

# Transport Architectures for an Evolving Internet

Keith Winstein

MIT Computer Science and Artificial Intelligence Laboratory

March 5, 2014



Joint work with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan.

# The Internet evolves

In 20 years, computer networks have seen dramatic change:

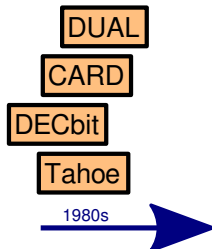
- ▶ Wi-Fi
- ▶ Cellular networks
- ▶ Datacenters
- ▶ 10 GigE
- ▶ Transoceanic links
  
- ▶ Ubiquitous mobility
- ▶ Huge amounts of streaming video

# Coping with change

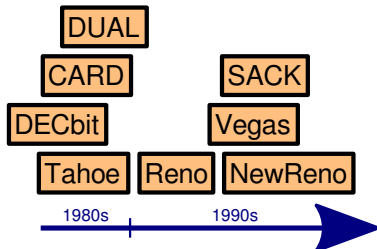
How should users deal with an evolving network?

One approach: design new protocols.

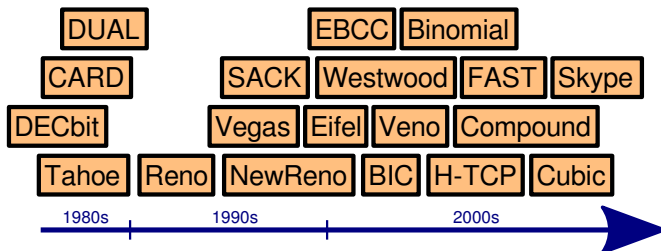
# The march of congestion-control protocols



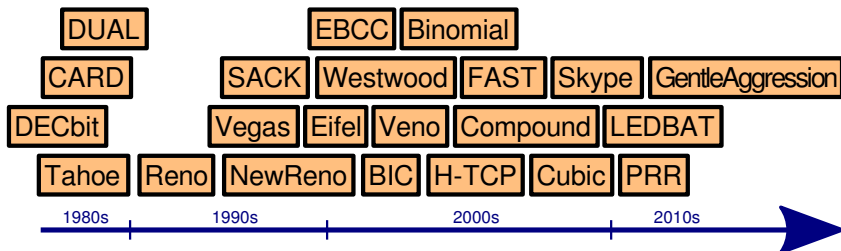
# The march of congestion-control protocols



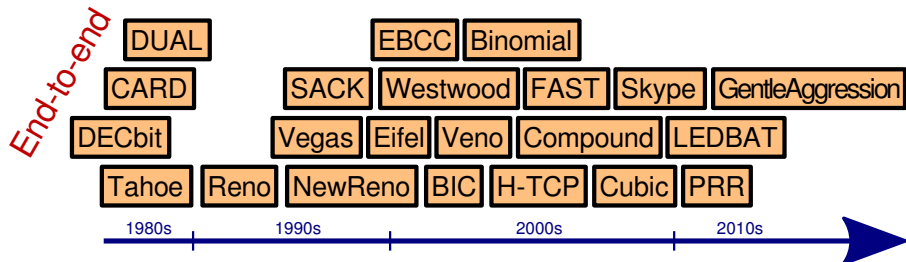
# The march of congestion-control protocols



# The march of congestion-control protocols

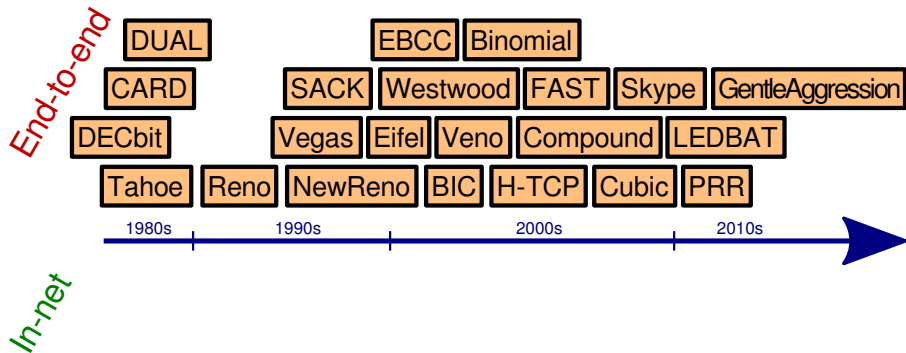


# The march of congestion-control protocols

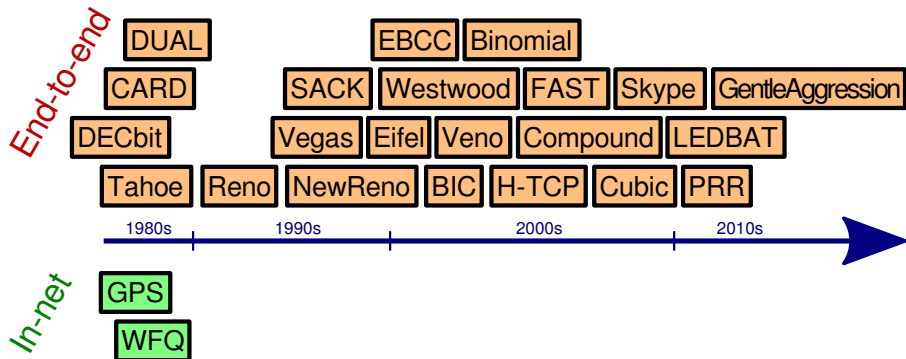




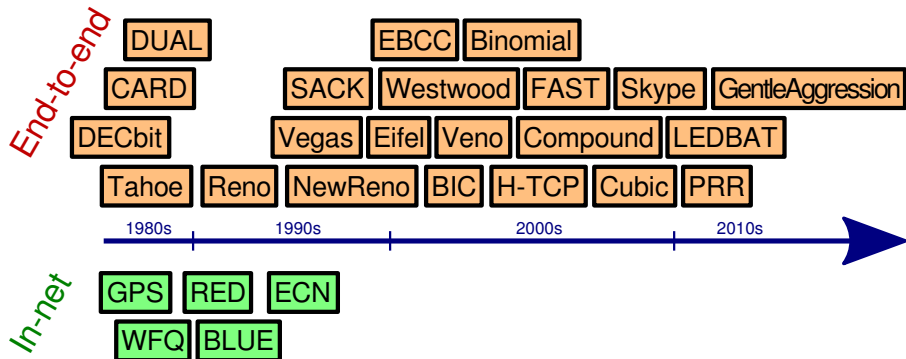
# The march of congestion-control protocols



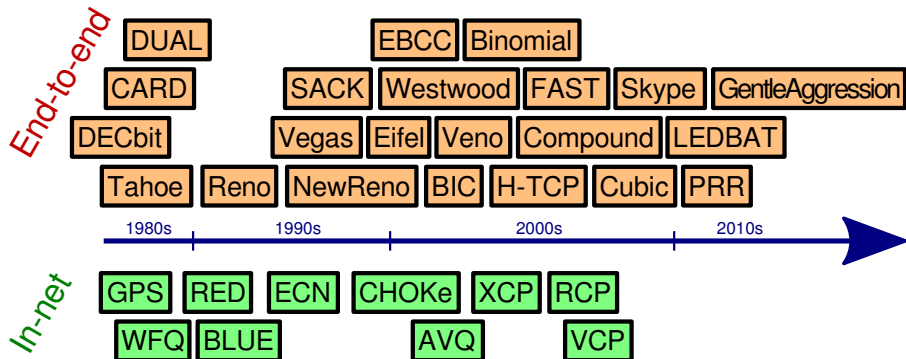
# The march of congestion-control protocols



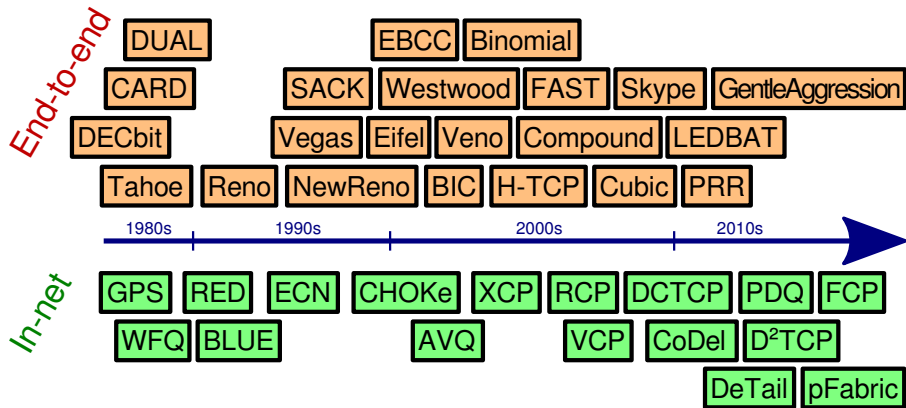
# The march of congestion-control protocols



# The march of congestion-control protocols



# The march of congestion-control protocols



# Declarative design

Systems with a **model** and a **mission**.

# Explicitness in systems design

**Model:** explicit statement of assumptions about the problem

**Mission:** objective that the application wants

Explicit design considerations → **freedom to make changes**

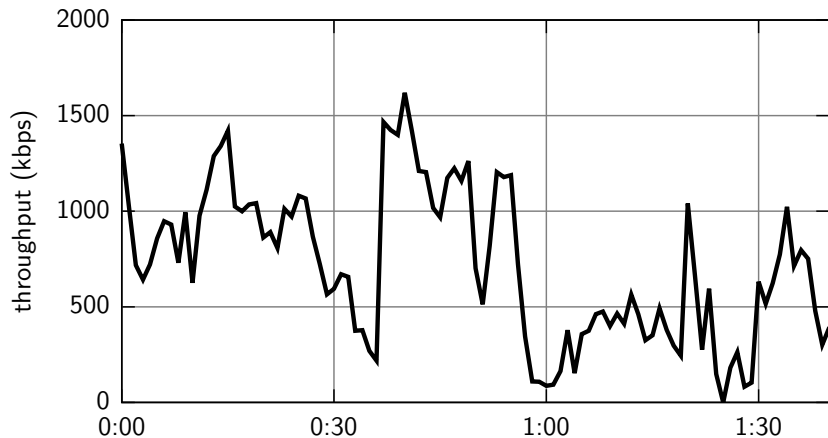
# Sprout: a transport protocol designed for variability

## *Observation:*

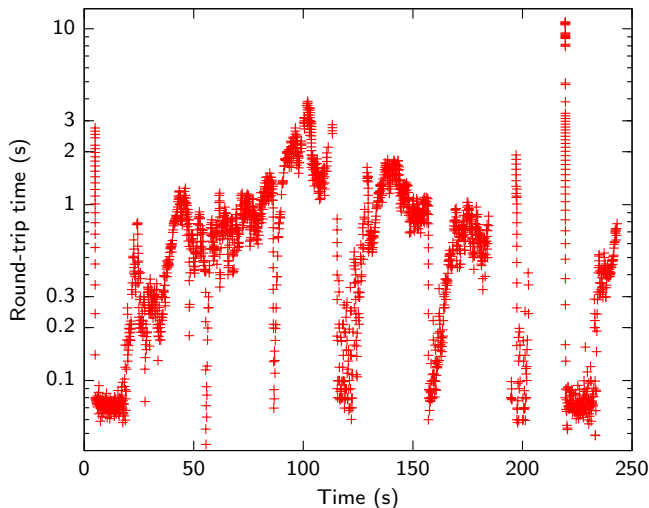
Videoconferences perform poorly over cellular networks.



# Verizon LTE uplink throughput



# Verizon LTE ping delay during one TCP download

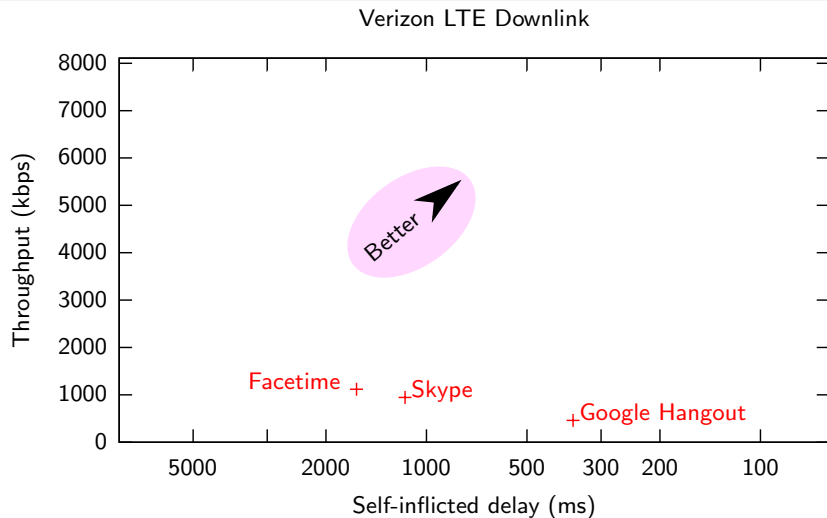


# Interactive apps work poorly

- ▶ We measured cellular networks while driving:
  - ▶ **Verizon LTE**
  - ▶ Verizon 3G (1xEV-DO)
  - ▶ AT&T LTE
  - ▶ T-Mobile 3G (UMTS)
- ▶ Then ran apps across replayed network trace:
  - ▶ **Skype** (Windows 7)
  - ▶ Google Hangouts (Chrome on Windows 7)
  - ▶ Apple Facetime (OS X)

# Skype's performance

# Performance summary



# What's wrong?

- ▶ Existing schemes **react** to congestion signals.
  - ▶ Packet loss.
  - ▶ Increase in round-trip time.
- ▶ Feedback comes too late.
- ▶ The killer: **self-inflicted queueing delay**.

# Sprout's mission

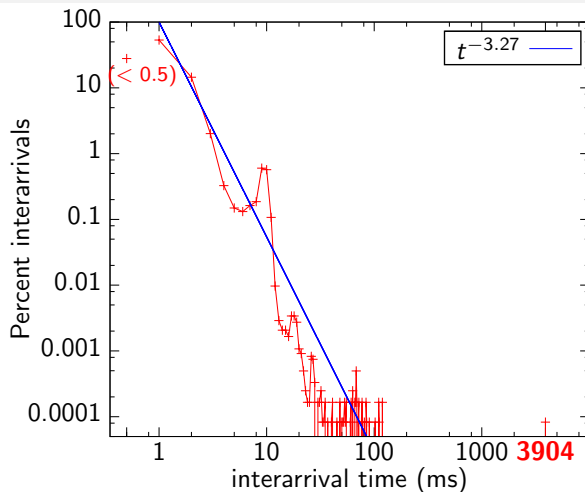
- ▶ Most throughput
- ▶ Bounded risk of delay  $> 100$  ms

# Bounded risk of delay

- ▶ **Model** variation in link speed
- ▶ **Infer** current link speed
- ▶ **Predict** future link speed
  - ▶ Don't wait for congestion
- ▶ **Control:** Send as much as possible, but require:
  - ▶ 95% chance all packets arrive within 100 ms

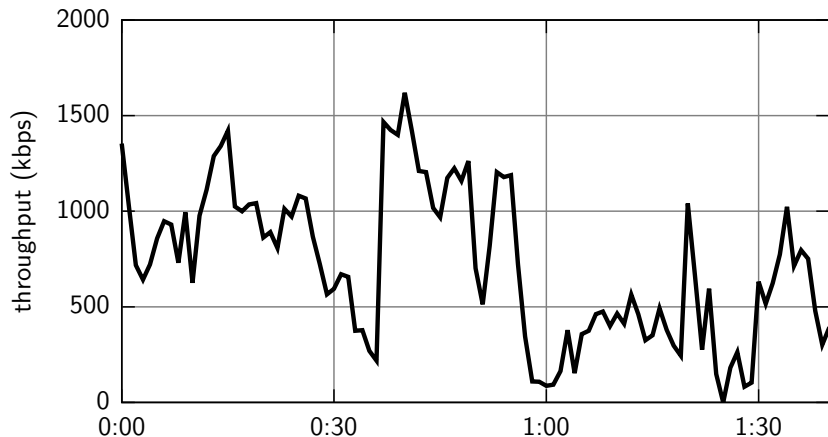


# Model: packet deliveries looks like flicker noise

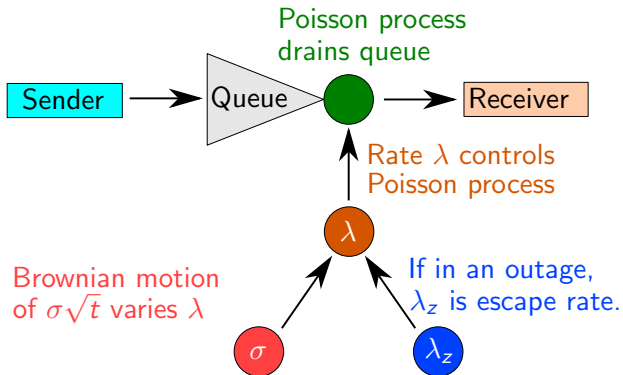


(Verizon LTE, phone stationary.)

## Model: average rate looks like random walk



# Sprout's model



## Sprout's model parameters

Volatility $\sigma$ : fixed @	$200 \frac{\text{pkts}/s}{\sqrt{s}}$
Expected outage time $1/\lambda_z$ :	$1 s$
Tick length ( $\tau$ ):	$20 ms$
Forecast length:	$160 ms$
Delay target:	$100 ms$
Risk tolerance:	$5\%$

All source code was **frozen before data collection began.**

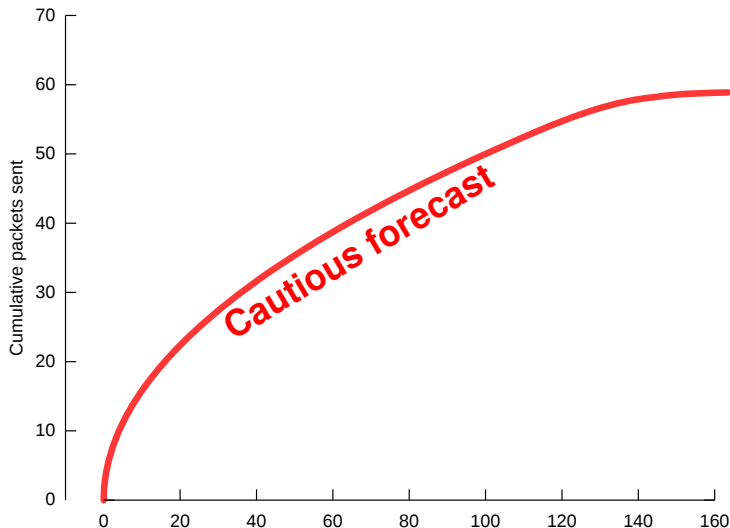
## Infer: current link speed

- ▶ **Observe** packets received every  $\tau$
- ▶ **Update**  $P(\lambda)$

## Predict: future link speed

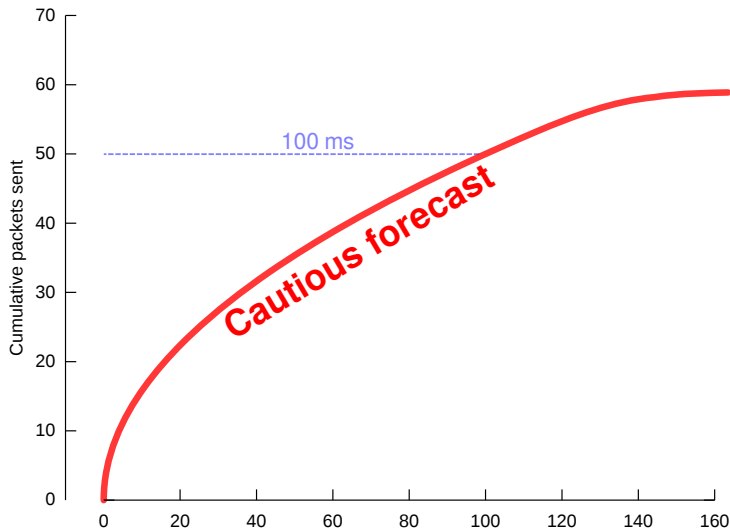
- ▶ **Evolve** model forward
- ▶ **Forecast** 5th percentile cumulative packets

## Control: fill up 100 ms forecast window



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

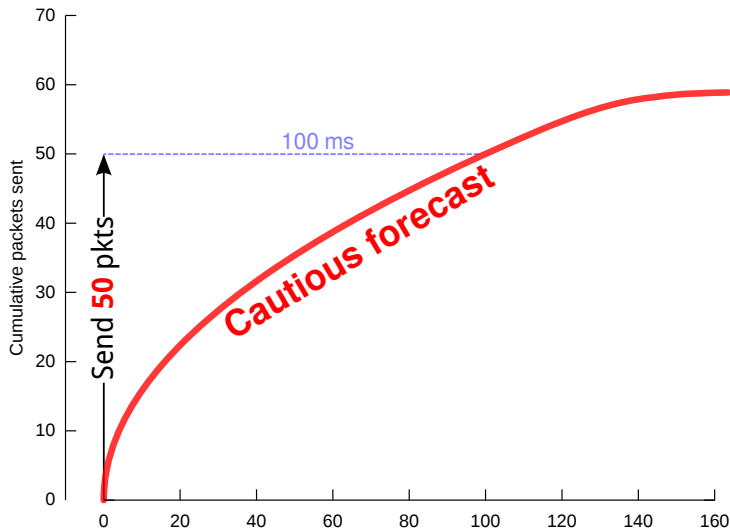
## Control: fill up 100 ms forecast window



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

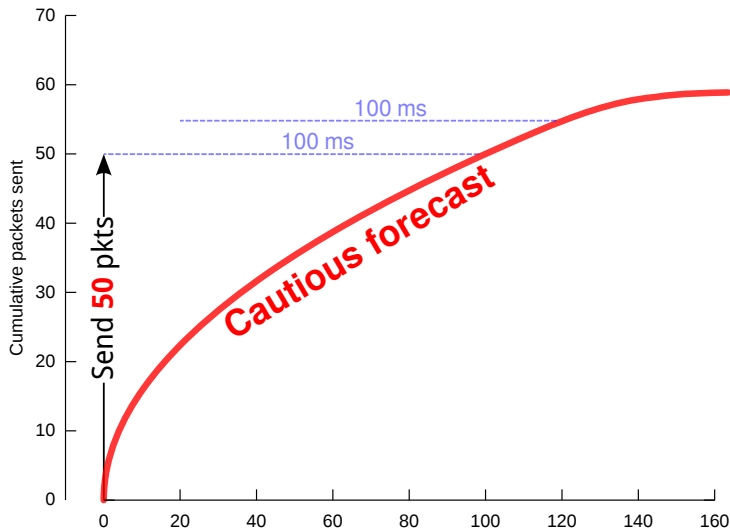


## Control: fill up 100 ms forecast window



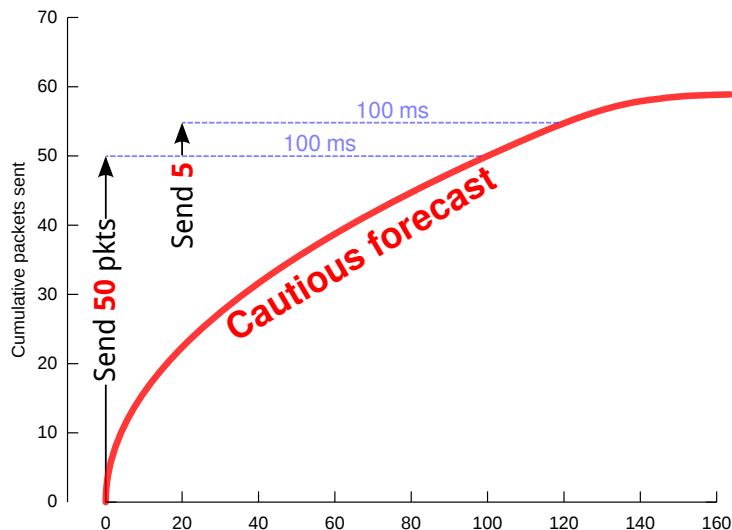
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Control: fill up 100 ms forecast window



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

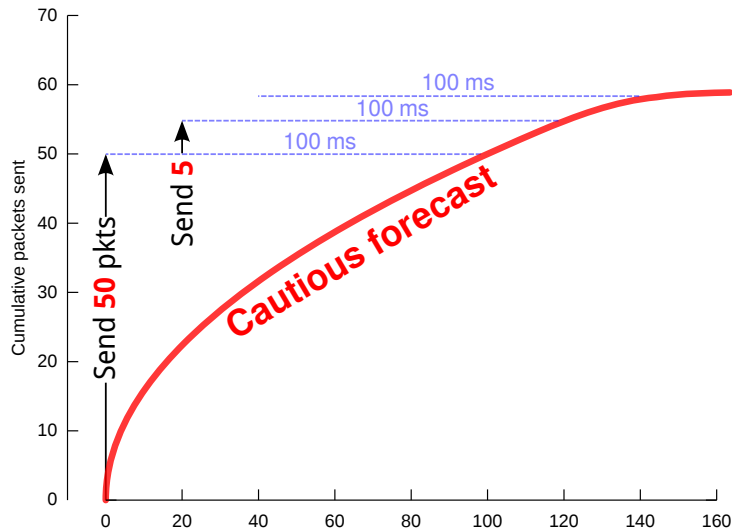
## Control: fill up 100 ms forecast window



