TANGO: Secure Collaborative Route Control across the Public Internet

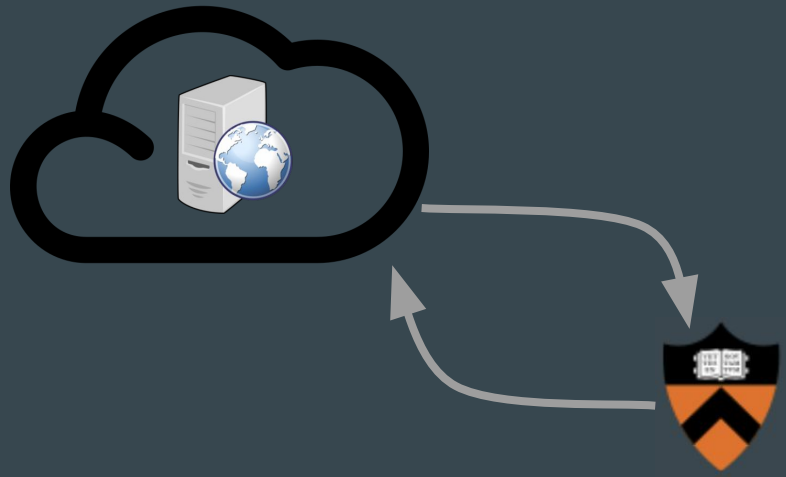
Henry Birge-Lee,
Sophia Yoo, Benjamin Herber,
Jennifer Rexford, Maria Apostolaki

PRINCETON UNIVERSITY

Is today's Internet good enough for performance-critical applications?

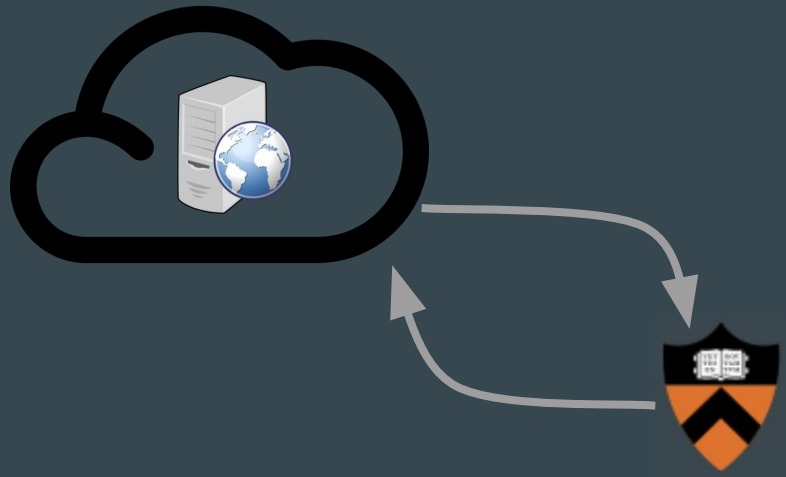
Scenario: University working with a small cloud provider

- Small cloud provider offers best price and capabilities
- Performance critical apps need to be reached from campus
- Campus and cloud communicate over the public Internet



Problem: Internet doesn't offer performance guarantees

- Latency can be too large or inconsistent
- Loss can be unacceptably high
- Reliability suffers



One Approach: Network performance is offered as a paid service!

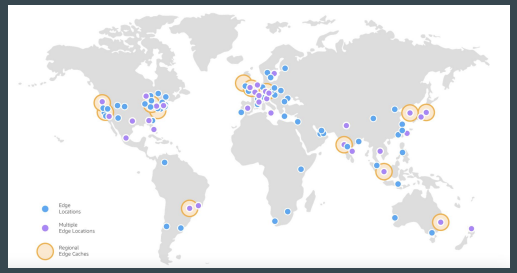
Edge computing: performance-critical services placed close to edge networks

Network-as-a-Service (NaaS): on-demand products offering reliable, reserved bandwidth point-to-point links

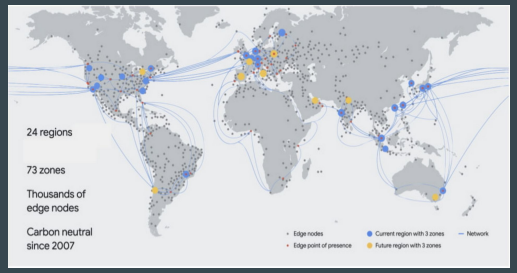
But small networks and underserved regions cannot compete

Only hypergiants can afford vast numbers of edge nodes

AWS:

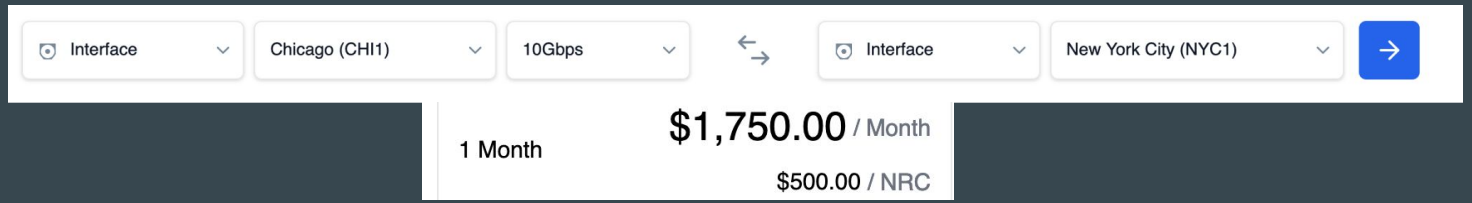


Google:



Only large organizations can afford network-as-a-service

Packet Fabric:



What would it take for the Internet to serve performance-critical apps?

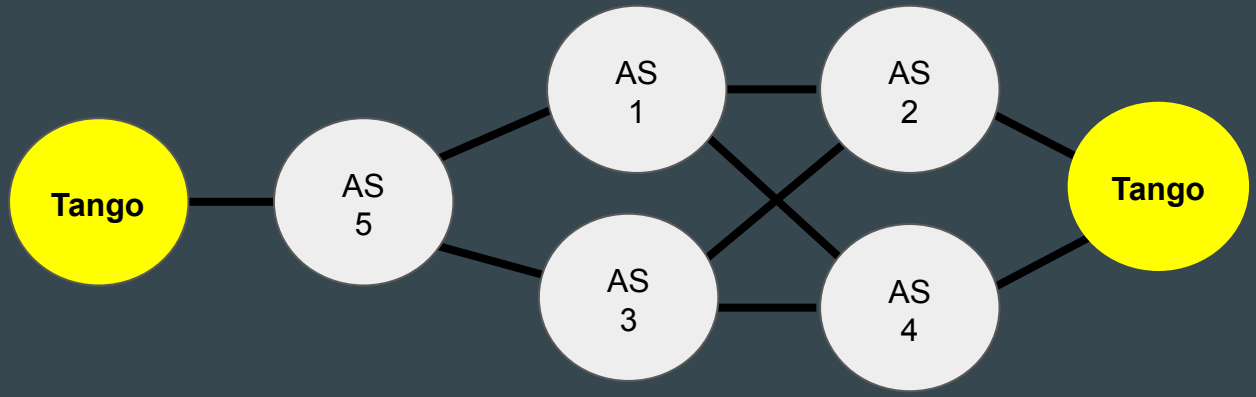
- We need to know what paths are available (Path Diversity)
- We need to accurately measure performance along those paths (Measurements)
- We need to dynamically route traffic down the best path (Route Control)

