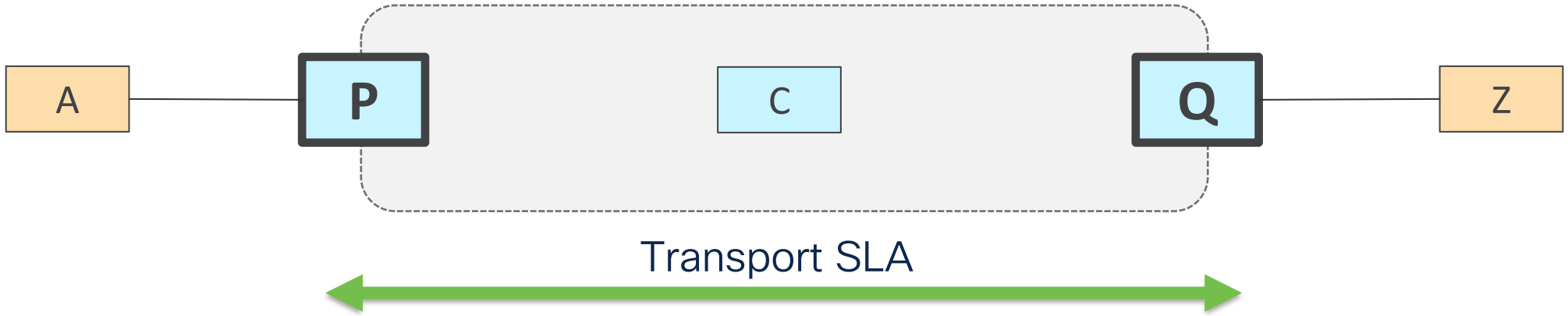
IP Measurements

Bart Van de Velde

Sr. Director, Engineering

Networking CTO Office

Problem statement



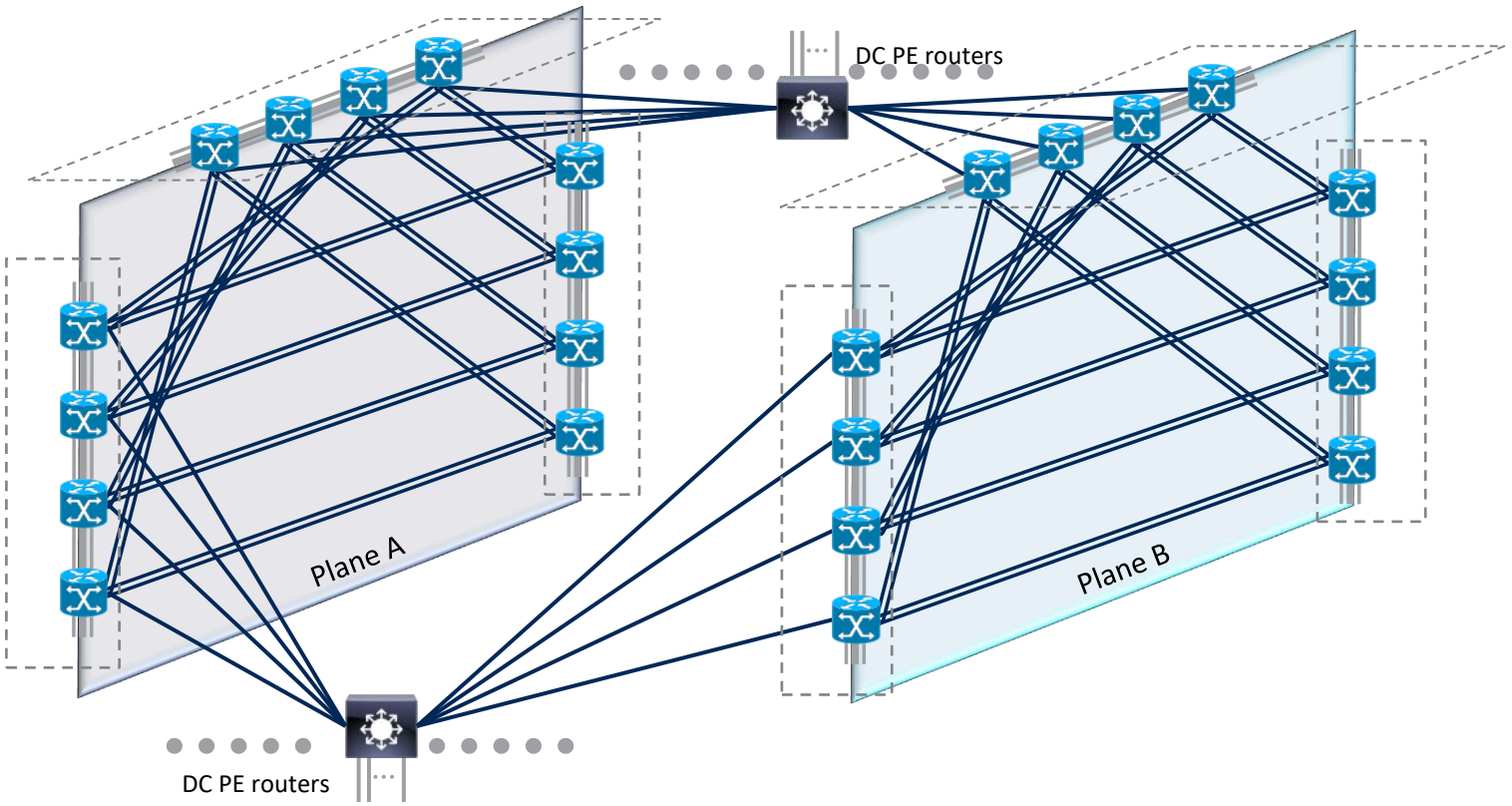
- The transport SLA is defined by three key metrics: **Latency, Loss, and Liveness**.
- A key challenge in IP networks is the inherent ECMP nature of the fabric.
- Measuring transport SLAs becomes increasingly complex at scale:
  - 500 edge nodes
  - 16 ECMP paths
  - 1 probe every millisecond

