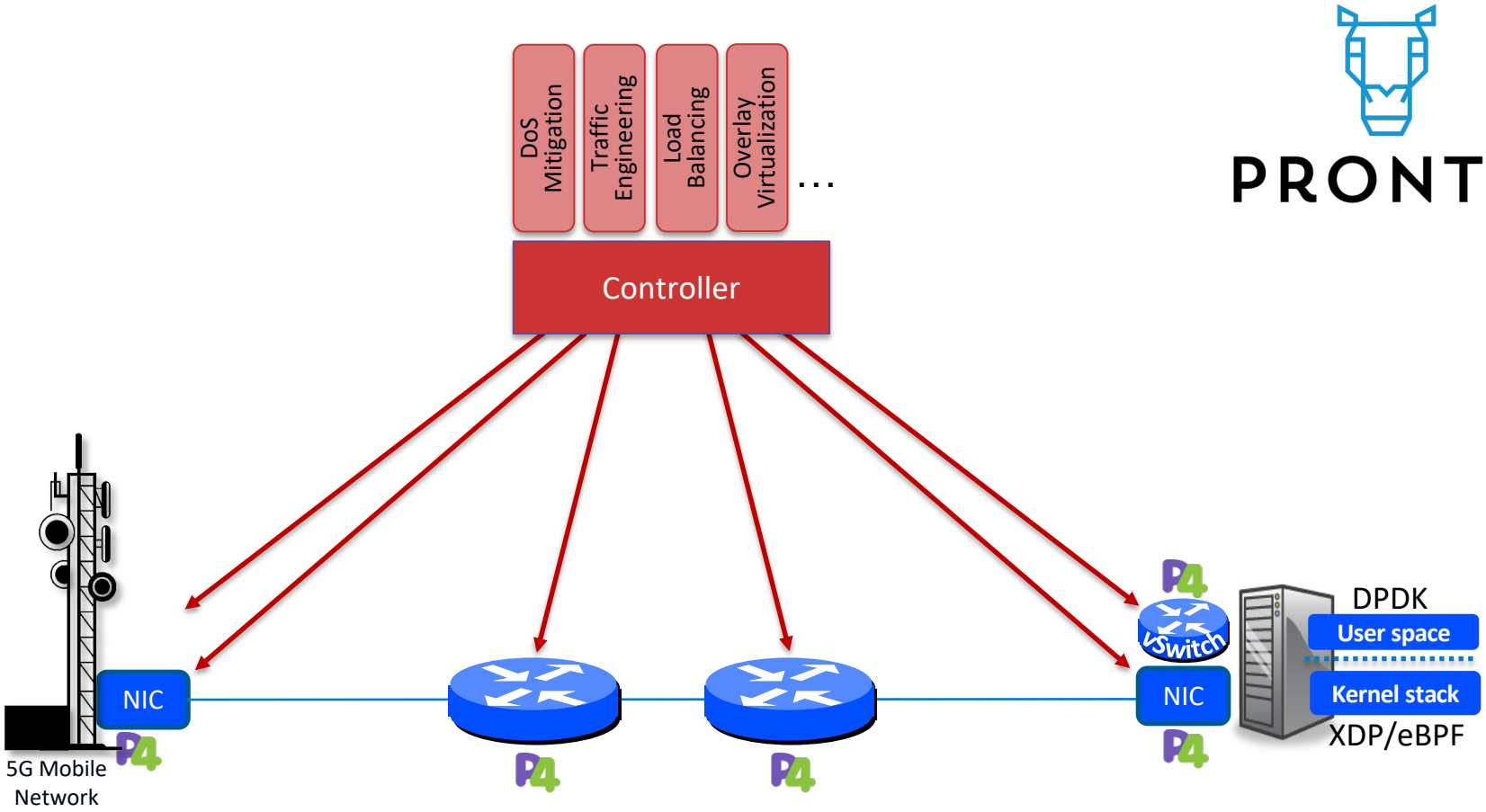
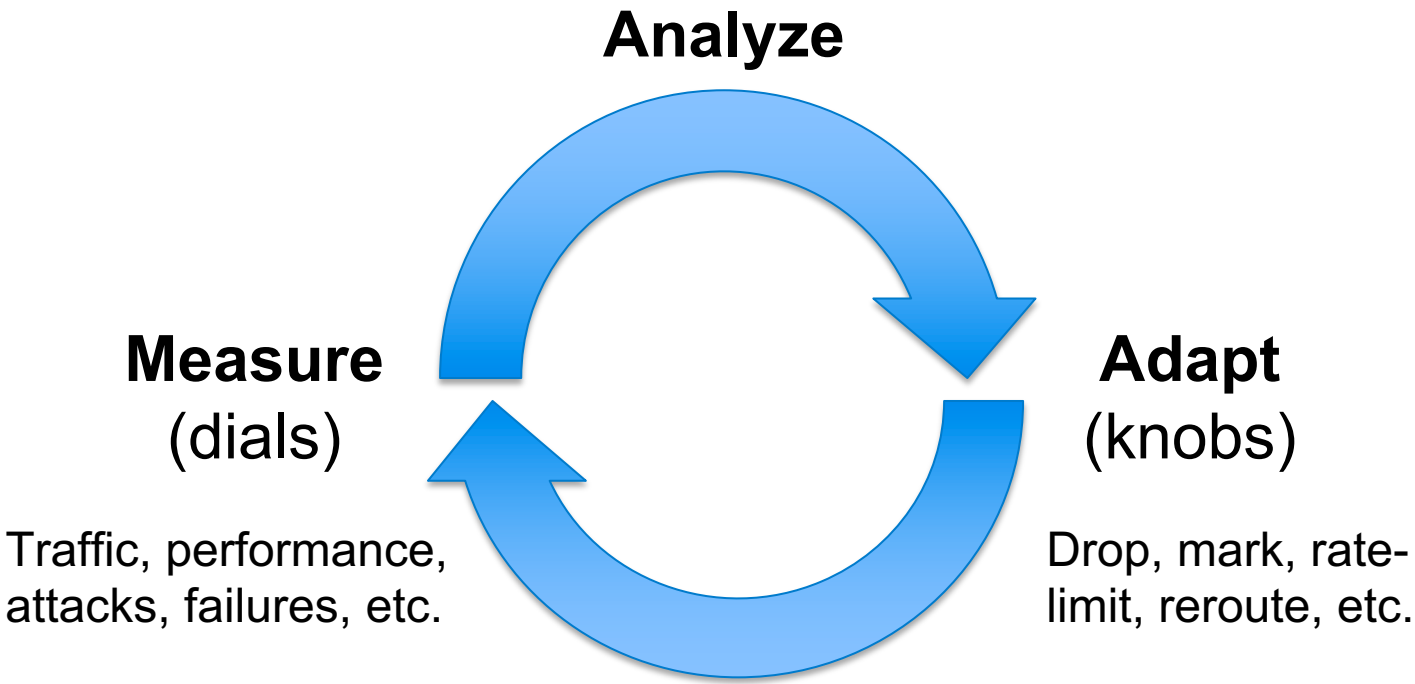
Higher-Order Telemetry in the Data Plane