Overcoming the challenges of Internet performance with Tango

Tango

Overcoming the challenges of Internet performance with Tango

Tango

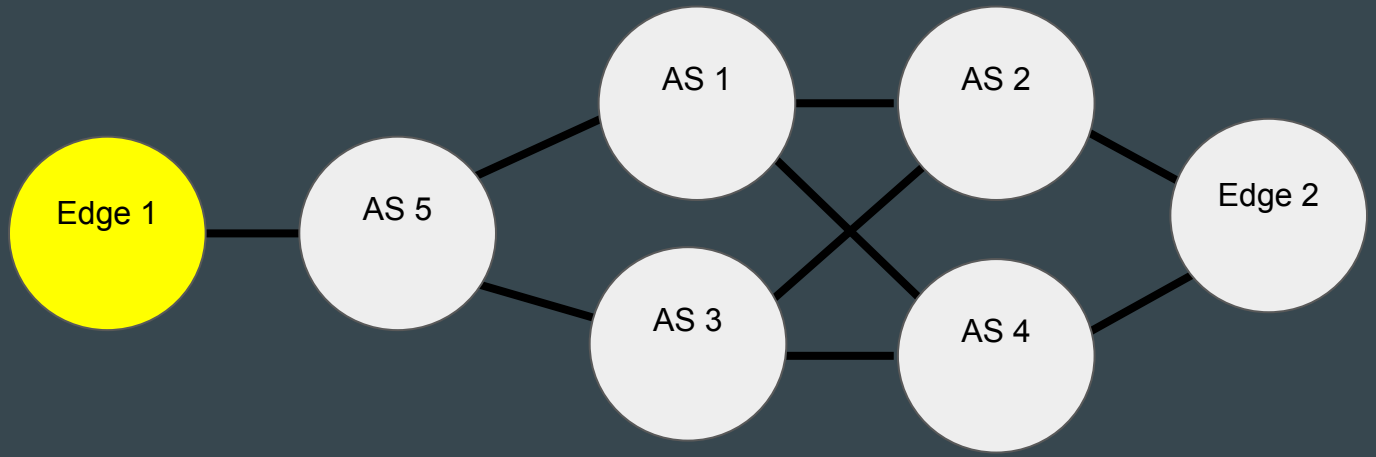


Cooperative and secure edge-to-edge routing with Tango

What would it take for the Internet to serve performance-critical apps?

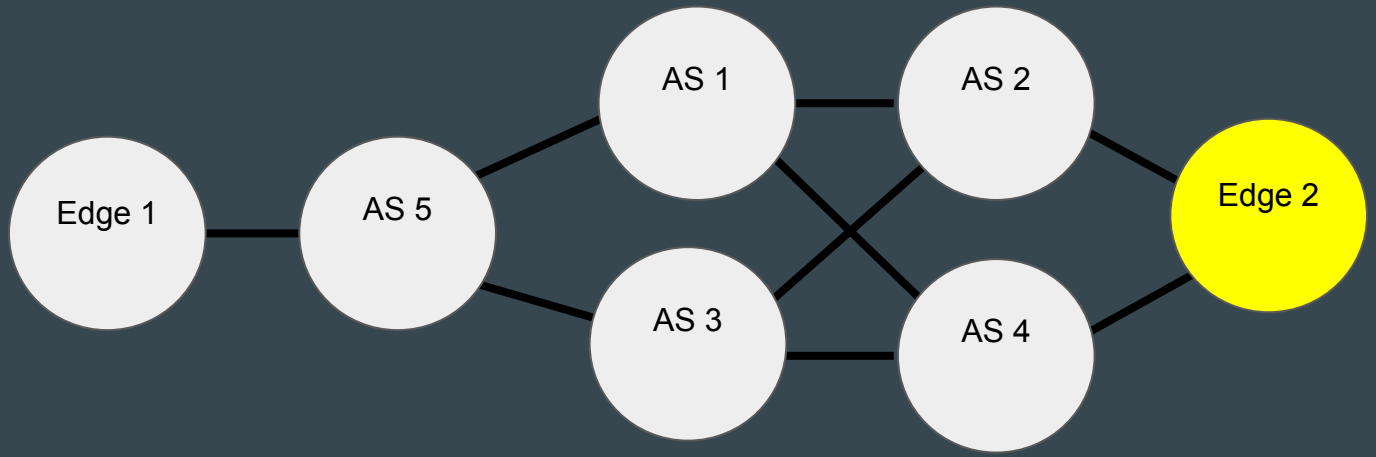
- **Path Diversity** 
- Measurements
- Route Control

Why Path Diversity is hard:
Traditional Internet Routing (BGP) does not expose all paths



- Edge 1 only has a single upstream thus a single path exported by AS 5

Why Path Diversity is hard:
Traditional Internet Routing (BGP) does not expose all paths



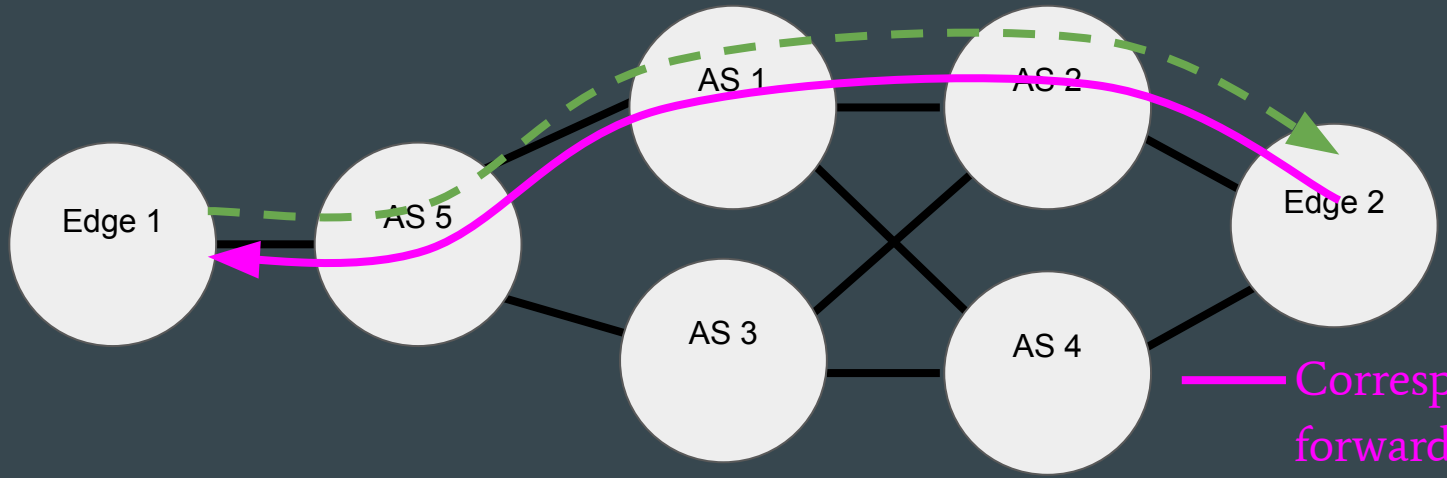
- Edge 2 is multi-homed but only influences a single hop, not the whole route

Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation
prefix: abcd:1::/48

BGP Pathfinder has no knowledge of the topology:

Begin with the default path

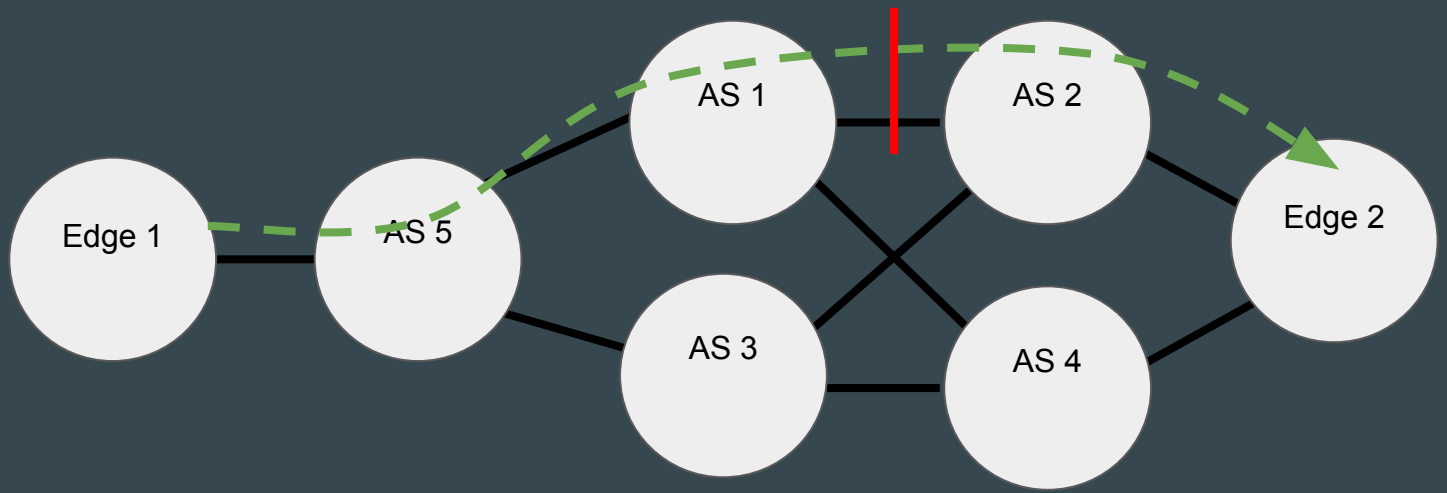


Corresponding traffic forwarding path

Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation
prefix: abcd:1::/48
communities: AS1:No_Export_to_AS2

