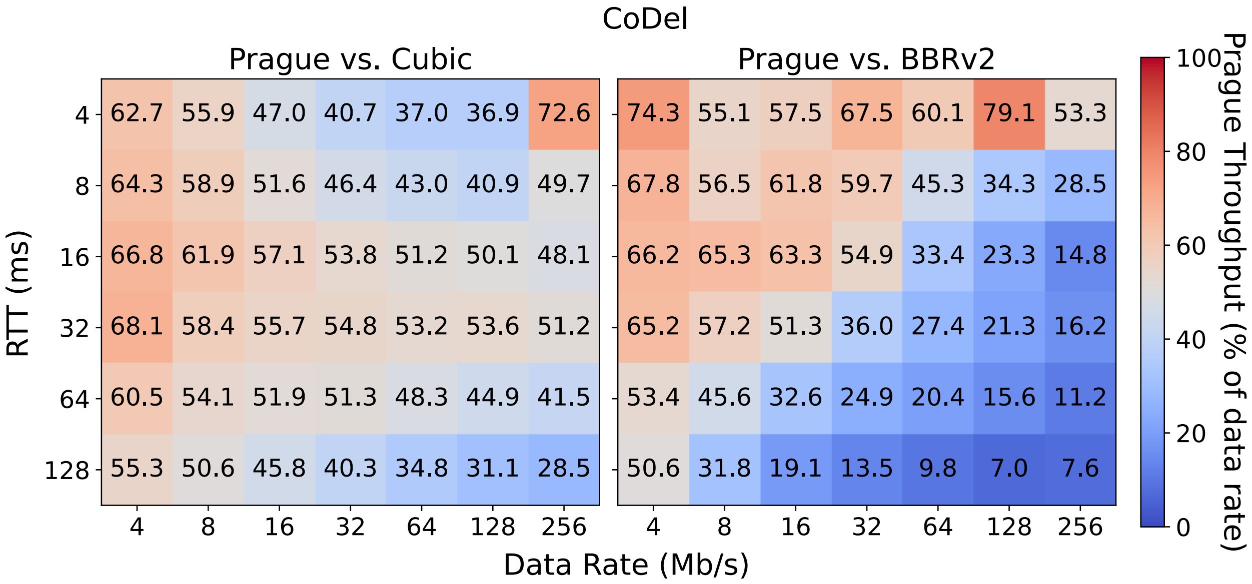
Evaluation of the Corrected Fallback: CoDel

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow on a classic AQM



Evaluation of the Corrected Fallback: CoDel

- Mininet Dumbbell Topology with different link rates and RTTs and an L4S bottleneck
- Measuring Prague's throughput against a single classic flow on a classic AQM



- Prague still shares the bottleneck fairly

Conclusion

We

Conclusion

We

- discovered an issue with Prague's fallback implementation

Conclusion

We

- discovered an issue with Prague’s fallback implementation
- explained the discrepency in the past literature

Conclusion

We

- discovered an issue with Prague's fallback implementation
- explained the discrepency in the past literature
- compared the rate sharing capabilities of Prague with Cubic and BBRv2 on CoDel and DualPI2

Questions?

References

[1] B. Briscoe and A. S. Ahmed, "TCP Prague Fall-back on Detection of a Classic ECN AQM," *arXiv preprint arXiv:1911.00710*, 2019.

[2] F. B. Sarpkaya, A. Srivastava, F. Fund, and S. Panwar, "To switch or not to switch to TCP Prague? Incentives for adoption in a partial L4S deployment," in *Proceedings of the 2024 Applied Networking Research Workshop*, 2024, pp. 45–52.

[3] F. B. Sarpkaya, F. Fund, and S. Panwar, "To adopt or not to adopt L4S-compatible congestion control? Understanding performance in a partial L4S deployment," in *International Conference on Passive and Active Network Measurement*, 2025, pp. 217–246.