Jennifer Rexford
Princeton University

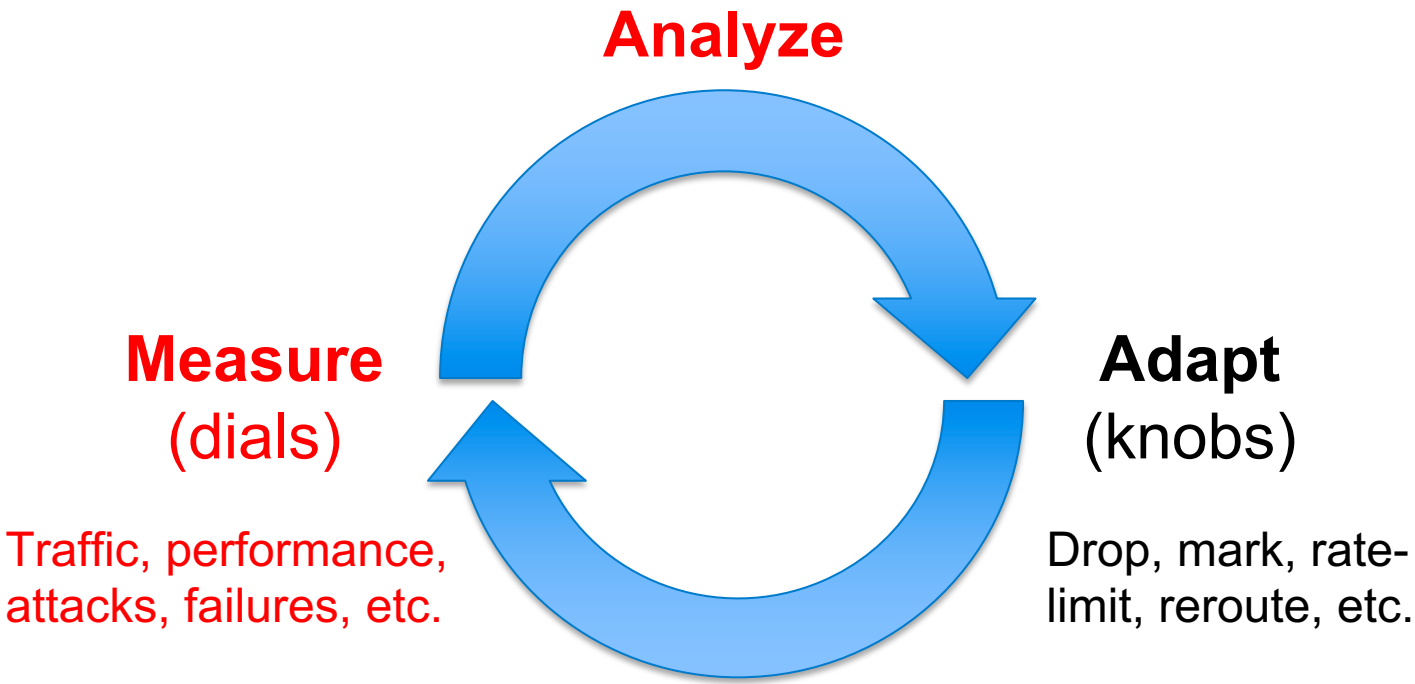
Programmability Top-to-Bottom and End-to-End