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

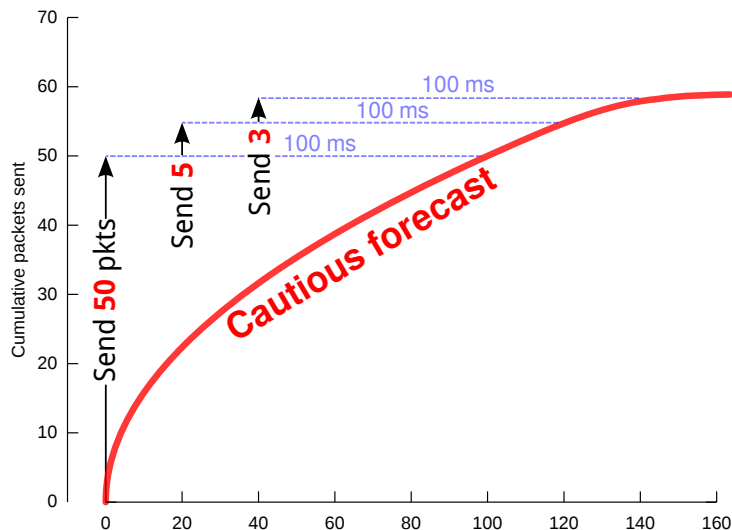
## Control: fill up 100 ms forecast window



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

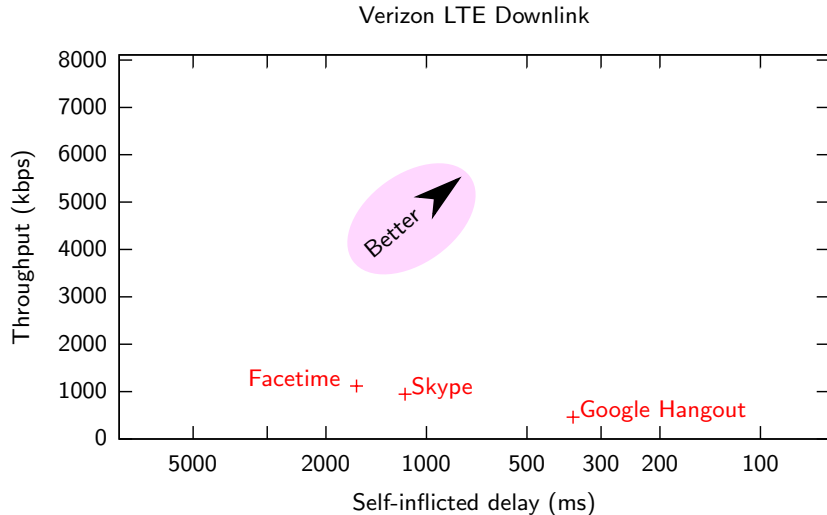
## Control: fill up 100 ms forecast window

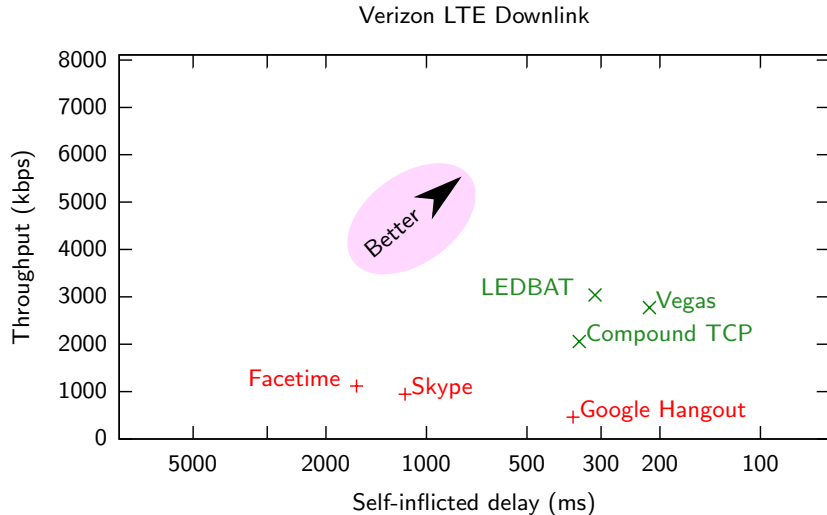


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

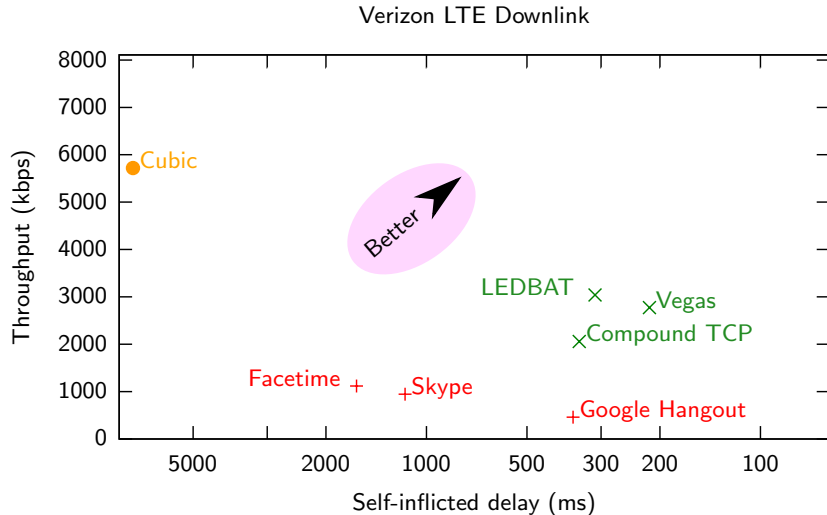
Transport Architectures for an Evolving Internet

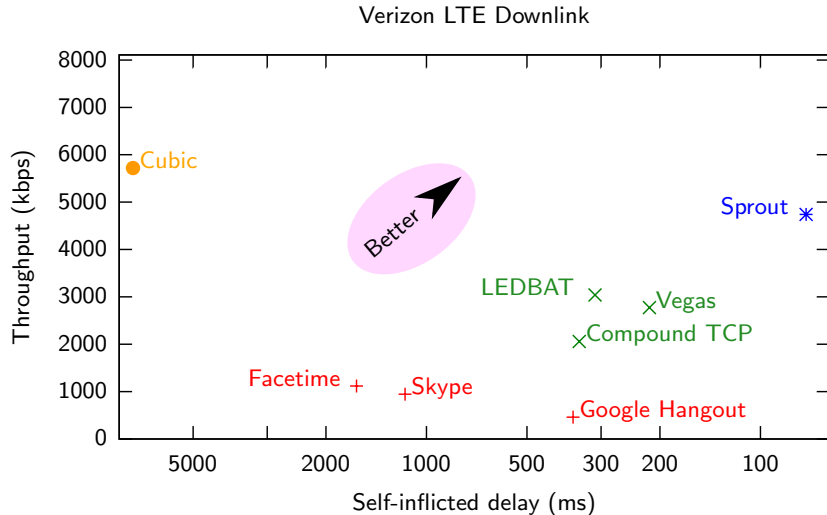
# Sprout's results

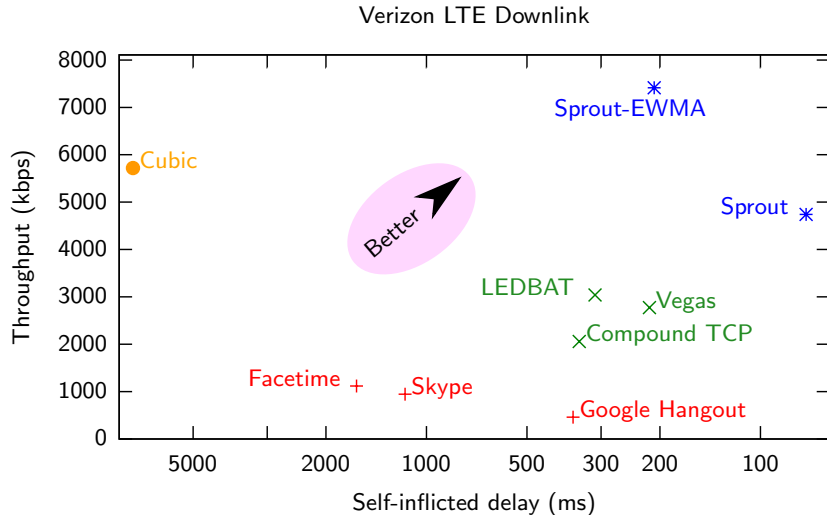


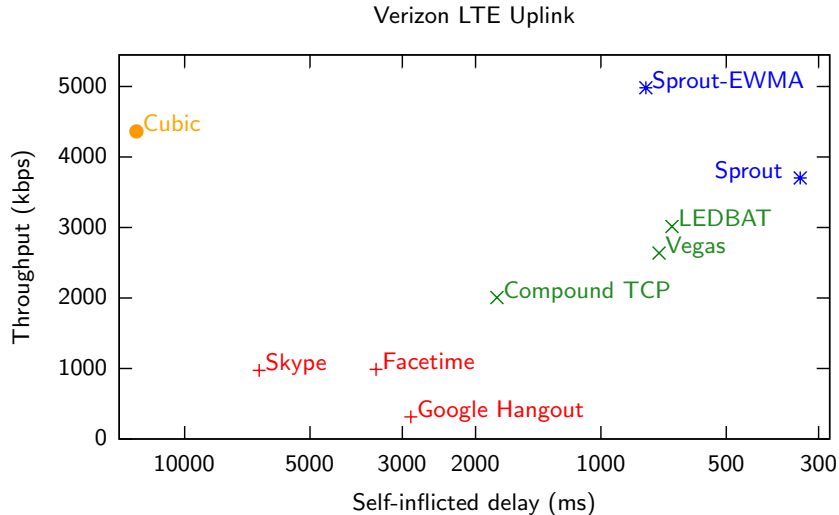










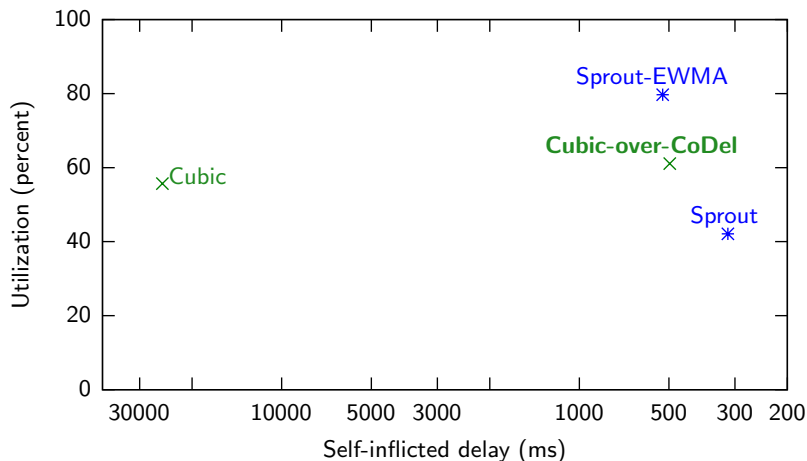


## Overall results on 8 links

Verizon 3G/LTE, AT&T LTE, T-Mobile 3G uplink and downlink:

Sprout vs.	Avg. speedup	Delay reduction
Skype	2.2×	7.9×
Hangout	4.4×	7.2×
Facetime	1.9×	8.7×
Compound	1.3×	4.8×
TCP Vegas	1.1×	2.1×
LEDBAT	Same	2.8×
Cubic	0.91×	79×

# Sprout is end-to-end, but comparable to in-net control



## M.I.T. 6.829 contest (March–April 2013)

- ▶ Turnkey network emulator, evaluation
- ▶ Sender, receiver run in Linux containers
- ▶ **Mission**: maximize throughput/delay
- ▶ 4th prize: \$20
- ▶ 3rd prize: \$30
- ▶ 2nd prize: \$40
- ▶ (If beat Sprout) 1st prize:

## M.I.T. 6.829 contest (March–April 2013)

- ▶ Turnkey network emulator, evaluation
- ▶ Sender, receiver run in Linux containers
- ▶ **Mission**: maximize throughput/delay
- ▶ 4th prize: \$20
- ▶ 3rd prize: \$30
- ▶ 2nd prize: \$40
- ▶ (If beat Sprout) 1st prize: **Co-authorship on future paper**

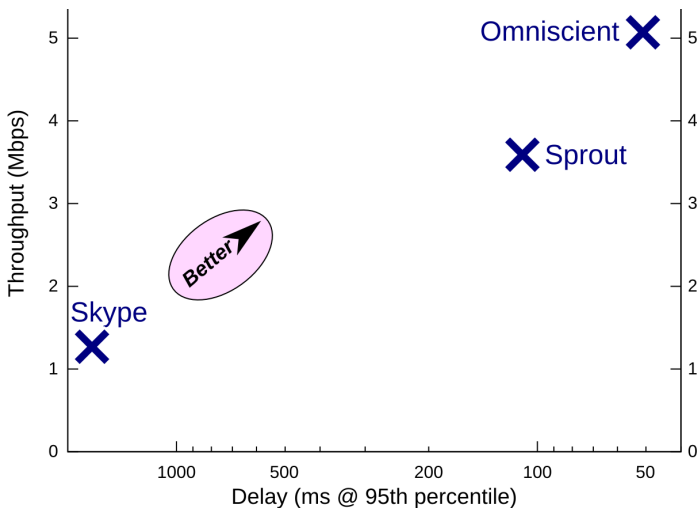


## M.I.T. 6.829 contest (March–April 2013)

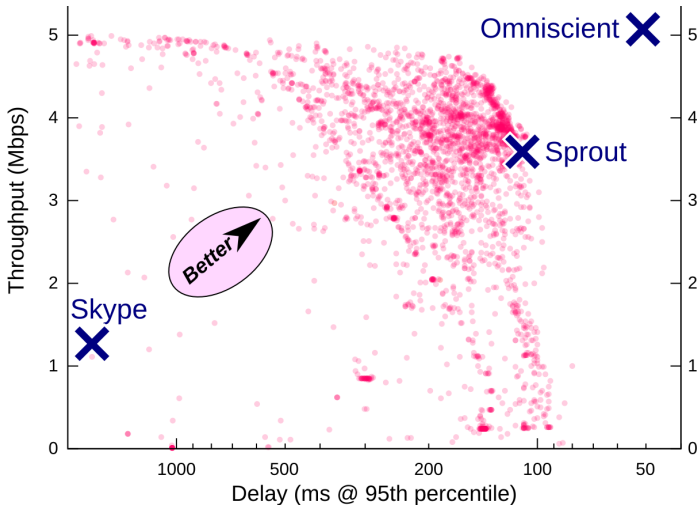
- ▶ Turnkey network emulator, evaluation
- ▶ Sender, receiver run in Linux containers
- ▶ **Mission**: maximize throughput/delay
- ▶ 4th prize: \$20
- ▶ 3rd prize: \$30
- ▶ 2nd prize: \$40
- ▶ (If beat Sprout) 1st prize: **Co-authorship on future paper**

Anirudh Sivaraman, KW, Pauline Varley, Somak Das, Joshua Ma, Ameesh Goyal, João Batalha, and Hari Balakrishnan, **Protocol Design Contests**, *in submission*

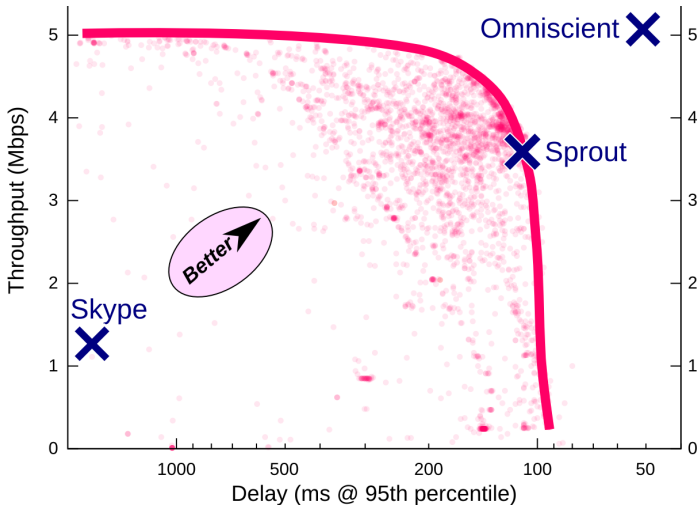
# Baseline



# Land of 3,000 student protocols



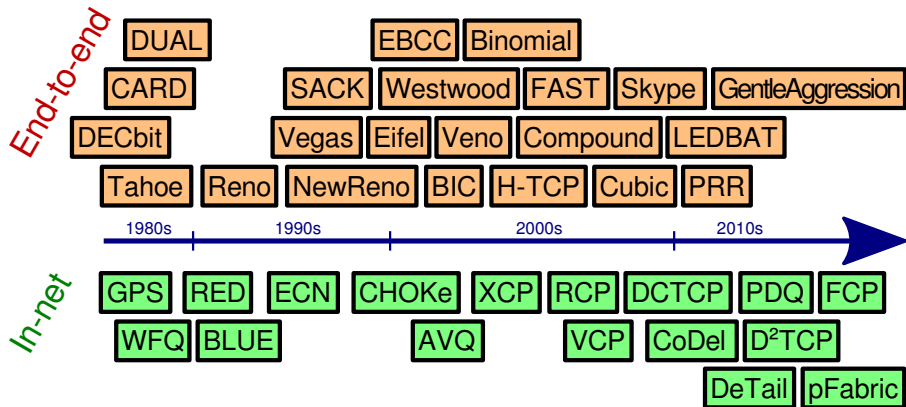
# Sprout was on the frontier



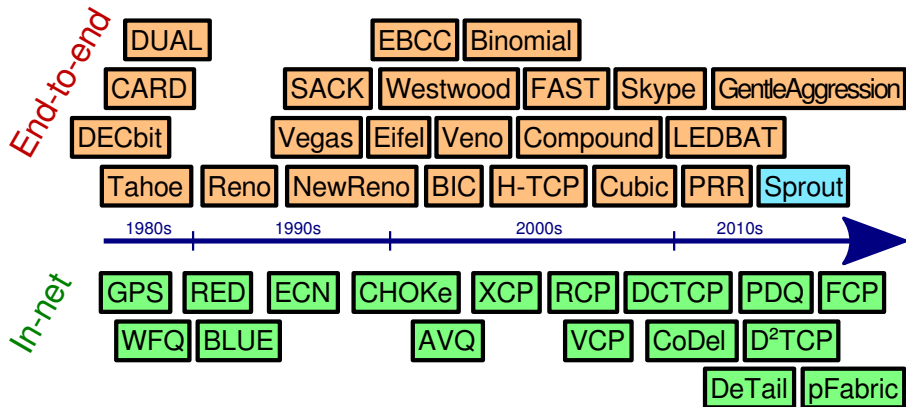
# Limitations

- ▶ Sprout wants to control all of the traffic on a queue.
  - ▶ Cells generally have **per-user** queues. . .
  - ▶ . . . but Wi-Fi and wired networks usually don't.
- ▶ What if cell link *isn't* the bottleneck?
- ▶ Assumption: application always has data to send

# Sprout's mark



# Sprout's mark



## Now that we have 40+ algorithms. . .

- ▶ Sprout for cellular networks?
- ▶ Wireless-TCP for Wi-Fi?
- ▶ High-BDP-TCP for transoceanic links?
- ▶ Datacenter-TCP for datacenters?
- ▶ CoDel for cable modems?
- ▶ **TBA-TCP for tomorrow's networks?**



# Rational choice of scheme is challenging

**Cubic** vs. **Compound**

- ▶ Different missions?
- ▶ Different assumptions about network?
- ▶ One scheme just plain better?

# Networks constrained by a fuzzy idea of TCP's assumptions

- ▶ Mask stochastic loss
- ▶ Bufferbloat
- ▶ Mask out-of-order delivery
- ▶ No parallel/multipath routing

*Advice for Internet Subnetwork Designers*  
(RFC 3819) is 21,000 words!

# Apps hack around TCP

- ▶ Open lots of flows
- ▶ Goose slow start
- ▶ Add pacing
- ▶ Give up and do it yourself

# Apps hack around TCP

- ▶ Open lots of flows
- ▶ Goose slow start
- ▶ Add pacing
- ▶ Give up and do it yourself



# Apps hack around TCP

- ▶ Open lots of flows
- ▶ Goose slow start
- ▶ Add pacing
- ▶ Give up and do it yourself



Google MICROSOFT

# Apps hack around TCP

- ▶ Open lots of flows
- ▶ Goose slow start
- ▶ Add pacing
- ▶ Give up and do it yourself



Google MICROSOFT

YouTube

# Apps hack around TCP

- ▶ Open lots of flows
- ▶ Goose slow start
- ▶ Add pacing
- ▶ Give up and do it yourself



Google MICROSOFT

YouTube

Chrome (QUIC)  
BitTorrent ( $\mu$ TP)  
Mosh (SSP)  
IBM Aspera (fasp)

# Idea: computer-generated protocols

Transport layer should adapt to **whatever**:

- ▶ network does
- ▶ application wants



# Idea: computer-generated protocols

Transport layer should adapt to **whatever**:

- ▶ network does (**model**)
- ▶ application wants (**mission**)

# What we built

**Remy:** a program that generates congestion-control schemes offline

## Input:

- ▶ Assumptions about network and workload (**model**)
- ▶ Application's objective (**mission**)

**Output:** CC algorithm for a TCP sender (**RemyCC**)

**Time:** hours to days

# The basic question of congestion control

At this moment, do I:

- ▶ send a packet
- ▶ not send a packet?

# Missions of congestion control

## Maximize

- ▶  $\sum_i \log [\text{throughput}_i]$  (proportionally fair throughput)

# Missions of congestion control

## Maximize

- ▶  $\sum_i \log [\text{throughput}_i]$  (proportionally fair throughput)

- ▶  $\sum_i \log \left[ \frac{\text{throughput}_i}{\text{delay}_i} \right]$  (proportionally fair throughput/delay)

# Missions of congestion control

## Maximize

▶  $\sum_i \log [\text{throughput}_i]$  (proportionally fair throughput)

▶  $\sum_i \log \left[ \frac{\text{throughput}_i}{(\text{delay}_i)^\delta} \right]$  (proportionally fair throughput/delay)

# Missions of congestion control

## Maximize

- ▶  $\sum_i \log [\text{throughput}_i]$  (proportionally fair throughput)
- ▶  $\sum_i \log \left[ \frac{\text{throughput}_i}{(\text{delay}_i)^\delta} \right]$  (proportionally fair throughput/delay)
- ▶  $\min_i \text{throughput}_i$  (max-min throughput)

## Minimize

- ▶ mean flow completion time
- ▶ page load time

## Prevent

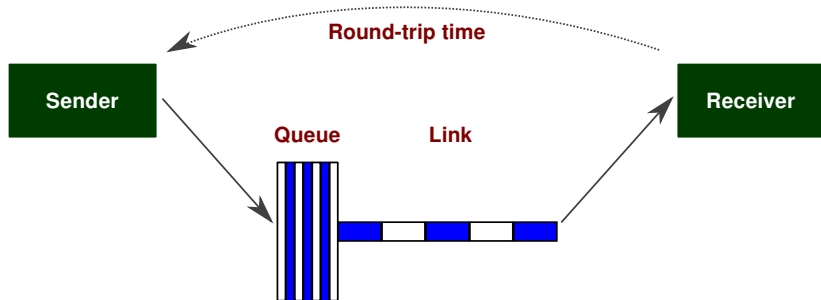
- ▶ pathological behavior
- ▶ congestion collapse

# Encoding the designer's prior assumptions

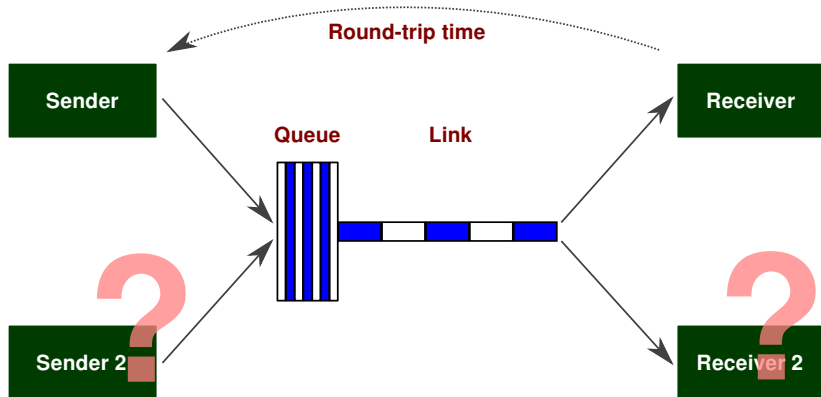
- ▶ **Model** of network uncertainty
  - ▶ Link speed distribution
  - ▶ Delay distribution
  - ▶ Topology distribution
- ▶ **Model** of workload
  - ▶ Web browsing
  - ▶ MapReduce
  - ▶ videoconferencing
  - ▶ streaming video (YouTube/Netflix)



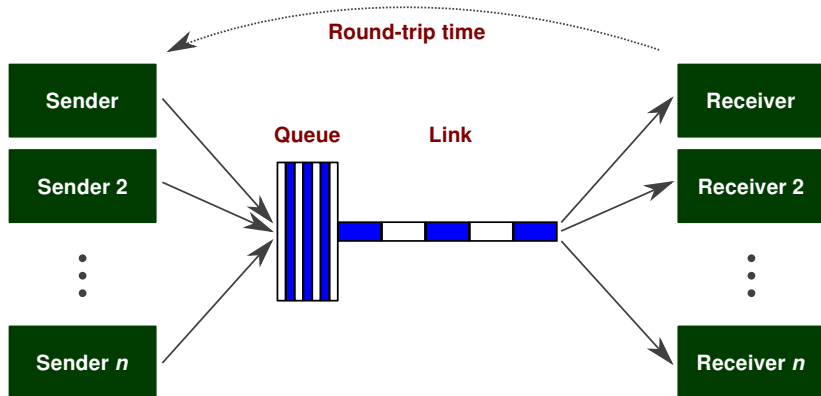
# Dumbbell network



# Dumbbell network



# Dumbbell network



# Superrational congestion control

At this moment, \* do I:

- ▶ send a packet
- ▶ not send a packet?

