



**ATLANTA, GA**  
**OCTOBER 11-14**

**SCTE**  
a subsidiary of CableLabs®

# UNLEASH THE POWER OF LIMITLESS CONNECTIVITY



**2021 Fall  
Technical Forum**  
SCTE • NCTA • CABLELABS





SCTE  
a subsidiary of CableLabs®

## Operational Transformation

# Optimizing DOCSIS 3.0 Configuration in the Upstream through Applied Reinforcement Learning

**Kevin Dugan**

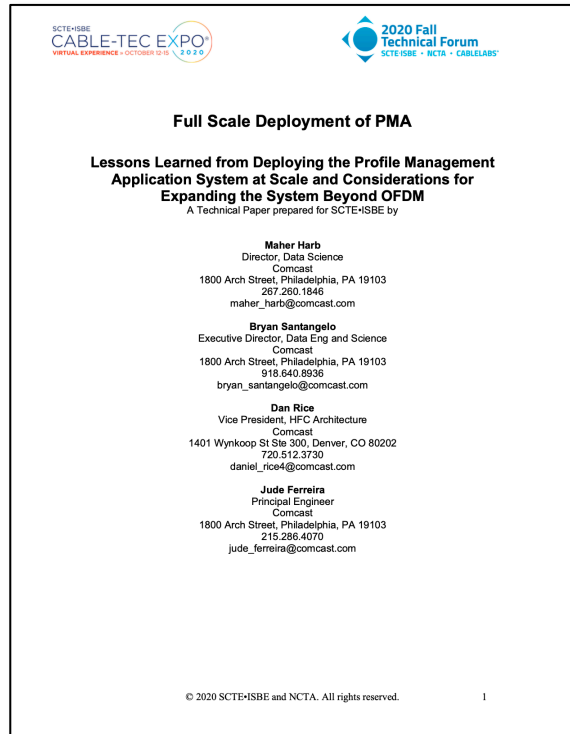
Scientist 3, Enterprise Data Analytics & Data Intelligence  
Comcast Cable



**VIRTUAL EXPERIENCE  
OCTOBER 11-14**

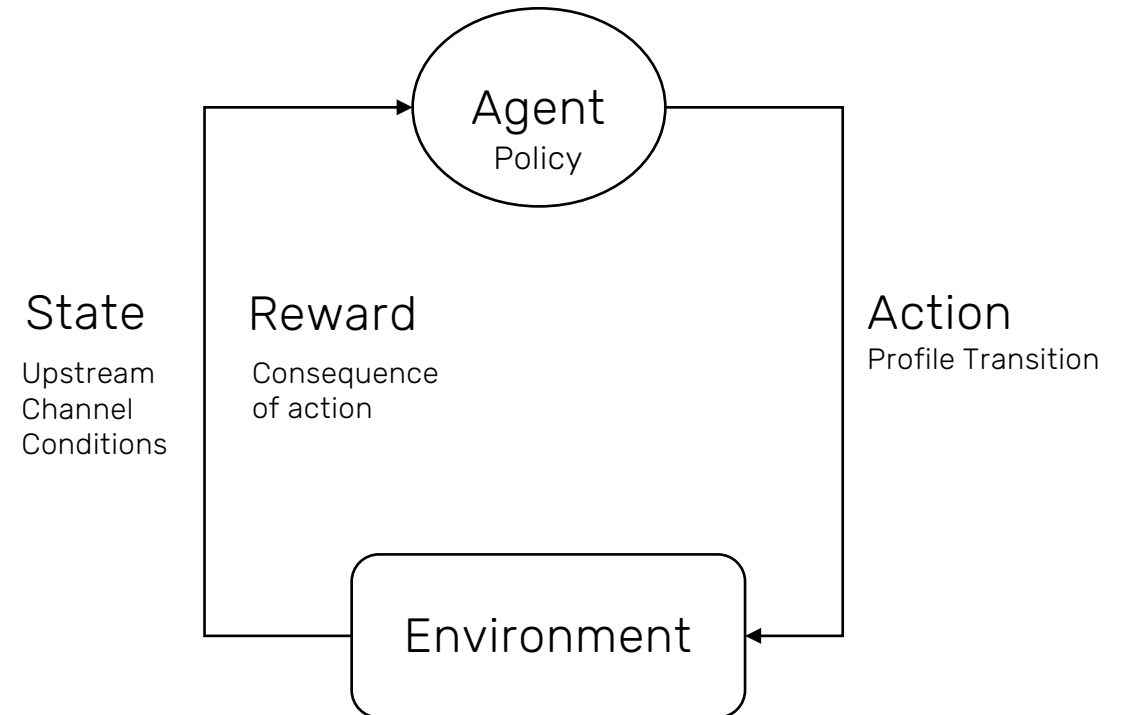
## Background

### D3.0 Upstream PMA



## Reinforcement Learning

### RL



# State-Action Value Function

SARSA – on-policy learning

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2 \dots$$

Temporal Difference – TD(0)

TD Error

Difference between estimated & actual reward

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha [R_{t+1} + \gamma Q(S_{t+1}, A_{(t+1)}) - Q(S_t, A_t)]$$

$\alpha$  = learning rate,  $\gamma$  = discount rate

## RL States, Actions, and Rewards for D3.0 US PMA

### States

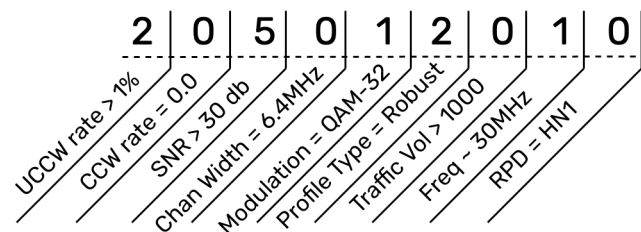
Category	Attribute	# Bins
Telemetry	Uncorrectable Codewords (UCCW)	3
	Correctable Codewords (CCW)	3
	Signal to Noise Ratio (SNR)	6
Channel Configuration	Channel Width	3
	Modulation	5
	Profile Type	5
	Traffic Volume	2
	Channel Frequency	6
	CMTS	5
<i>Total # Possible States</i>		<i>243,000</i>