Closing the Control Loop

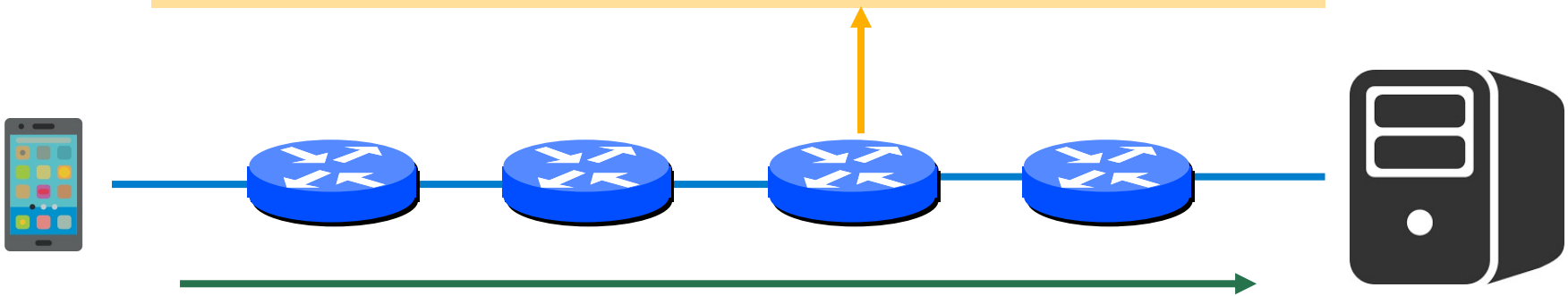


Key Enabler: Network Telemetry



Data-Plane Telemetry Today

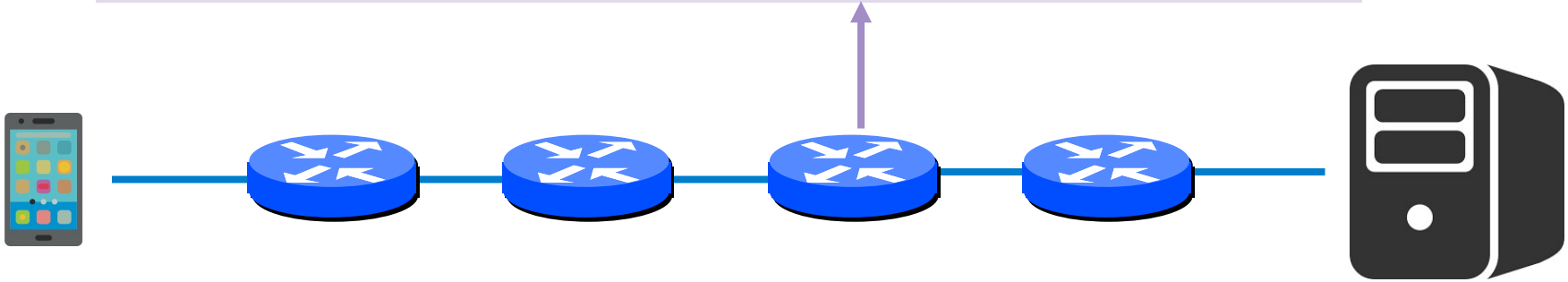
Counts based on network identifiers
“67,845 bytes between 1.2.3.4 and 5.6.7.8”



Path measurements of individual packets
“Queueing delay of 25 msec from 1.2.3.4 to 5.6.7.8”

Higher-Order Telemetry in the Data Plane

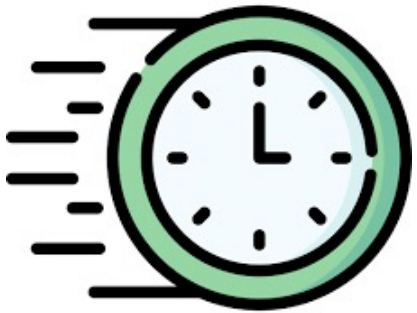
End-to-end performance for network services
"50 msec **round-trip time to Netflix**"



In the data plane for efficiency, privacy, and direct action!

Higher-Order Data-Plane Telemetry: Performance

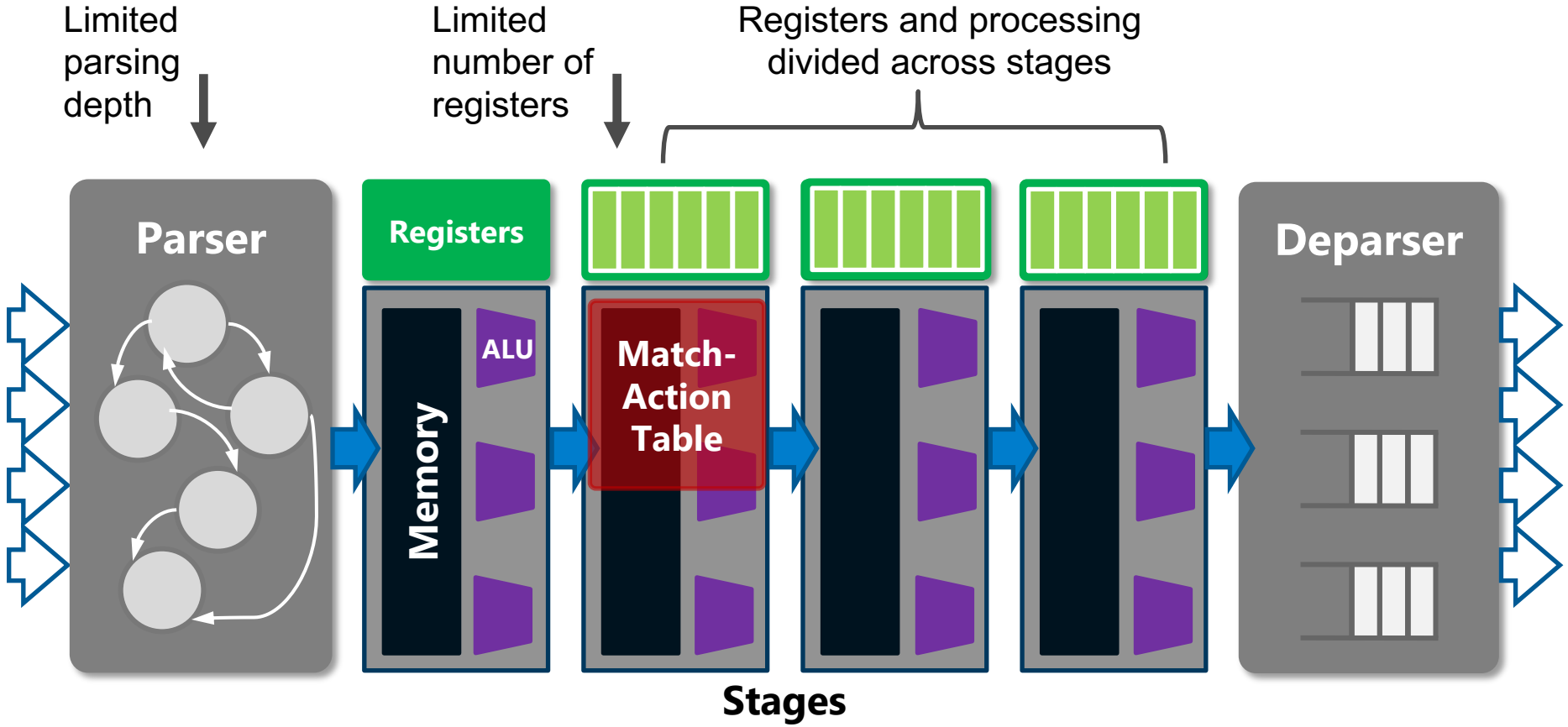
- From traffic counts to end-to-end performance
  - Traffic counts: # bytes, # packet losses, …
  - Performance: round-trip-time
  - *Challenge*: join a packet with its ACK
- Example use cases
  - Service Level Agreement violations
  - BGP interception attacks
  - CDN replica selection
  - Video QoE inference