# Superrational congestion control

At this moment,\* do I:

- ▶ send a packet
- ▶ not send a packet?

\* Assuming every node is running the same algorithm.

## Remy: tractable search for best policy

- ▶ Best decision given all history: not tractable
- ▶ Instead, **summarize the history**

# A RemyCC tracks four congestion signals

$r\_ewma_{\alpha}$ : **short-term** moving average of interval between acks  
*“How fast are packets arriving (now)?”*

$r\_ewma_{\beta}$ : **long-term** moving average of same  
*“How fast are packets arriving (smoothed)?”*

$s\_ewma$ : moving average of interval between acked timestamps  
*“How fast was I sending?”*

$rtt\_ratio$ : ratio of last RTT to smallest RTT so far  
*“How long is the queue?”*

## Why these four features?

- ▶ We can measure the benefit of each!
- ▶ Removing any one hurts
  - ▶ losing  $r\_ewma_{\alpha}$  hurts the most
- ▶ More signals increase search time
- ▶ ...but others might help on some networks



## A RemyCC maps each state to an action

$$\text{REMYCC}(r\_ewma_{\alpha\beta}, s\_ewma, rtt\_ratio) \rightarrow \langle m, b, \tau \rangle$$

$m$  Multiple to congestion window

$b$  Increment to congestion window

$\tau$  Minimum interval between two outgoing packets

# Runtime for a RemyCC

## On ack:

- ▶  $\langle m, b, \tau \rangle \leftarrow \text{REMYCC}(r\_ewma_{\alpha\beta}, s\_ewma, rtt\_ratio)$
- ▶  $wnd \leftarrow m \cdot wnd + b$

## Send packet if:

- ▶  $wnd > \text{FlightSize}$ , and
- ▶ last packet sent  $> \tau$  ago

## Remy's job

Find piecewise-continuous `REMYCC()` that optimizes expected value of objective function

# Remy example: 2D state space

**On ack:**

$$\langle m, b, \tau \rangle \leftarrow \text{REMYCC}(s\_ewma, r\_ewma_{\alpha}, r\_ewma_{\beta}, rtt\_ratio)$$

# Remy example: 2D state space

**On ack:**

$$\langle m, b, \tau \rangle \leftarrow \text{REMYCC}(s\_ewma, r\_ewma_{\alpha}, \text{[redacted]})$$

## Remy example: **model**

Quantity	Distribution	Units
Link speed	Uniform(10, 20)	Mbps
RTT	Uniform(100, 200)	ms
$n$	Uniform(1, 16)	
“On” process	$\exp[\mu = 5]$	seconds
“Off” process	same	

## Remy example: mission

$$\sum_i \log \left[ \frac{\text{throughput}_i}{\text{delay}_i} \right]$$

One action for all states. Find the best value.

r\_ewma

$\langle ?, ?, ? \rangle$

s\_ewma



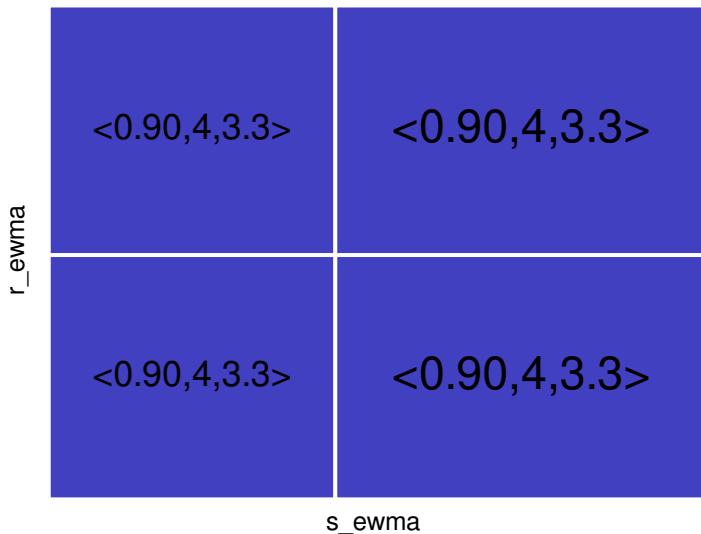
The best (single) action. Now split it on median.

r\_ewma

$\langle 0.90, 4, 3.3 \rangle$

s\_ewma

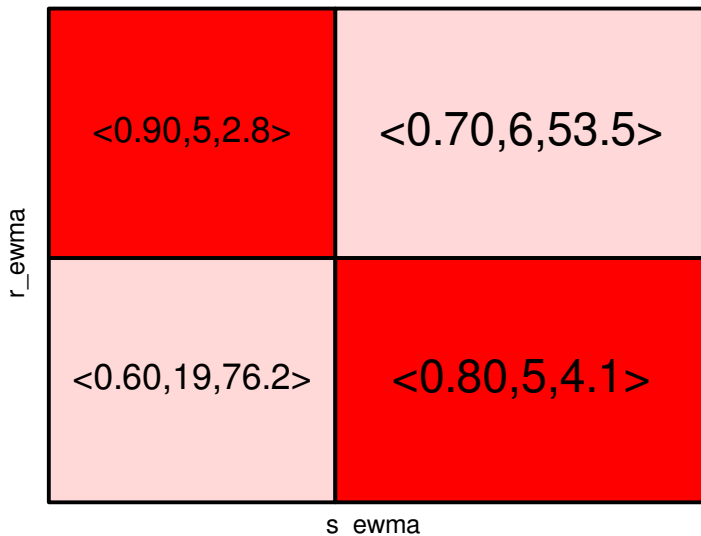
# Simulate



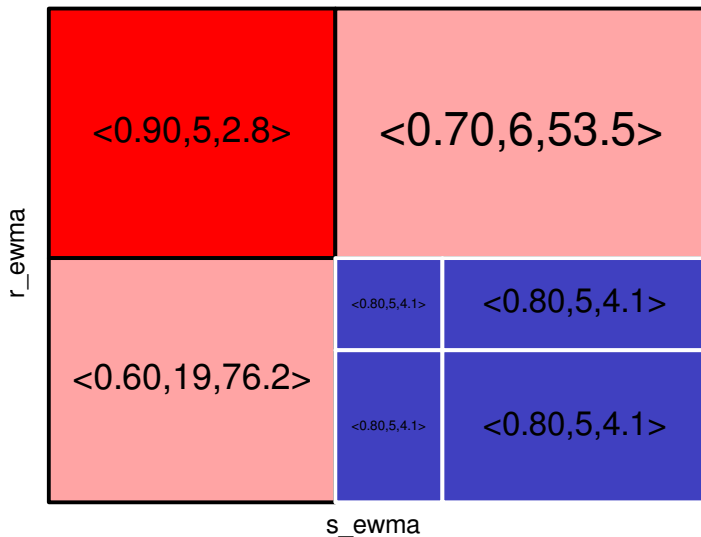
## Optimize each of the new actions

r_ewma	<0.90,4,3.3>	<0.90,4,3.3>
	<0.90,4,3.3>	<0.90,4,3.3>
		s_ewma

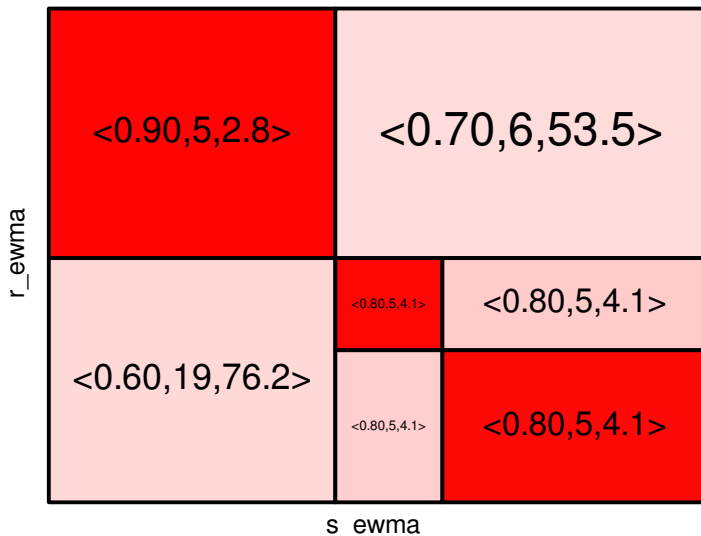
## Now split the most-used rule



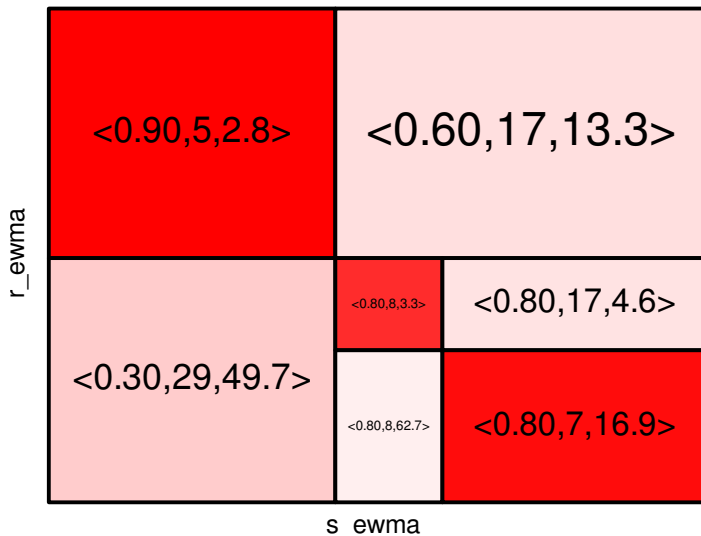
# Simulate



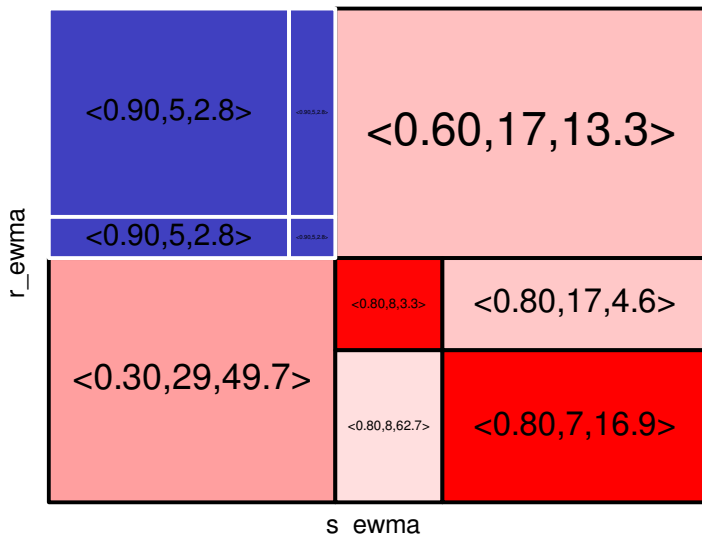
# Optimize



# Split

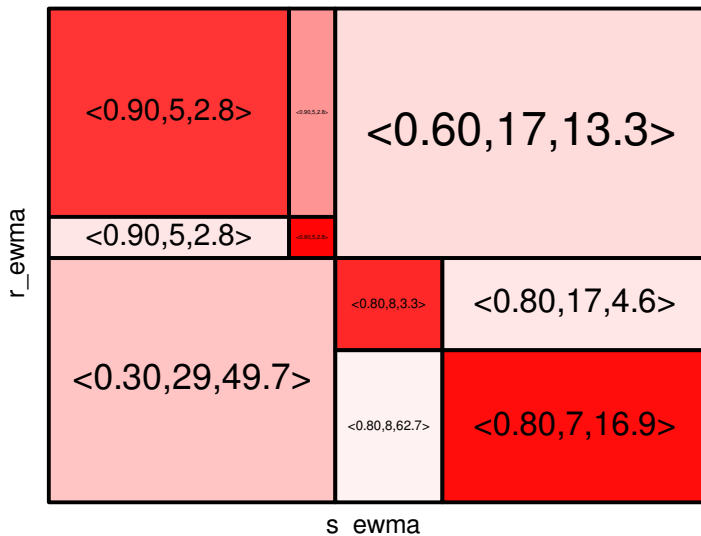


# Simulate

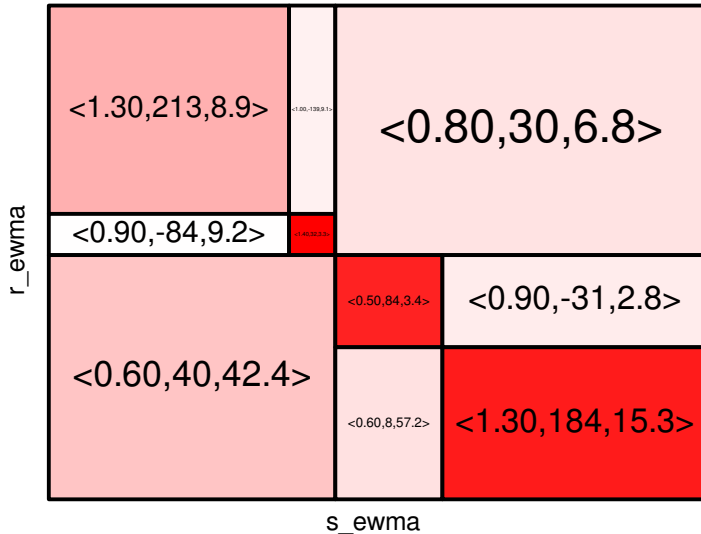




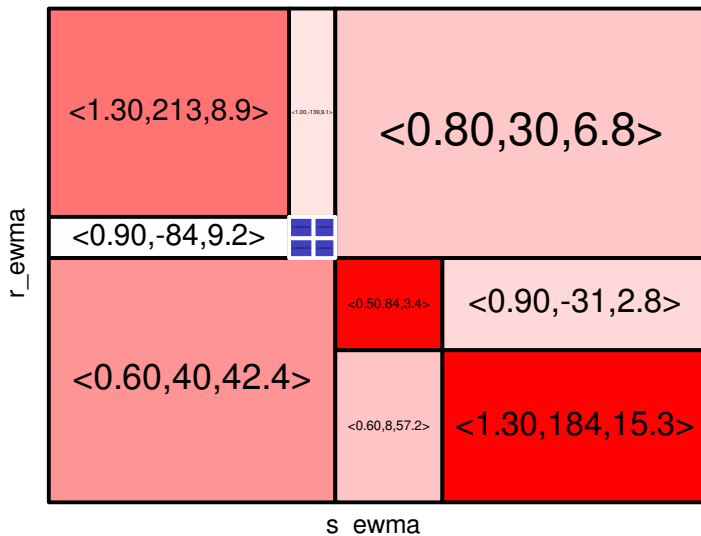
# Optimize



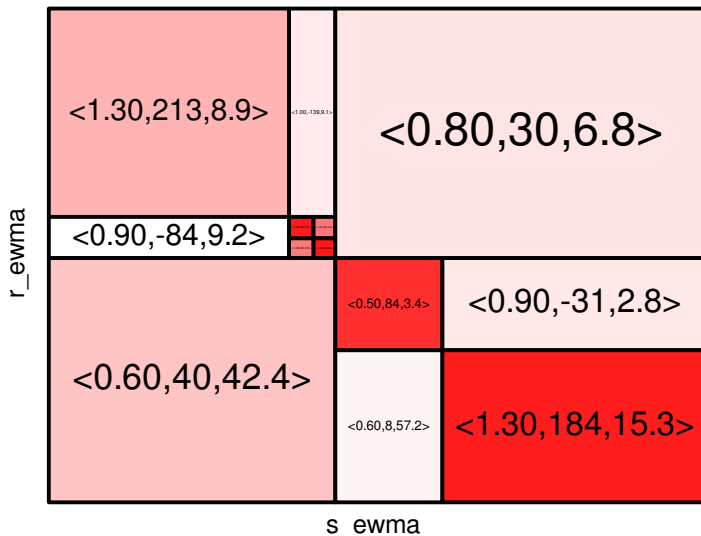
# Split



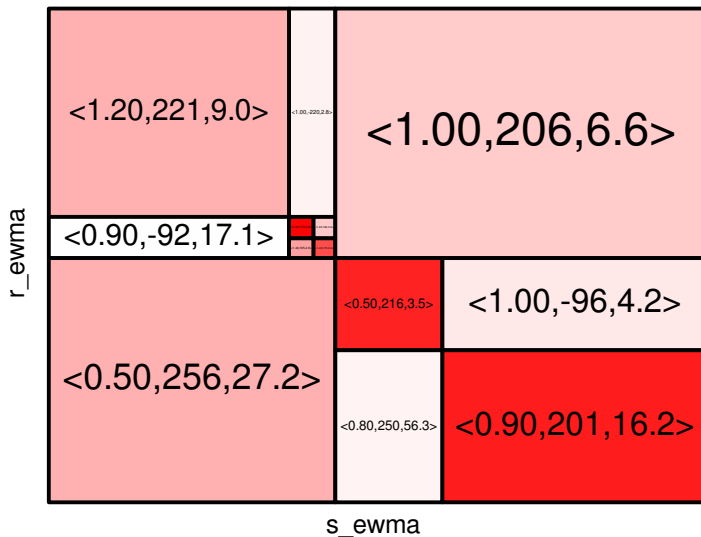
# Simulate



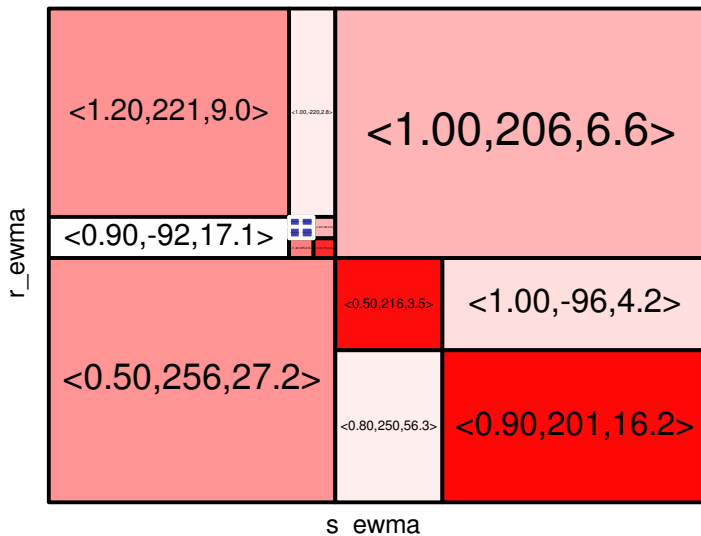
# Optimize



# Split



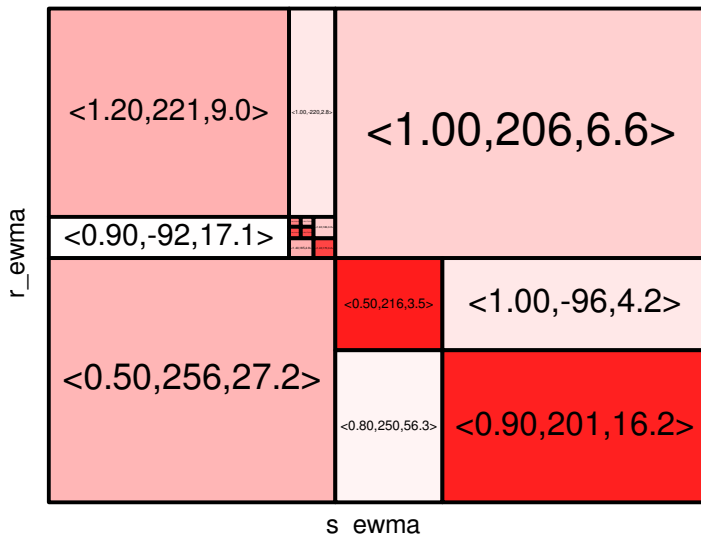
# Simulate



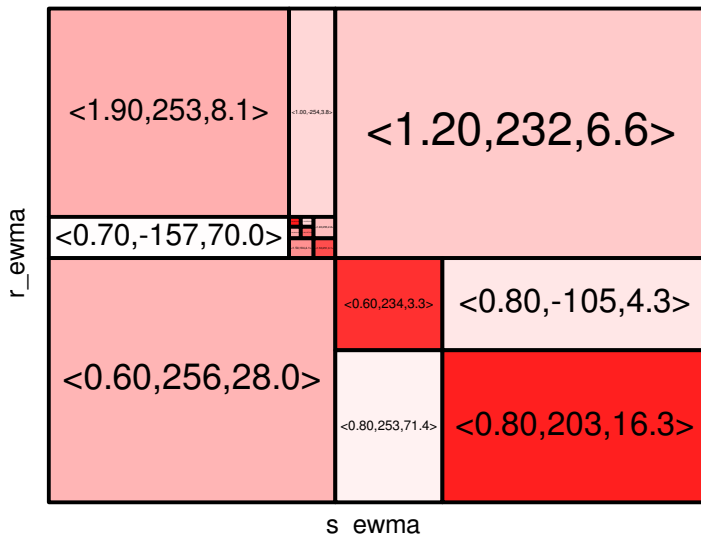
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Optimize

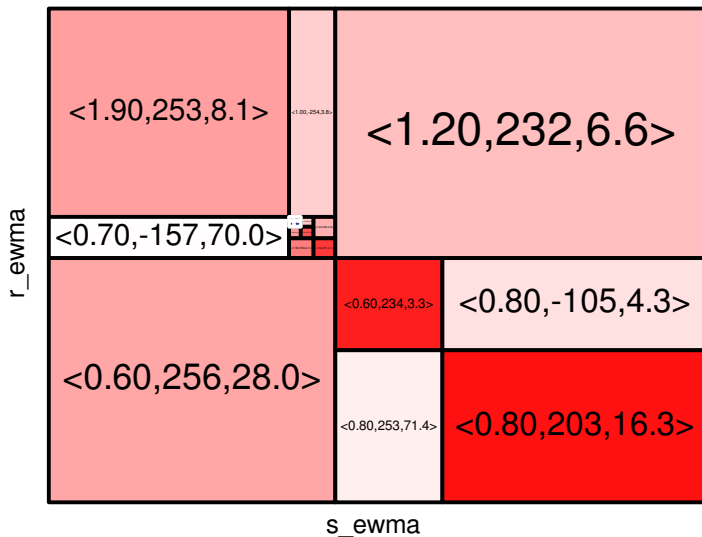


# Split

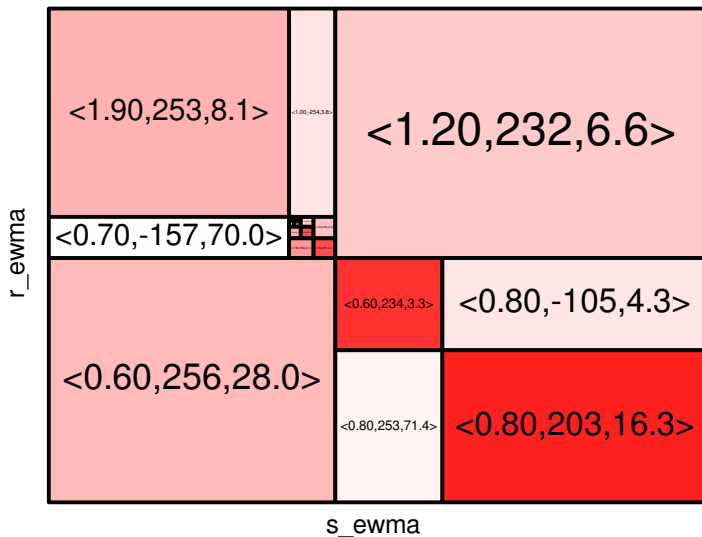




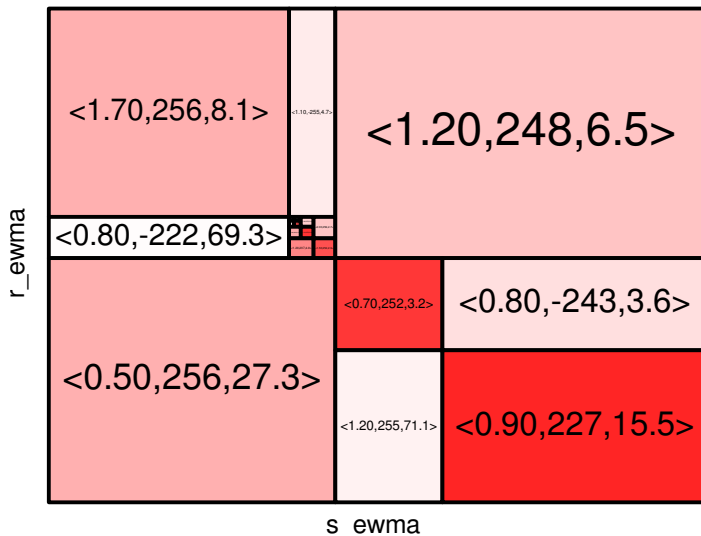
# Simulate



# Optimize



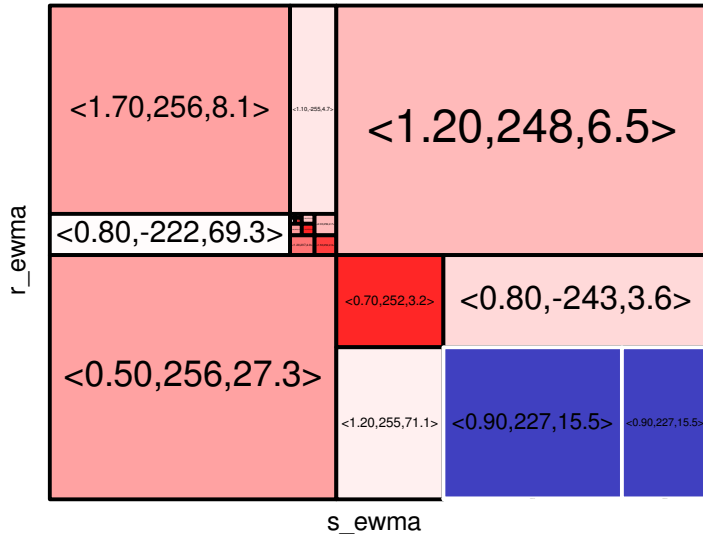
# Split



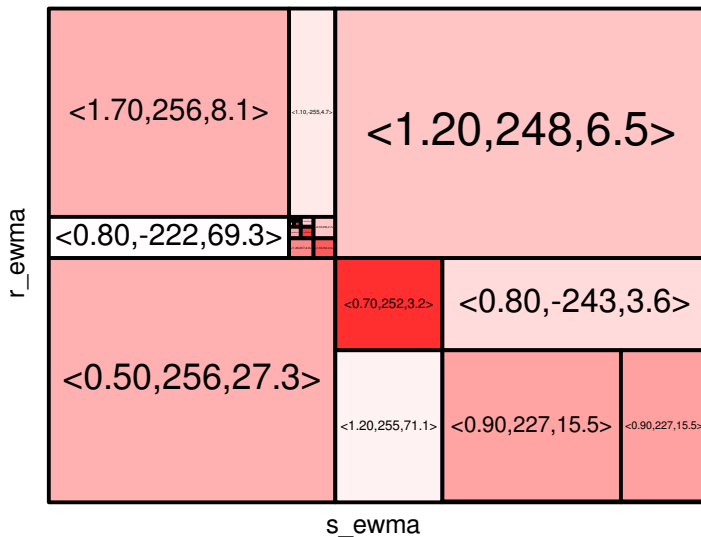
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Simulate