### Actions

- Upgrade, Downgrade
- Same
- On / Off Transient

### Rewards

- 1 + Profile Speed Gain, if UCCW < 1%
- -10, if UCCW 1%



## Global Static Policy

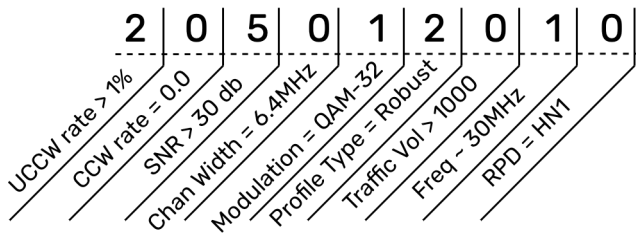
- Static decision criteria
  - Telemetry thresholds
- Manual updates
- One size fits all

## Dynamic Policy

- Flexible decision criteria
  - Learns continuously
  - Needs experience
- Automated updates
- Tailored to systems

## Building Dynamic Policies

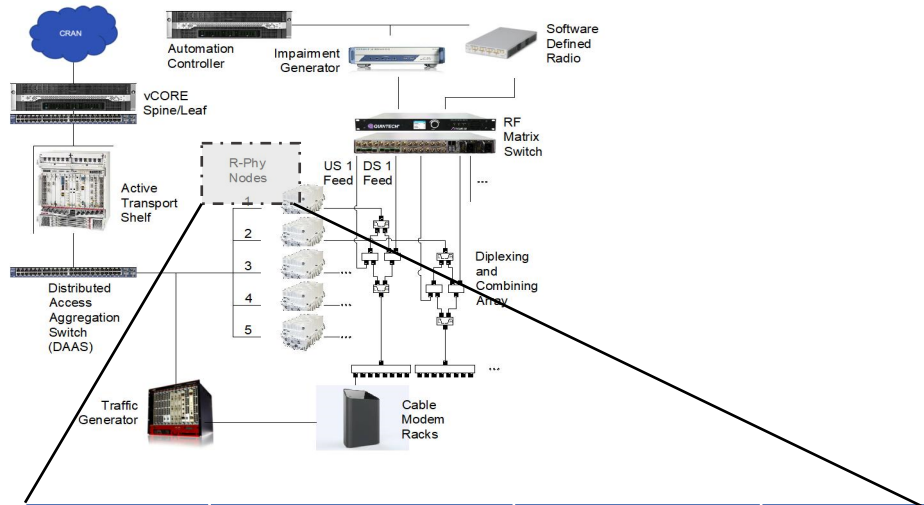
- Methodology used to calculate state-value pairs for the TD(0) equation
- Delayed reward
- Next state and next action become current state/action on next time step



	t		t + 1		
Time Step	State	Action	Reward	Next State	Next Action
24	005004020	upgrade 3	3	005001020	upgrade 1
25	005001020	upgrade 1	2	005000020	same
26	005000020	same	-10	205000020	downgrade 2
27	205000020	downgrade 2	-2	005002020	upgrade 1

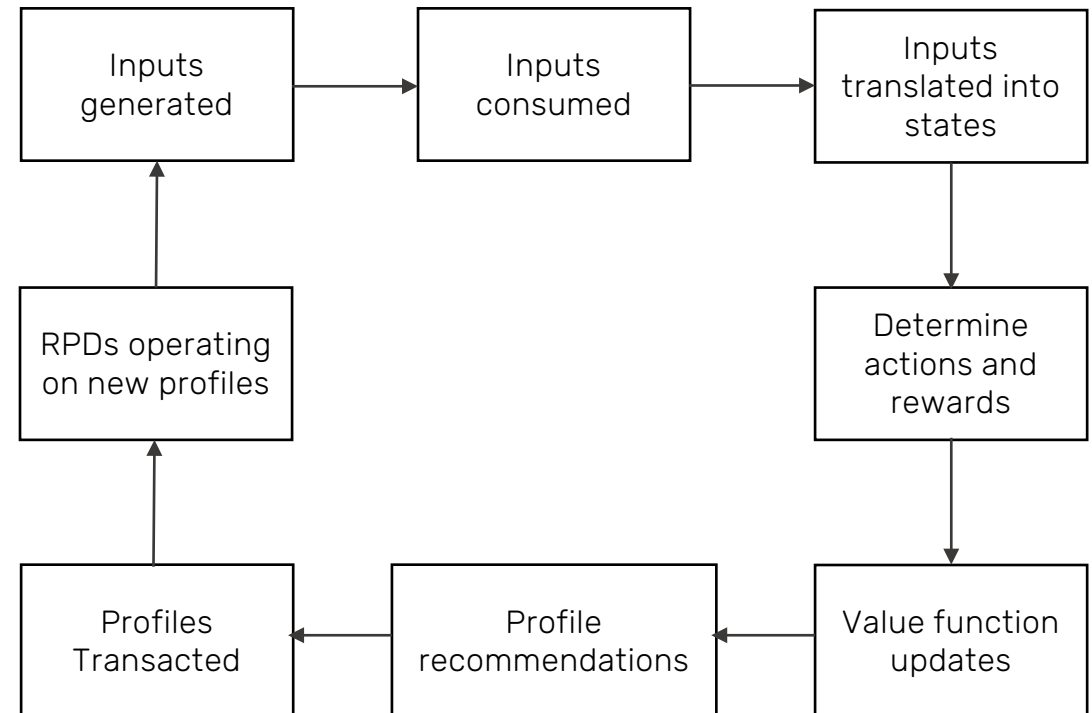


## Lab Systems



RPD Name	Number of Channels	Number of CMs	Vendor
AS2	6	6	CMTS X
AN1	4	8	CMTS X
CN1	4	7	CMTS Y
HN1	4	12	CMTS Z
HS3	6	8	CMTS Z

## E2E Closed Loop





## Trial Design

- 5 individual trials
  - Profiles set to baseline
  - Single policy makes profile recommendations for all RPDs per trial
    - 1 static policy
    - 4 dynamic policies
  - 25 time steps / iterations per trial
  - Random lab impairments
- 5 RPDs
  - 4 or 6 D3.0 US channels
  - 24 total channels