Agenda

- The Problem Statement & Proposed Solution for measuring transport SLA:
  - Scale & HW integration
  - Metrics
  - Standard probe format
  - Cost (CAPEX & OPEX)
  - Routing Correlation
- Example Operators' use-case
  - Goldman Sachs: Hardware Integrated Performance Measurements
  - Colt: Routing Correlated Analytics
- We welcome your feedback
  - The problem statement and the proposed solution
  - Facing the same issues/challenges

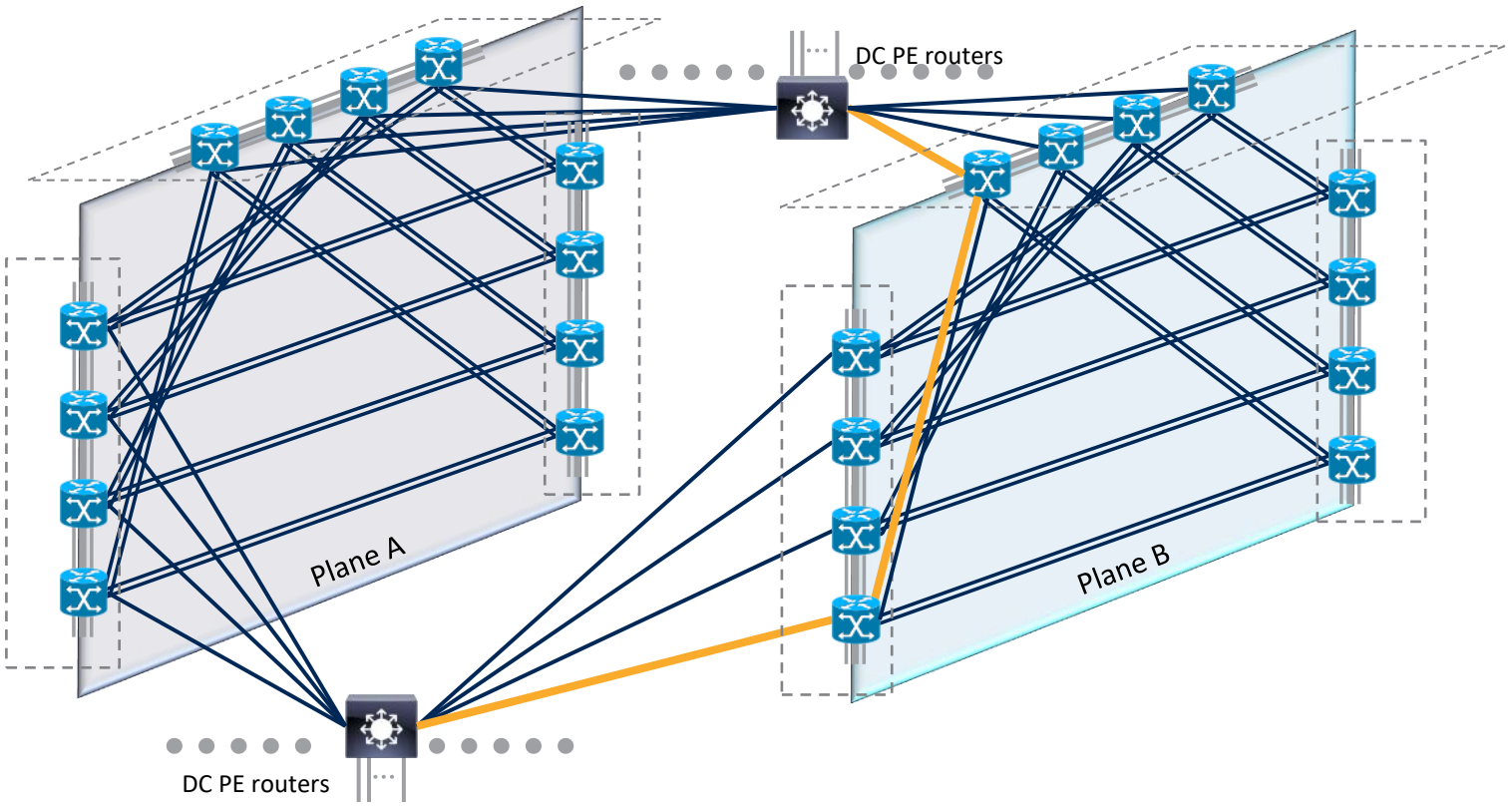
Scale

The nature of IP is ECMP



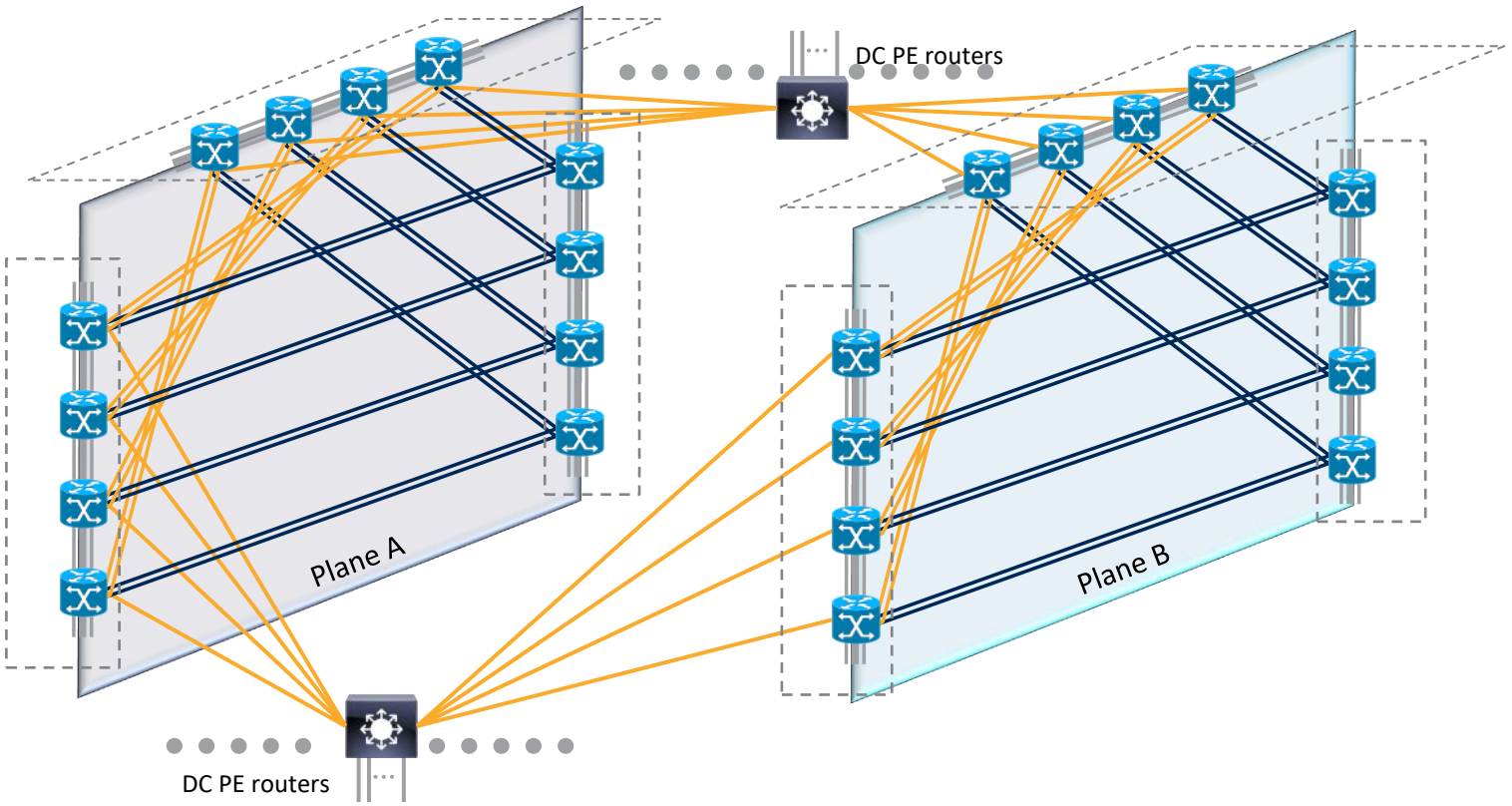
- Simple topologies already exhibit 16 ECMP paths
- The number of ECMP paths grows significantly with larger topologies