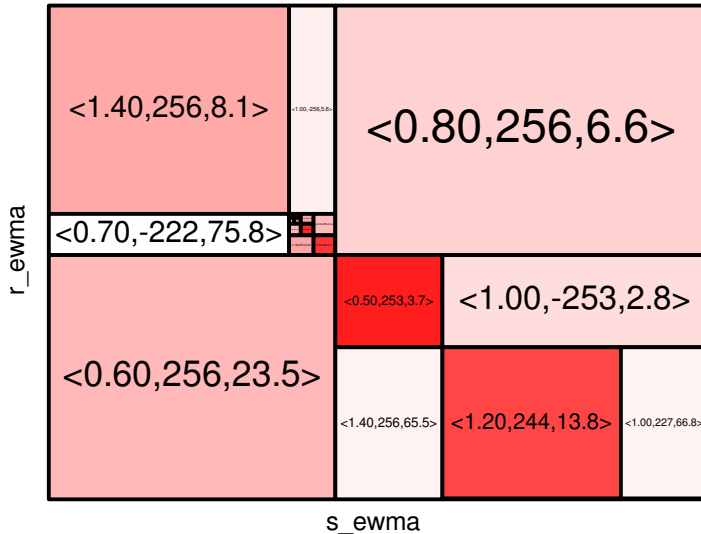
# Optimize



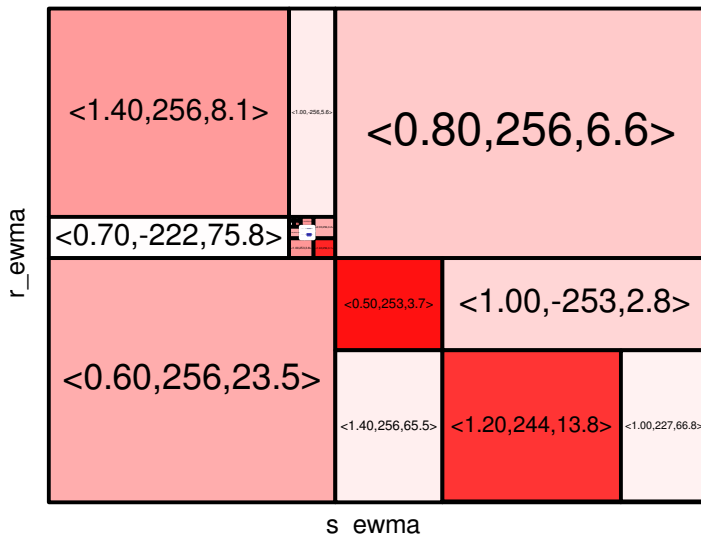
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

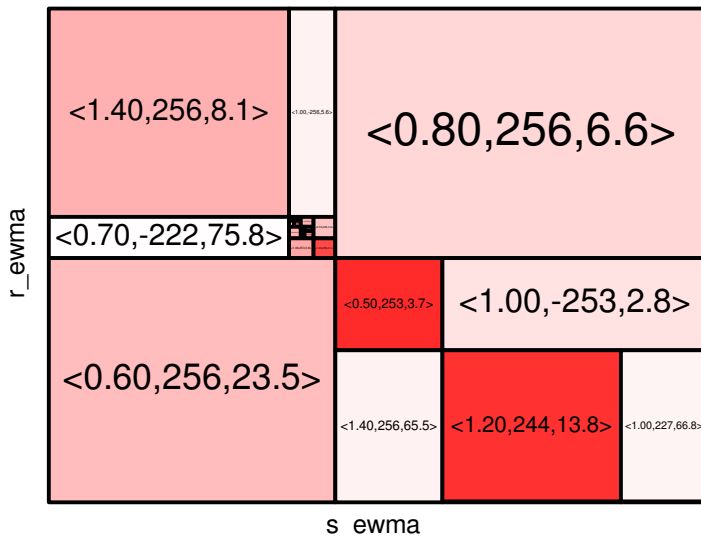
# Split



# Simulate

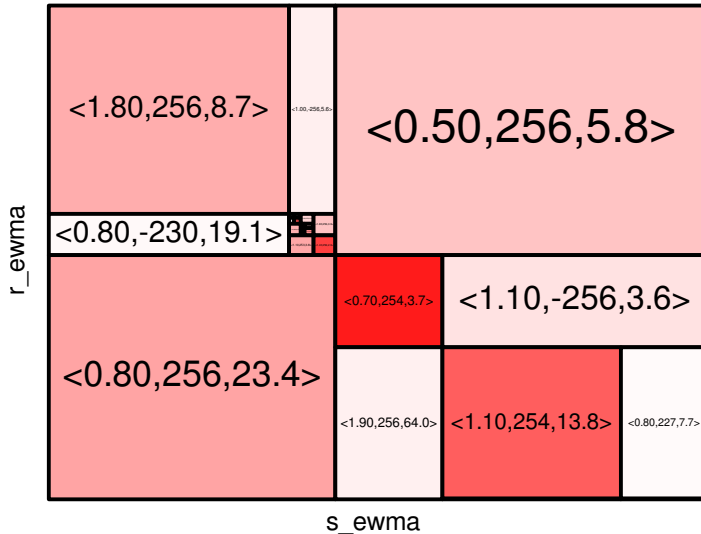


# Optimize

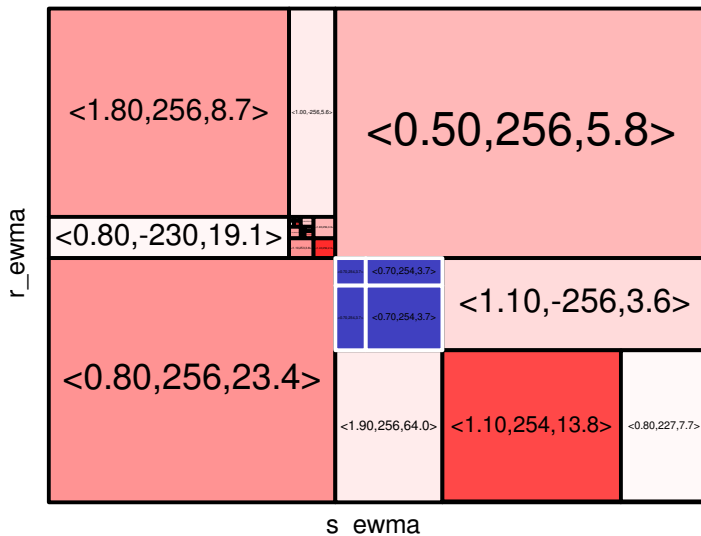




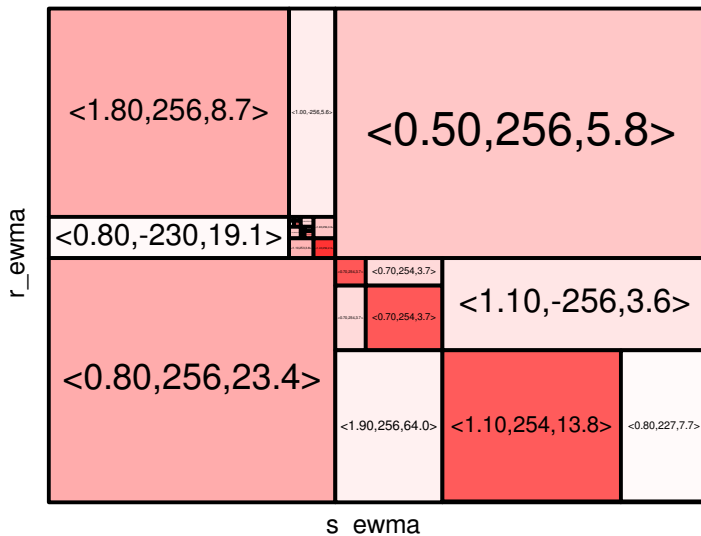
# Split



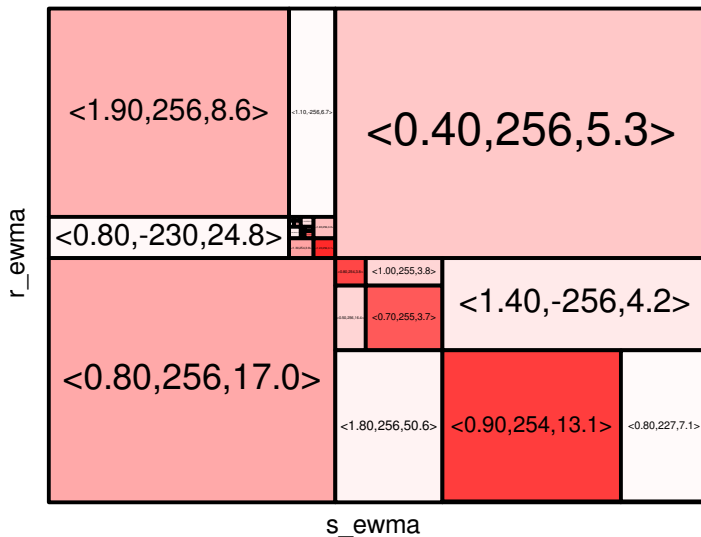
# Simulate



# Optimize



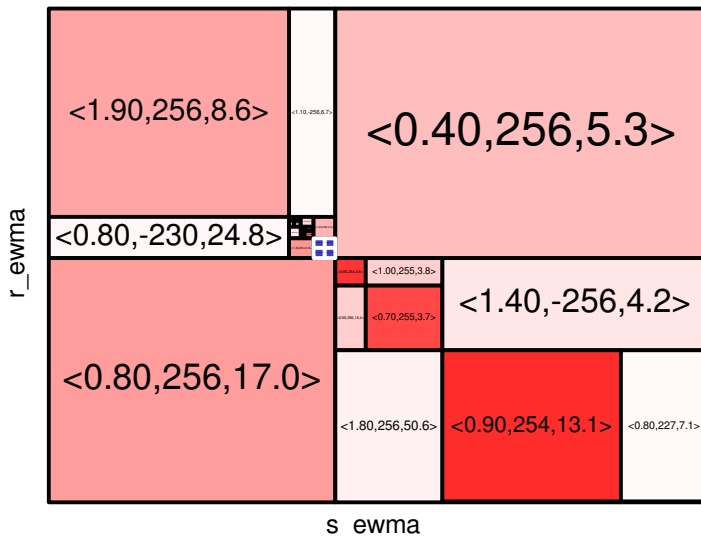
# Split



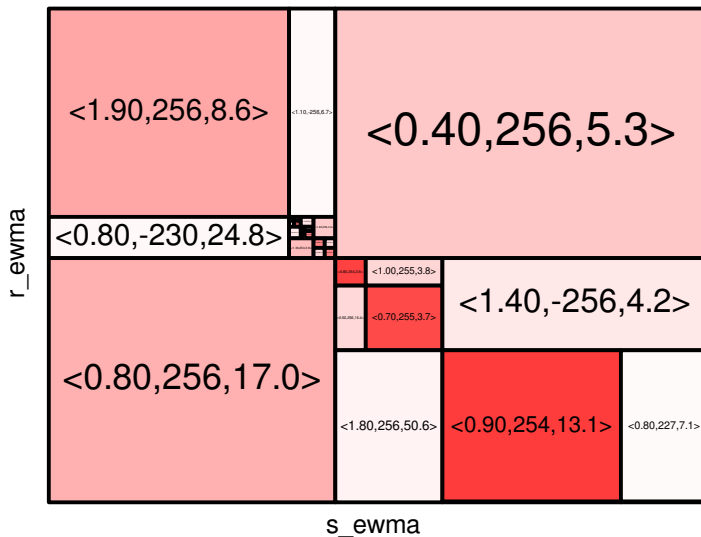
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

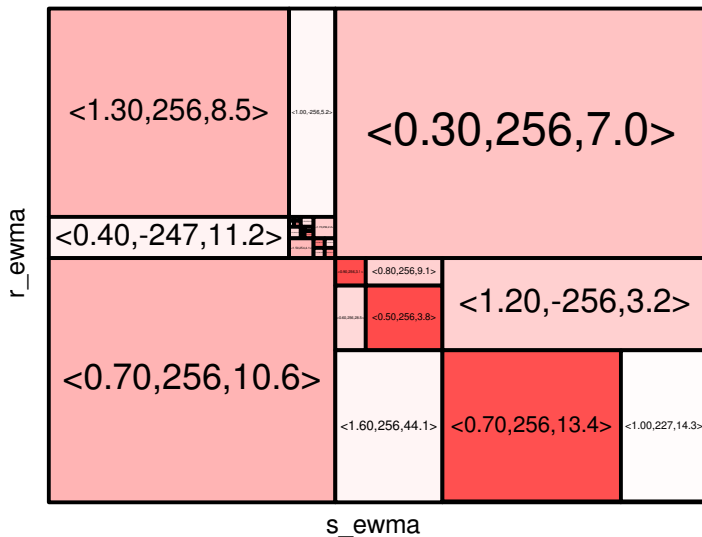
# Simulate



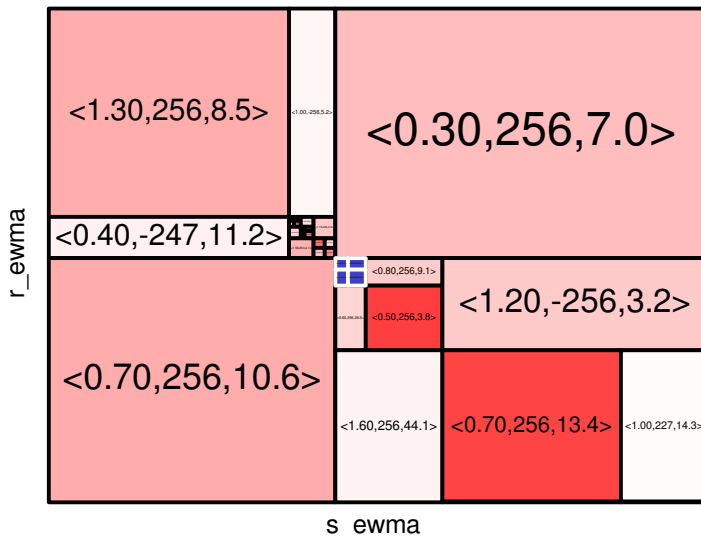
# Optimize



# Split



# Simulate

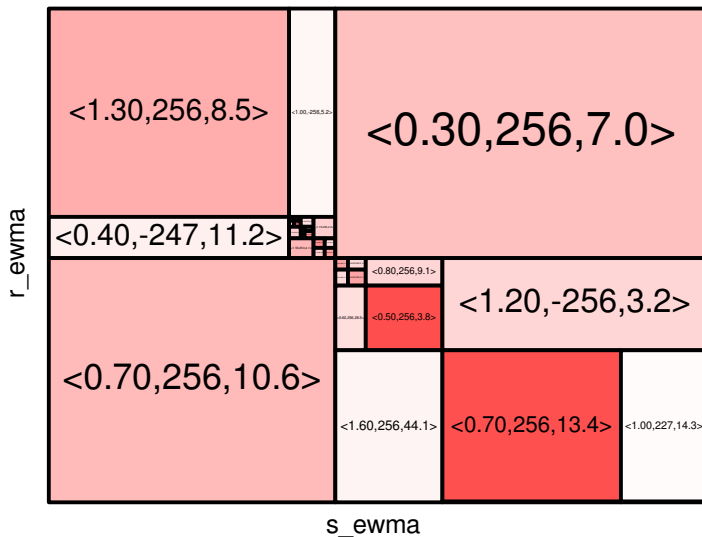


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet



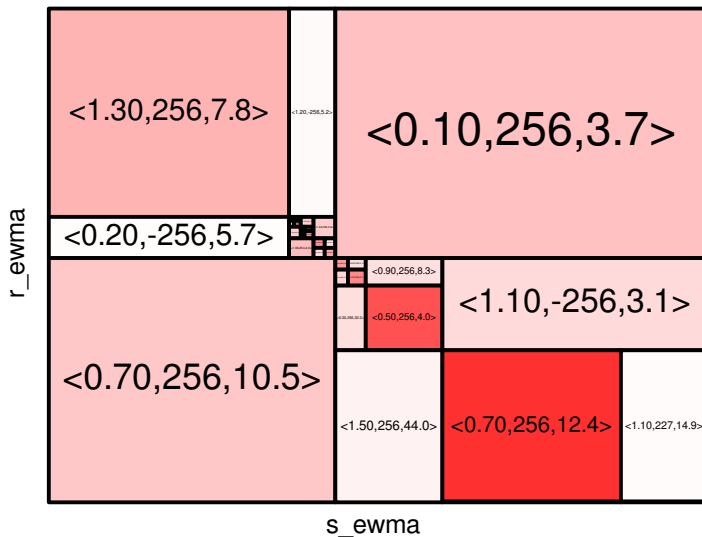
# Optimize



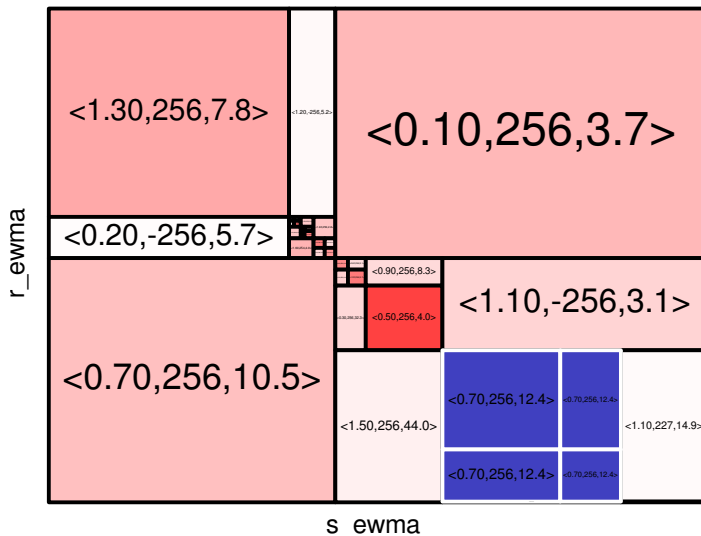
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

