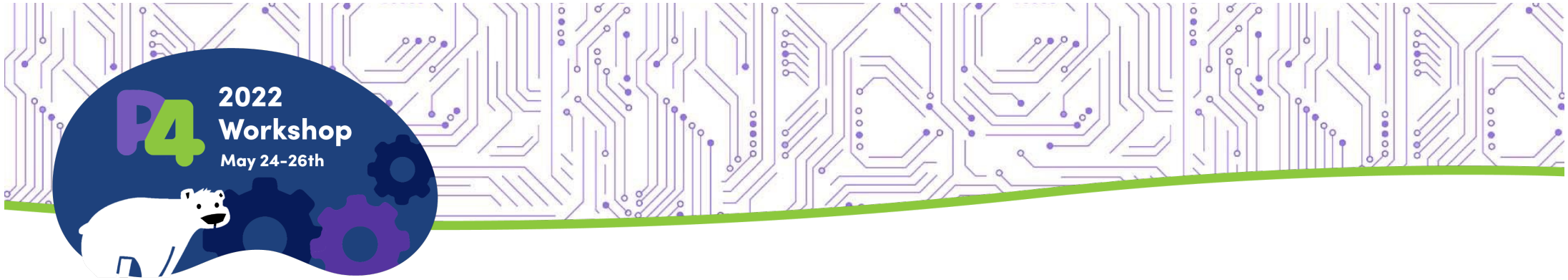
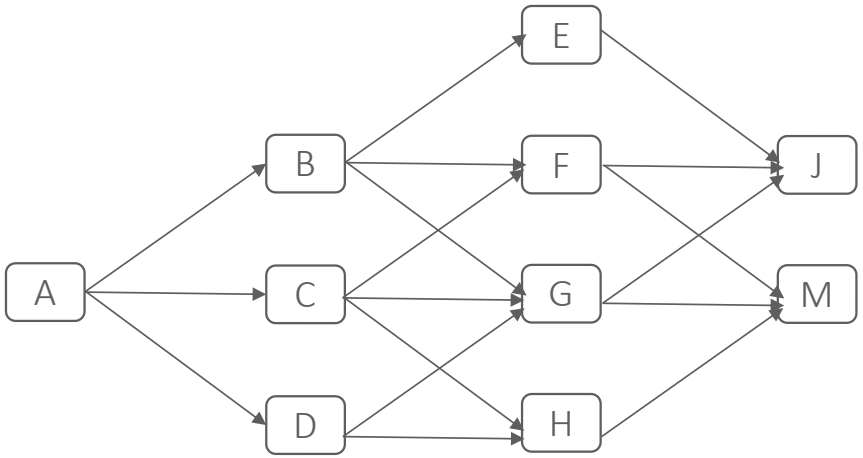