Probing neglects the ECMP reality



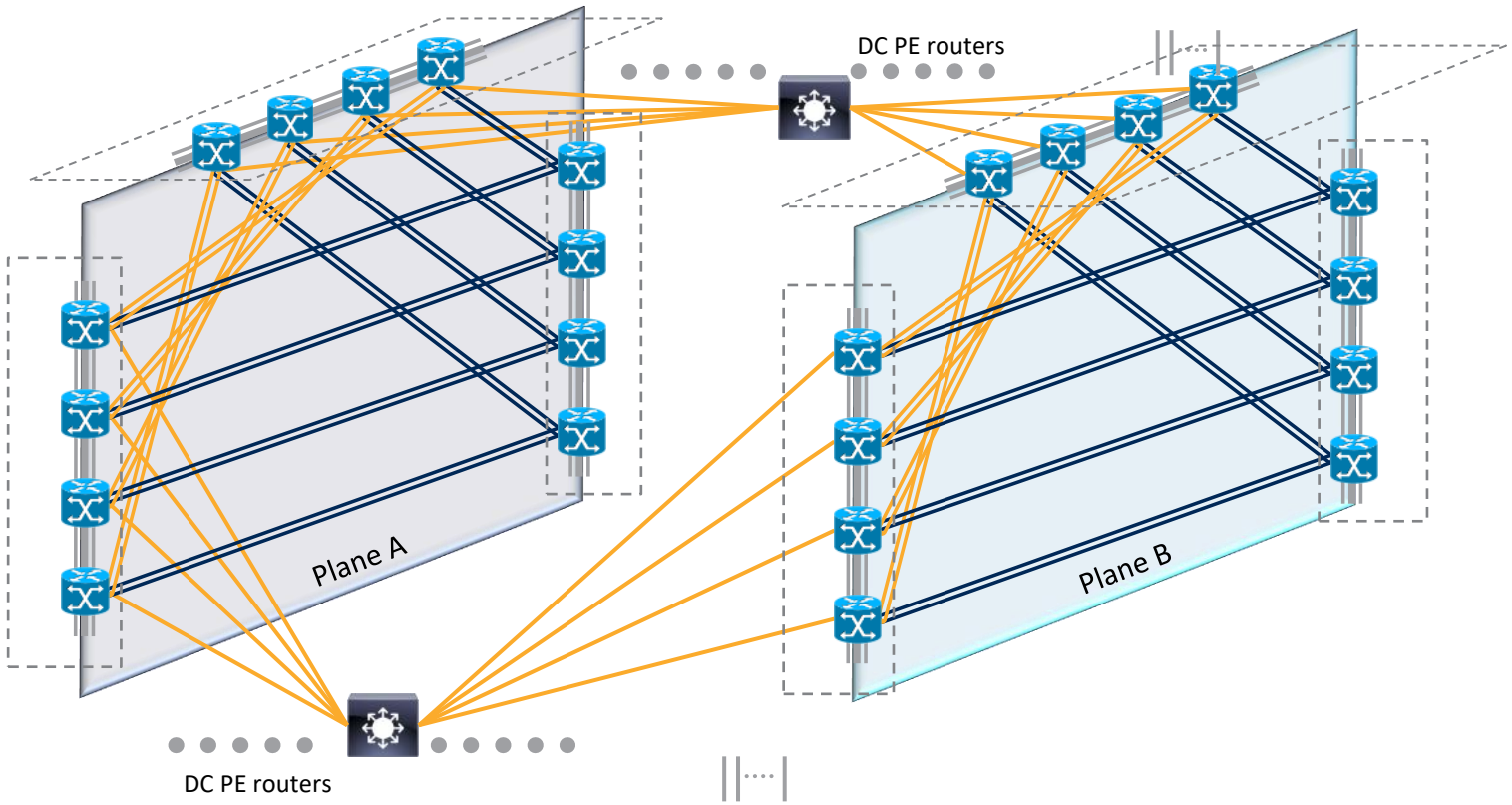
- Lack of scale to measure all ECMP paths
- Simple probe from Source to Destination.
  - A single 5-tuple. Whatever hashing may be.