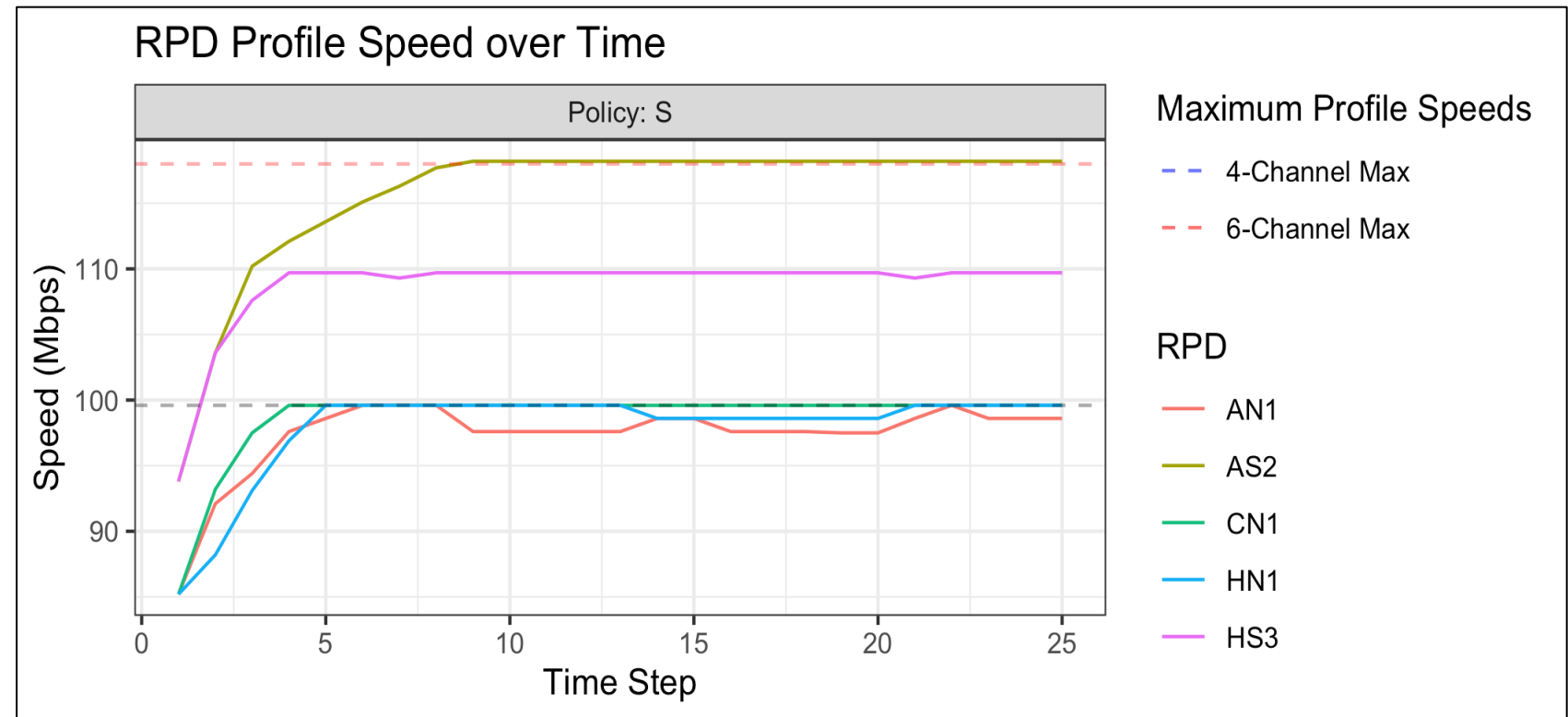
## Policy Evaluation

- Profile speeds achieved
  - per RPD (bonding group of channels)
  - considerations for impairments
- UCCW rates
  - Policy response to adverse UCCW
- Latency to best available profile
  - Number of time steps to best possible profile

## Profile Speeds Achieved

### Summary

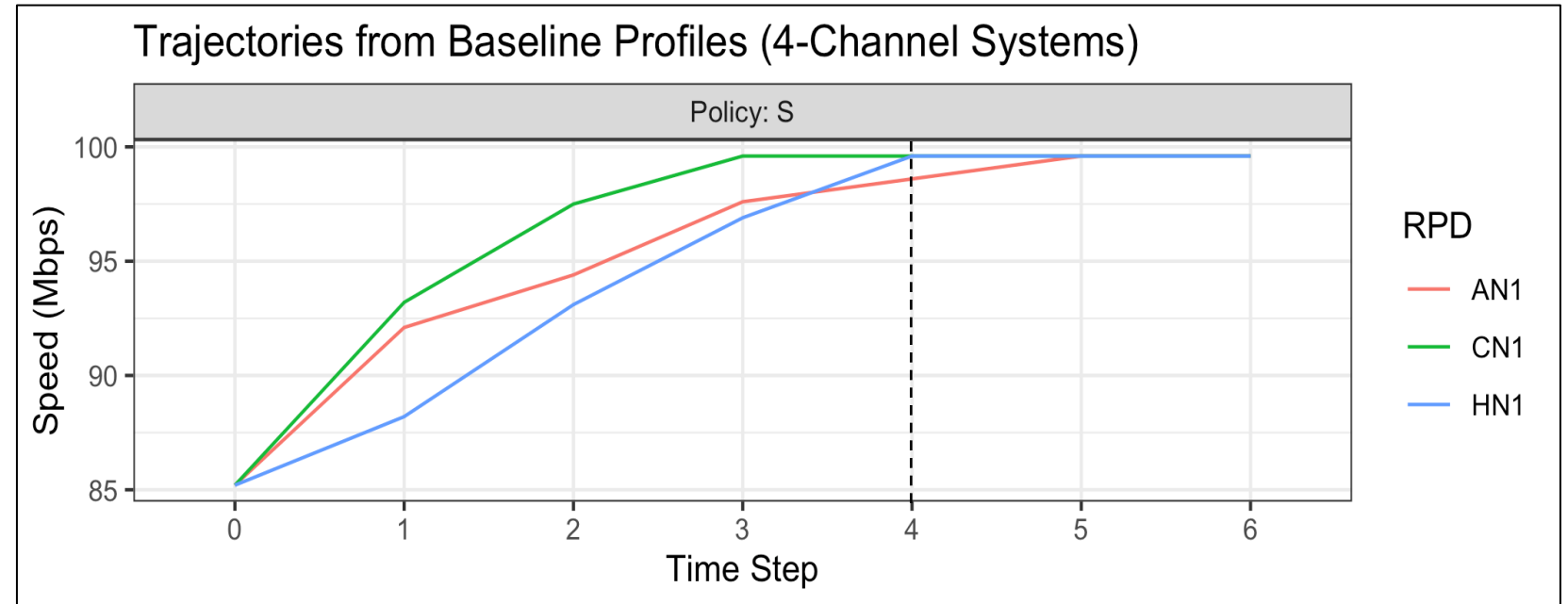
- Most channels achieved optimal profiles
- Some on best-available profiles
- AN1 profile speed dips
- Absence of adverse telemetry
- HS3 channel reporting error