Path Tracing

Ahmed Abdelsalam

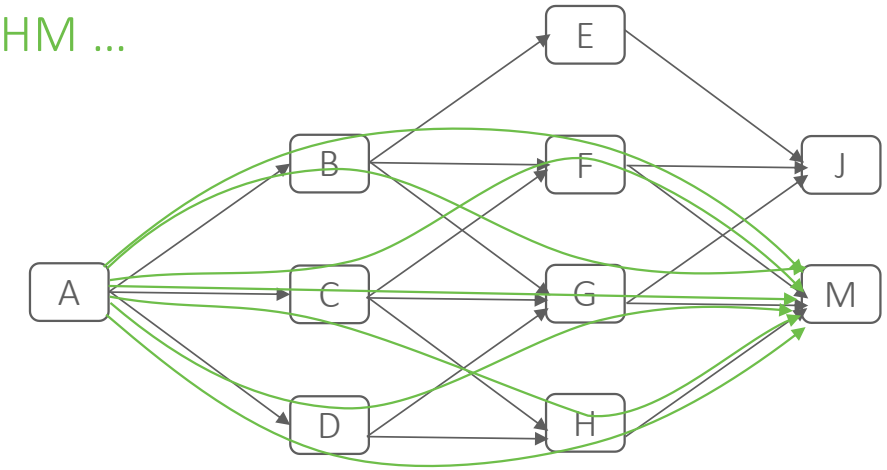
Software Engineer, Cisco Systems

The exact path from A to M is unknown



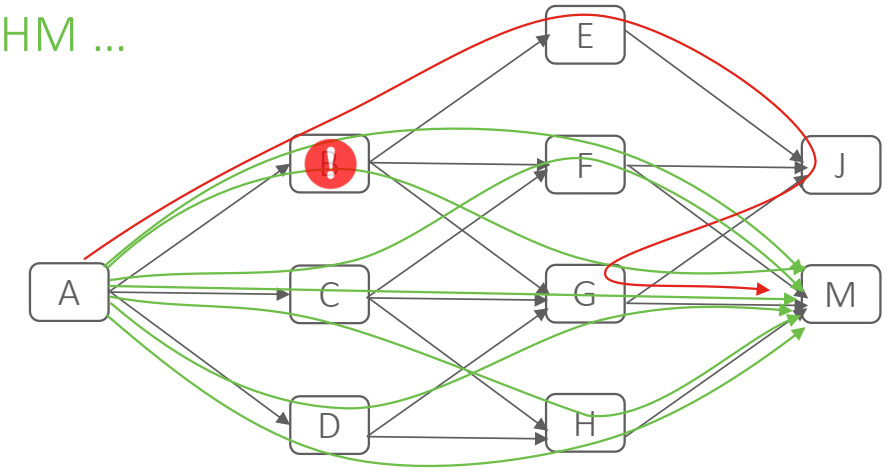
The exact path from A to M is unknown

- 7 possible “valid” ECMP path
  - ABFM, ABGM, ACFM, ACGM, ACHM, ADGM, ADHM …



The exact path from A to M is unknown

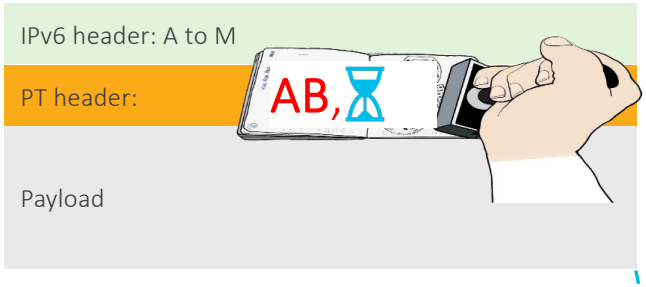
- 7 possible “valid” ECMP path
  - ABFM, ABGM, ACFM, ACGM, ACHM, ADGM, ADHM …
- The path may be invalid
  - Routing or FIB corruption @ B
- Timestamp at each hop
- Interface Load at each hop