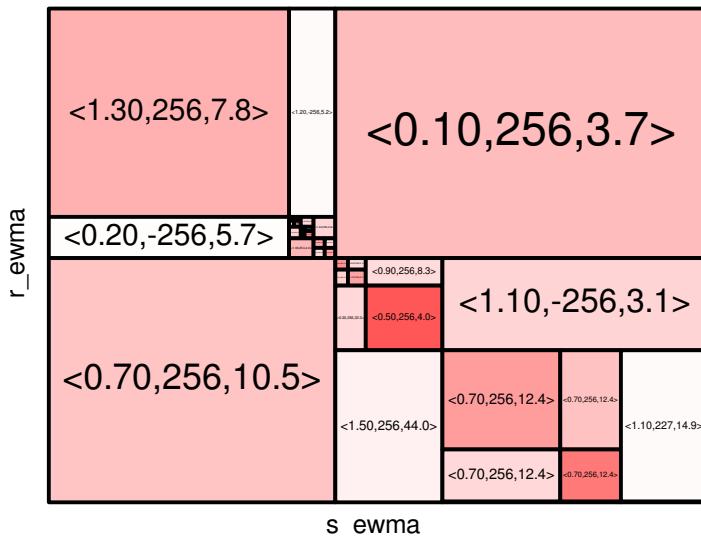
# Split



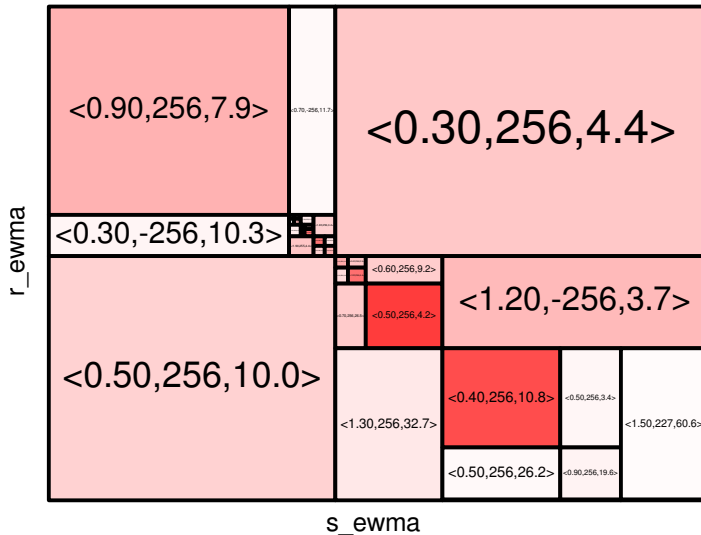
# Simulate



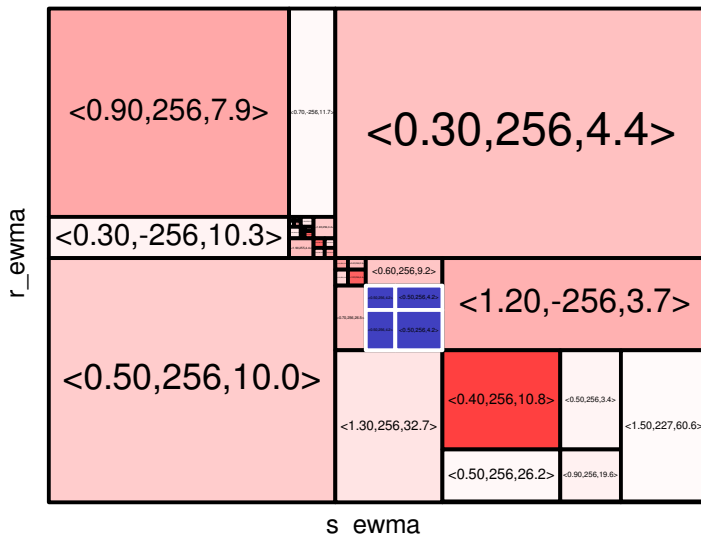
# Optimize



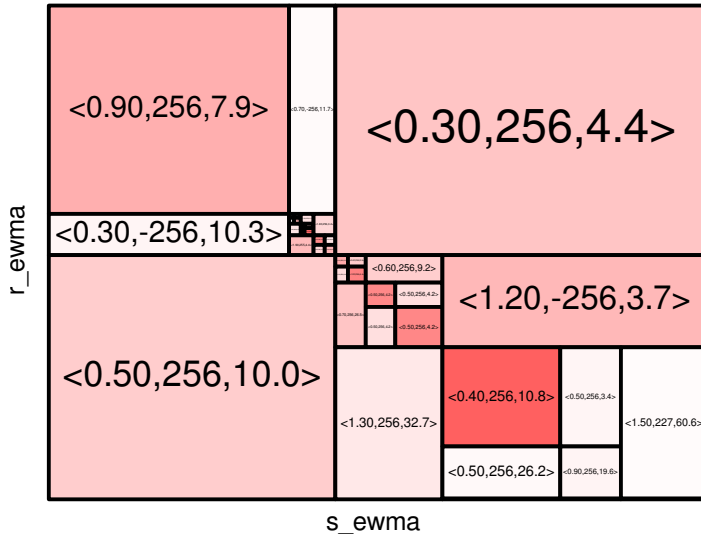
# Split



# Simulate



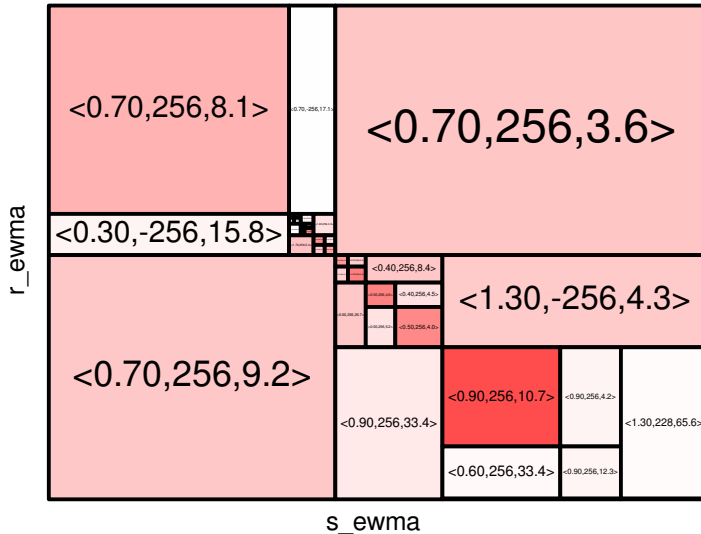
## Optimize



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

# Split

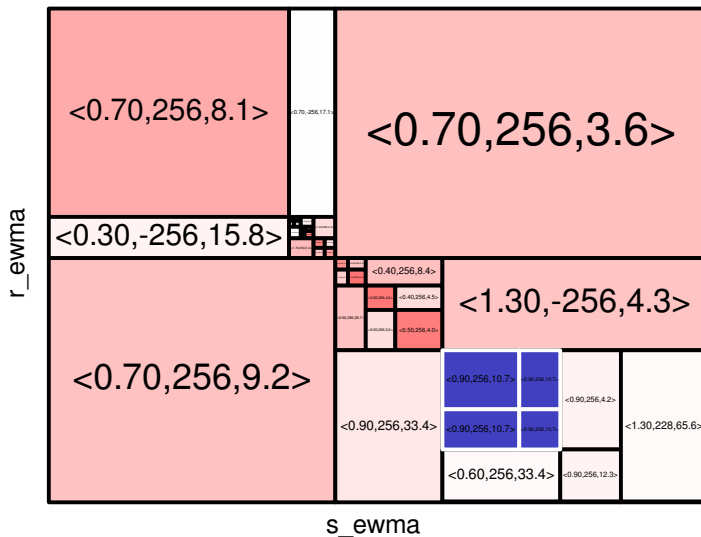


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

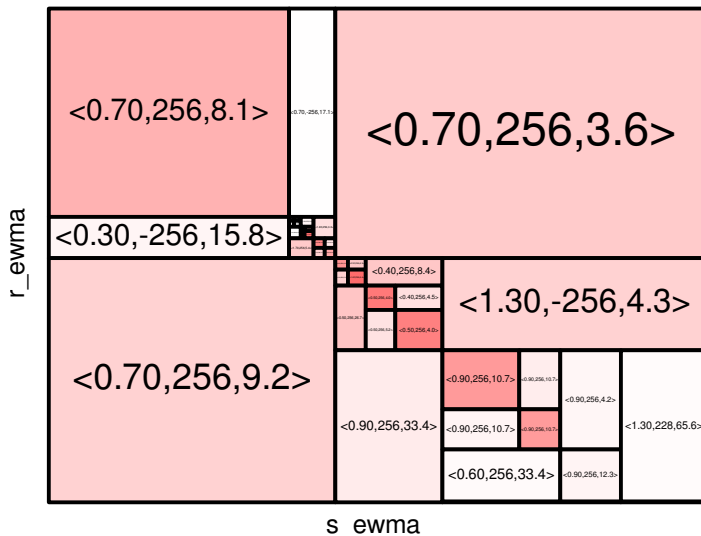
Transport Architectures for an Evolving Internet



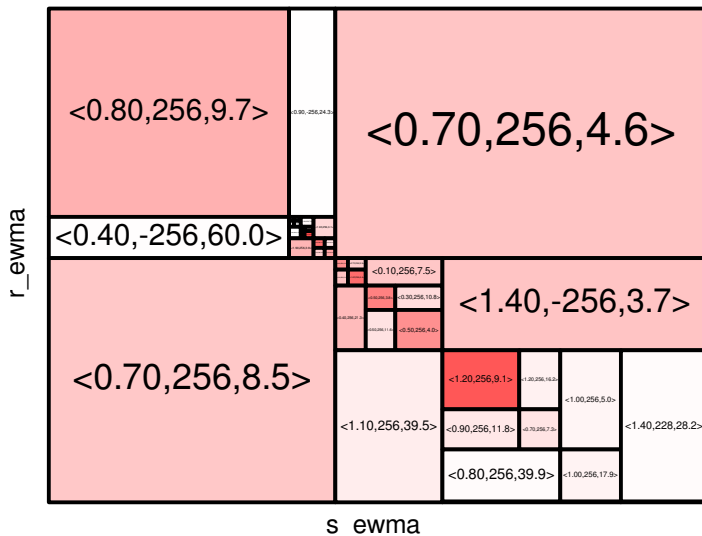
# Simulate



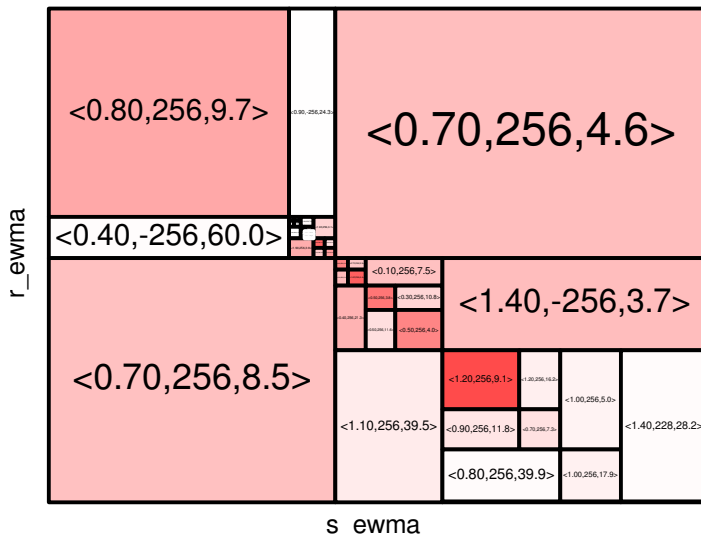
# Optimize



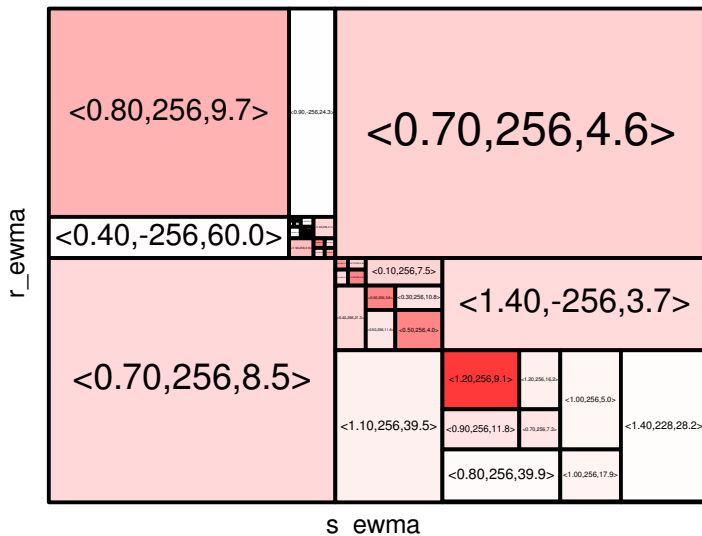
# Split



# Simulate



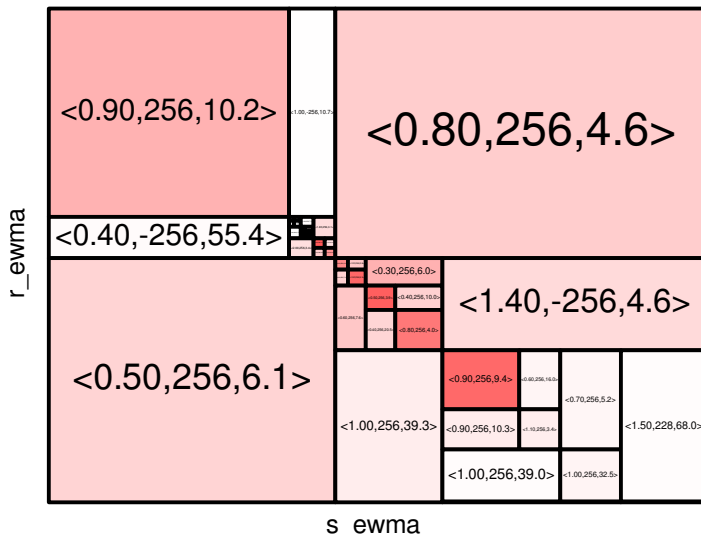
# Optimize



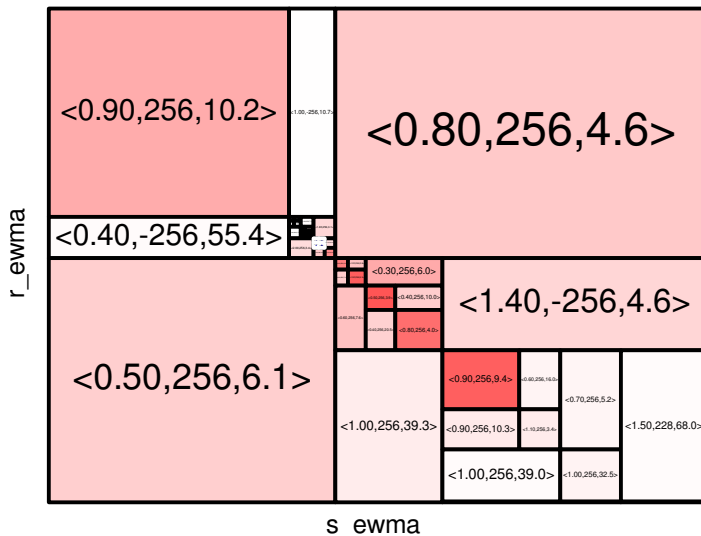
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Split



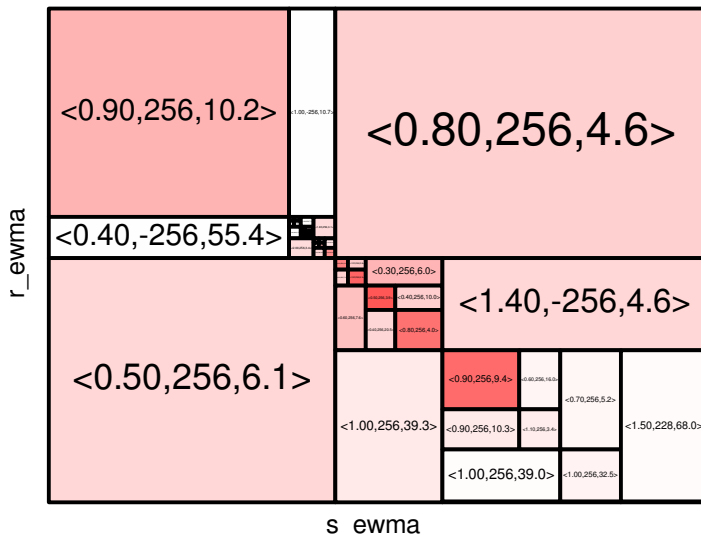
# Simulate



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

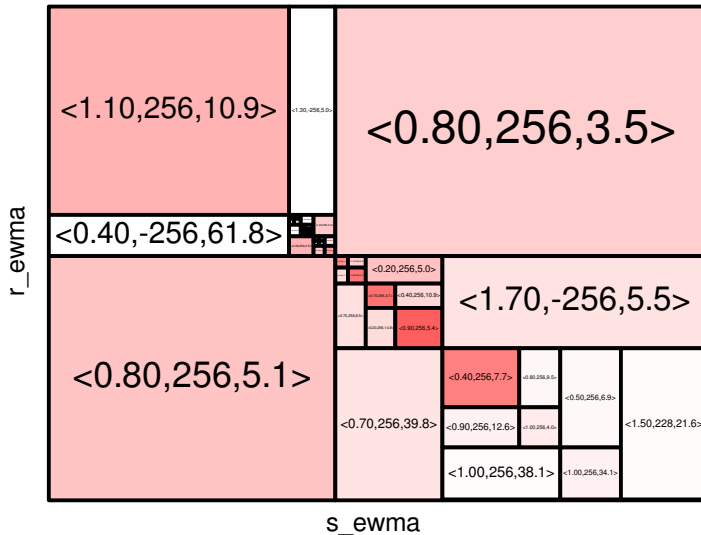
Transport Architectures for an Evolving Internet

# Optimize





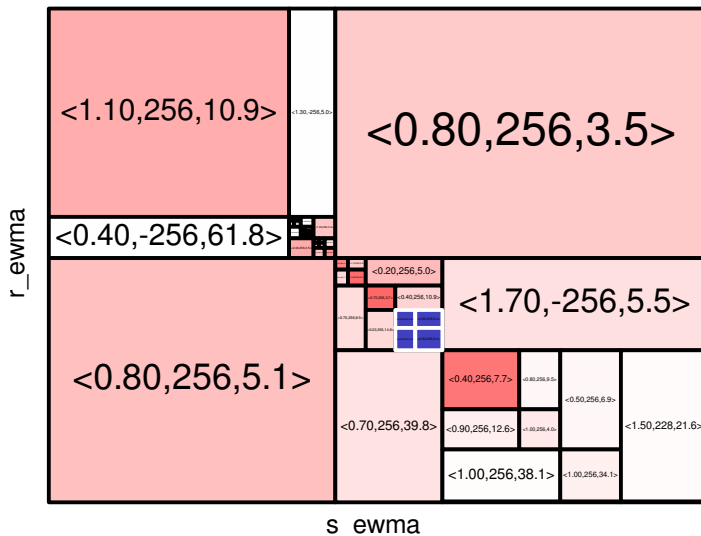
## Split



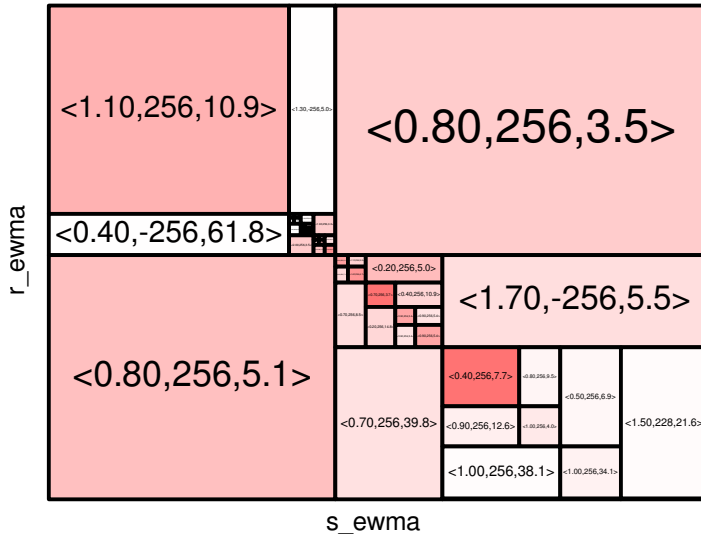
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

# Simulate



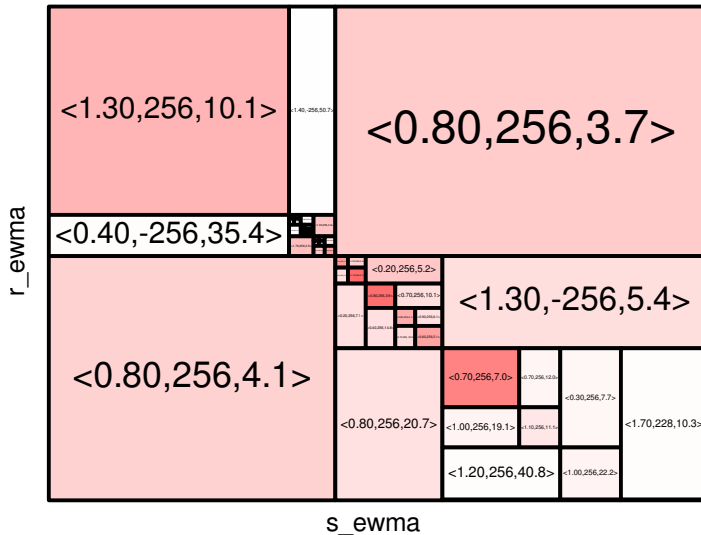
## Optimize



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

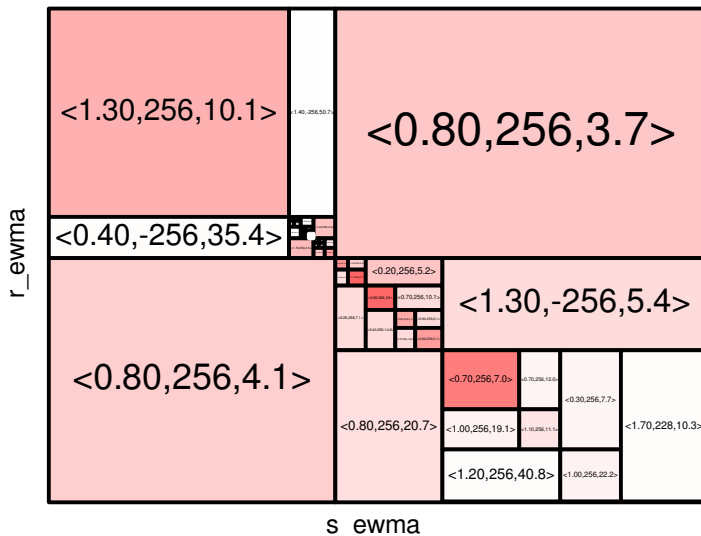
## Split



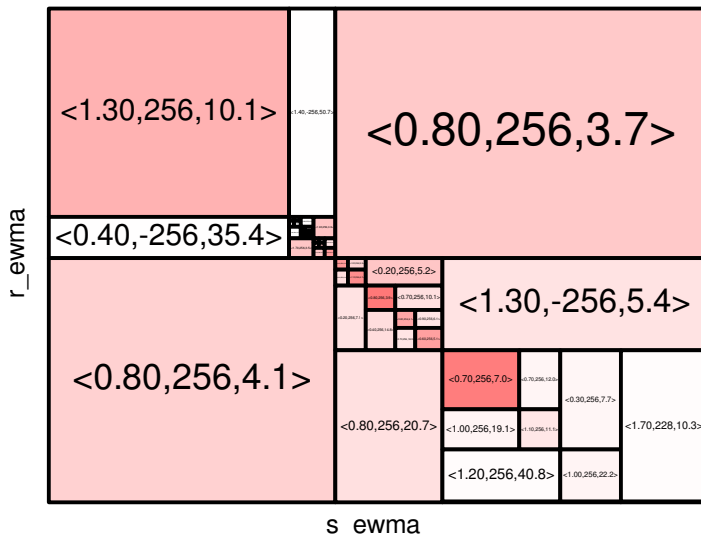
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

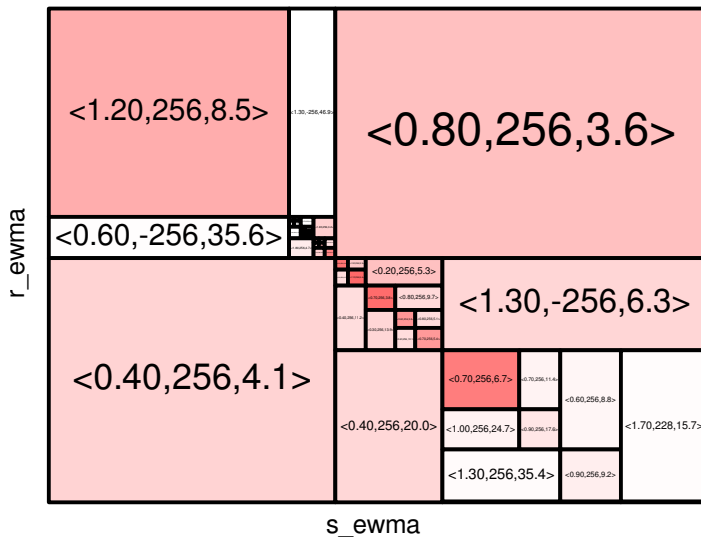
# Simulate



# Optimize



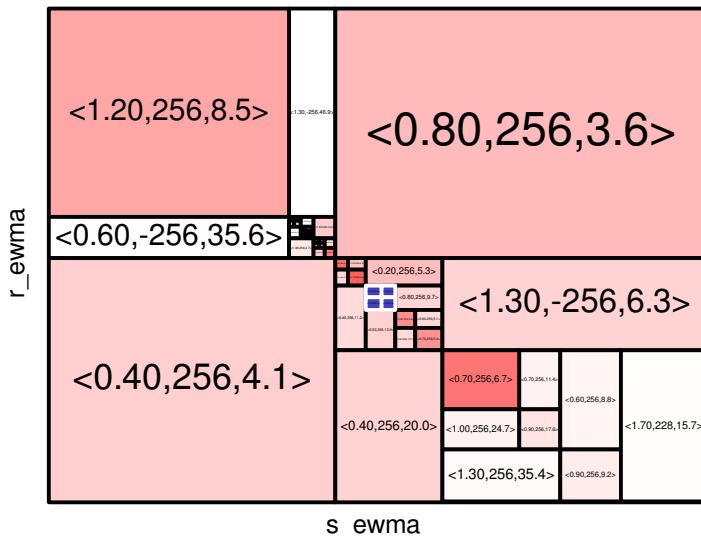
# Split



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Simulate

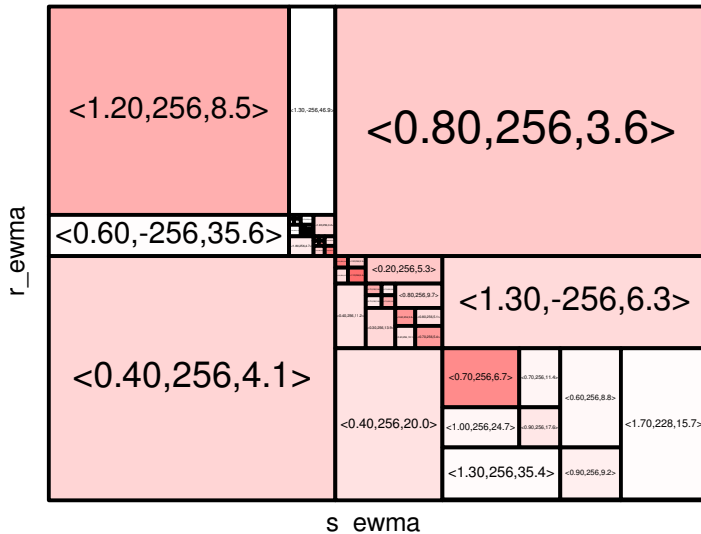


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet



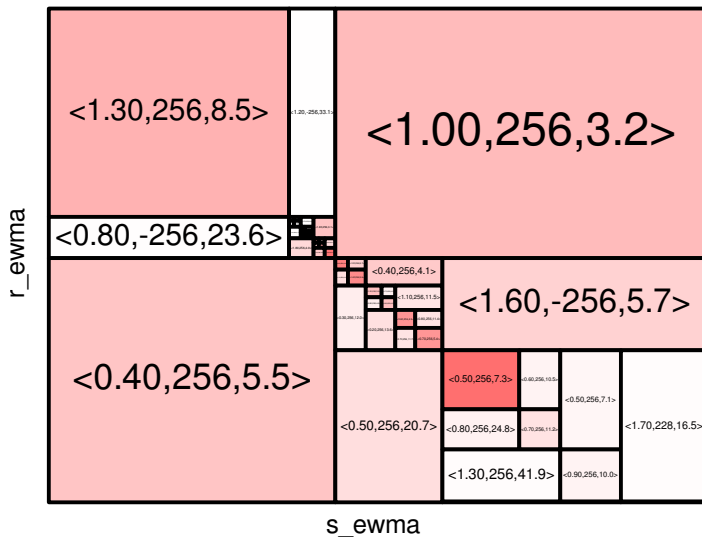
# Optimize



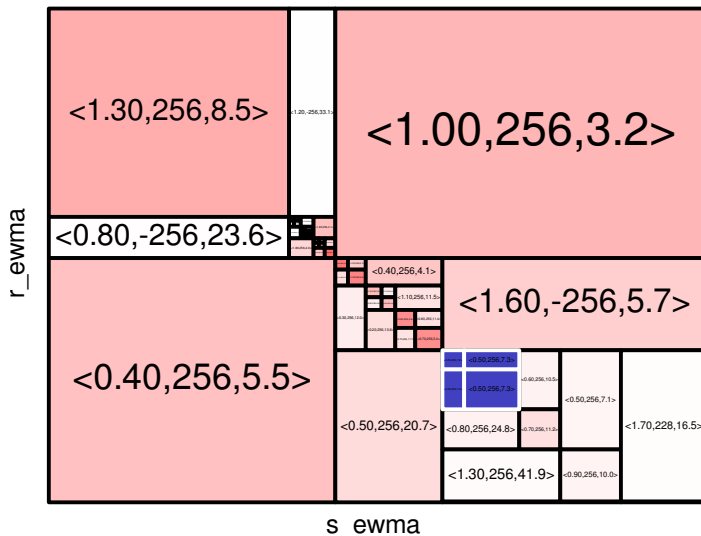
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

