

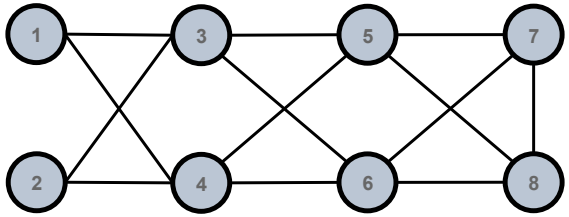

# Data Plane Monitoring in Segment Routing Networks

Faisal Iqbal – Cisco Systems (faiqbal@cisco.com)

Clayton Hassen – Bell Canada (clayton.hassen@bell.ca)

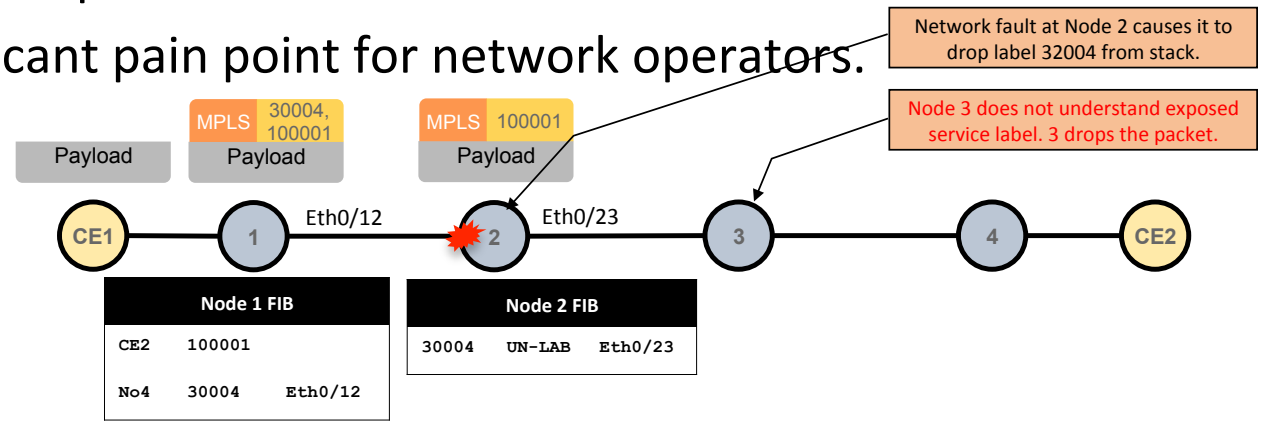
# Reference Topology & Conventions

- SR control plane is IS-IS and data plane is MPLS.
- Node k Prefix SID: 1.1.1.k/32
- Node k Prefix SID label: 16000+k
- MPLS label nth adjacency between XY: 24nXY
- Traditional MPLS labels: > 30000
- Service labels: > 100000



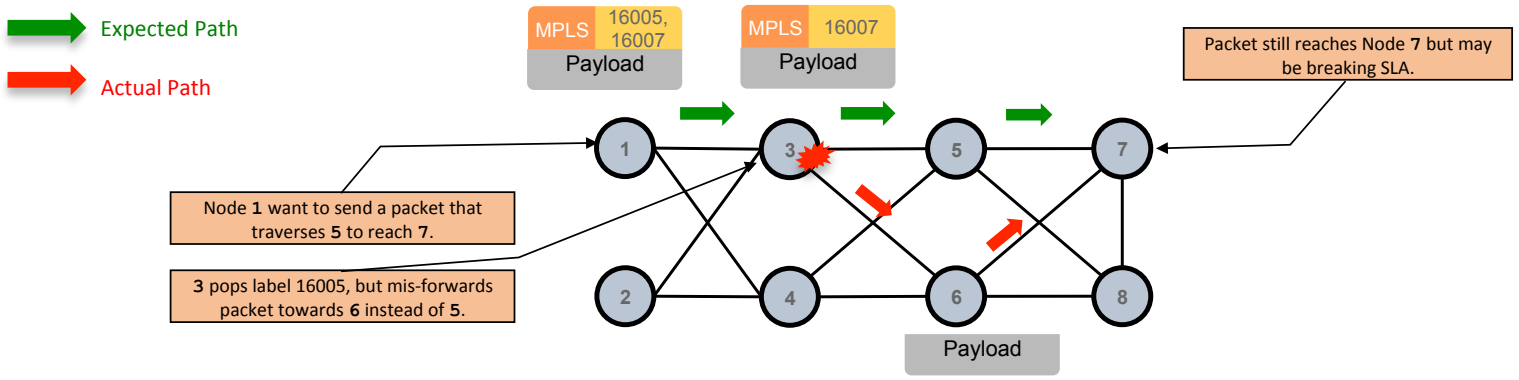
# Traffic Black holes

- Router starts dropping traffic against one or more prefixes or labels.
- May occur due to:
  - Forwarding inconsistency
  - Out of sync neighbors
  - User configuration error
  - Driver/hardware issues
- Dropped packets could be core or VPN traffic.
- Significant pain point for network operators.