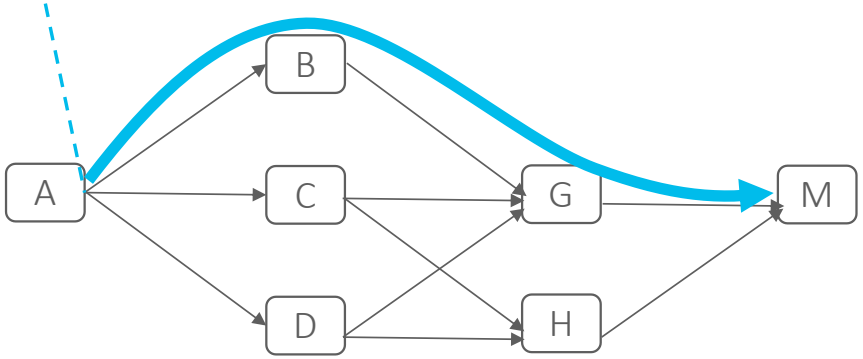


The Path Tracing idea

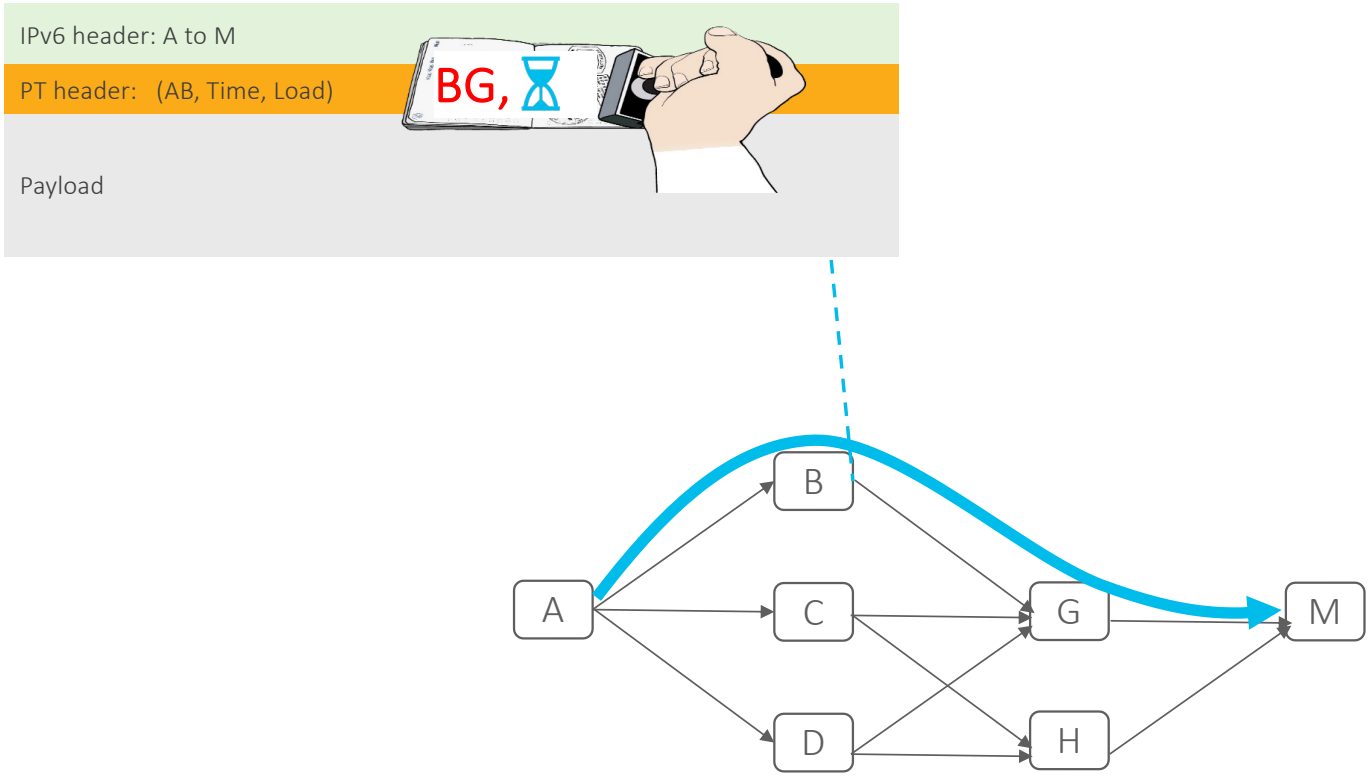
- Stamping in the Packet Header
- Implemented at linerate in the most basic HW pipeline
  - Linerate for any packet
  - No punting to CPU, no offload to co-processors
- Ultra-MTU-efficient: only 3 bytes per hop!
  - 12-bit Interface, 8-bit Timestamp, 4-bit Load
- For IPv6, with or without SRH
  - MPLS solution also designed
- Interwork with legacy node