## Latency to Best Available Profiles

### Summary

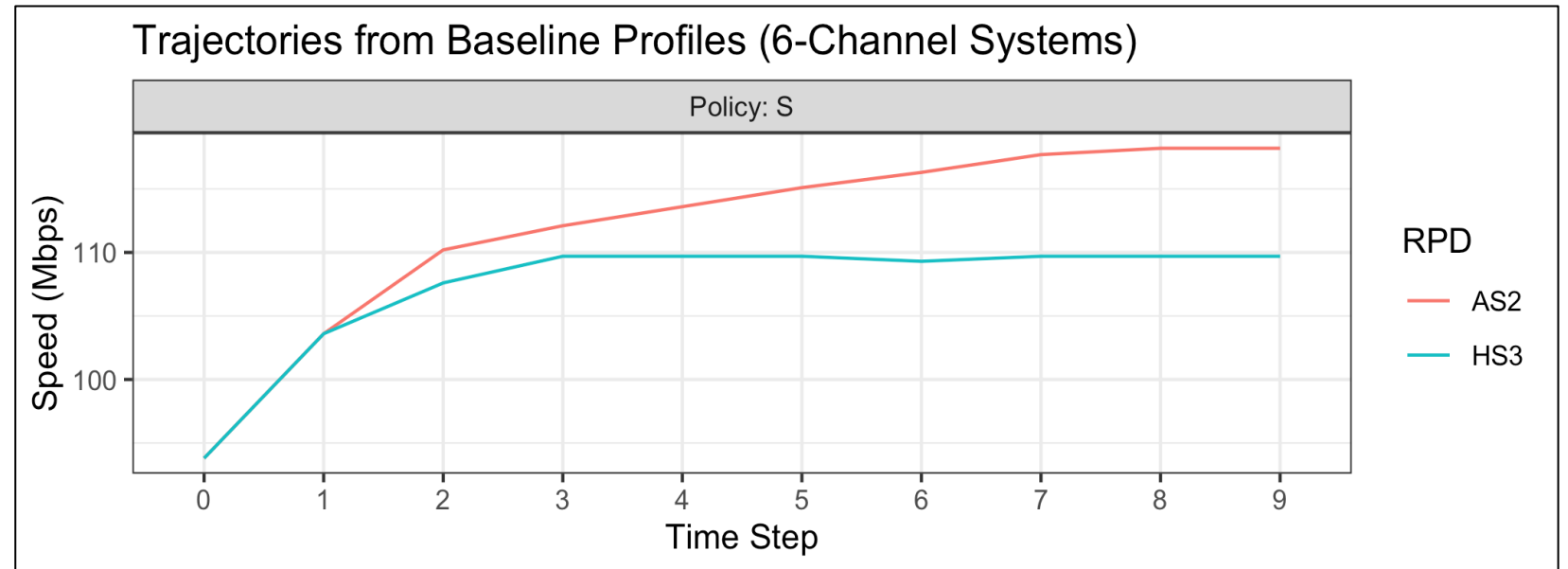
- 4-channel systems
- Avg 4 time steps to steady state profile
- Cautious steps from baseline
- All achieved optimal profiles by 5th time step



## Latency to Best Available Profiles

### Summary

- 6-channel systems
- HS3 reached best available in 3 time steps
- AS2 reached best available in 8 time steps