# Path Divergence

- Due to global nature of SR labels, packet still gets label switched and may reach the destination.
- Network traffic diverges from the expected path specified by imposed label stack or IGP metrics.
- May result in congestion, delay, or breakage of SLA with customers.
- Fault may remain undetected by most probe based monitoring tools.



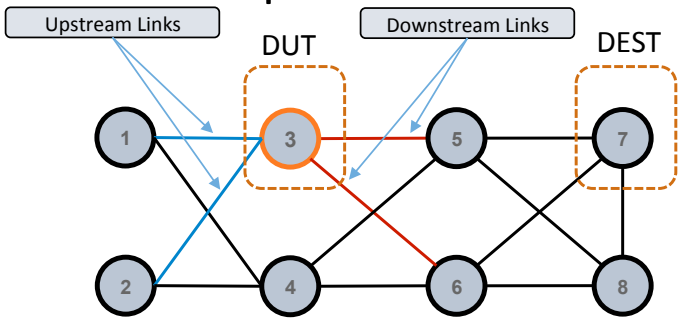
# Existing Fault Detection Mechanisms

- Proactive vs. Reactive Detection
  - Proactive – Continuously monitor the entire network to detect a failure before customer observes service disruption. Detection on the order of tens of ms to a few minutes.
    - Difficult to monitor entire network within suitable detection period.
  - Reactive – Internal or external customer alerts operator of traffic loss. Operator performs reactive diagnostics to identify the fault and point of failure.
- Local vs. End-to-end Detection
  - Local – Perform local consistency check tools (LCC/RCC), proactively or on-demand, to detect a network inconsistency.
  - End-to-end – Verify the health of an LSP from one endpoint to another. Uses probe packets for proactive (BFD) or reactive (on-demand LSP echo requests) detection.
- Incomplete detection is limited and only identifies a sub-set of network faults. It may fail to detect some faults due to nature of the tool, implementation constraints, or both. Most existing tools including BFD, LSP Ping, and LCC/RCC are incomplete.

**Segment Routing Data Plane Monitoring (SR-DPM)** is an independent solution for a device to validate its own data plane for any SR-MPLS traffic received from any incoming interface towards any downstream ECMP.

# Introduction

- A distributed but localized data plane monitoring mechanism based on source routing paradigm.
  - Each node proactively validates forwarding health of itself and its neighbors.
  - Localized to a node and its neighbors, while validating actual traffic path.
  - Distributed processing/validation at each node, overcoming scale limitations of end-to-end verification tools.
- Verifies each neighbor link and each destination prefix using specially crafted LSP Ping packets.
- Reports detected faults to the operator.



# Detection Process

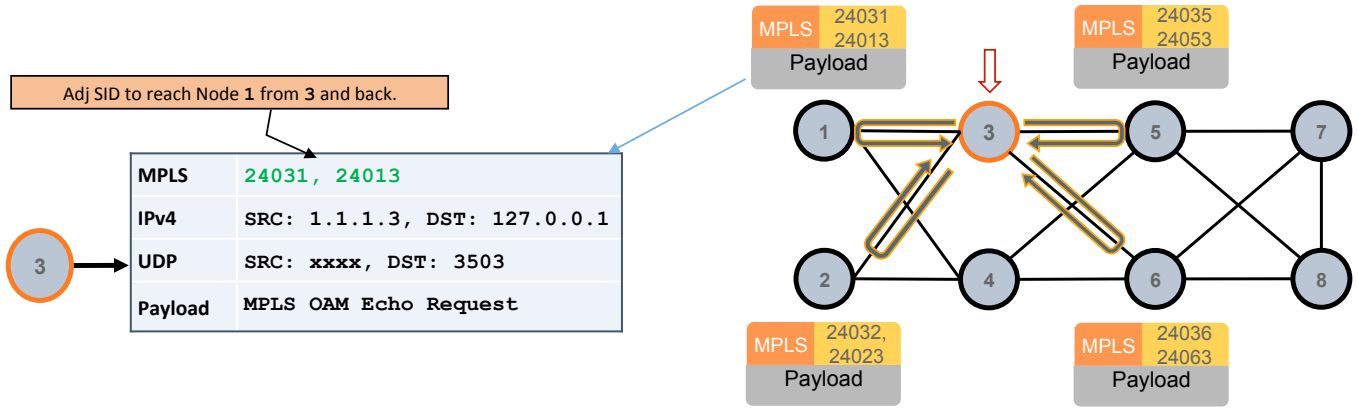
- Two stage process

- **Phase 1 – Adjacency Validation**
  - Specially crafted MPLS echo request packets to validate MPLS forwarding capability of each link to and from every neighbor.