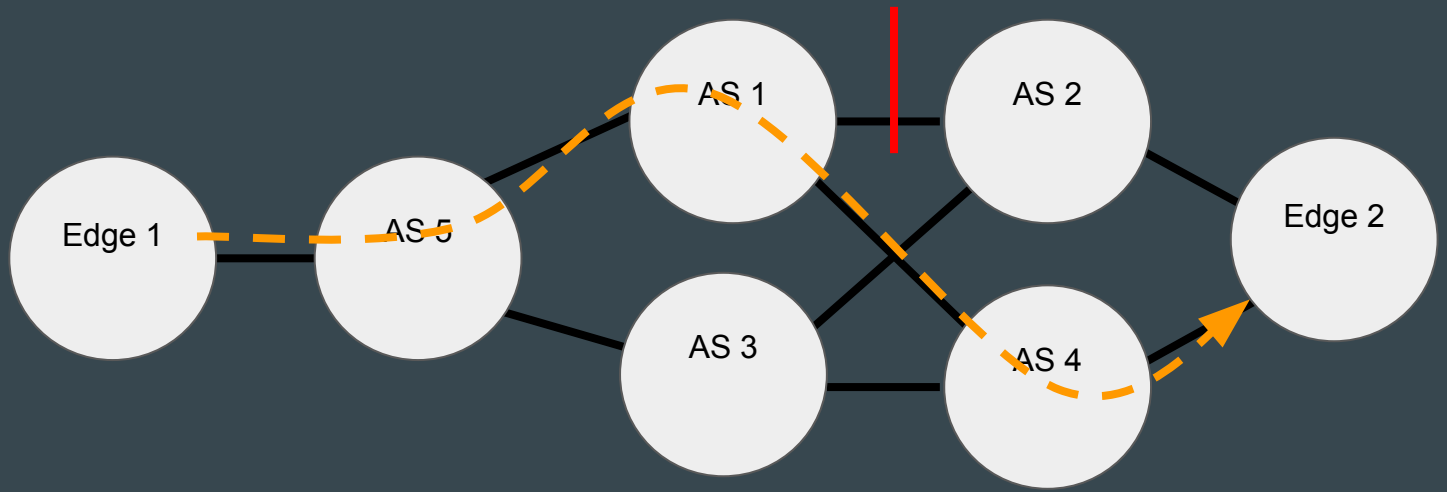
Suppress the default path



Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation
prefix: abcd:1::/48
communities: AS1:No_Export_to_AS2

Find the next path



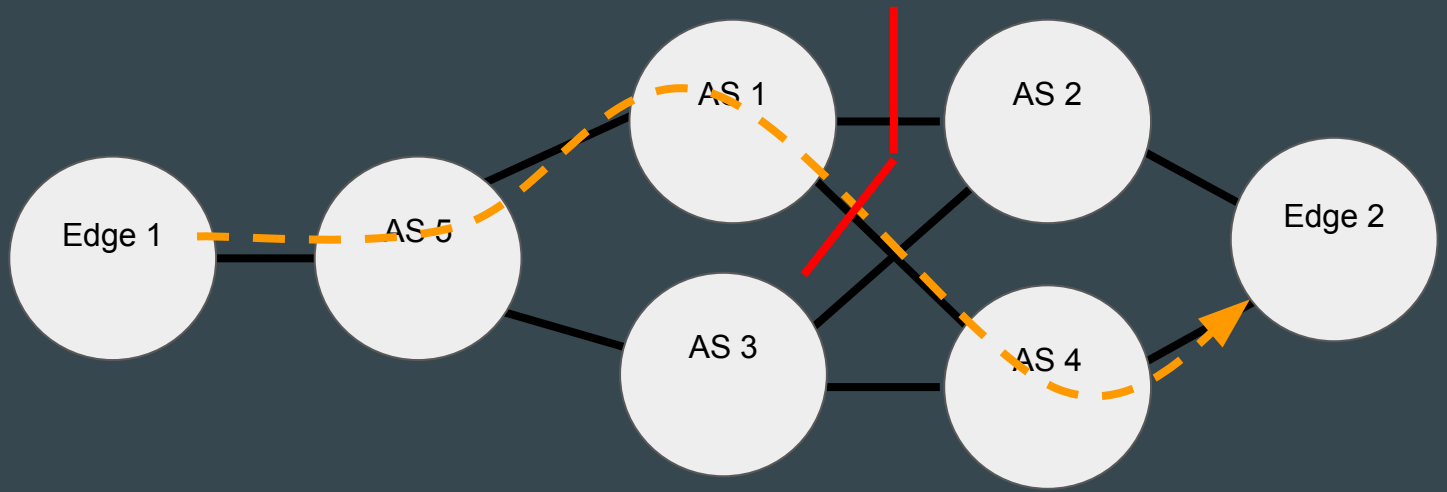
Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation

prefix: abcd:1::/48

communities: AS1:No_Export_to_AS2, AS1:No_Export_to_AS4

Suppress the new path



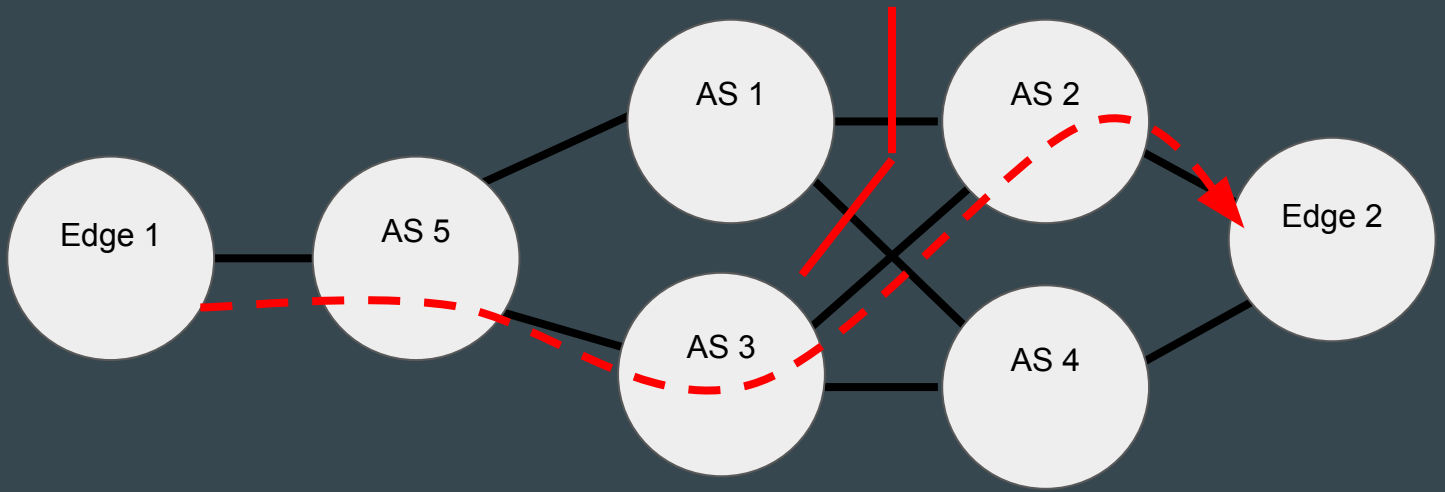
Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation

prefix: abcd:1::/48

communities: AS1:No_Export_to_AS2, AS1:No_Export_to_AS4

Find the next path



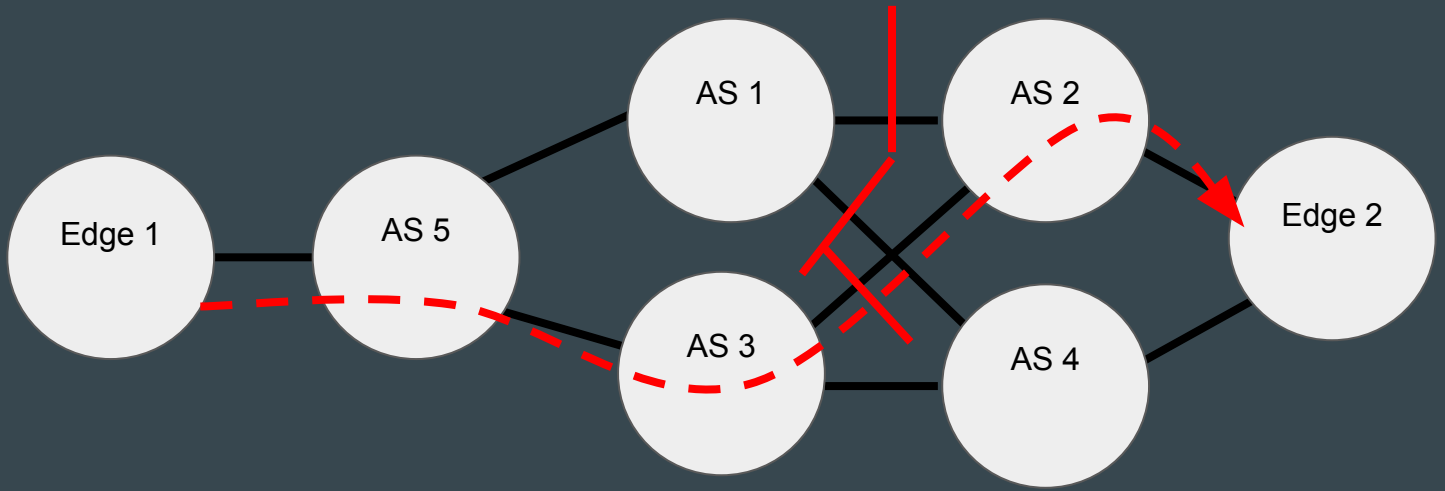
Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation

Repeat

prefix: abcd:1::/48

communities: AS1:No_Export_to_AS2, AS1:No_Export_to_AS4, AS3:No_Export_to_AS2



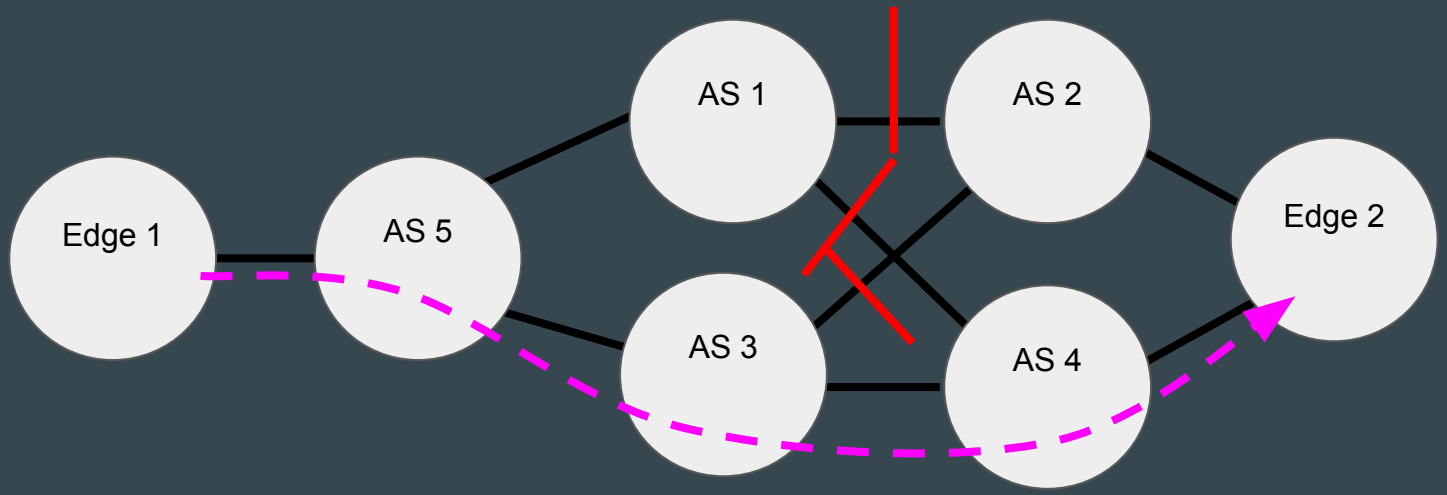
Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation

prefix: abcd:1::/48

communities: AS1:No_Export_to_AS2, AS1:No_Export_to_AS4, AS3:No_Export_to_AS2

Repeat...



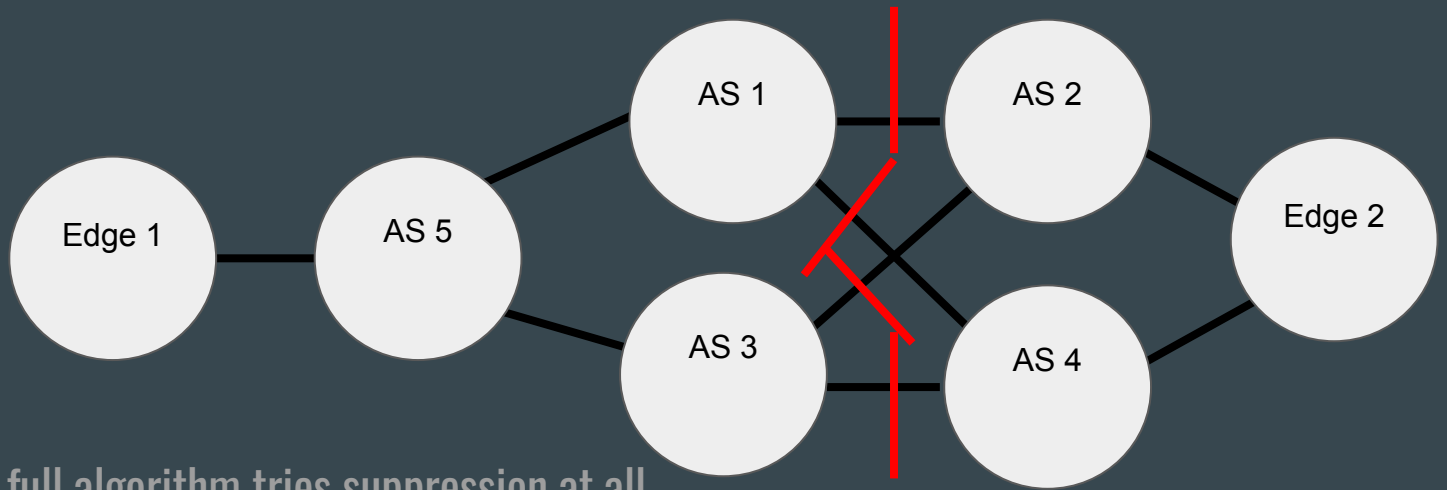
Solution: automatically exposing Path Diversity with BGP pathfinder

BGP announcement propagation

prefix: abcd:1::/48

communities: AS1:No_Export_to_AS2, AS1:No_Export_to_AS4, AS3:No_Export_to_AS2, AS3:No_Export_to_AS4