A measurement solution MUST:



- Discover and monitor all ECMP paths:
- Provide enough Probes Per Second (PPS) to measure all ECMP paths
- Report accurately across ECMP paths

How many probes per second are needed?



- 500 edges
- 16 ECMP paths
- 1 measurement every msec

8M probes per sec

Problem statement

- Relying on router (or host) CPU for Probe generation/processing:
  - Few thousands of probes per second
  - Lack of scale to measure all ECMP Paths
  - typically, 1000 times not scalable enough
  - coverage is < 0.1%
  - As a result, Operators learn outages from clients (client has been already impacted)

Solution: HW integration

- New HW generation are equipped with offload engines (e.g., dedicated NPU, OAM co-processors)
- Leverage the offload engine for probe generation/processing
- Significant performance boost
- Million of probes per second compared to few thousands in case of CPU

- As an example, Cisco Silicon One provides up to 14 MPPS
  - Generation
  - Receive & Process

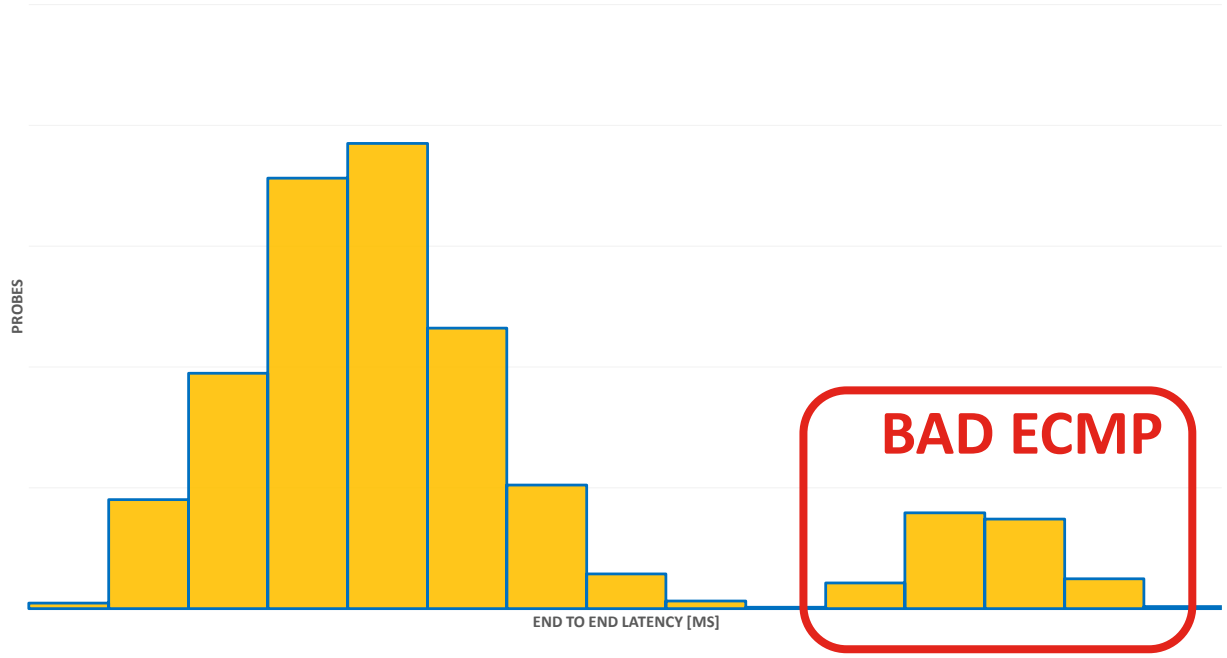
Metrics

Problem statement

- Relying on metric such as min/max/average:
  - DO NOT reflect the experience of the whole client spectrum
- Example:
  - 1 bad path out of 8 ECMP
  - 12.5% of the clients impacted
  - Average hides the issue (only little increase in the Avg)