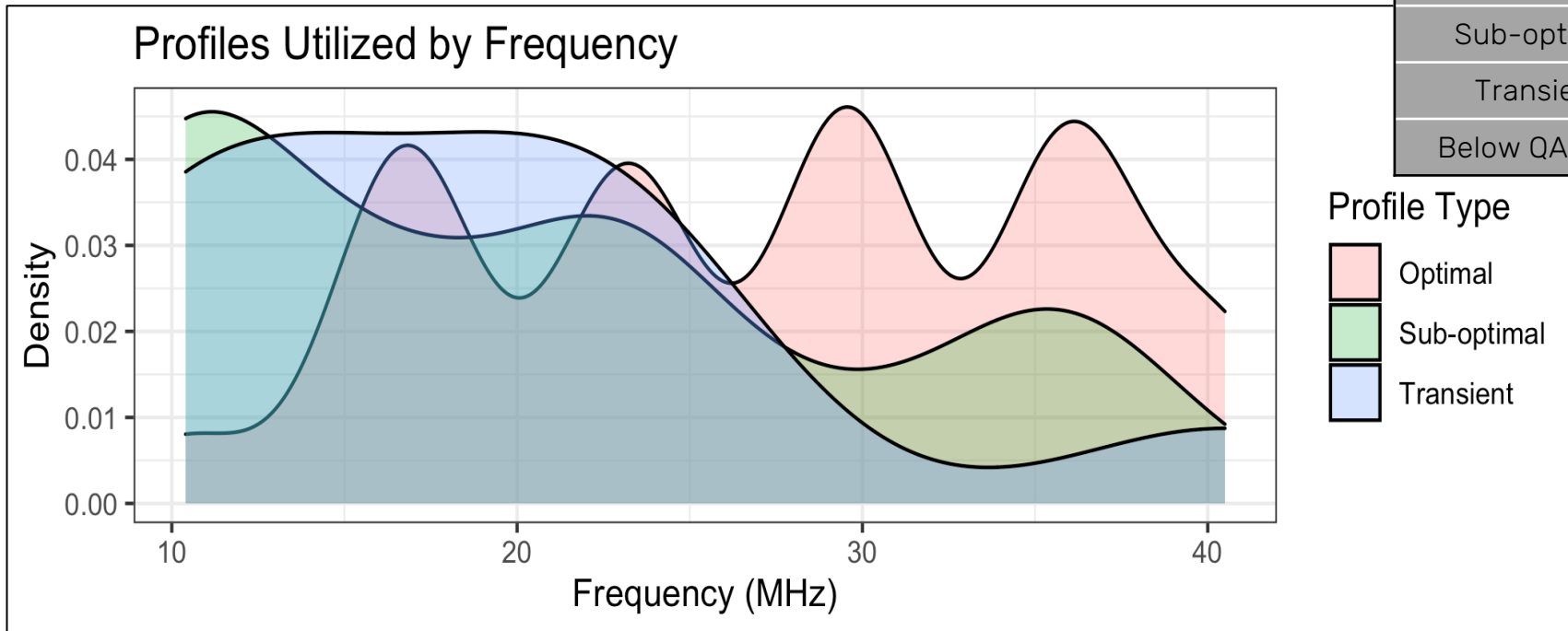




# Profile Density by Frequency

Avg of 104.3 Mbps per iteration (across all RPDs)

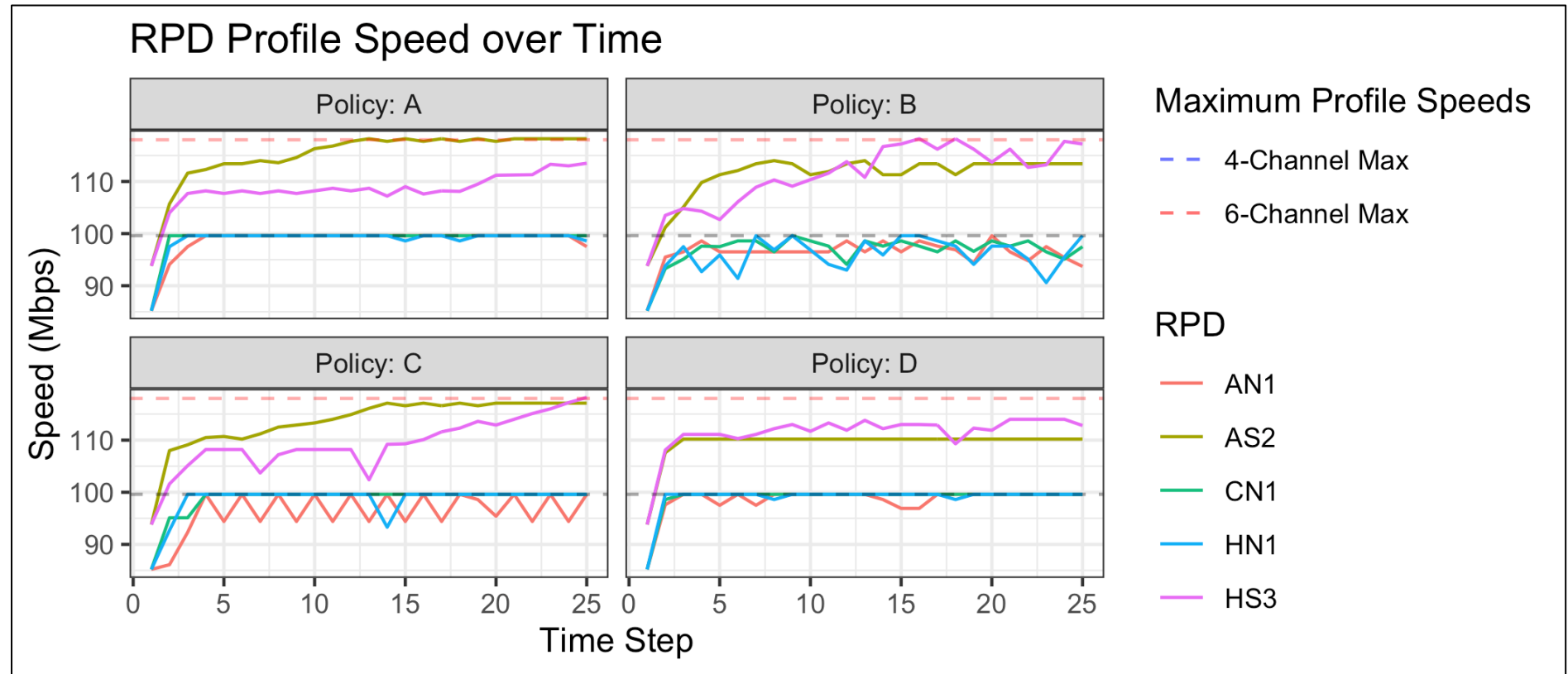
Profile Type	% of Total Speed	% Profile Occurrences
Optimal	88.95%	85.15%
Sub-optimal	8.83%	8.44%
Transient	1.37%	2.24%
Below QAM-64	0.84%	4.17%