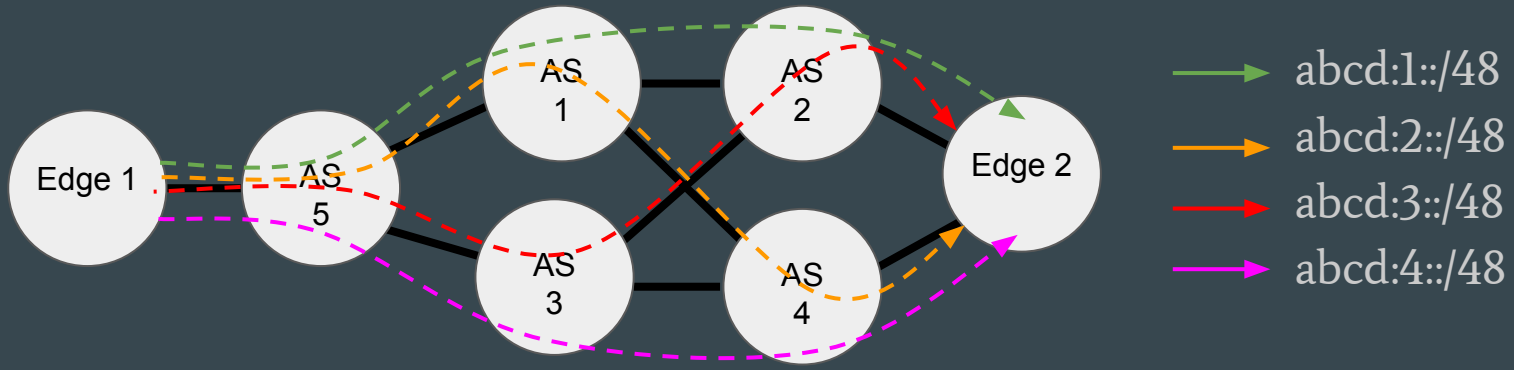
Stop when no paths remain



Note: full algorithm tries suppression at all hops along path

NO Paths Available

Different prefixes are announced along different paths

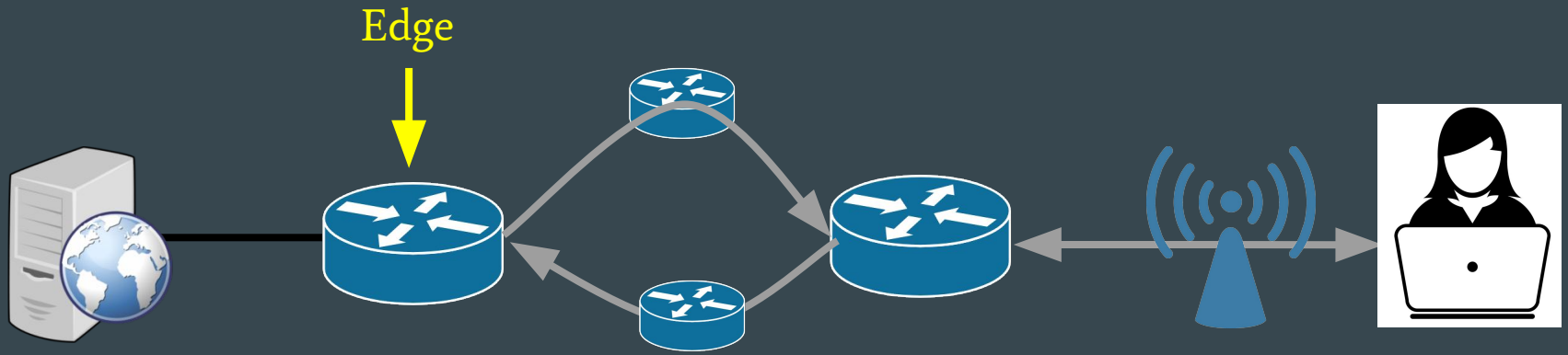


- ***Destination* Tango nodes** can announce multiple IP prefixes along different paths using BGP communities
- ***Source*** Tango node can select which path to use by selecting a prefix to reach the destination
- *BGP announcements* are stable, BGP pathfinder only needs to be rerun periodically

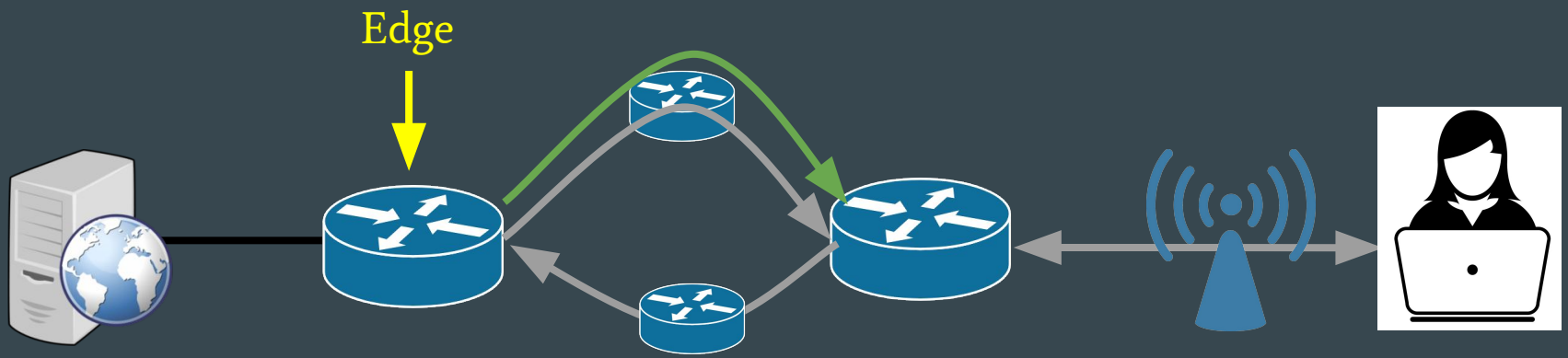
What would it take for the Internet to serve performance-critical apps?