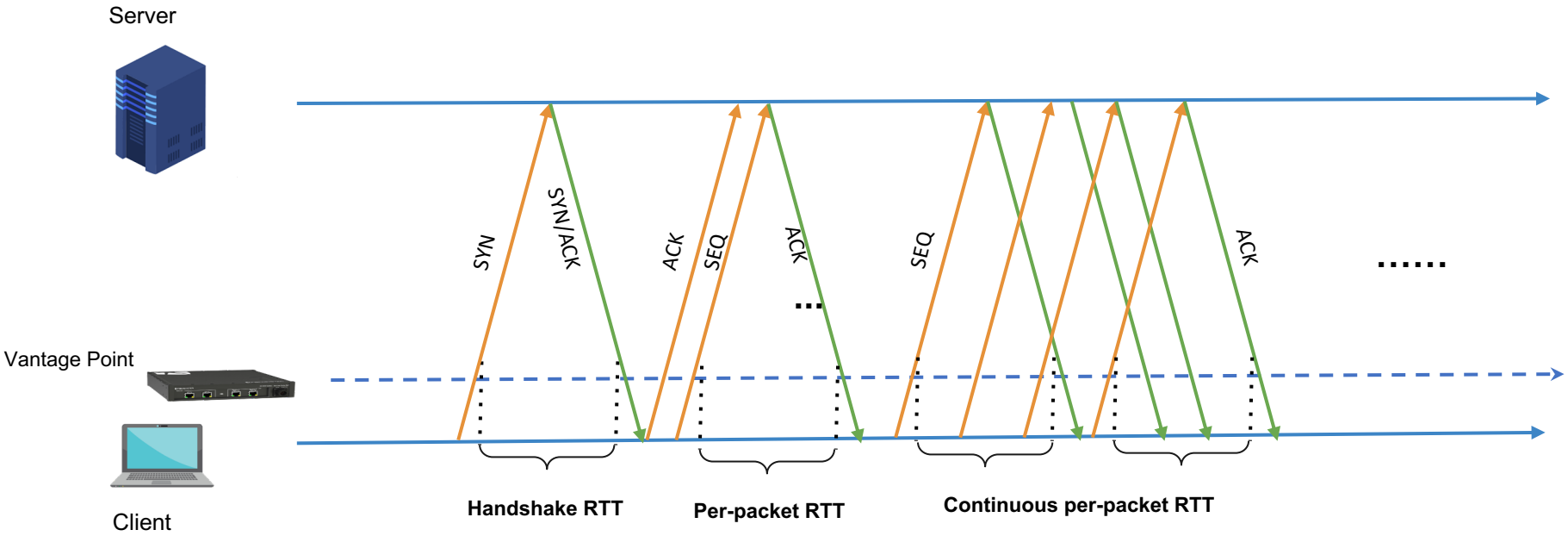
Higher-Order Data-Plane Telemetry: Names

- From network identifiers to high-level names
  - Identifiers: IP addresses, TCP/UDP port numbers, …
  - Names: domain names (e.g., *.netflix.com)
  - *Challenge*: join DNS response with subsequent data packets
- Example use cases
  - Traffic volume by site name or class
  - Intrusion Detection System bypass
  - IoT device fingerprinting
  - DNS tunneling detection