## Profile Speeds Achieved

### Summary

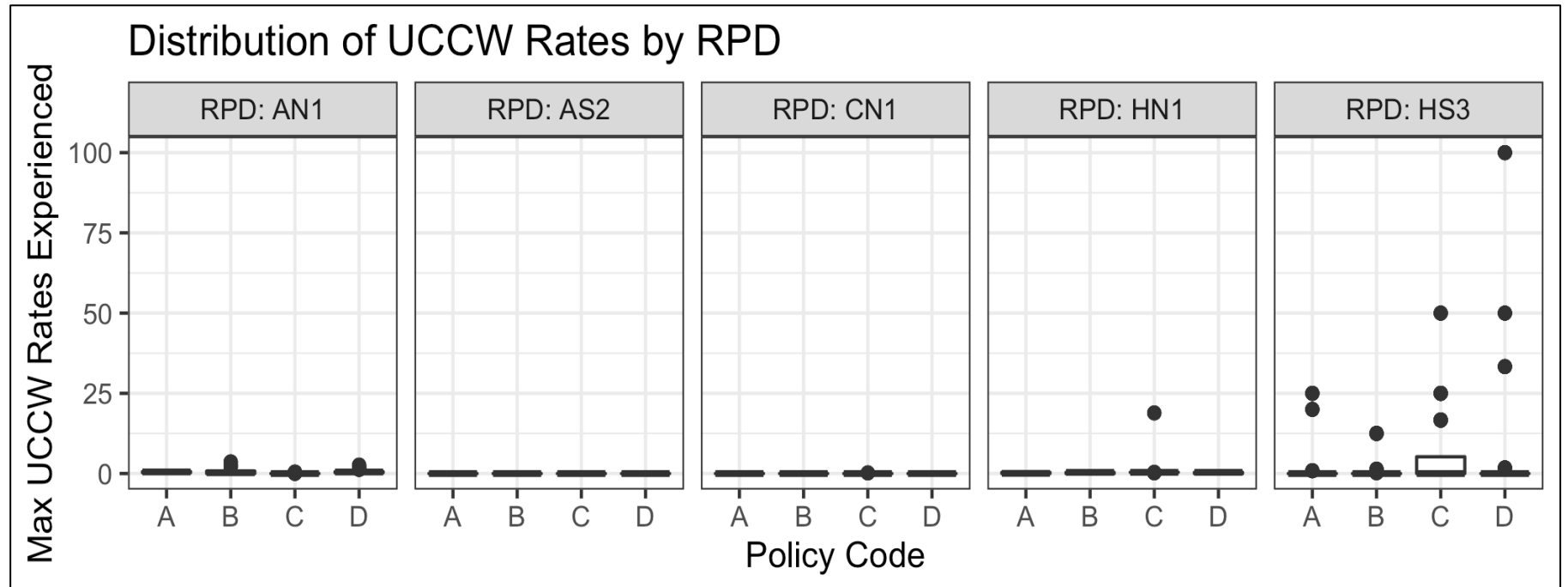
- A & D steady on 4-channel systems
- More UCCW events = more profile movement



## UCCW Rate Evaluation

### Summary

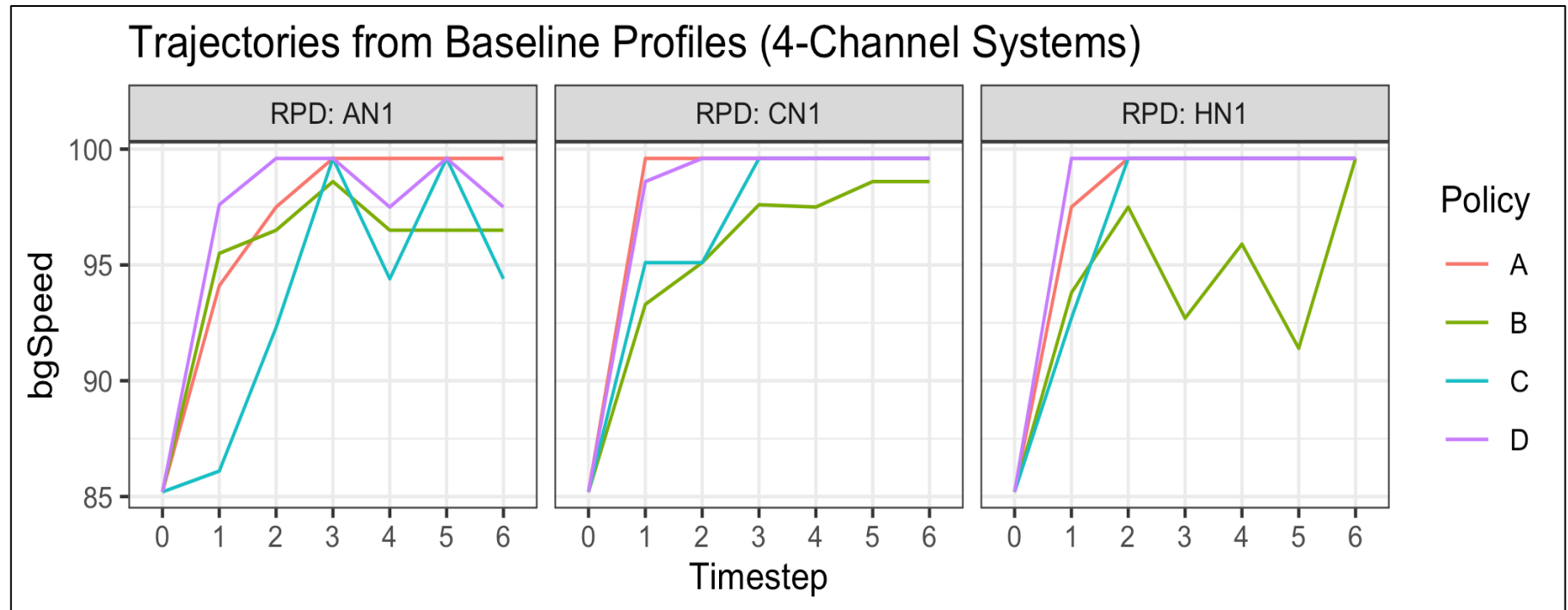
- 4-channel systems mostly noise-free
- HS3 experienced several adverse rates



## Latency to Best Available Profiles

### Summary

- 4-channel systems
- A & D avg 2 time steps
- B & C difficulty reaching optimal profiles
- Policy indecisiveness