Solution: Compute a Latency Histogram

- Show the latency distribution across all ECMP Paths.
- Reports the experience of the whole population
- Multi-mode = One/More Bad ECMP



Problem statement

- Round trip time (RTT) measurement
  - Latency = RTT/2
- Network Path are asymmetric
  - Upstream path is different from downstream
  - London-to-Rome path NOT Equal Rome-to-London path
- We can’t detect if the issue is in the upstream or down stream path

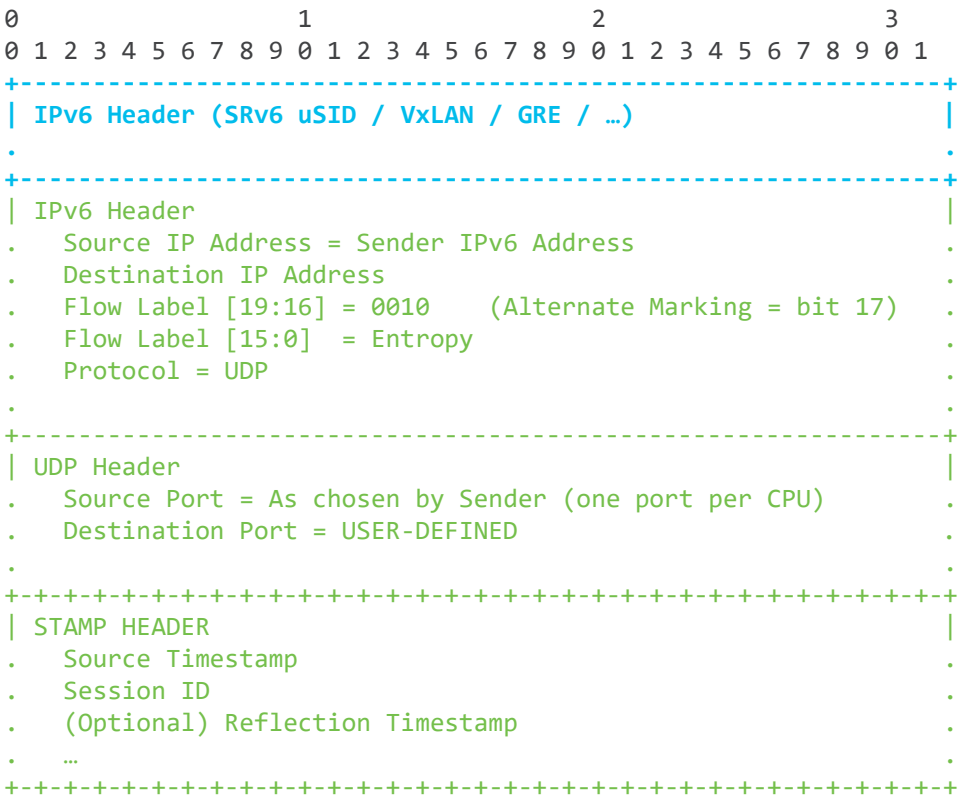
Solution: One-way measurement

- Latency of upstream and downstream paths are measured independently
  - Time synchronization between the edge nodes.
- Allows the pinpoint of the issue (upstream vs downstream)

Standard

Standard Protocol & Packet format

- Leverage IETF standard protocols
  - RFC8762 – Simple Two-way Active Measurement Protocol
  - RFC8972 – STAMP extensions
- No custom non-standards protocol
- Packet Format:
  - Outer Encapsulating header:
    - Any IP Encapsulation
    - SRv6 uSID/GRE/VxLAN/…
  - STAMP measurement packet:
    - Alternate Marking bit as part of Flow Label
    - Flexible STAMP header supporting several deployment models