Data-Plane Challenges

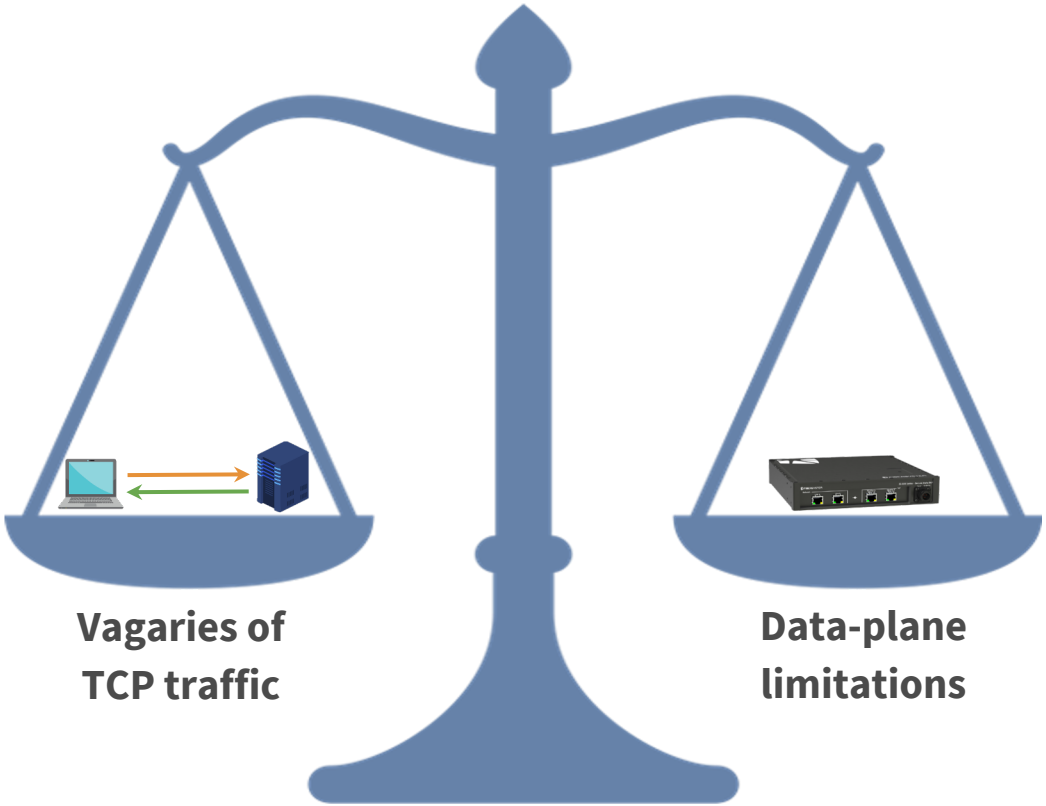


Round-Trip Time Monitoring

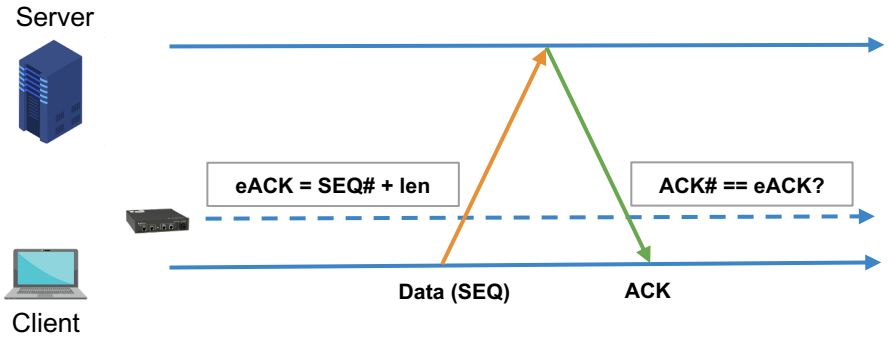


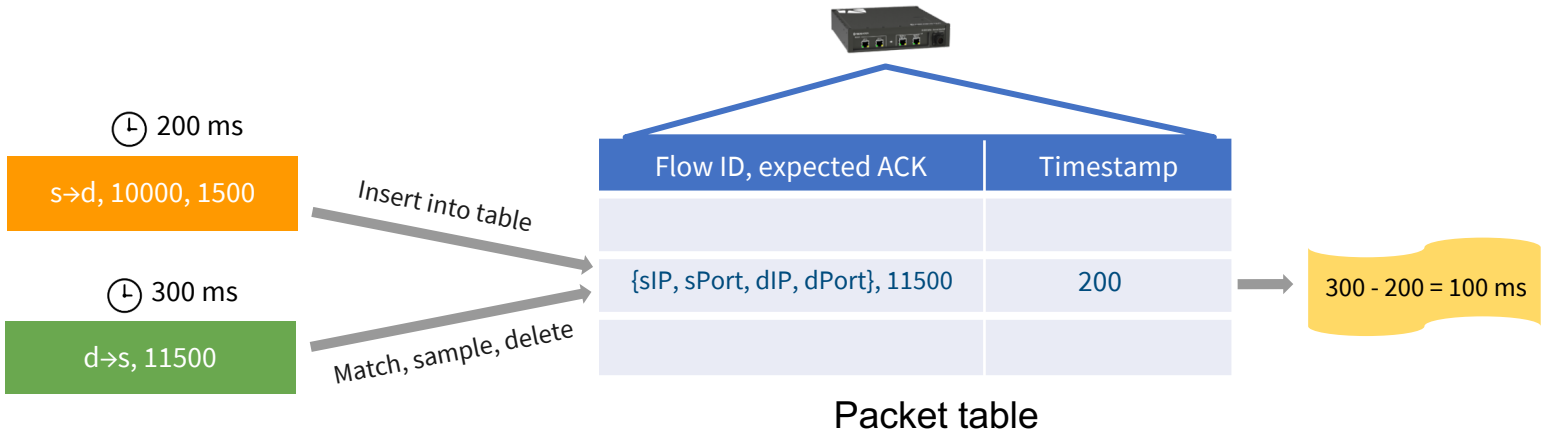
RTT Challenges: Correctness and Efficiency

- Packet retransmission
- Packet reordering
- Multiple packets in flight
- Packets never ACKed
- Long round-trip times
- SYN floods and port scans
- Many active connections
- …



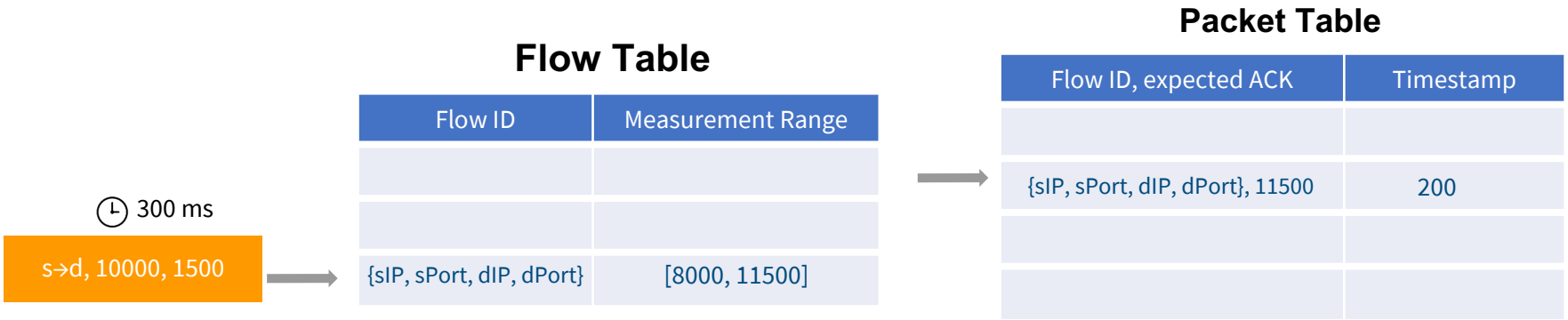
Matching a Packet With its ACK





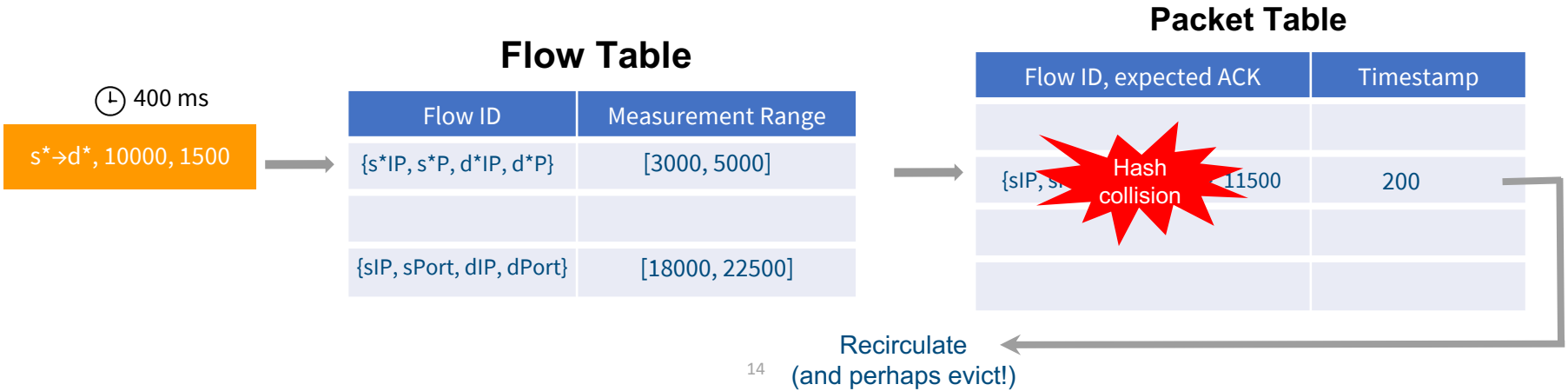
Ensuring Correctness

- Flow table
  - Track sequence number range for valid samples
  - Avoid taking inappropriate RTT measurements