- Path Diversity
- **Measurements** 
- Route Control

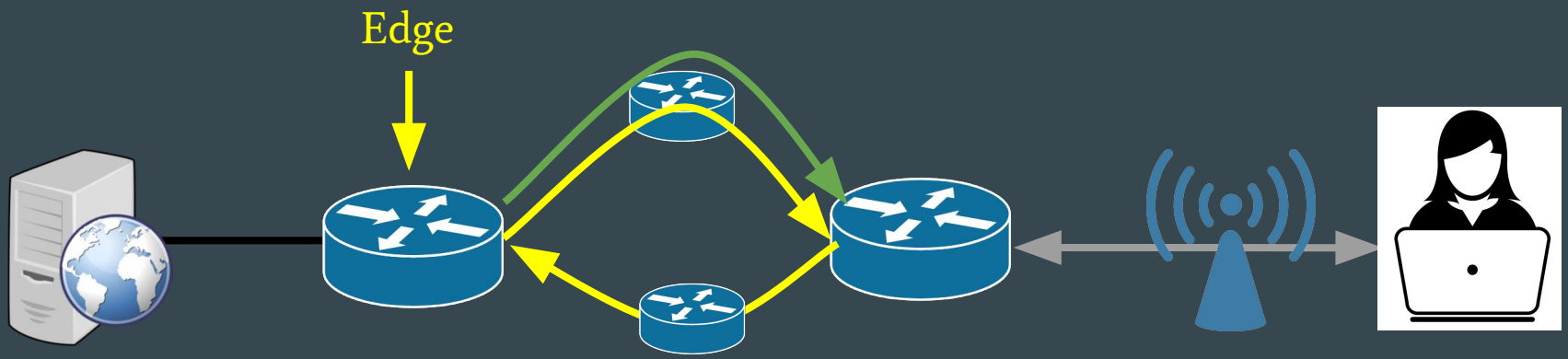
Why Measurements are hard



Why Measurements are hard

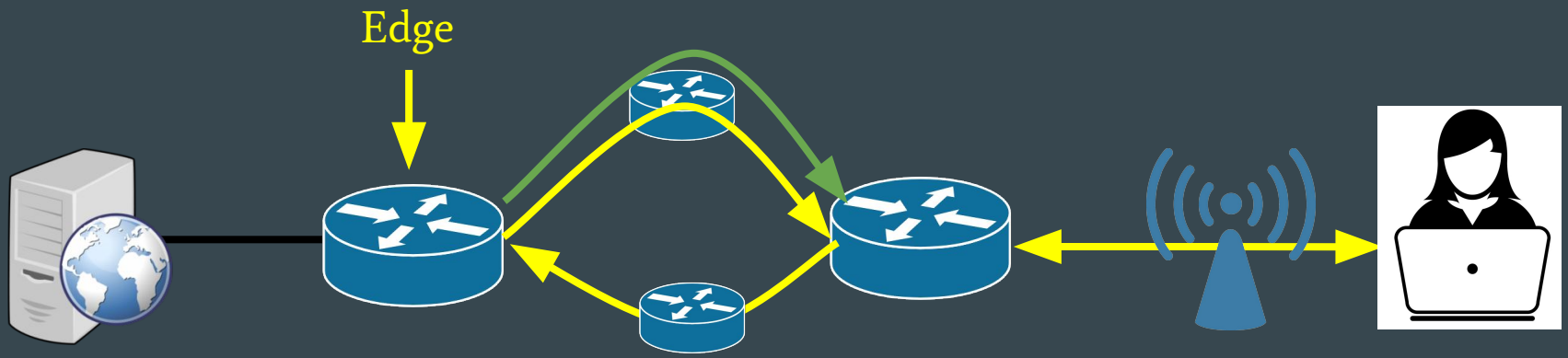


Why Measurements are hard



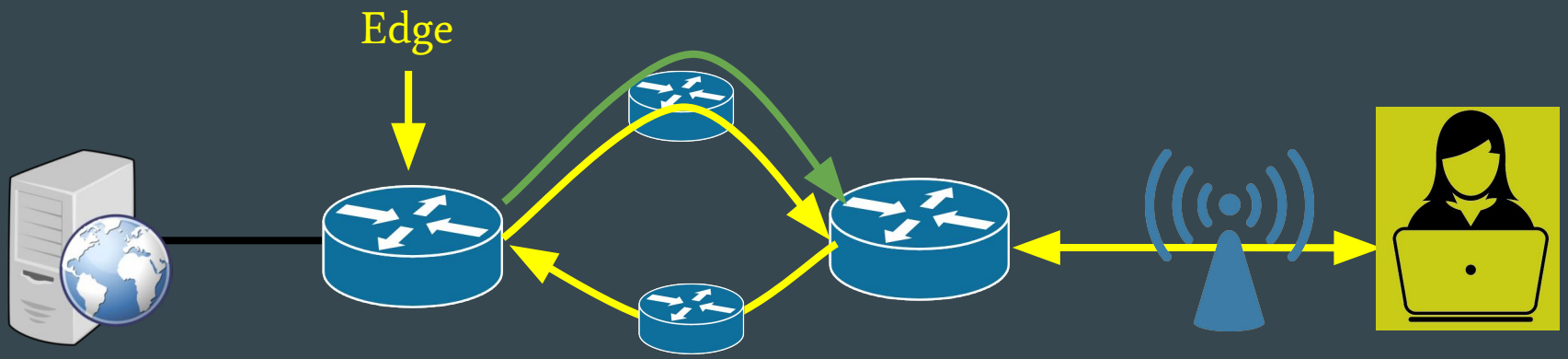
- An edge network only sees round-trip-time (RTT) not one-way-delay

Why Measurements are hard



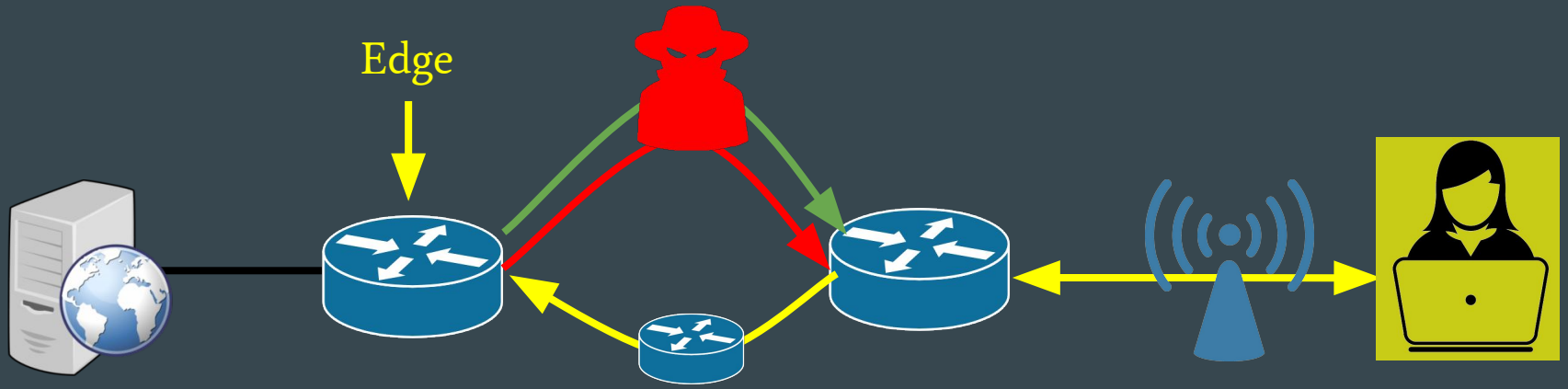
- An edge network only sees round-trip-time (RTT) not one-way-delay
- RTT includes edge network delays

Why Measurements are hard



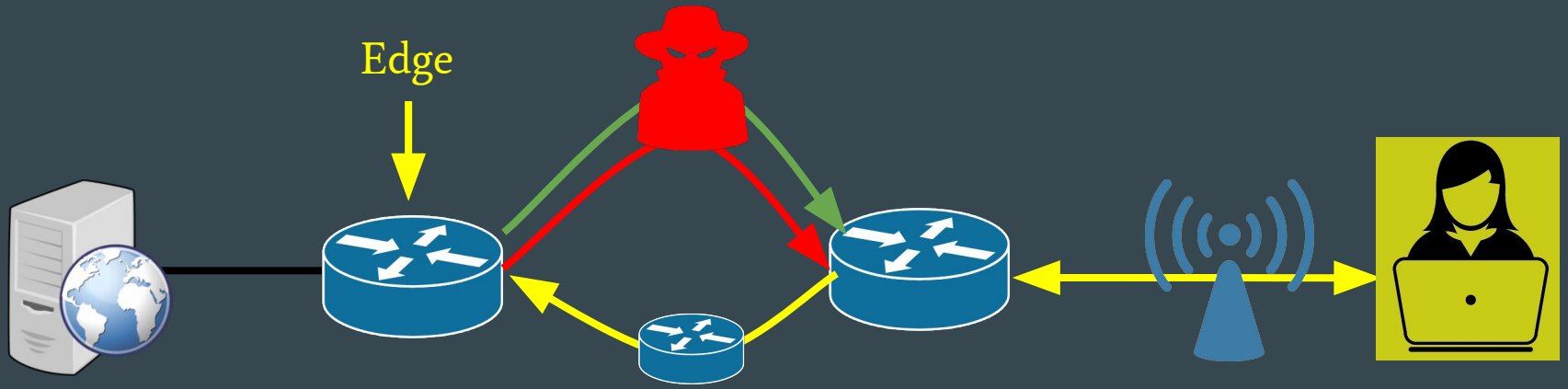
- An edge network only sees round-trip-time (RTT) not one-way-delay
- RTT includes edge network delays
- RTT monitoring requires understanding end host and L4 protocol behavior

Why Measurements are hard



- An edge network only sees round-trip-time (RTT) not one-way-delay
- RTT includes edge network delays
- RTT monitoring requires understanding end host and L4 protocol behavior
- An attacker might try to persuade us that her RTT is lower

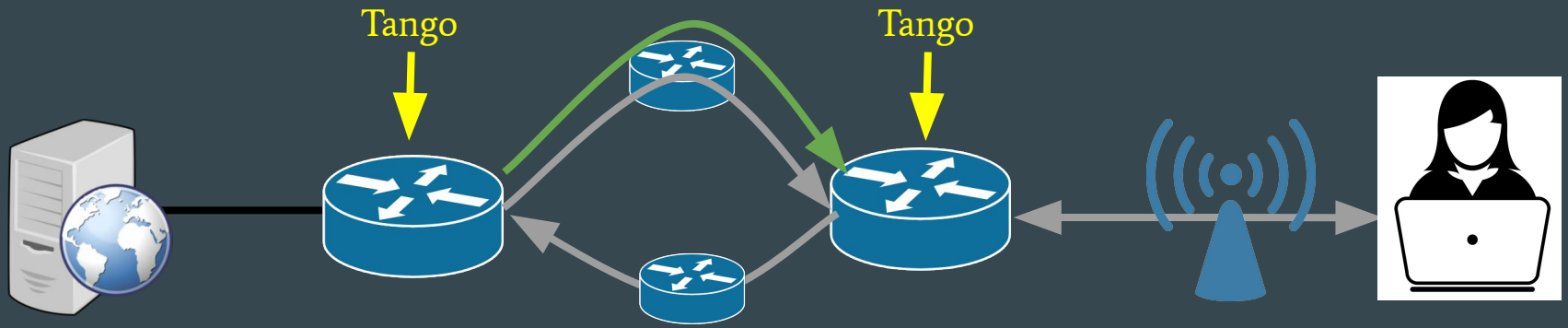
Why Measurements are hard



- An edge network only sees round-trip-time (RTT) not one-way-delay
- RTT includes edge network delays
- RTT monitoring requires understanding end host and L4 protocol behavior
- An attacker might try to persuade us that her RTT is lower
- Has to run at line rate

Improved Measurements through cooperation

- Custom Tango header added at the edge of the network using programmable switches or eBPF
- Measurements only include one-way, wide-area component
- Measurements do not rely on application behavior



Protecting measurements from adversaries

- An adversary could manipulate measurements to hijack traffic by making her route look more preferable

How to protect billions of sequence numbers a second?