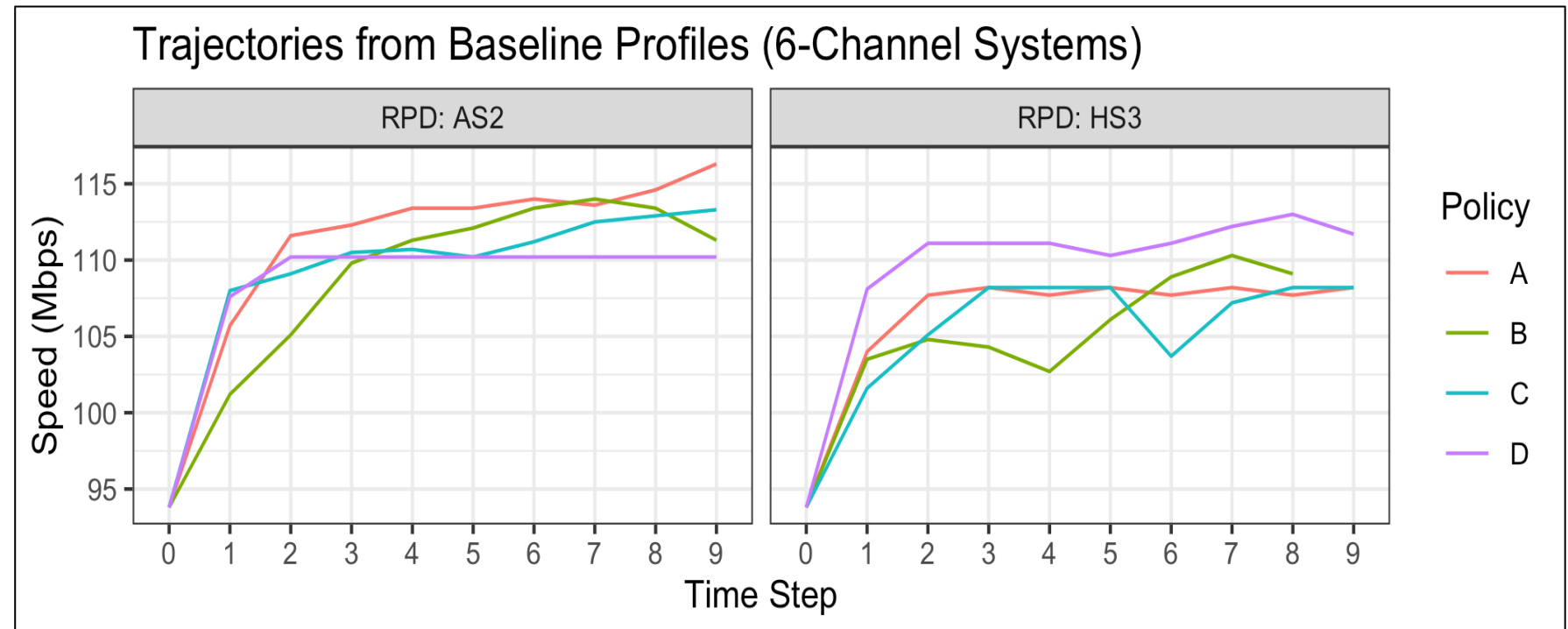




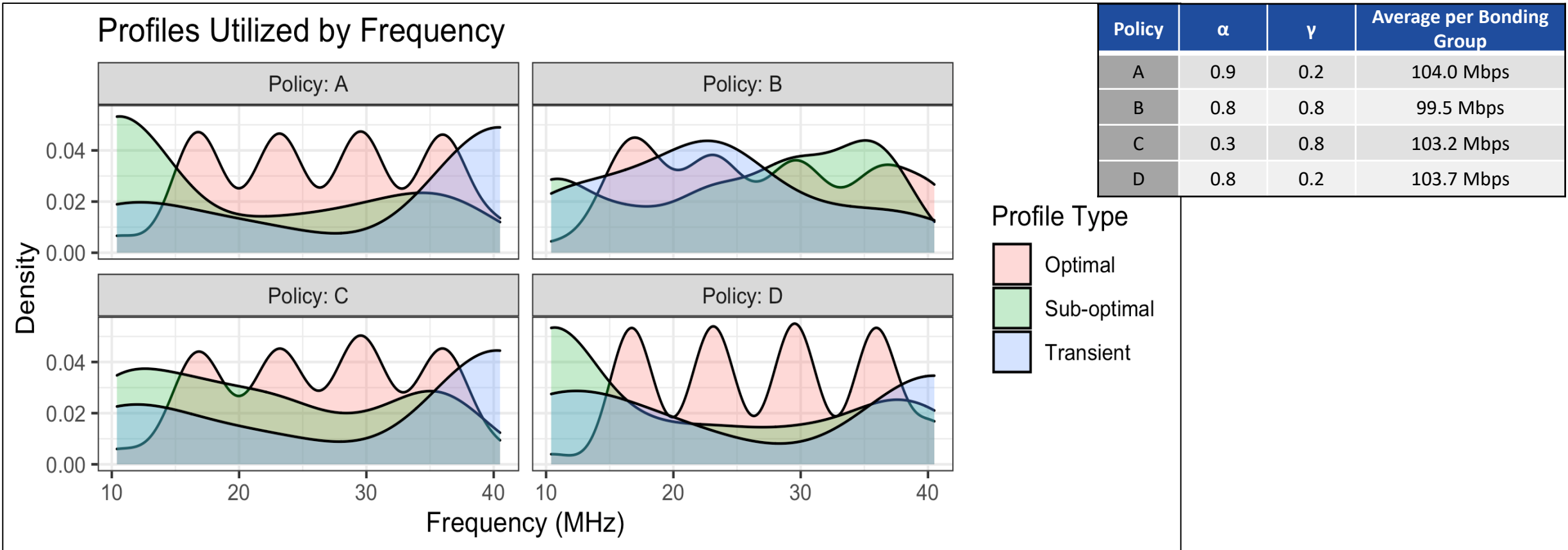
## Latency to Best Available Profiles

### Summary

- 6-channel systems
- Policy A steady growth on AS2, 9 time steps to max
- HS3 UCCW rate impacted policies
- Policy A & D reached best-available profiles in 2 steps on HS3