Ensuring Efficiency

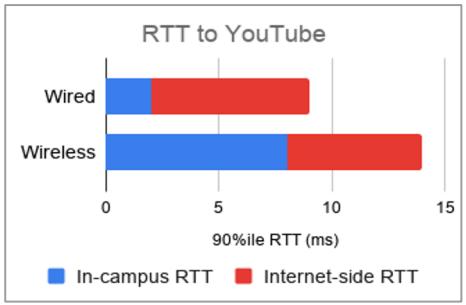
- Reducing memory pressure
  - Lazy eviction of packets that never produce RTT samples
  - Lazy eviction of inactive flows from flow table
  - No RTT samples for handshake packets (e.g., SYN flood)



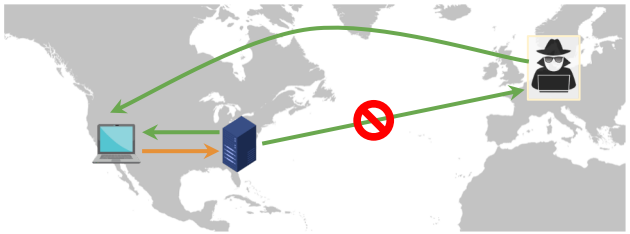
Round-Trip Time Monitoring in the Wild

- Princeton campus traffic (p4campus.cs.princeton.edu)
  - Collects vast majority of viable RTT samples (vs. tcptrace)
- Prototype in progress on the Intel Tofino switch
- Evaluating practical use cases