- Protected packet fields:
  - Sequence numbers: protects against adversary hiding lost packets
  - Timestamps: protects against adversary manipulating latency
  - Route control messages: protects against adversary forging ctrl messages

Signing sequence numbers

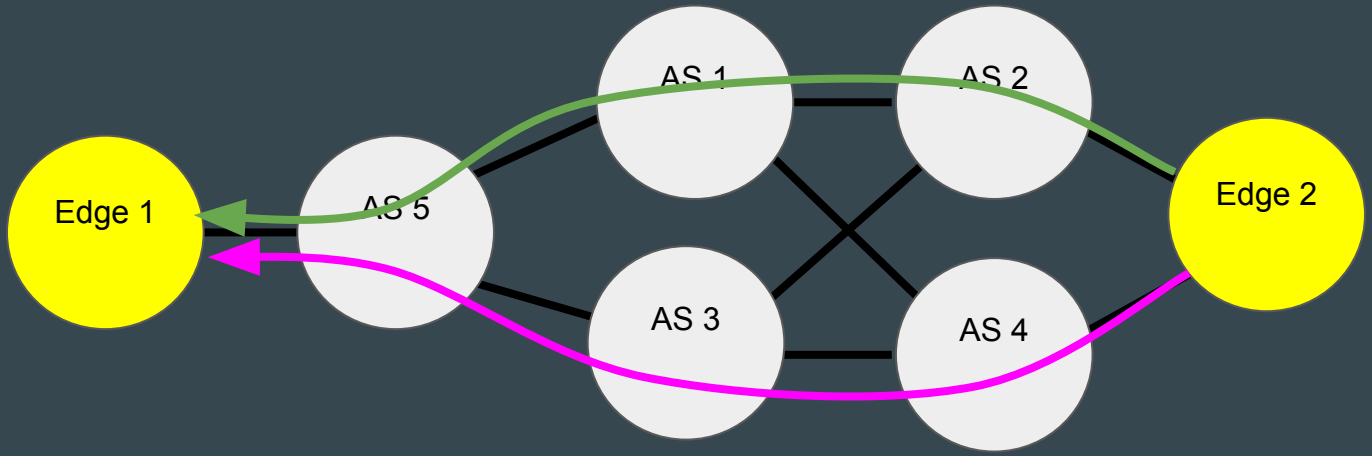
- Each sequence number **corresponds to an index** in a cipher book shared by the sender and receiver
- Packets contain the sequence number and **a single bit "signature"** from the corresponding book index to keep up with throughput
- Adversary has a 50/50 chance of guessing a signature, but **needs to guess many signatures** to meaningfully affect loss
- **Likelihood of avoiding detection decreases exponentially** with each guess

What would it take for the Internet to serve performance-critical apps?

- Path Diversity
- Measurements
- **Route Control** 

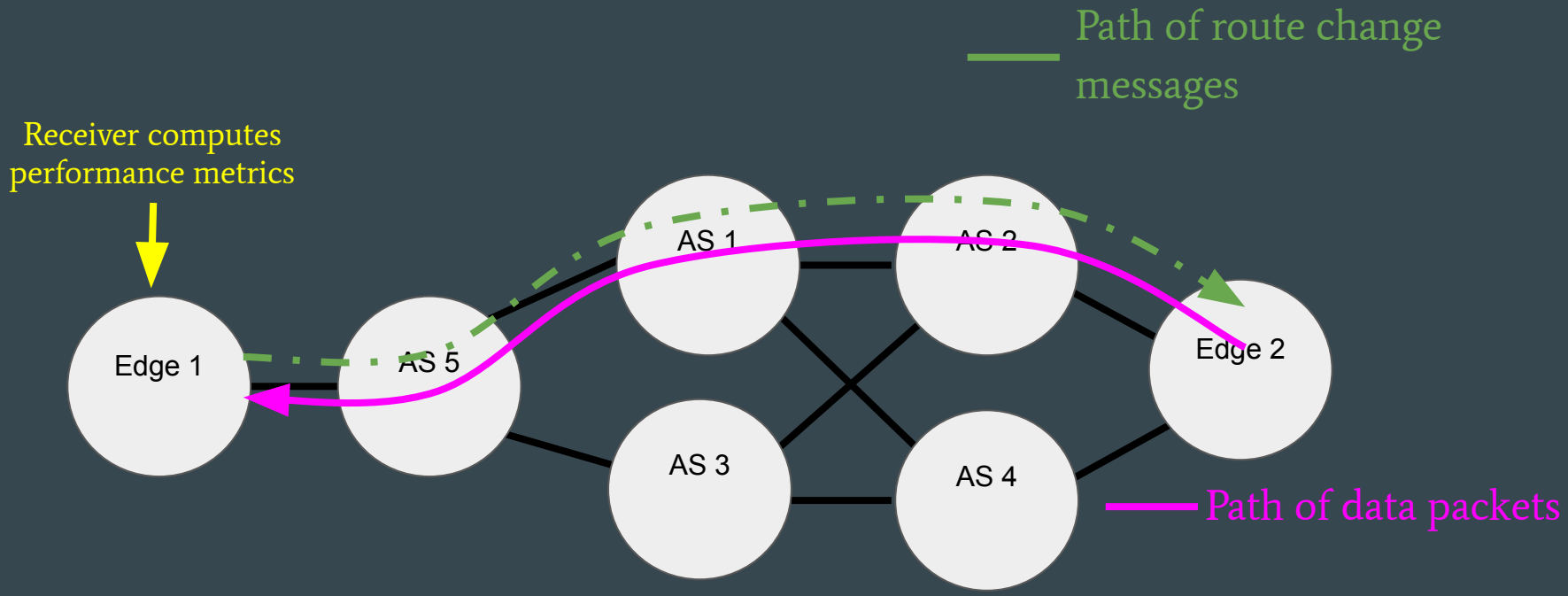
Challenges With Route Control

Consider traffic from Edge 2 to Edge 1



- Edge 1 sees one-way-delay data
- Edge 2 needs to know how to route packets

Real-time route control with Tango

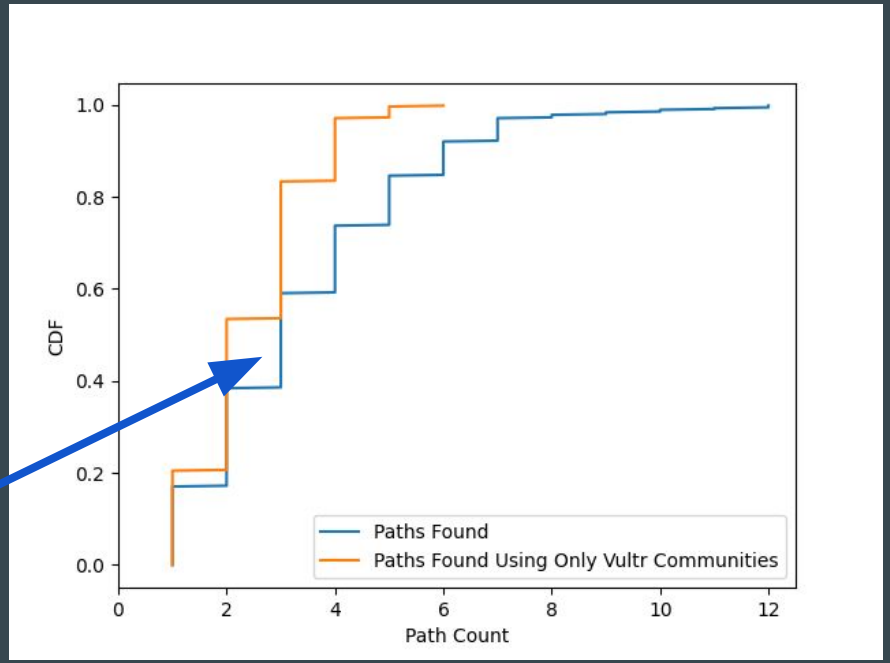


Evaluation

- **How many paths can we find?** 
- Can we beat the performance of the default path?
- Can we run measurements and crypto at line rate?