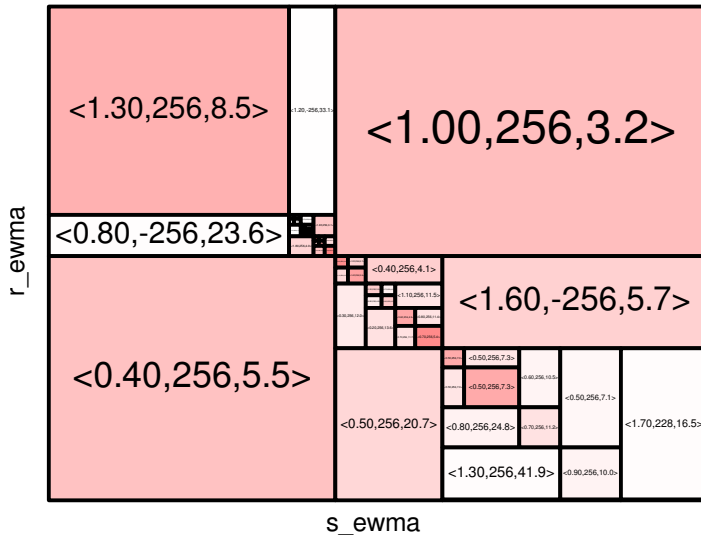
# Split



# Simulate



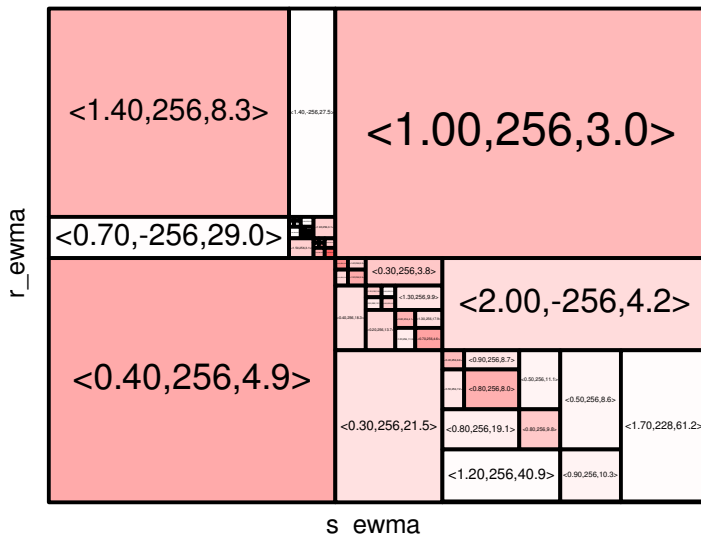
## Optimize



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

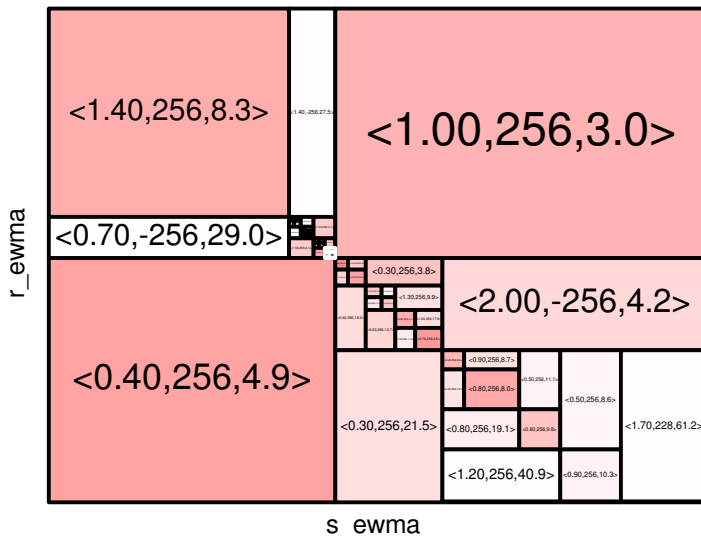
# Split



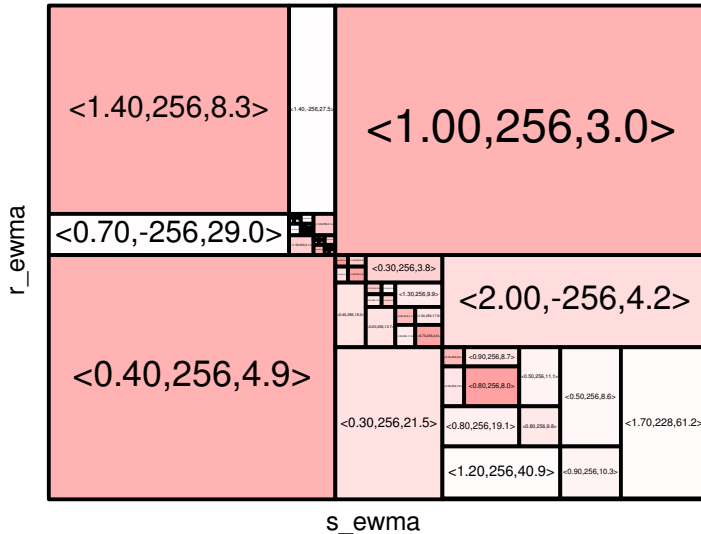
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Simulate



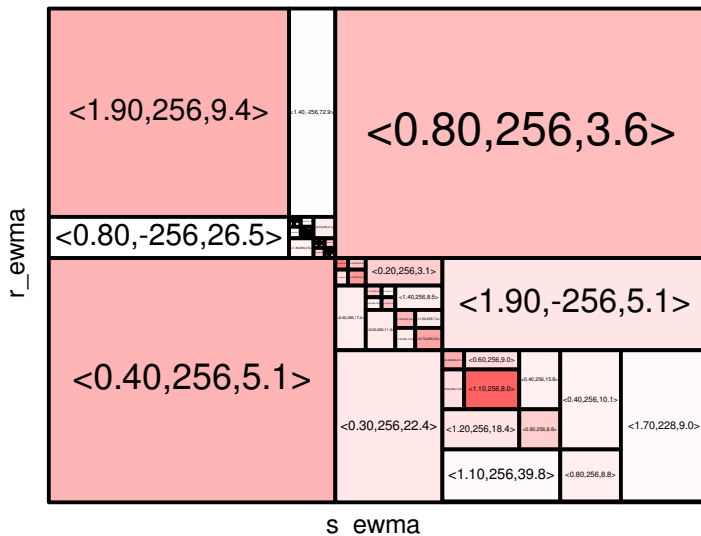
## Optimize



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

# Split

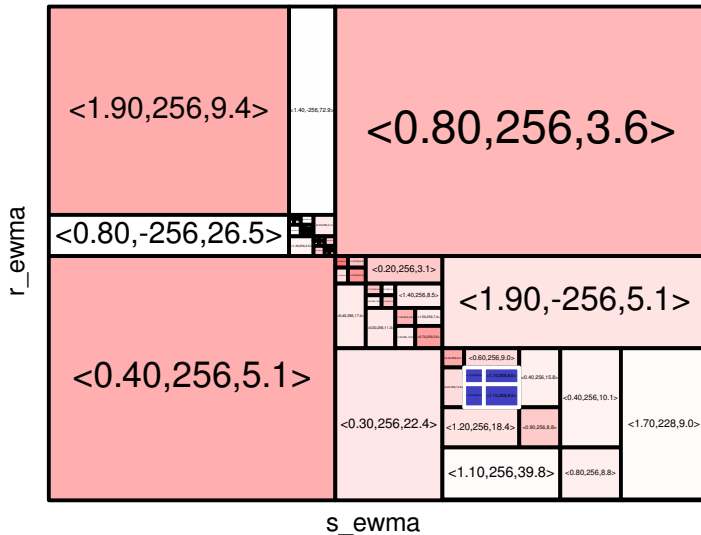


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet



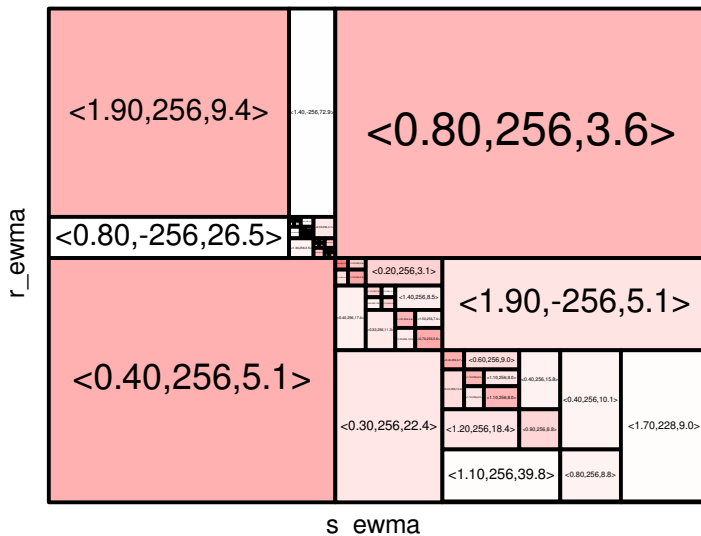
## Simulate



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

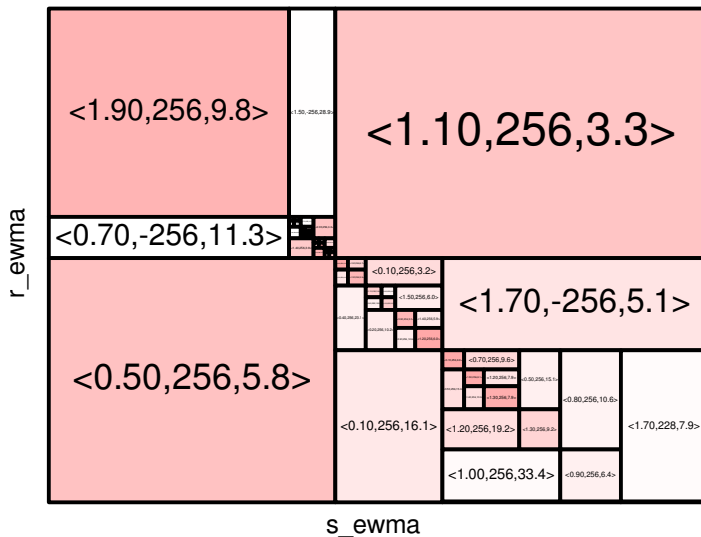
# Optimize



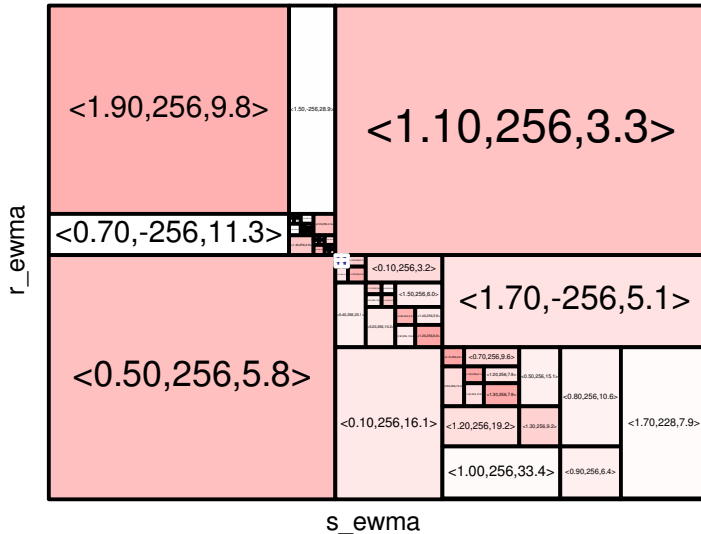
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Split



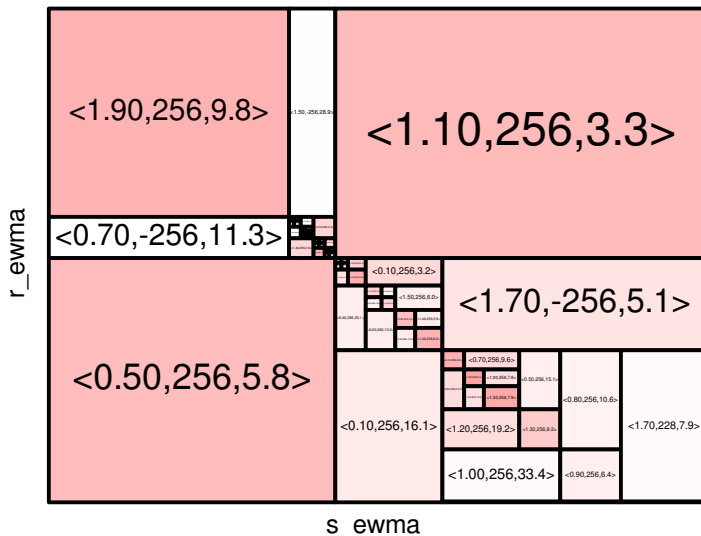
## Simulate



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

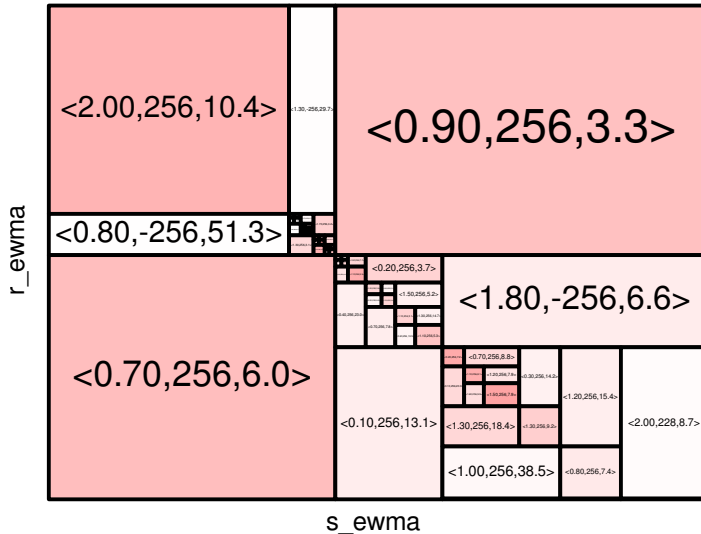
# Optimize



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

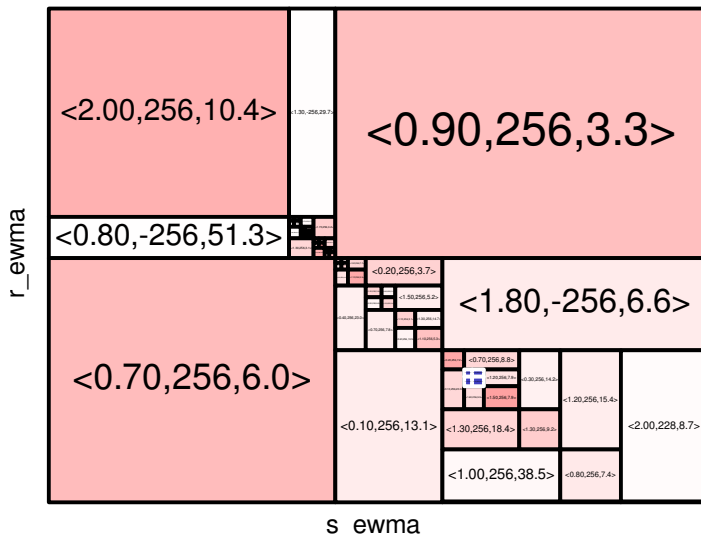
## Split



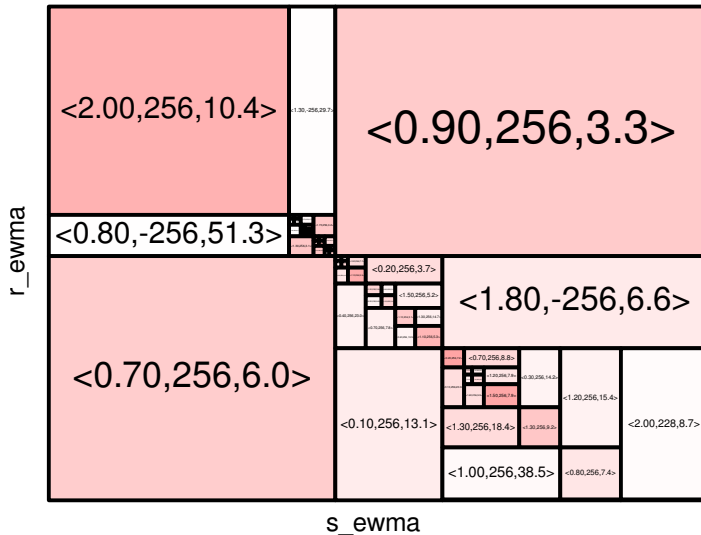
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

# Simulate



## Optimize

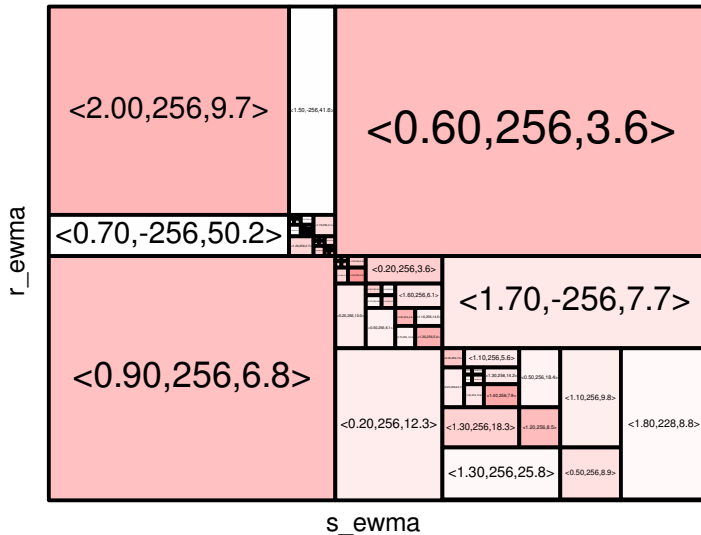


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet



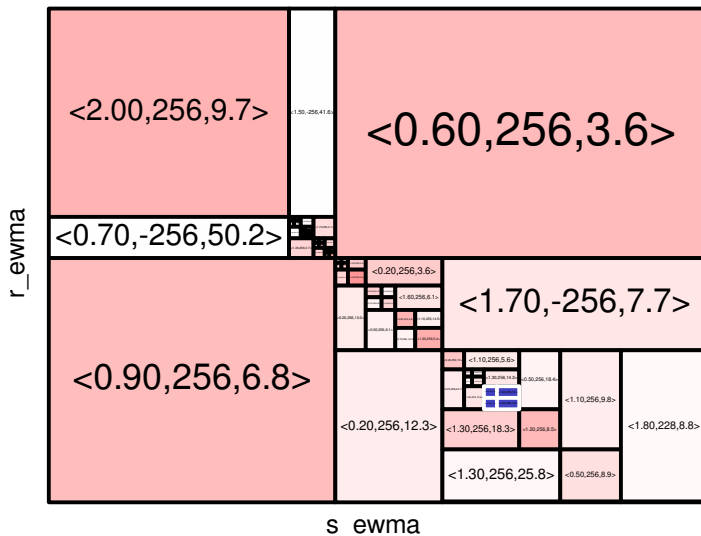
## Split



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

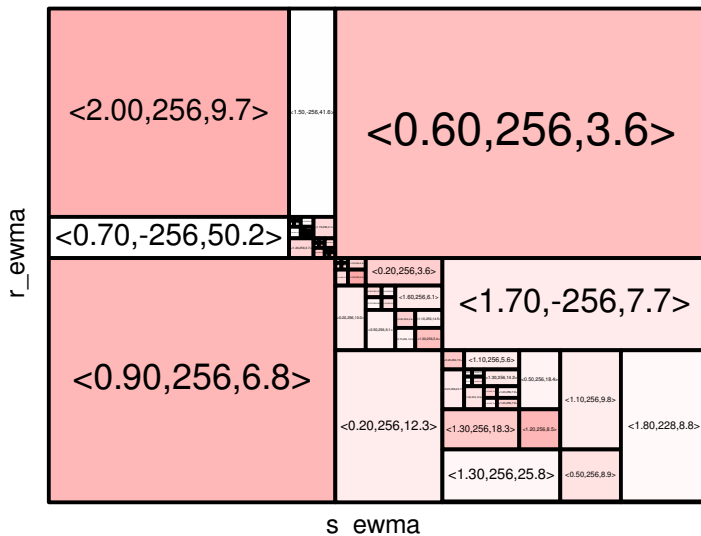
# Simulate



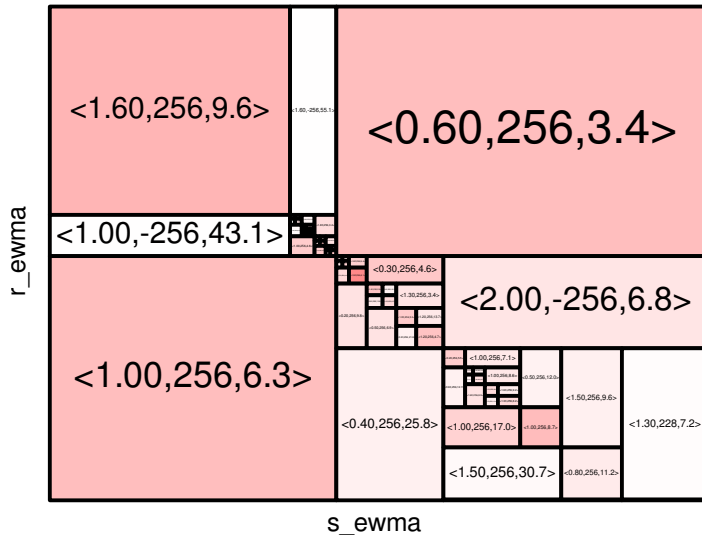
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

## RemyCC



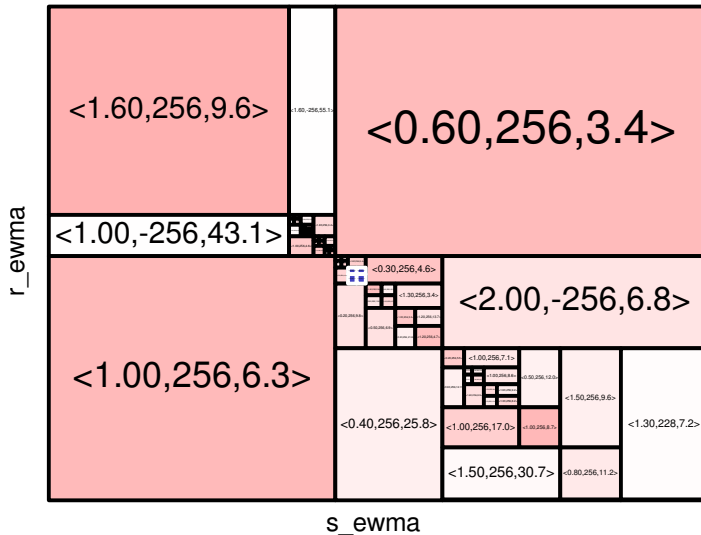
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

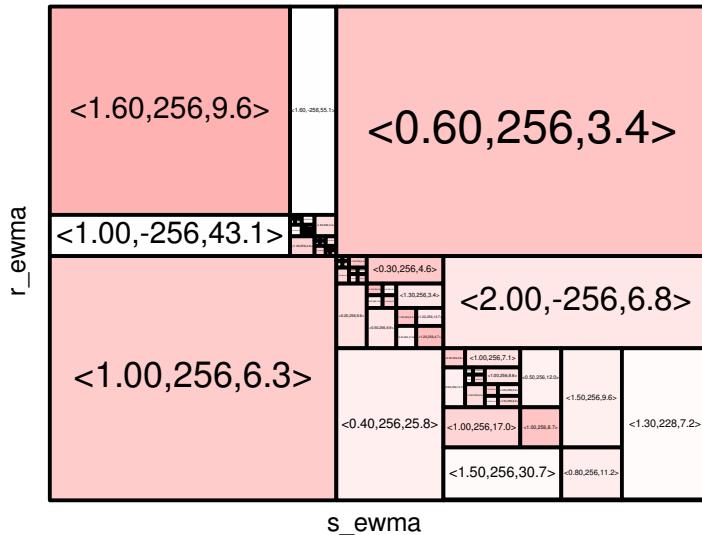
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