Campus WiFi performance

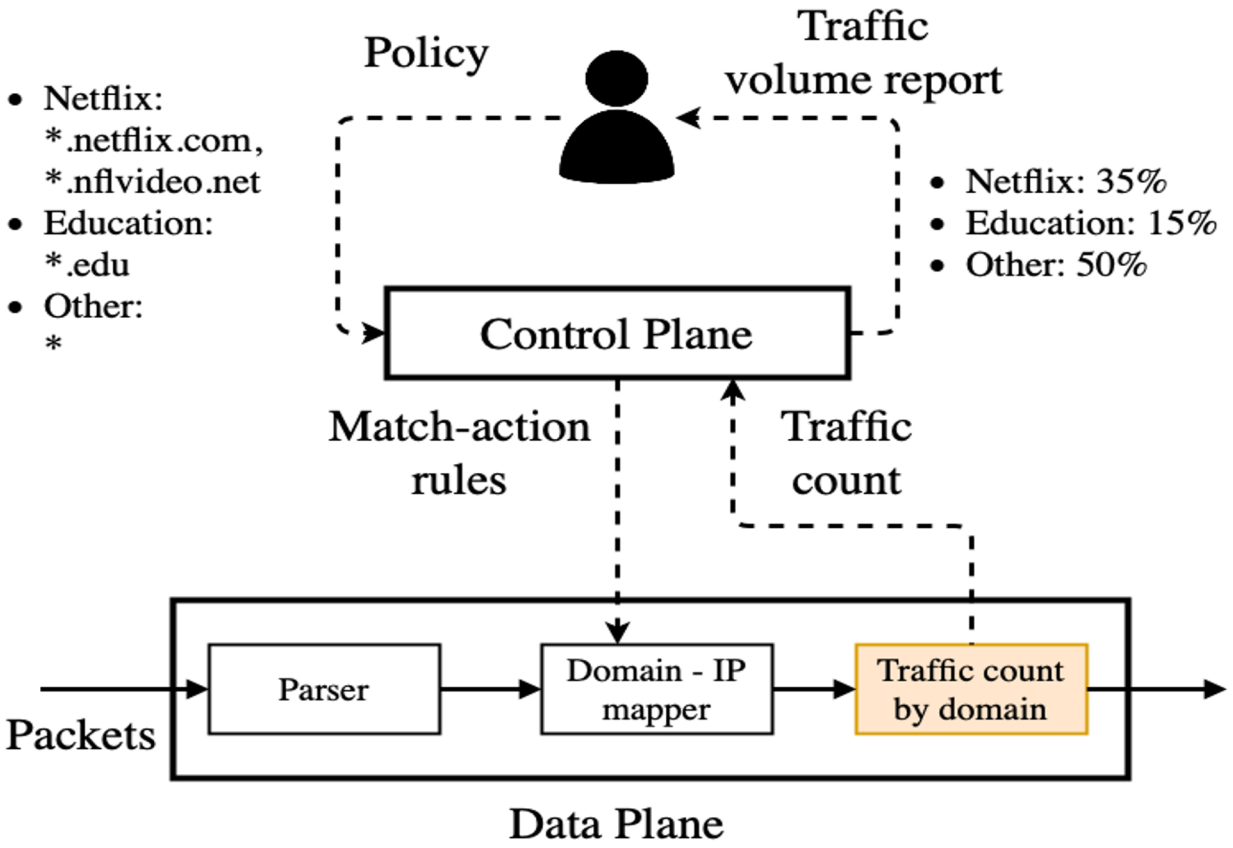


Detecting BGP interceptions

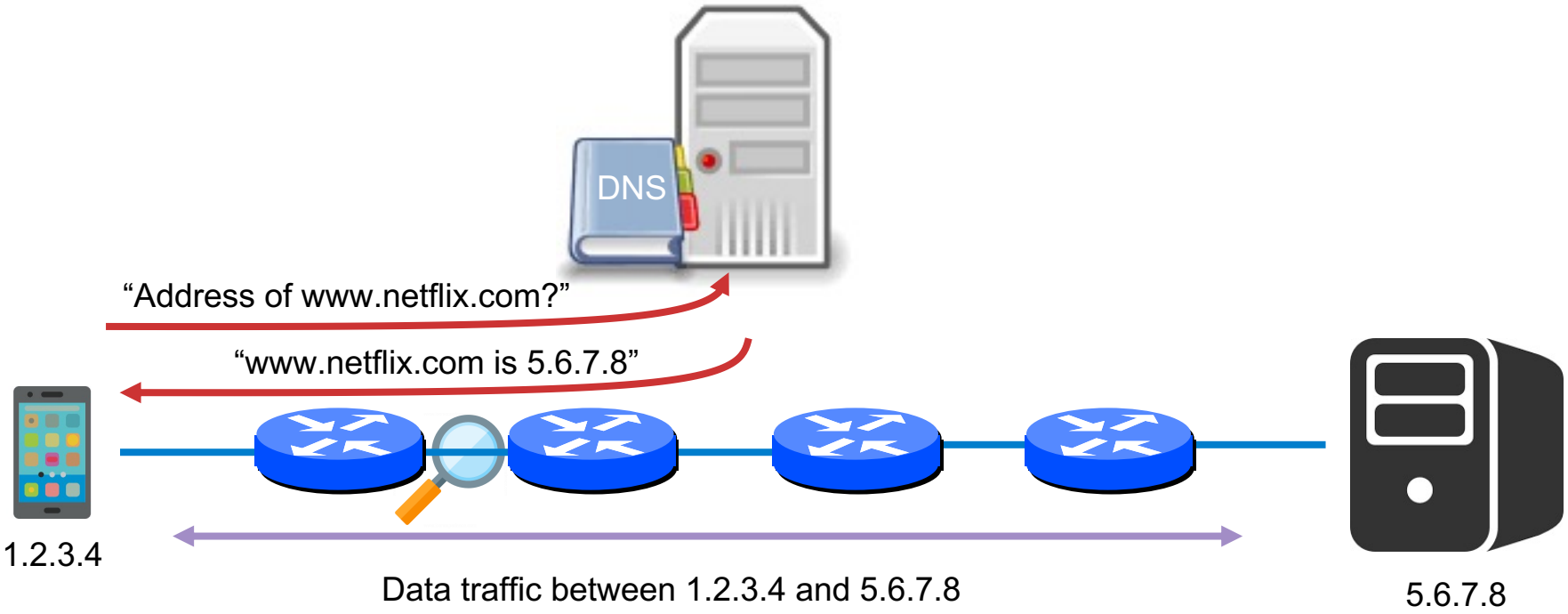


Network Telemetry by Domain Name

Network Telemetry by Domain Name

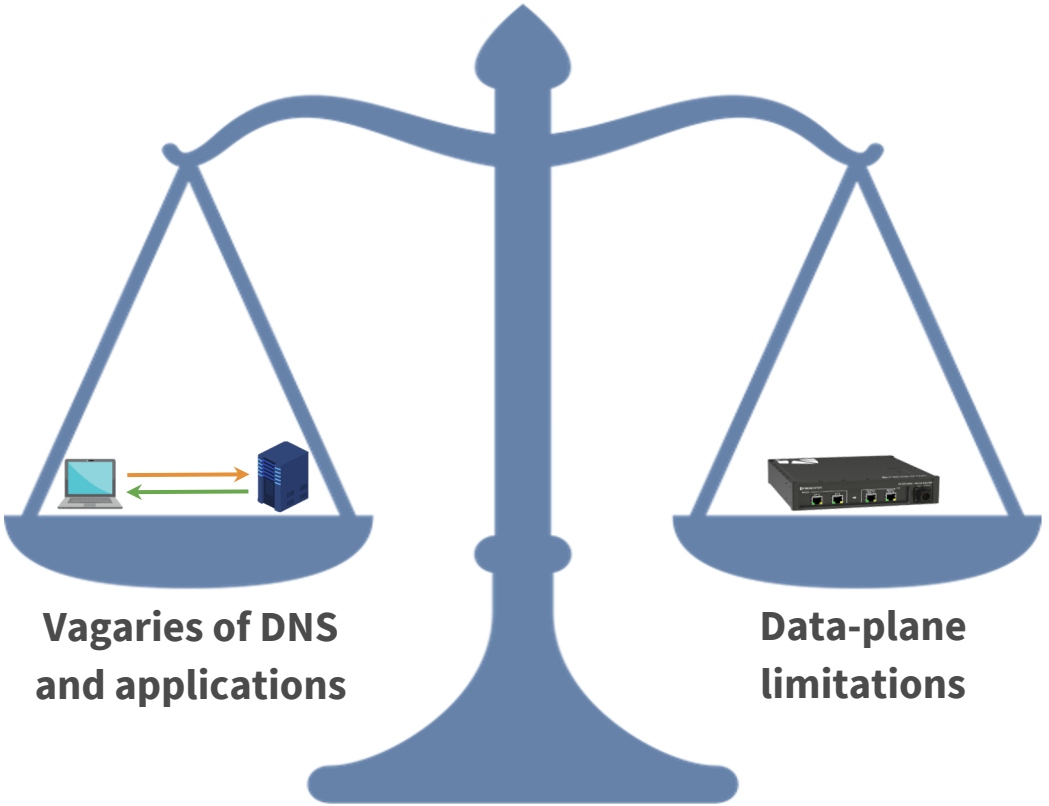


DNS Response Message and Application Traffic

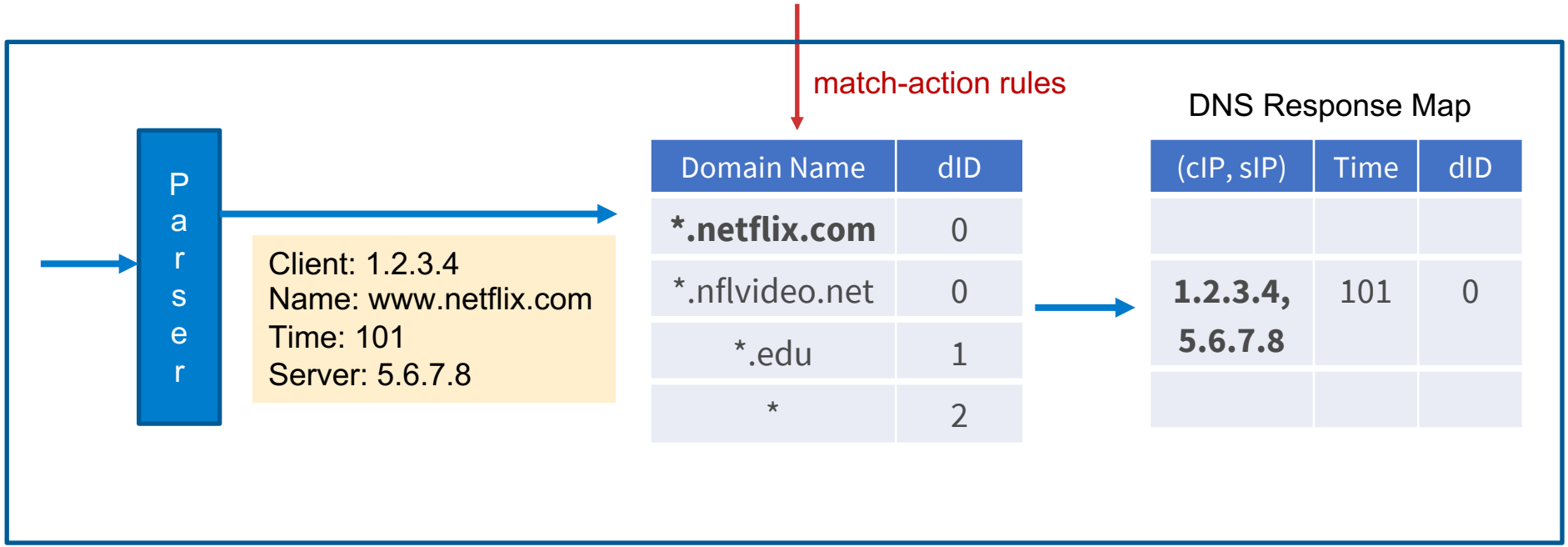


DNS Telemetry Monitoring Challenges

- Long domain names
- Variable-length names
- Browser DNS caching
- Long-lived application sessions
- Many clients and services



Storing DNS Response Data

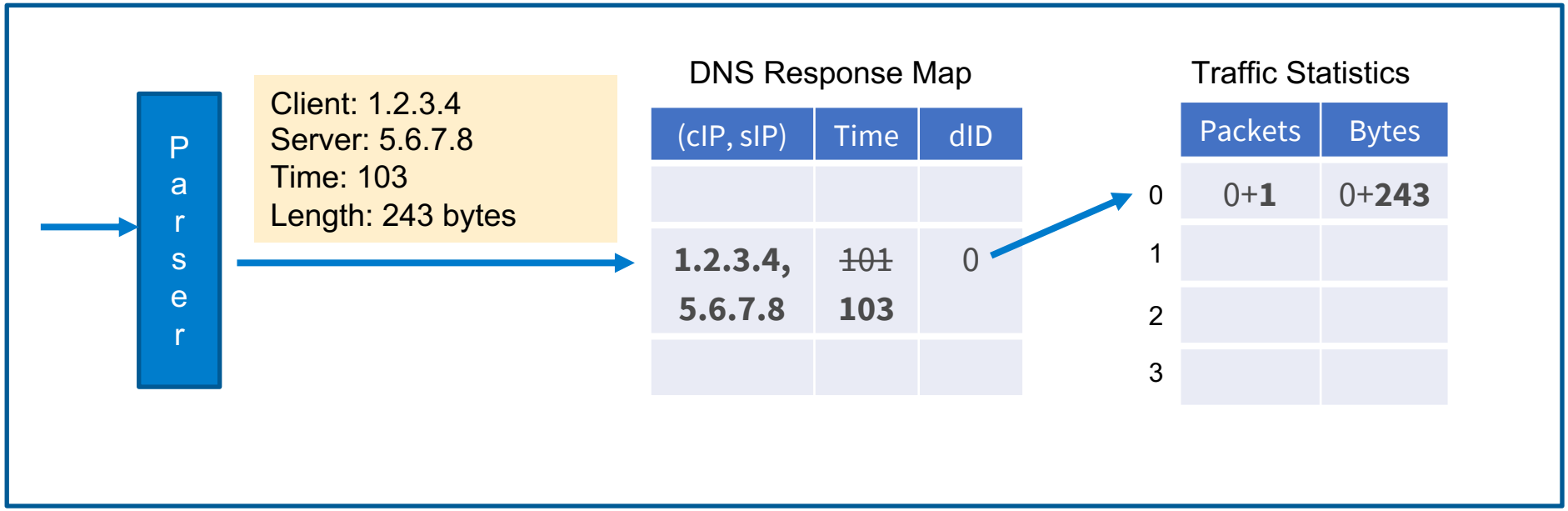


Limits on domain name parsing
<15char>.<15char>.<15char>.<15char>
(four parts, each at most 15 characters)

Domain IDs: Never store the domain names in DNS responses!

Lazy eviction of stale DNS Response data (100 sec timeout)

Counting Data Traffic by Service



Maintaining freshness!

Counting traffic by service!

Tofino Prototype and Deployment

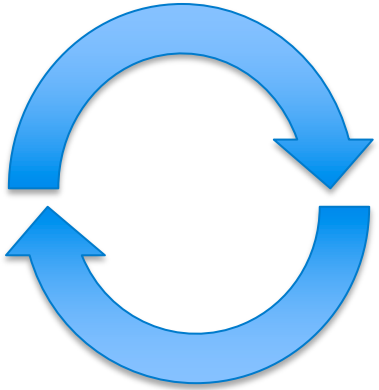
- Tofino prototype ()github.com/jkim117/Meta4)
  - Parses up to four 15-byte name labels
  - 100-second DNS response timeout