How many paths can we find?

- Ran BGP pathfinder between 503 globally-distributed nodes from the cloud provider Vultr
- Took ~30min per pair (can be parallelized)
- By default Vultr only exported a single path
- 80% of nodes had additional paths
- **BGP Pathfinder can expose 3 paths for the median pair**
- Some nodes had as high as 10-12 paths



Evaluation

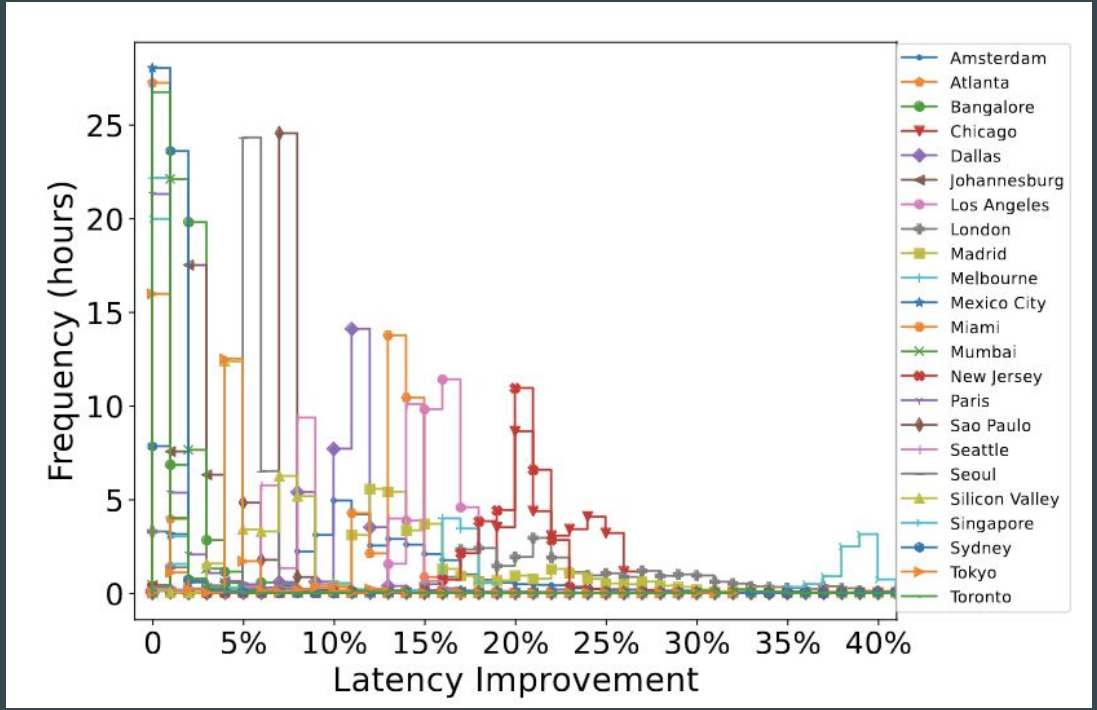
- How many paths can we find?
- **Can we beat the performance of the default path?** 
- Can we run measurements and crypto at line rate?

Tango performance measurements

- Took measurements from 25 global nodes
- Routed traffic over different paths to two destinations: LA and Stockholm
- Took latency and loss measurement every 10ms

Outperforming the Default Path

- For many src,dst pairs, the optimal Tango path had 22% lower latency than BGP default path
- In some cases, Tango saw a 39% improvement

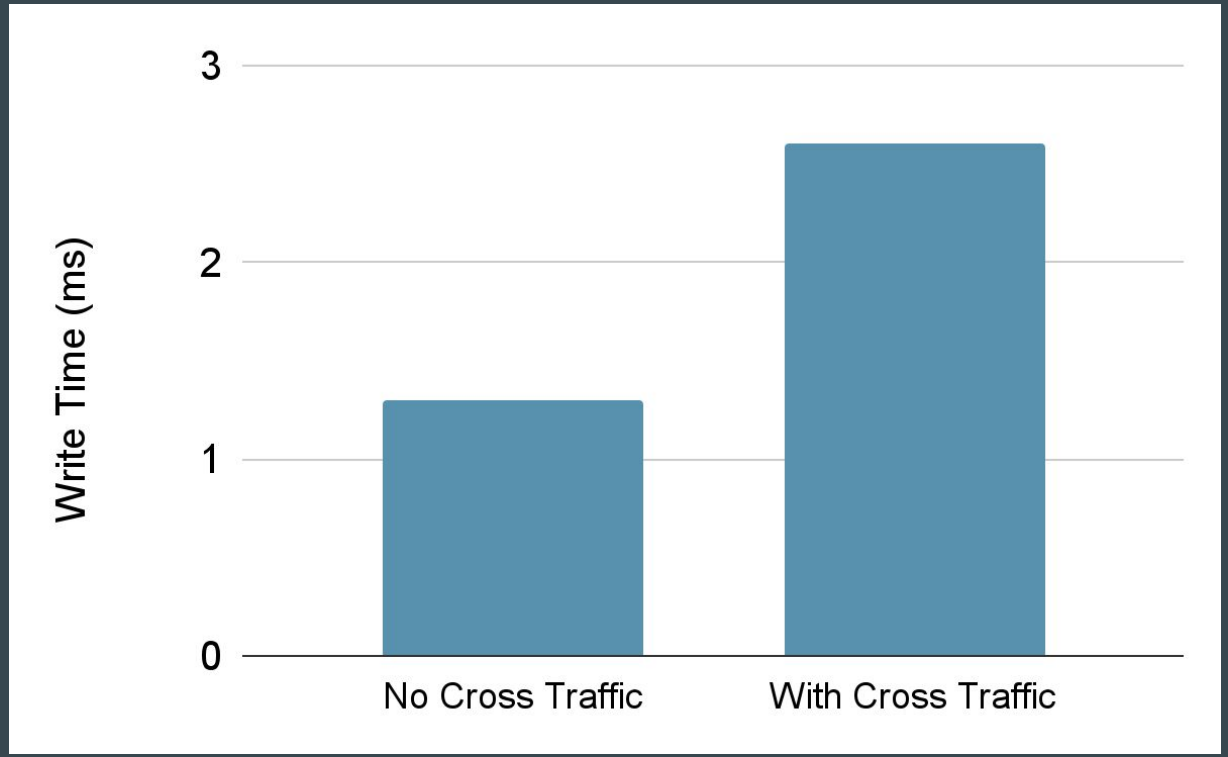


Evaluation

- How many paths can we find?
- Can we beat the performance of the default path?
- **Can we run measurements and crypto at line rate?** 

Generating sequence number signatures at line rate

- Implemented on Tofino1 Programmable Switch
- Signature values sent to switch in dataplane
- Switch recirculated packets and wrote signatures to data-plane registers
- Wrote 2^{20} signatures in 2.6ms even with cross traffic
- Keeps up with 100Gbps line rate



Conclusion

Tango

3 surprising finds from Tango

- We can find alternative paths through the public Internet
- These paths often have improved performance
- We can run trustworthy telemetry in the data plane

Questions?

Thank you for your time

Henry Birge-Lee

birgelee@princeton.edu

Offering Dynamic Route Control

- Several high-loss and high-latency events plague networks periodically
- Dynamically-switching to better paths can evade these events
- Often other unaffected paths exist