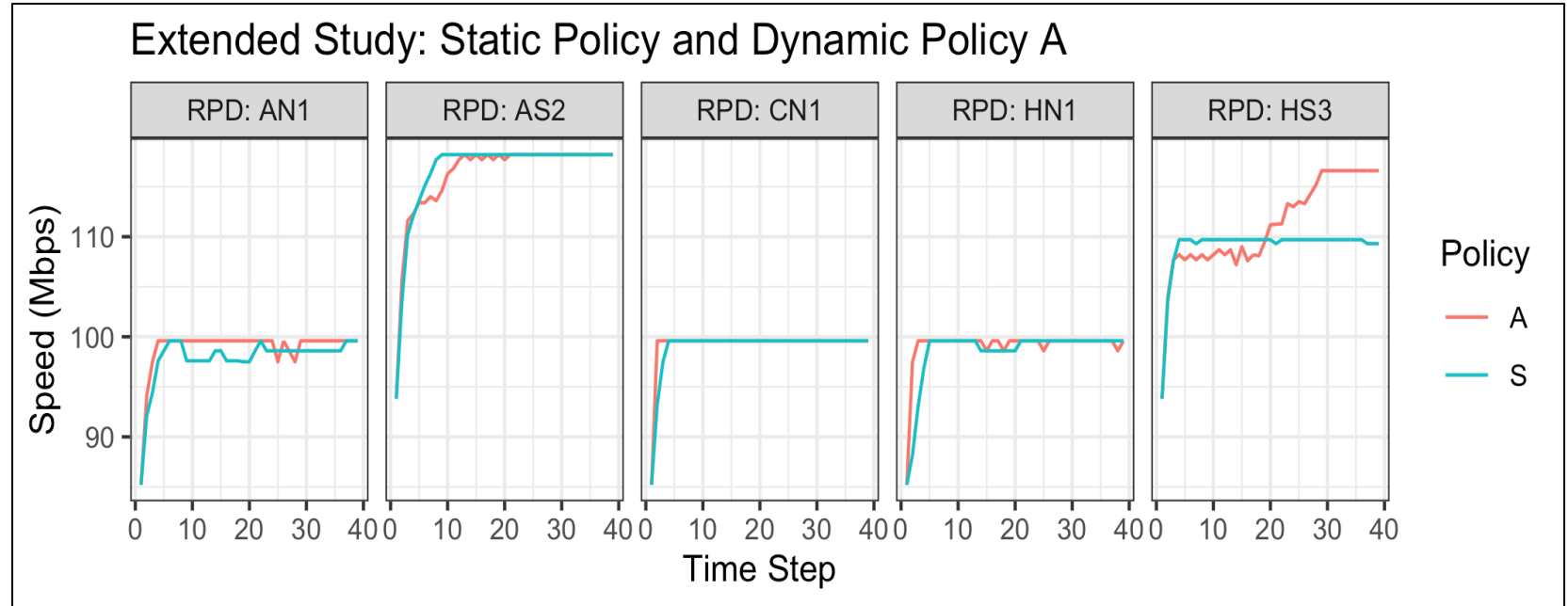
## Profile Density by Frequency



# Extended Study Results

## Summary

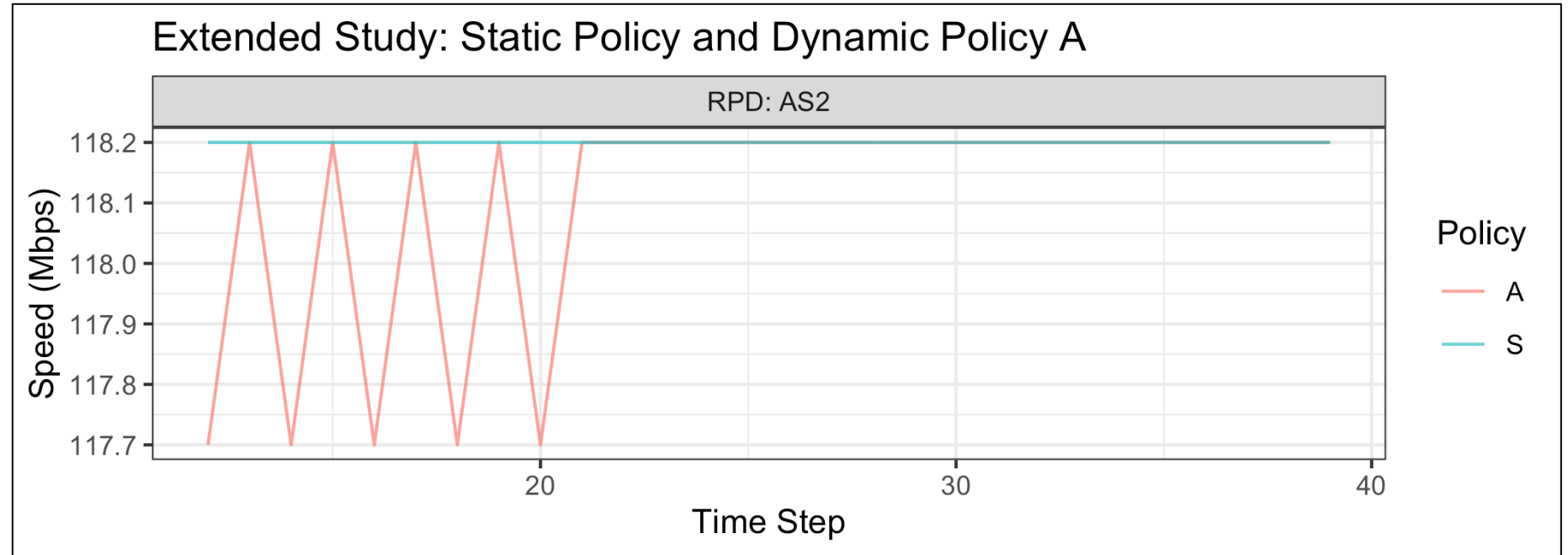
- Policy A vs static policy
- Similar behavior for most RPDs
- Policy A avg bonding group speed: 104.2 Mbps
- Static policy avg bonding group speed: 104.6 Mbps



## Example of RL Policy Learning New Action

### Summary

- Policy A fluctuated between optimal and sub-optimal
- Learned new action (stay in optimal profiles)





## Opportunities for Enhancement / Steps Forward

- Train initial policy from scratch
  - Eliminate influence of static policy patterns
- Synchronize impairments per time step
  - Cleaner comparison of policy behavior
- Consideration of using n-step TD prediction methods
  - Accounts for evaluating longer sequence of changes



ATLANTA, GA  
OCTOBER 11-14

SCTE<sup>®</sup>  
a subsidiary of CableLabs<sup>®</sup>

# Thank You!

**Kevin Dugan**

Scientist 3, Enterprise Data Analytics & Data Intelligence  
Comcast Cable  
[kevin\\_dugan2@comcast.com](mailto:kevin_dugan2@comcast.com)