  - 2-stage, 2^{16} entry DNS response table

- Princeton deployment highlights
  - 8-11am: 33% Skype/Microsoft Teams
  - 3-3:15pm: 16% Steam Games, 12% Facebook
  - DNS tunneling and IoT device fingerprinting use cases

Conclusion

PRONTO

- Network telemetry
  - Enables the network “control loop”
- Higher-order telemetry
  - Performance (e.g., RTT)
  - Services (e.g., domain names)
- Enables a rich set of use cases
- Possible in high-speed data planes!



Learn More

X. Chen, H. Kim, J. Aman, W. Chang, M. Lee, J. Rexford, "Measuring TCP round-trip time in the data plane," *ACM SIGCOMM Workshop on Secure Programmable Network Infrastructure*, SPIN 2020.
https://p4campus.cs.princeton.edu/pubs/rtt20_paper.pdf

S. Sengupta, H. Kim, J. Rexford, "Fast and accurate passive round-trip time measurement in the data plane," May 2021.
https://p4campus.cs.princeton.edu/pubs/P4-RTT.pptx

J. Kim, H. Kim, J. Rexford, "Analyzing traffic by domain name in the data plane," May 2021.
https://p4campus.cs.princeton.edu/pubs/Meta4_sigconf-latest.pdf

Thank You!

Jennifer Rexford
http://www.cs.princeton.edu/~jrex