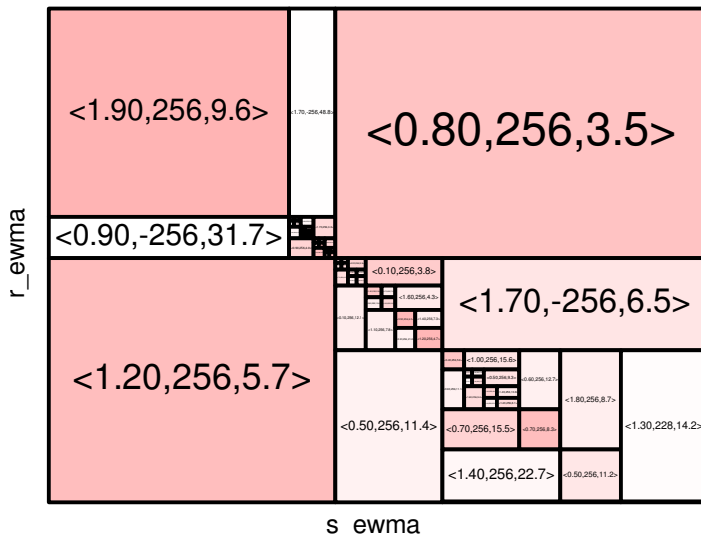
## RemyCC



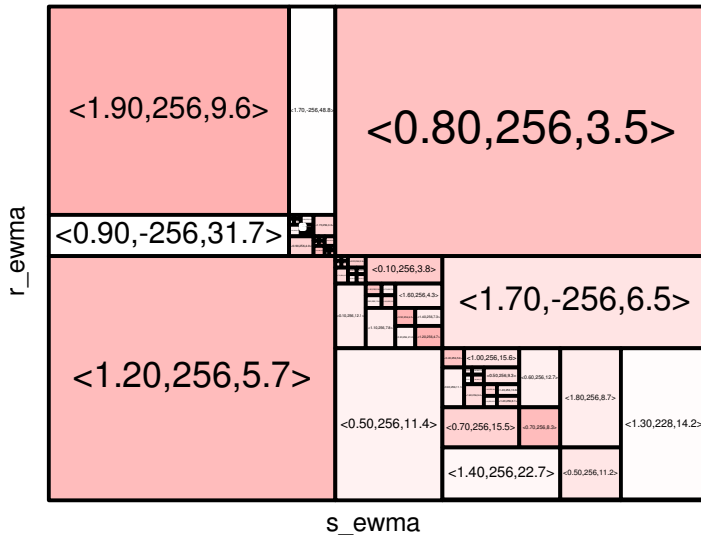
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

## RemyCC



## RemyCC

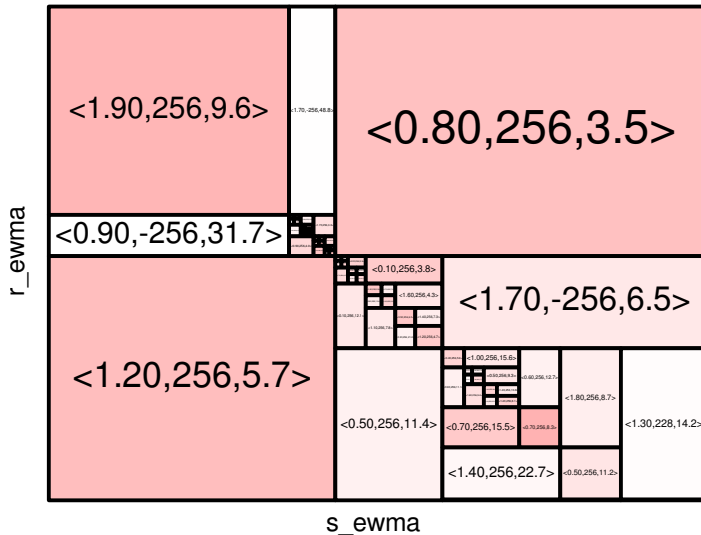


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet



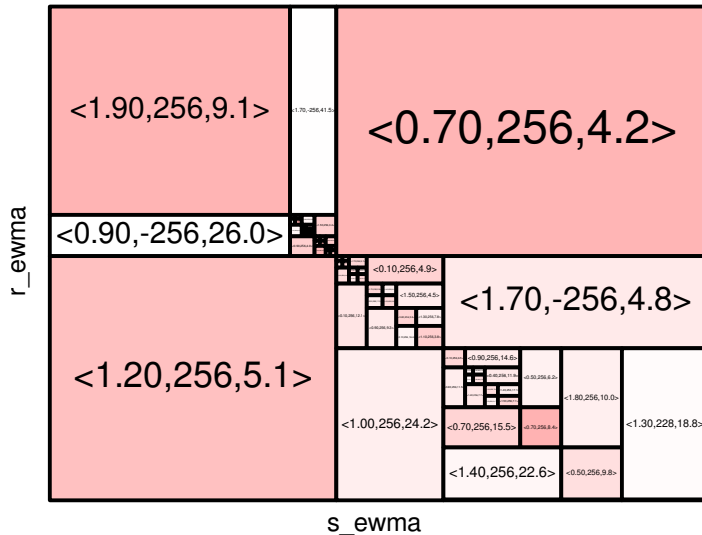
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

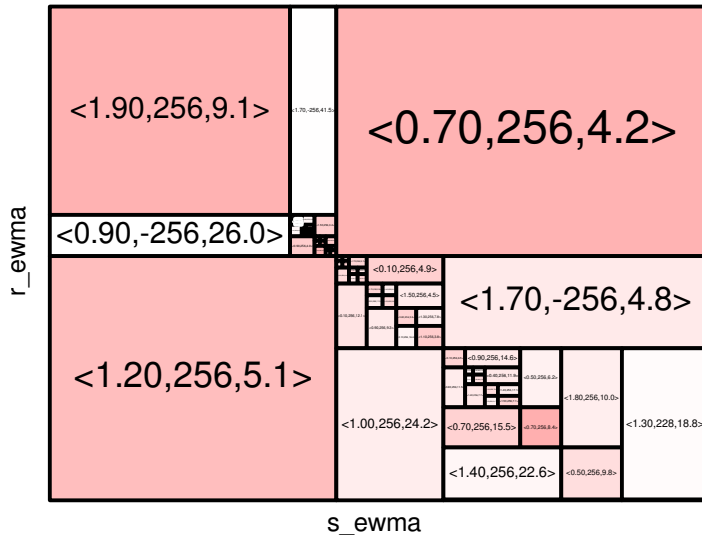
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

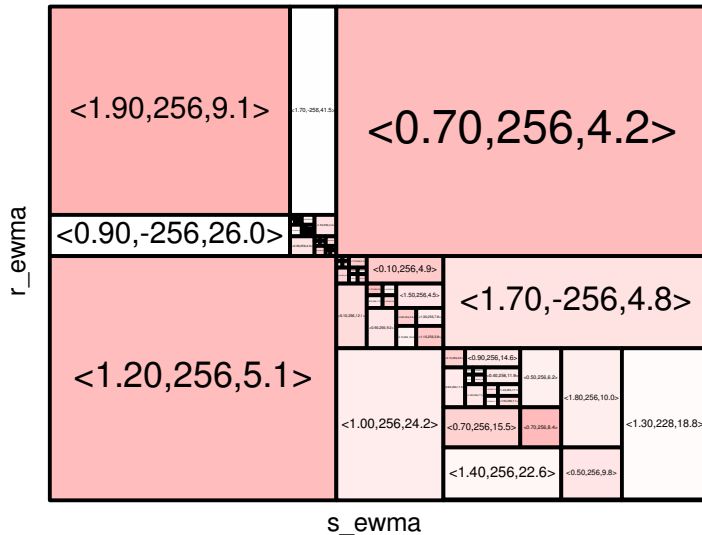
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

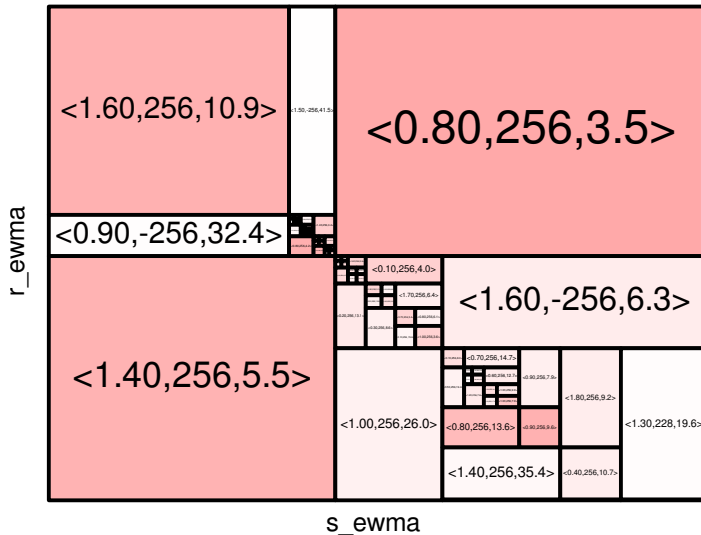
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

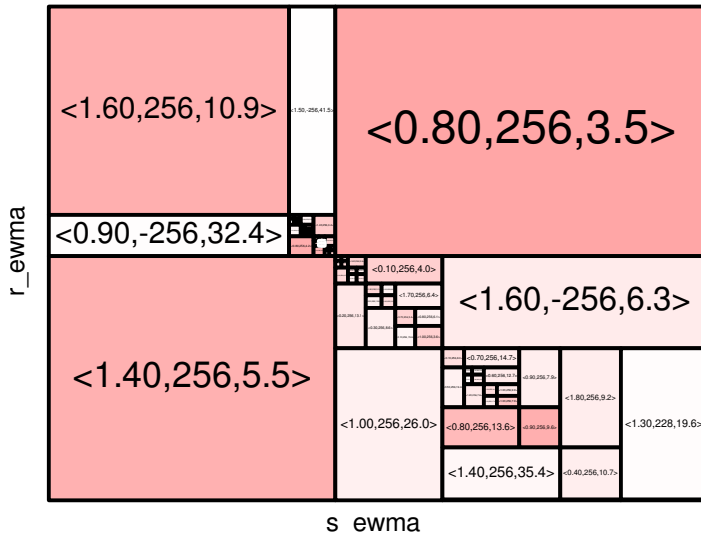
## RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet

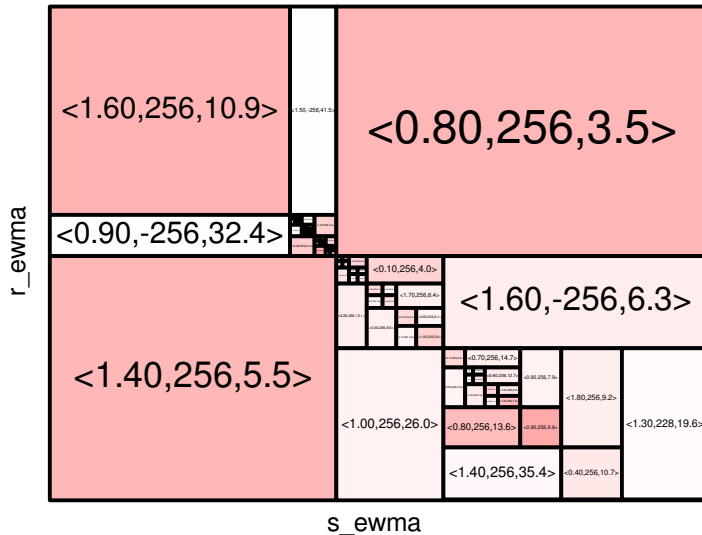
# RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

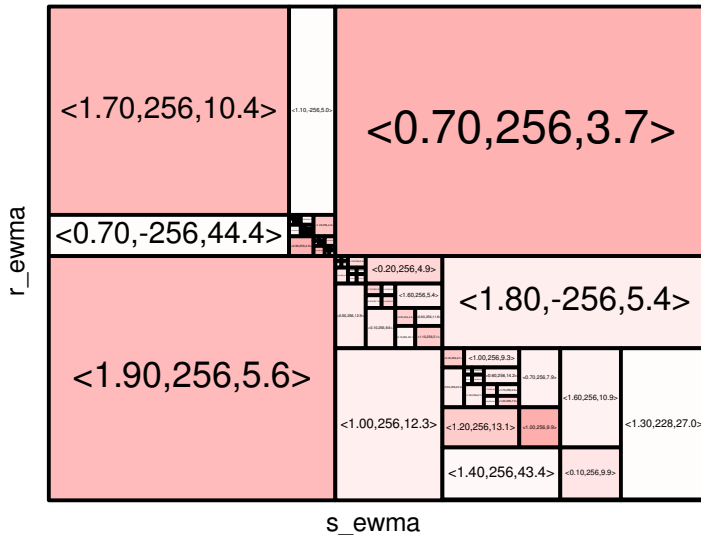
# RemyCC



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

## RemyCC



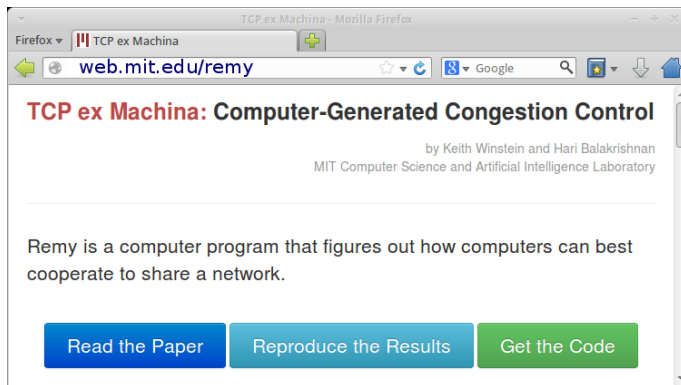
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

## Transport Architectures for an Evolving Internet



# Evaluation in ns-2

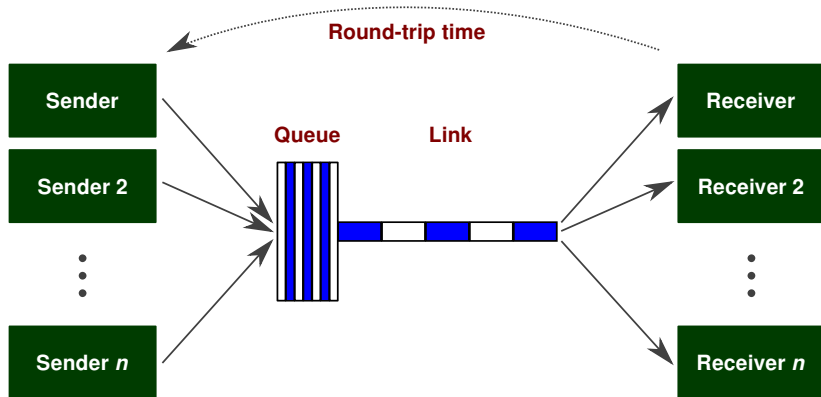
- ▶ End-to-end comparators: NewReno, Cubic, Compound, Vegas
- ▶ In-net comparators: Cubic-over-sfqCoDel, XCP
- ▶ Simulation setup published for replication



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Scenario 1: fixed-rate network, homogenous senders



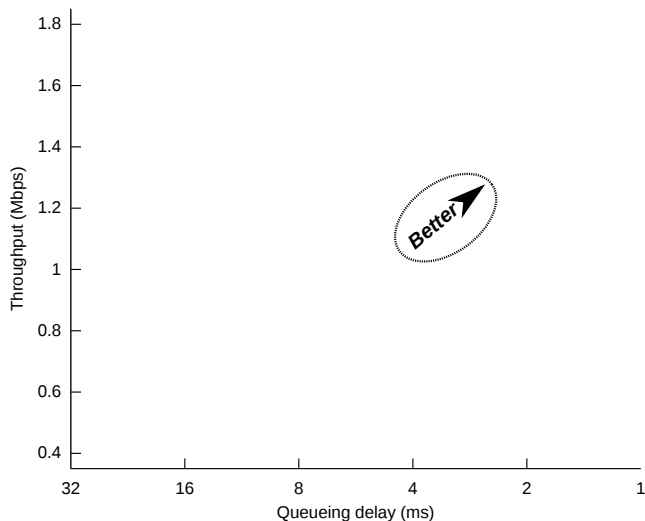
## Scenario 1: details

Quantity	Simulation parameter	Remy assumptions
Link speed	15 Mbps	Uniform(10, 20) Mbps
RTT	150 ms	Uniform(100, 200) ms
$n$	8	Uniform(1, 16)
“On” process	$\exp[\mu = 100] \text{ kB}$	$\exp[\mu = 5] \text{ s}$
“Off” process	$\exp\left[\mu = \frac{1}{2}\right] \text{ s}$	$\exp[\mu = 5] \text{ s}$

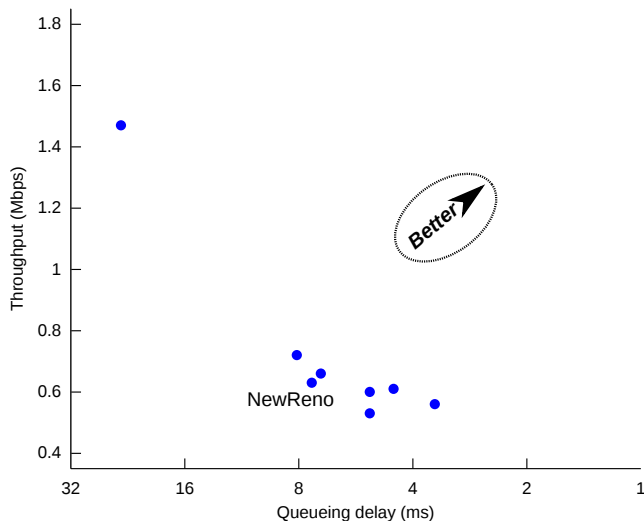
**Remy objective:**

$$\sum_i \log \left[ \frac{\text{throughput}_i}{(\text{delay}_i)^\delta} \right]$$
$$\delta \in \left\{ \frac{1}{10}, 1, 10 \right\}$$