Stamping Trajectory in PT Header

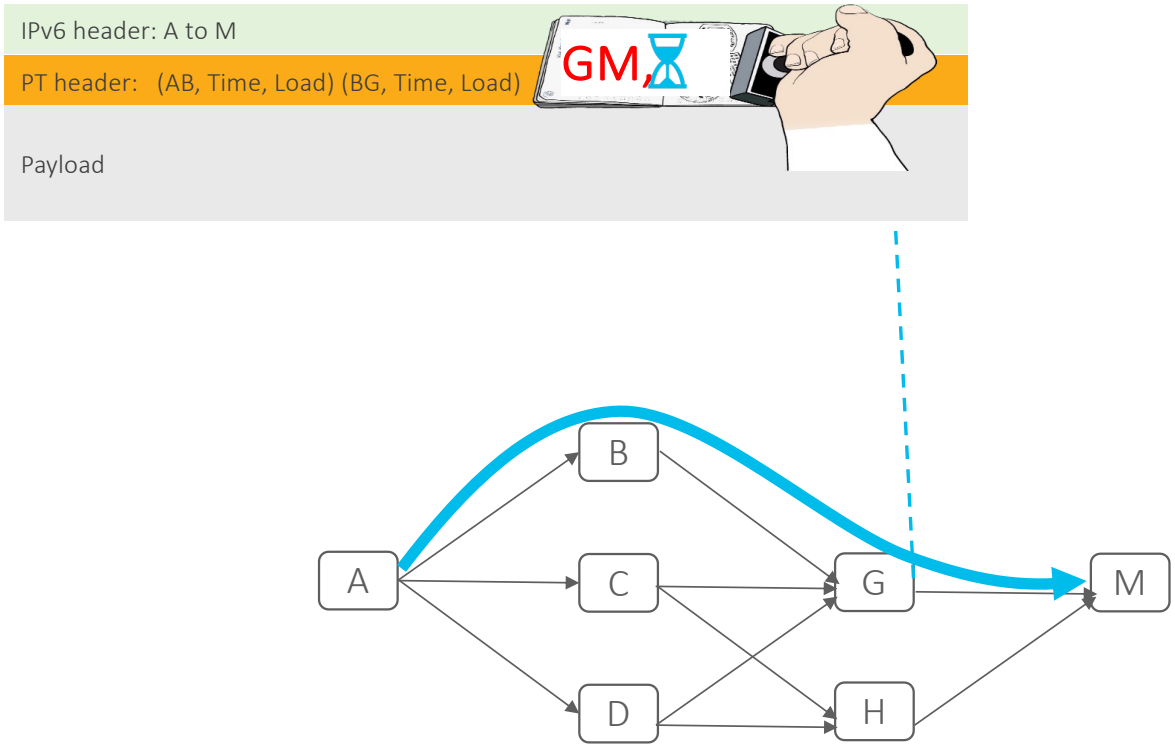




Stamping Trajectory in PT Header



Stamping Trajectory in PT Header

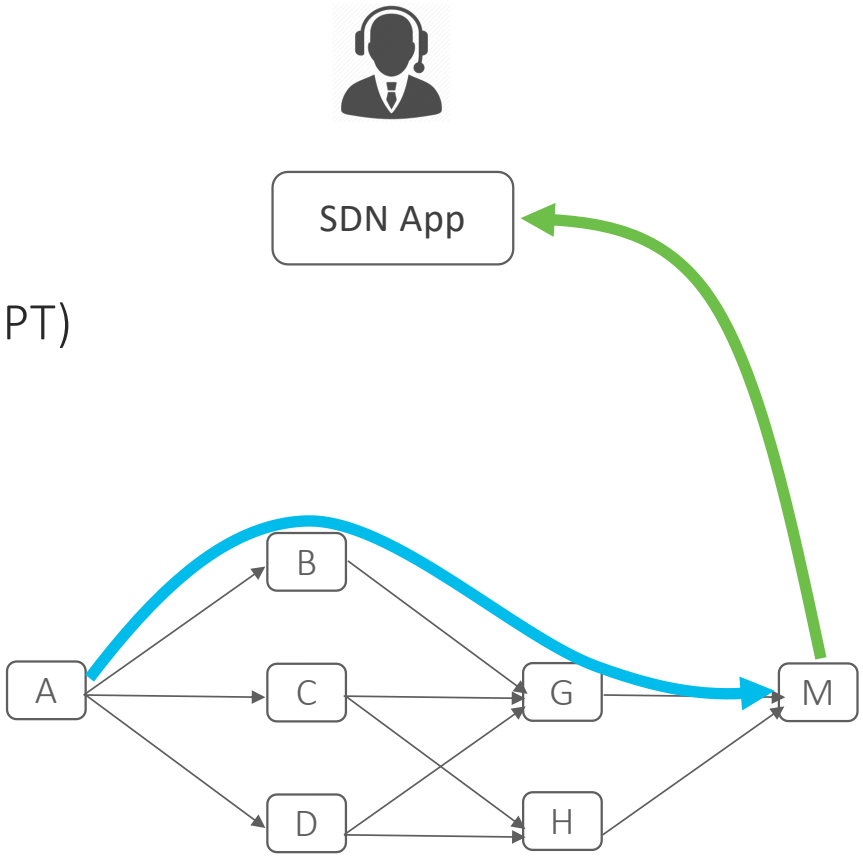


Dataplane Encapsulation

- Minimize NPU parsing
- Minimize # of Read/Write
- Minimize depth of Read/Write
- Maximize Read/Write at fixed positions
- Avoid Header Insert/Resize
- Minimize MTU

Minimize HW complexity by leveraging SDN analytics

- Analytics
  - translates the list of collected IDs into a path
  - deduces the timing and load history at each hop
  - Highlights hotspots
- Consistency check:
  - expected paths (PCE) vs actual forwarding paths (PT)
- Feedback loop to applications
  - Trigger a change of path (SR, MTCP)
  - Trigger a change of rate



Product, Deployment & EcoSystem

- Cisco Shipping in CY22
  - Demo and Training (link)
- Strong Operator Interest
- Rich Eco-System
- Rich Open-Source



- At IETF: draft-filsfils-spring-path-tracing
- NANOG85: https://www.nanog.org/events/nanog-85/
  - Mike Valentine, Goldman Sachs