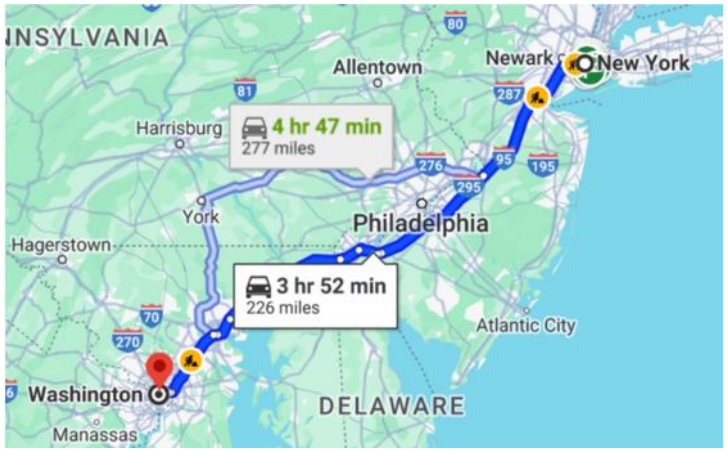
Cost

Cost reduction

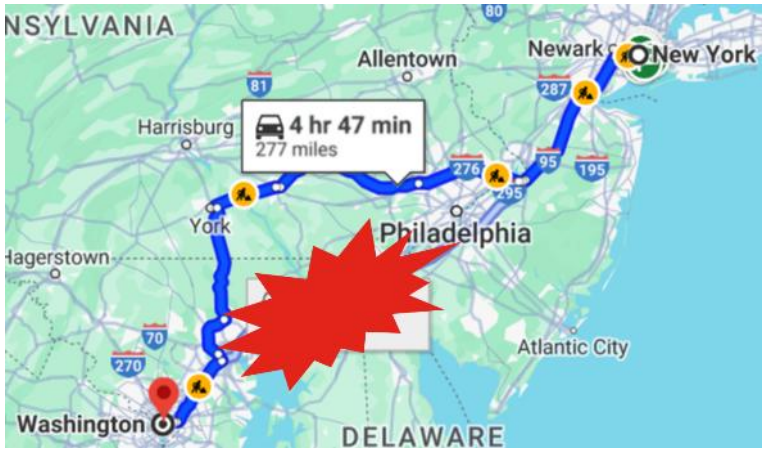
- Capex Elimination
  - SLA Appliance
  - Router port to appliance
- Opex Elimination
  - Rack Space
  - Power

Routing correlation

Continuous Correlation to Routing



Measured Latency
compared to best topology



Measured Latency
compared to current topology

- Time-series of Measurements from any P to any Q along any ECMP path
- Time-series of ECMP routed paths from any P to any Q

Inference

- Measurement (**PAR**, **MAD**) report SLA degradation (e.g., loss)
- Without any additional measurement, Routing Correlation allows to infer other (SRC, DST) pairs that are also impacted
  - BRU to MAD/LIS/SEV is impacted
  - LON to MAD/LIS/SEV is impacted

Measurements & Routing Correlation

- Time-series of Measurements from any P to any Q along any ECMP path
- Time-series of ECMP routed paths from any P to any Q
- Use-cases
  - Grade a measurement with respect to the current topology
  - Root Cause Analysis
  - Post-Mortem analysis
  - Clustering
  - Inference
  - Planned vs Unplanned Outage

Operators’ Use-cases

Goldman Sachs

- Presented at MPLS SD&AI NET WORLD – PARIS 2024
  - Michael Valentine, Technology Fellow, Network Architecture, Goldman Sachs (Link)

Requirements for ECMP-aware measurements

- Enough PPS
  - Need to spray probes across ALL ECMP paths
  - IP fabric with 100 edge nodes, 192 ECMP paths, and 100 measurements per second:
    - A fabric edge node needs to process 100 * 192 * 100 = **1.92 M probes per second**
  - Current solutions provide 100-1000 times less
- Accurate metrics
  - Current reporting based on (min, max, avg) is insufficient
  - Need to have a representation of the entire population, all ECMP paths
  - Percentile or Histogram based