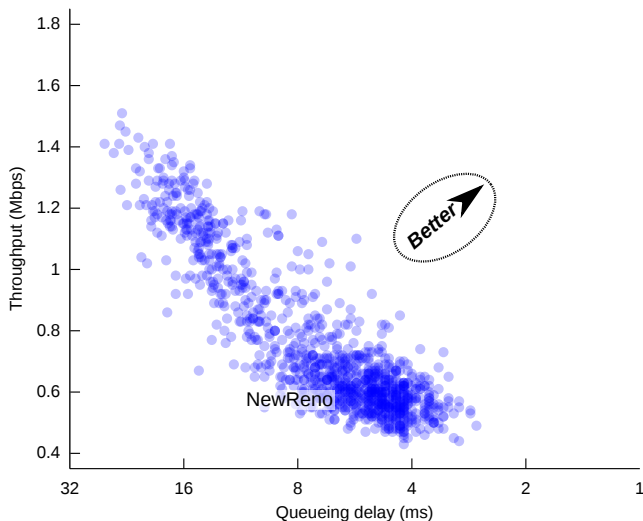
# Scenario 1: throughput-delay plot



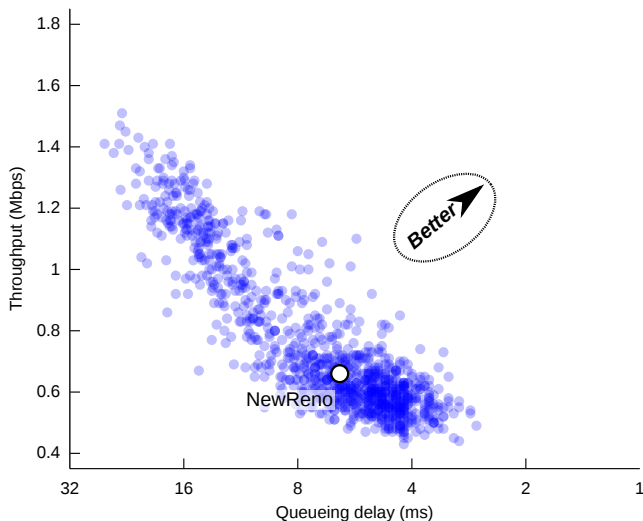
# Scenario 1: throughput-delay plot



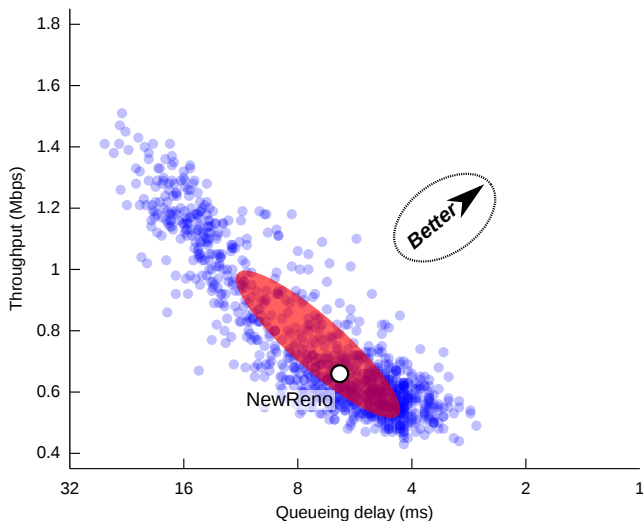
# Scenario 1: throughput-delay plot



# Scenario 1: throughput-delay plot

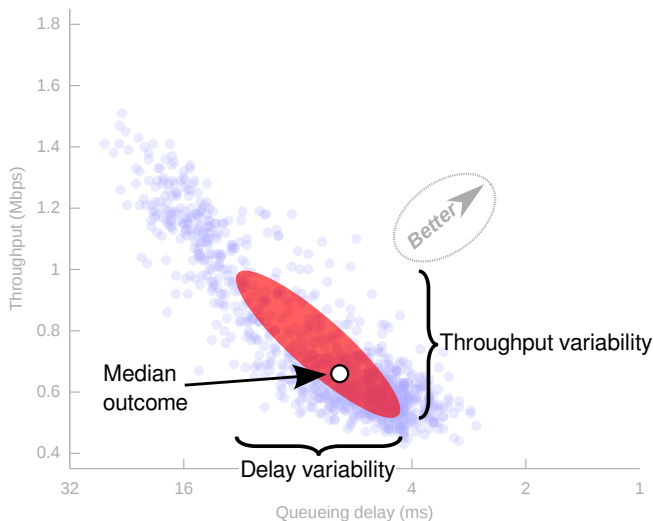


# Scenario 1: throughput-delay plot

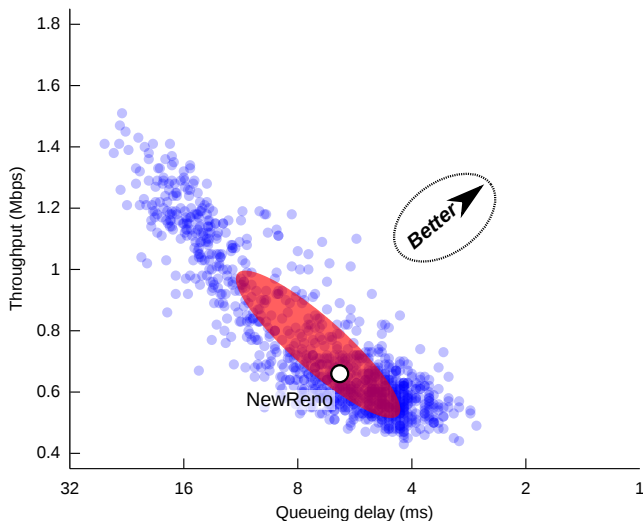




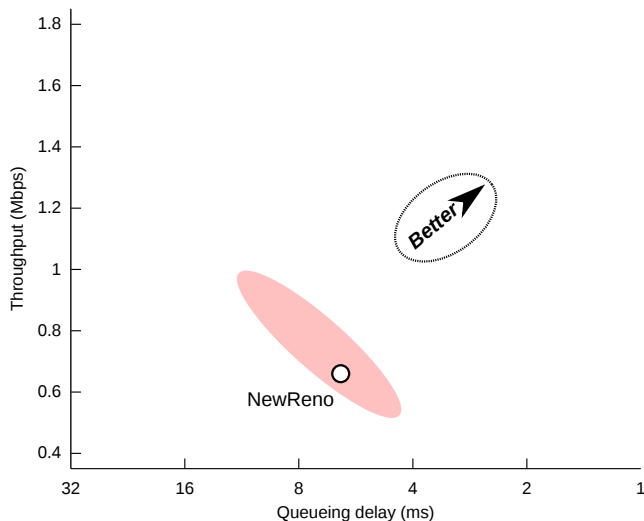
# Scenario 1: throughput-delay plot



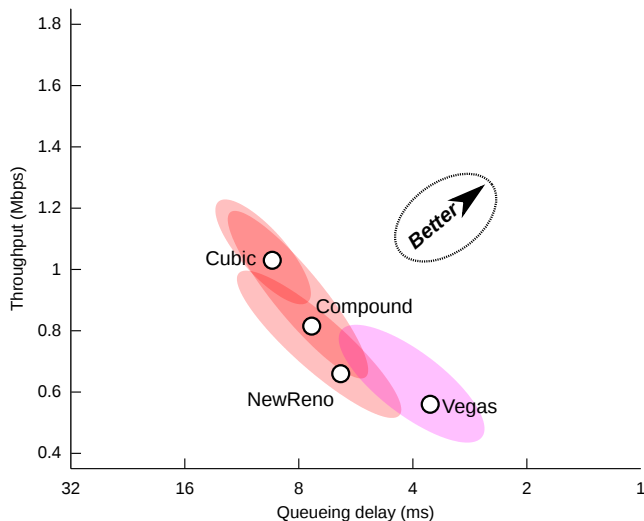
# Scenario 1: throughput-delay plot



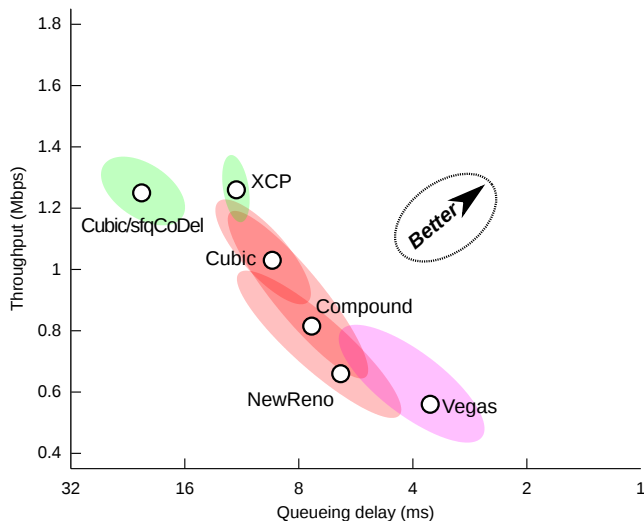
# Scenario 1: throughput-delay plot



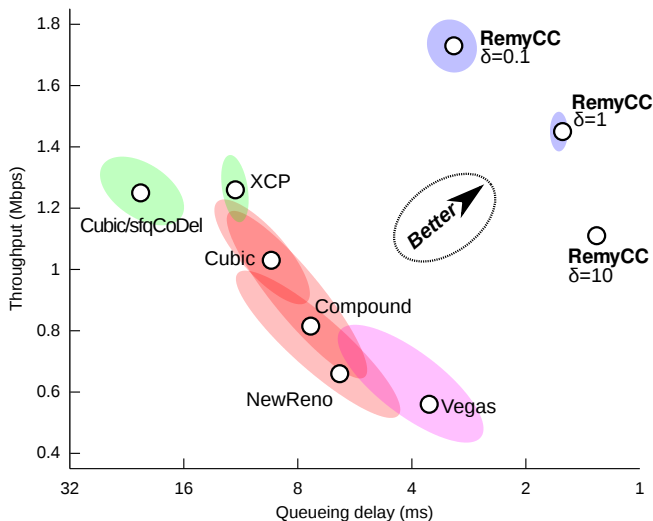
# Scenario 1: throughput-delay plot



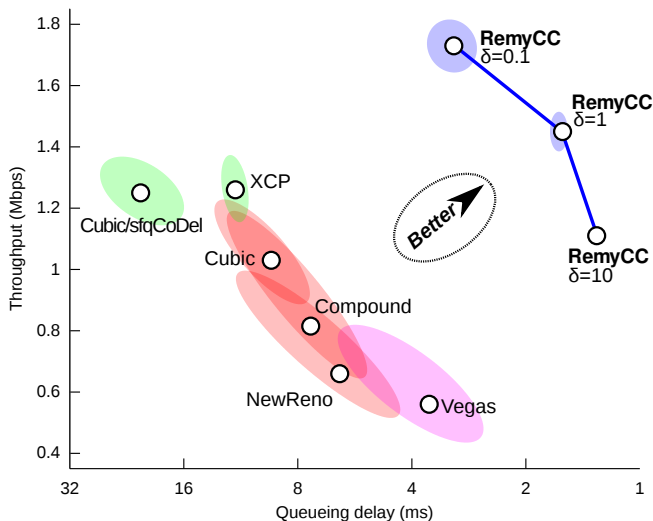
# Scenario 1: throughput-delay plot



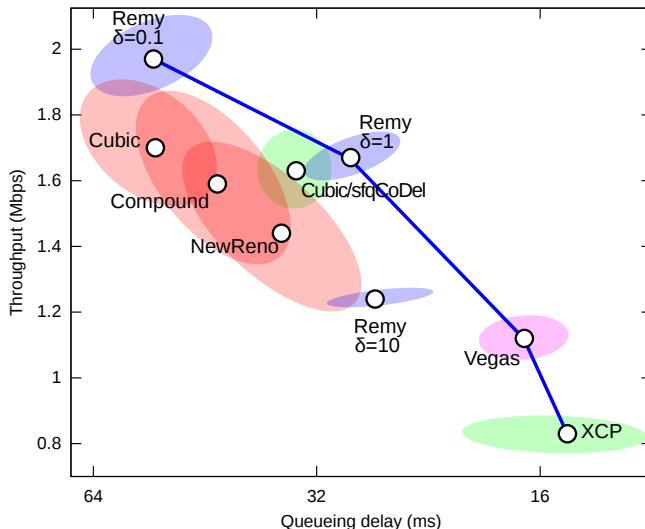
# Scenario 1: throughput-delay plot



# Scenario 1: throughput-delay plot



## Scenario 2: Verizon LTE, $n = 8$



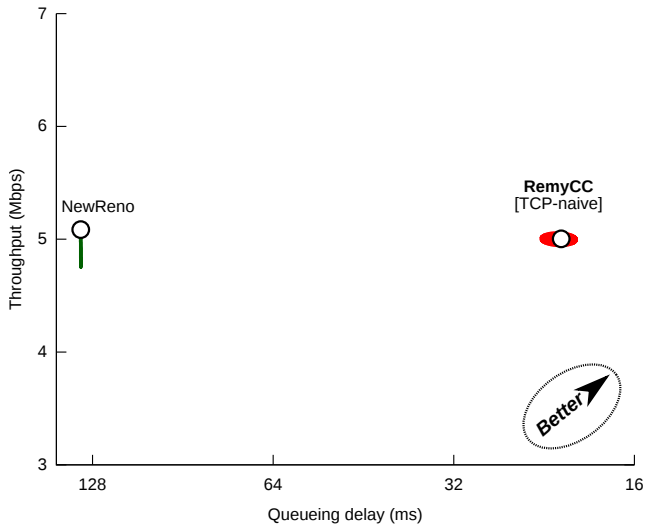


## Remy as an instrument to study network science

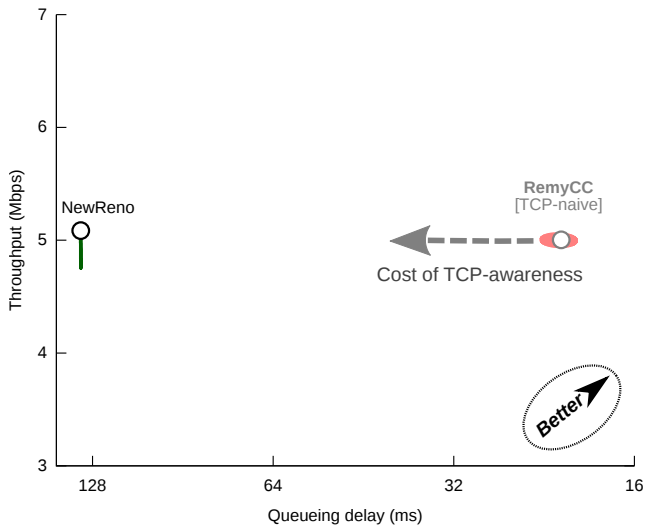
*From the perspective of an endpoint, **what does it help to know** about the network?*

*How difficult is it to learn a good protocol, given an **imperfect** model of the network?*

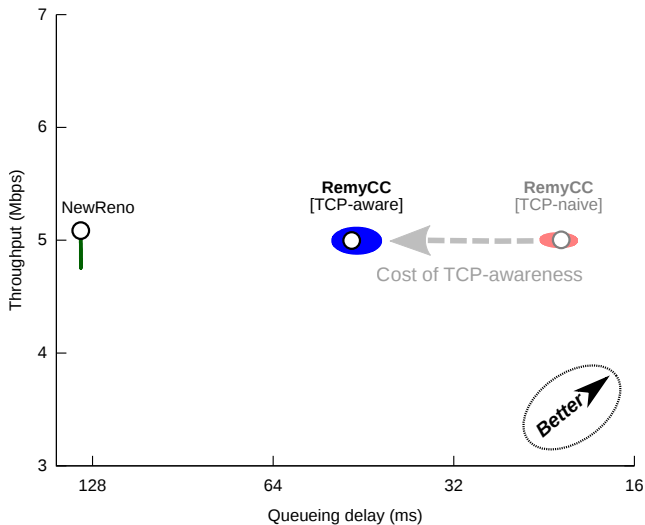
# RemyCC competing against itself



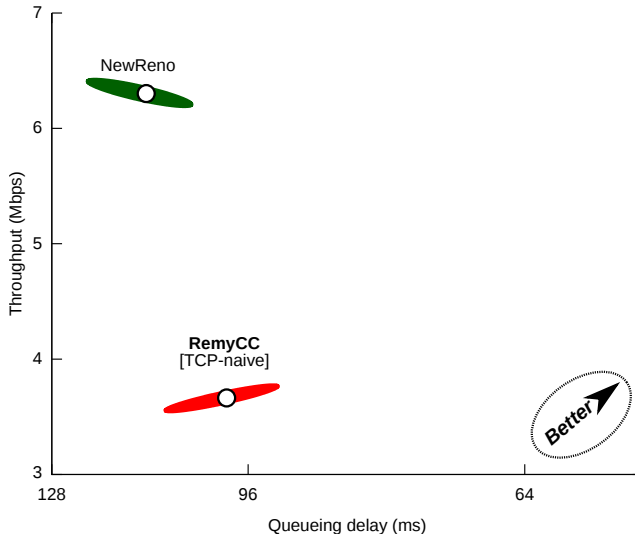
# RemyCC competing against itself



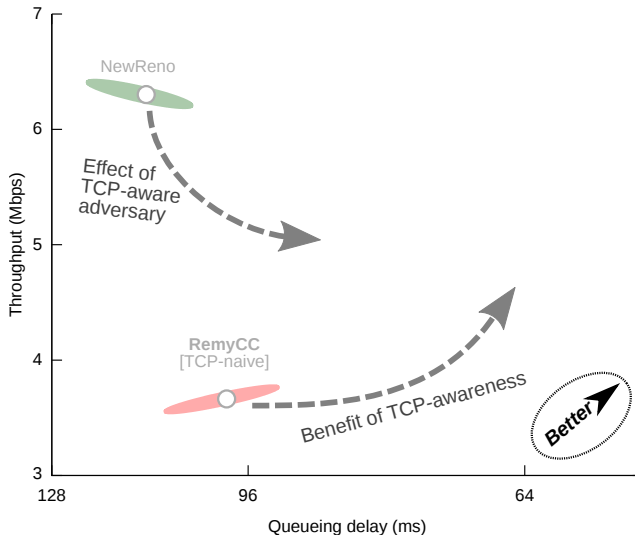
# RemyCC competing against itself



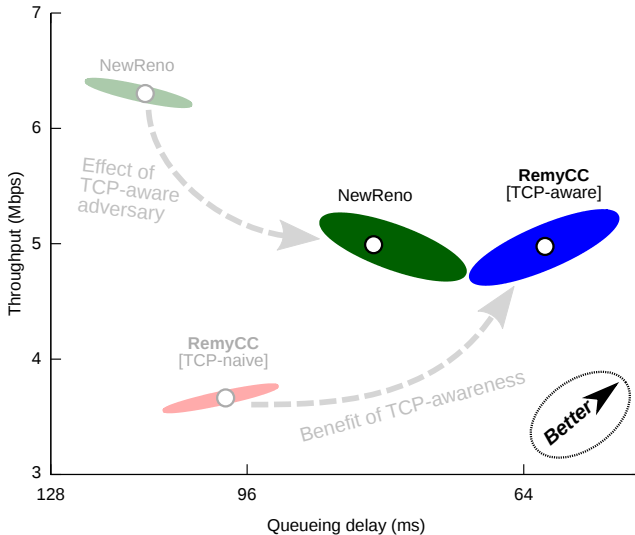
# RemyCC competing against TCP NewReno



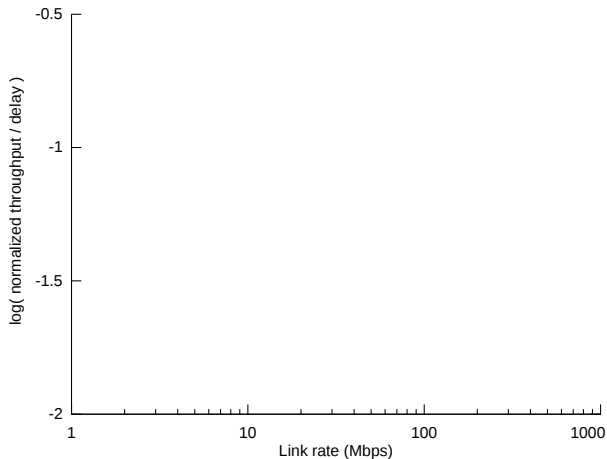
# RemyCC competing against TCP NewReno



# RemyCC competing against TCP NewReno



# The cost of generality

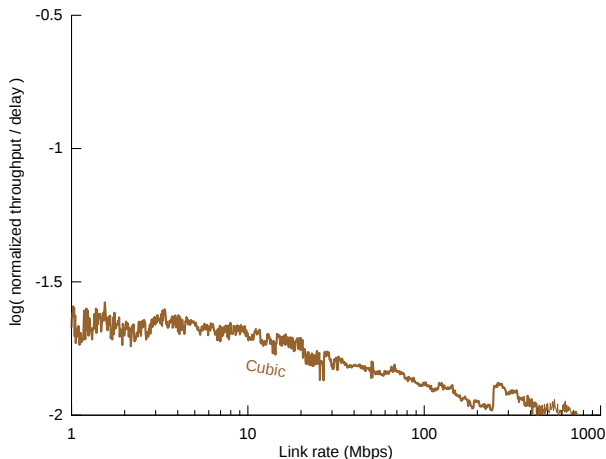


Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet



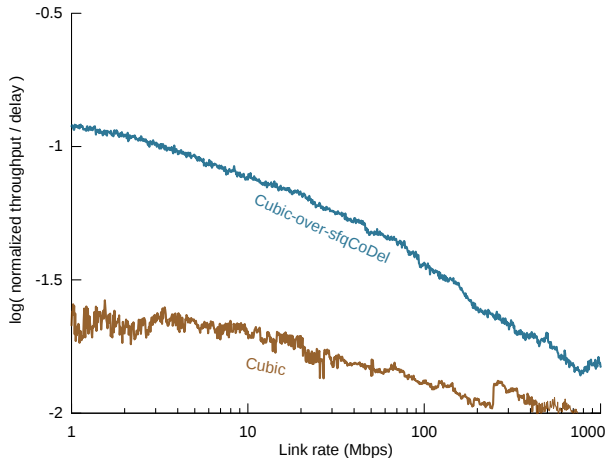
# The cost of generality



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

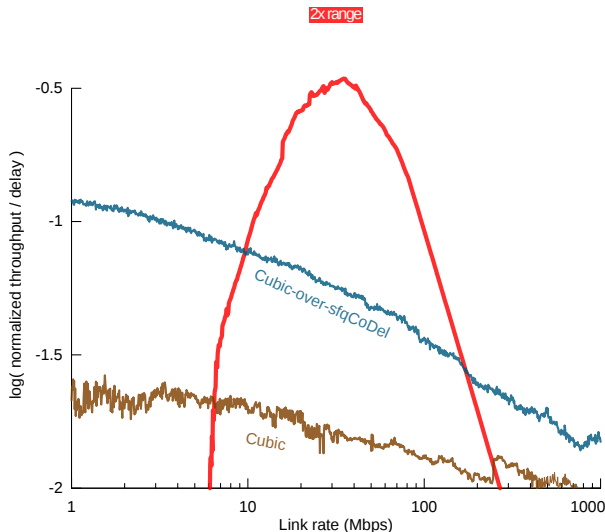
# The cost of generality



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

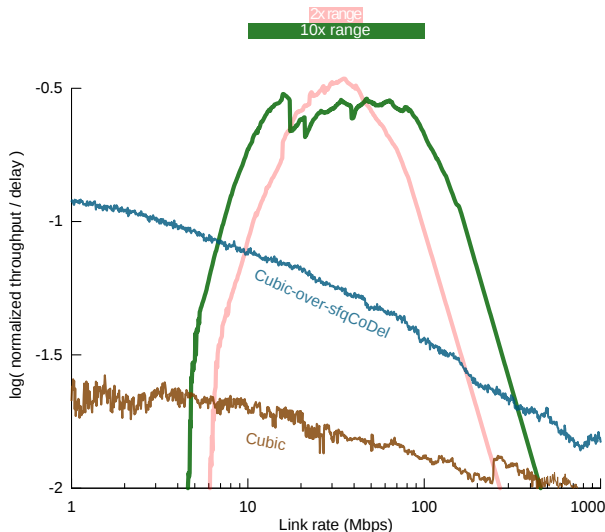
# The cost of generality



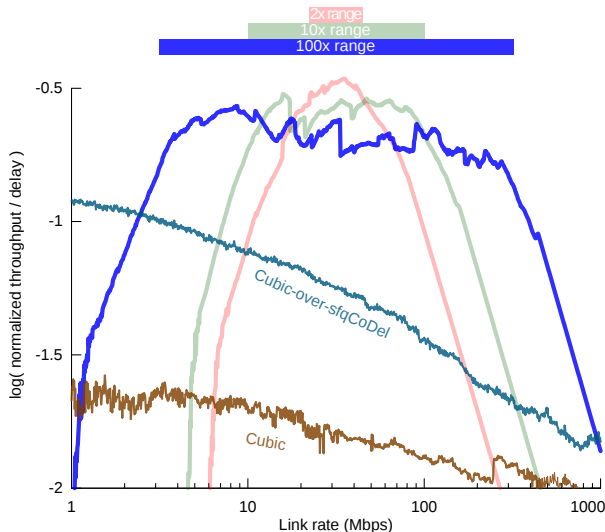
Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# The cost of generality



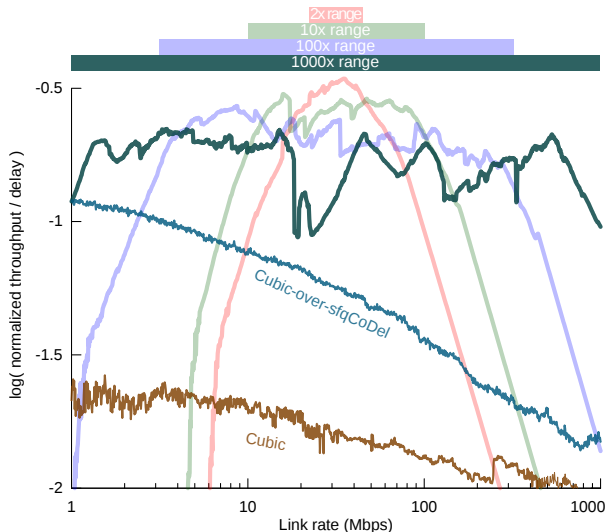
# The cost of generality



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# The cost of generality



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# Systems ex Machina

- ▶ Explicit design considerations → **freedom to make changes**
- ▶ “If this system is the answer, what’s the question?”

**Sprout** 2–4× the throughput and 7–9× less delay than Skype, etc.

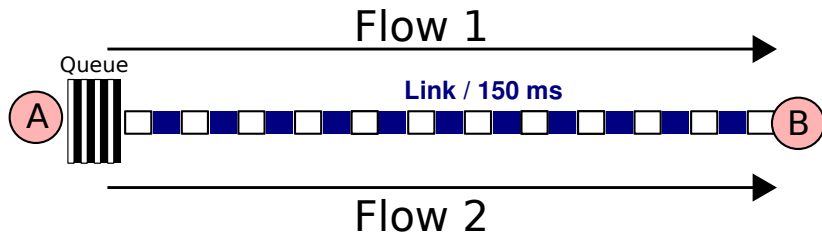
**Remy** computer-generated protocol design

keithw@mit.edu

<http://mit.edu/keithw>

# When the model is wrong about the topology

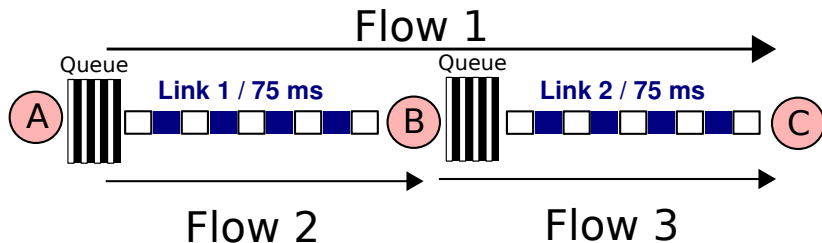
One bottleneck



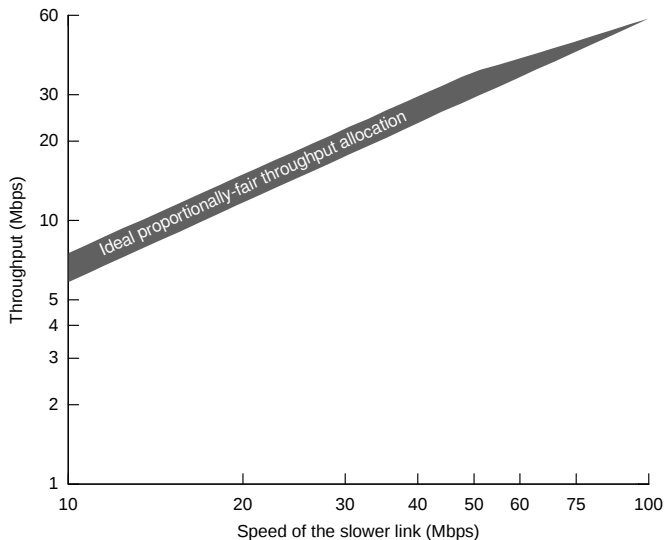


# When the model is wrong about the topology

Two bottlenecks



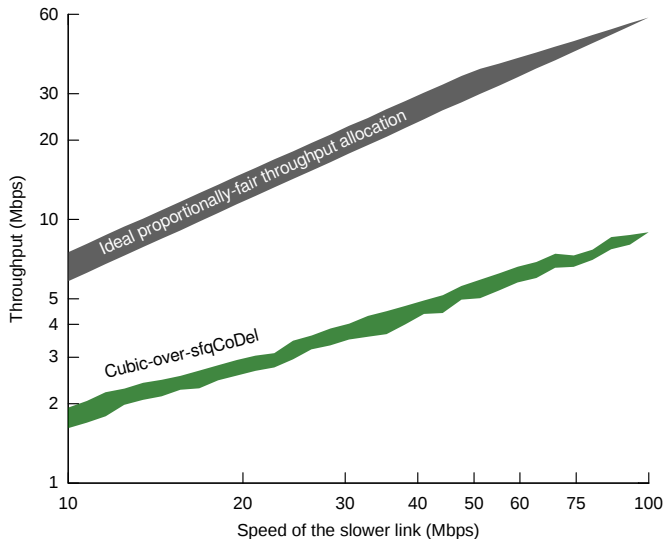
# When the model is wrong about the topology



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

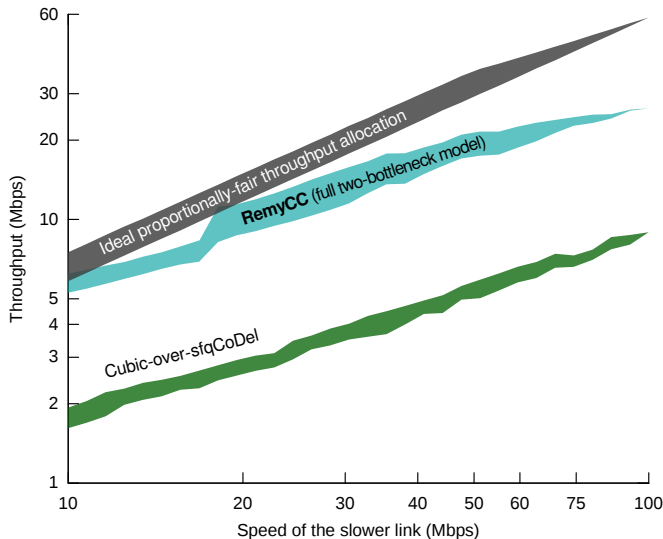
# When the model is wrong about the topology



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

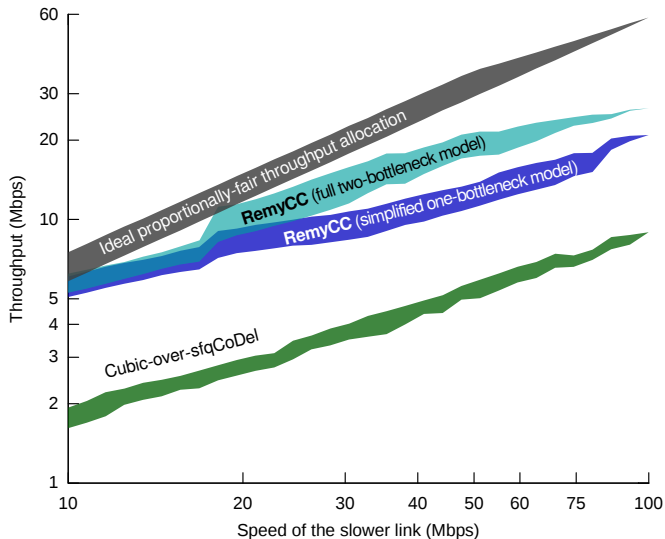
# When the model is wrong about the topology



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

# When the model is wrong about the topology



Keith Winstein (with Anirudh Sivaraman, Pratiksha Thaker, and Hari Balakrishnan)

Transport Architectures for an Evolving Internet

## Verizon 3G (1xEV-DO) Downlink