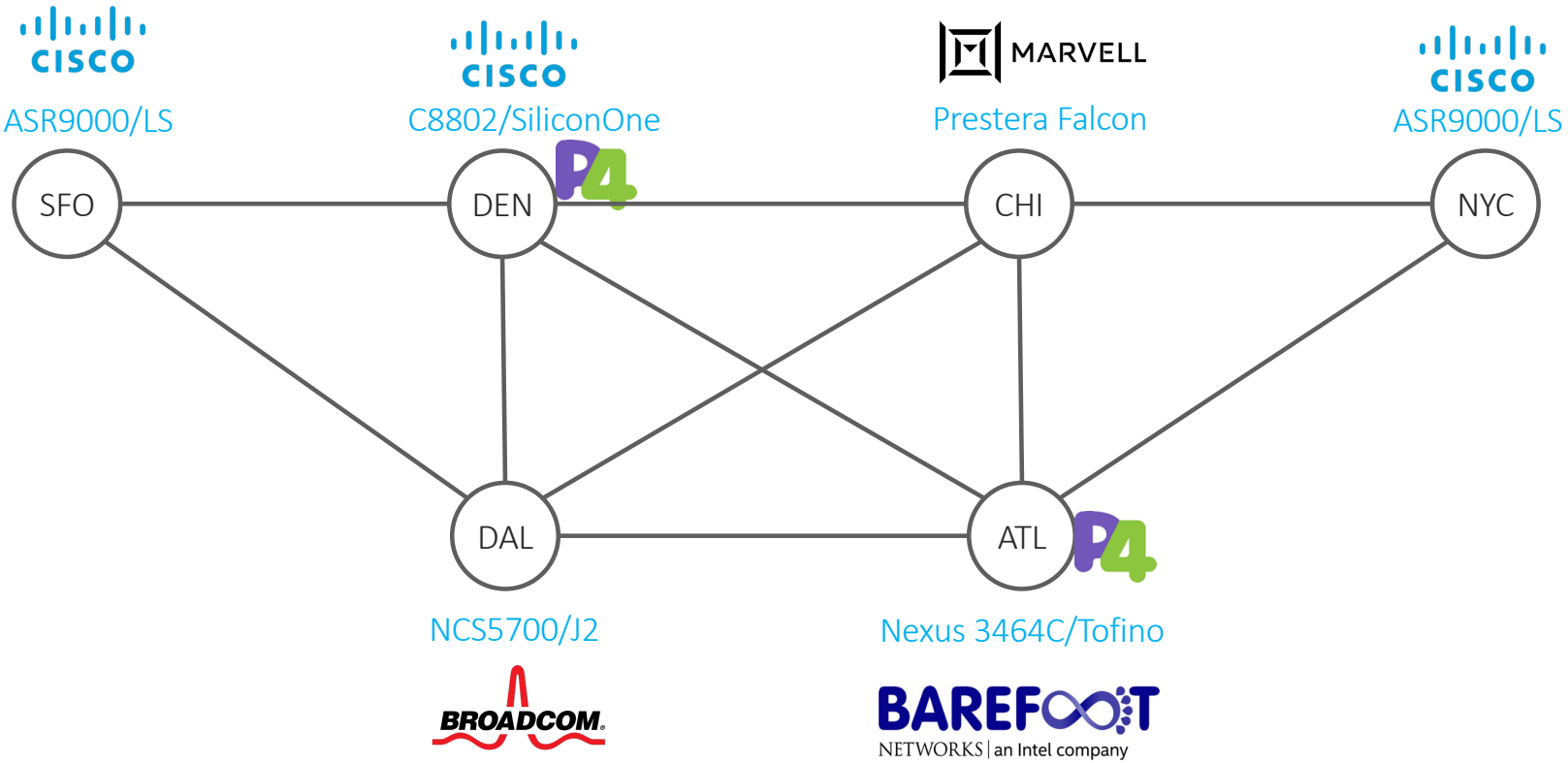
Path Tracing vs Alternatives

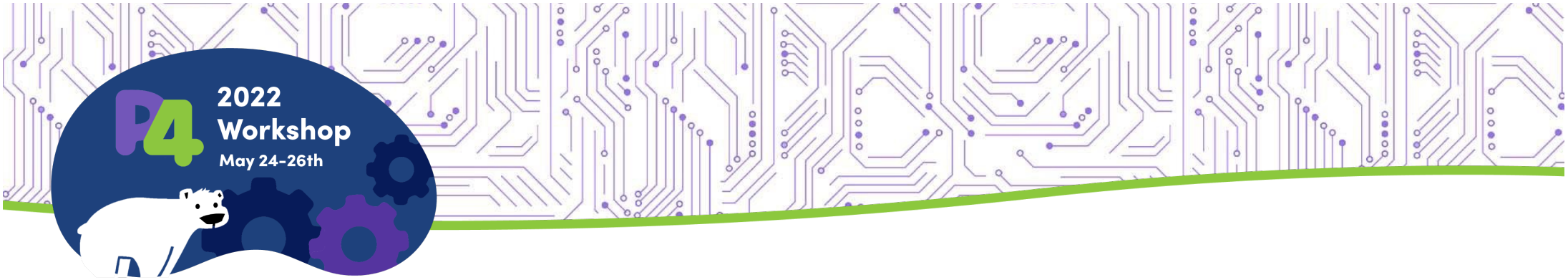
- Much Smaller Header Overhead
  - Collect 3byte per hop versus 20/32/20 for INT/IFA/iOAM
- Simpler Header Processing
  - Alternatives adjusts header size each hop
  - Alternatives' header location depends on packet type (VXLAN/NSH/UDP/GRE)
- HW linerate
  - Path Tracing already implemented in Cisco, Broadcom, Marvell, Others
  - Alternatives are difficult to implement at linerate
- Monitors the true packet HW pipeline
  - Monitoring through a different path (OAM assist, FPGA, LC CPU) has much less value
- Smoother deployment
  - Its simplicity enables legacy system leverage

Conclusion

- Simplicity Always Prevails
- Path Tracing
  - Deterministic Per-Packet Tracing
  - Implemented at linerate in the base HW pipeline
  - Ultra-MTU-Efficiency
- Product, Deployment & Ecosystem
  - Rich Eco-System (Cisco, Broadcom, Marvell, others)
  - Strong Operator Interest
  - Rich Open-Source
  - PT is being standardized at IETF

Demo


ASR9000/LS
C8802/SiliconOne
Prestera Falcon
ASR9000/LS
SFO
DEN
CHI
NYC
DAL
ATL
NCS5700/J2
Nexus 3464C/Tofino
BROADCOM
BAREFOOT
NETWORKS | an Intel company



Thank You

https://www.segment-routing.net/path-tracing