Integrated Performance Measurement

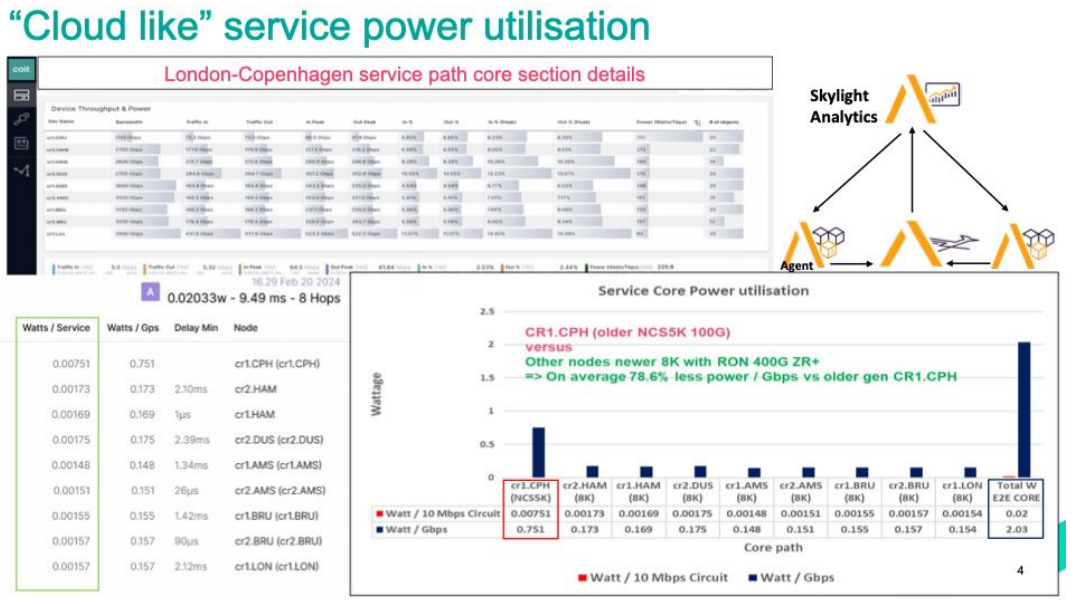
- Built-in measurement from the routers using dedicated engine in Silicon 1 ASIC
- Scale, Richness, Cost Reduction
- 14MPPS generation & 14 MPPS ingestion
- Standard-based proving: STAMP (RFC8762/RFC8972)
- 3L's for each probe:
  - one-way latency (usec accuracy, histogram)
  - liveness (sub-2ms)
  - absolute loss (instead of deductions)
- Currently testing in the lab

Goldman Sachs – Takeaways

- Legacy Performance monitoring solutions neglect the ECMP reality
  - 100’s of ECMP paths cannot be measured with 1 single probe
- Before extra routers we deployed to be used to do probing (IP SLA)
  - Difficult to mange, lots of touch points and products
  - Over the top app probes give vague answers, problem between DC1 and DC2, etc...
- Requirements for ECMP-aware measurements
  - Enough PPS: measure all edges to all edges across all ecmp paths
  - Accurate metrics: current reporting based on (min, max, avg) is insufficient
- IP Measurements:
  - Built-in measurement from the routers using dedicated engine in Silicon 1 ASIC
  - 14MPPS generation & 14 MPPS ingestion
  - Standard-based: STAMP (RFC8762/RFC8972)
  - 3L’s for each probe: one-way latency, liveness, Loss
- **Currently testing in the lab**

Colt

- Presented at MPLS SD&AI NET WORLD – PARIS 2024
  - Bart Janssens, Senior Specialist Packet Architecture, Colt Technology Services (Link)



Colt – Takeaways

- Routing Analytics is a key for Colt's Journey for Network as a service (NAAS). The journey includes:
  - Deploying Network probes at different parts in the network
  - Adopting new generation silicon for more efficient power
  - Routing Analytics: new set of apps that collect information from the network and deliver intelligence
  - Skylight Analytics: the portal where all the data are brought together.
- Colt 1st use-case: Service consumption visualization
  - Routing Analytics:
    - Collects BGP-LS topology real time
    - Augment the topology with the IP Measurement data and other telemetry data such as wattage
    - Computes the Path and associated latency hop-by-hop and other
  - Skylight Analytics:
    - Visualize the service path and associated latency and wattage hop-by-hop
  - **Deployed in production at Colt**
- This is a one step in Colt journey to deliver cloud-like services and using AI for operating the network.

Conclusion

- Legacy performance monitoring solutions:
  - Lack of scale to measure all ECMP Paths
  - Lack of accurate metrics: min, max, avg are not sufficient and does not show the problem
  - Capex and Opex cost
- IP Measurements:
  - HW Built-in measurement from the routers
  - Standard-based: STAMP (RFC8762/RFC8972)
  - Latency histogram, Absolute Loss, Liveness in HW
  - Routing Analytics: Routing and Measurement correlation
- Customer deployments and use-cases:
  - Goldman Sachs - Presented at MPLS SD&AI NET WORLD – PARIS 2024
  - Colt - Presented at MPLS SD&AI NET WORLD – PARIS 2024


CISCO