- **Phase 2 – Prefix Validation**
  - Building on top of Phase 1, validates the control plane and forwarding path health of each destination prefix locally, and at every downstream neighbor.

# Phase 1 - Adjacency Validation

- Node **3** wants to validate its adjacencies with its neighbors (**1, 2, 5, 6**).
- For each neighboring link, Node **3** does the following:
  a. Constructs echo packet with adjacency labels such that packet would transit to the neighbor and return on same link.
  b. Neighbor simply switches the packet based on adjacency SID label in hardware.
  c. Node **3** verifies that returned packet is received on the targeted link.
- If the validation fails for any link, **3** marks the link in DB and notifies operator.
- No hardware/software change is required at neighbors.

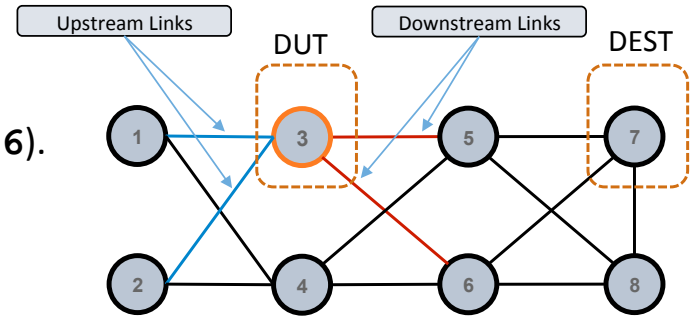


# Phase 1 – Adjacency Validation

- Allows SR-DPM processing node to identify malfunctioning adjacencies towards or from any neighboring link.
- Malfunction may include:
  - MPLS traffic drop over the link
  - Forwarding towards an incorrect node from DUT or the neighbor.
  - Switching the traffic towards an incorrect link against a given adjacency label at DUT or neighbor.
- Malfunctioning links are excluded from further usage in Phase 2 to avoid false positive.
- Also serves as a building block for Phase 2.

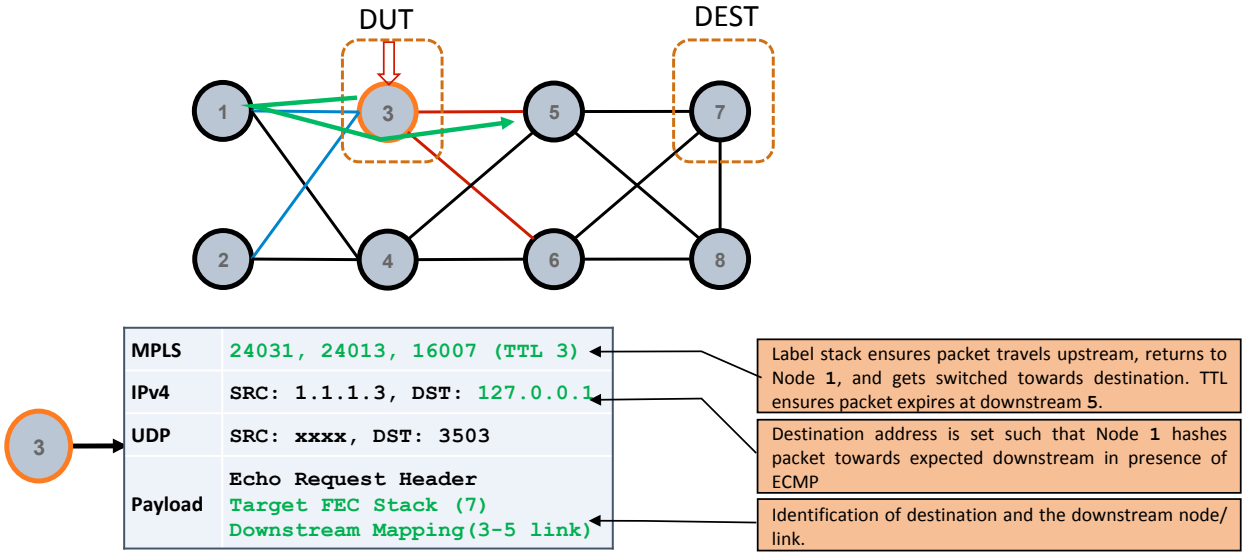
# Phase 2 – Prefix Validation

- Detects network inconsistencies in the SR-DPM processing node or any of its immediate neighbors.
- By looping the traffic from upstream, detection process simulates customer traffic path, validating the consistency of entire ingress + egress forwarding chain at SR-DPM processing node.
  - Packet path for locally generated probes in BFD/LSP Ping only validates egress forwarding path from the originator.
- Example: DUT (**3**) wants to validate its forwarding for all SR Prefix SIDs reachable using its IGP database.
- For each destination prefix SID (e.g. **7**), Node **3**:
  - Identifies potential upstream nodes and links (**1**, **2**).
  - Identifies potential downstream nodes and links (**5**, **6**).
  - Removes any malfunctioning links.



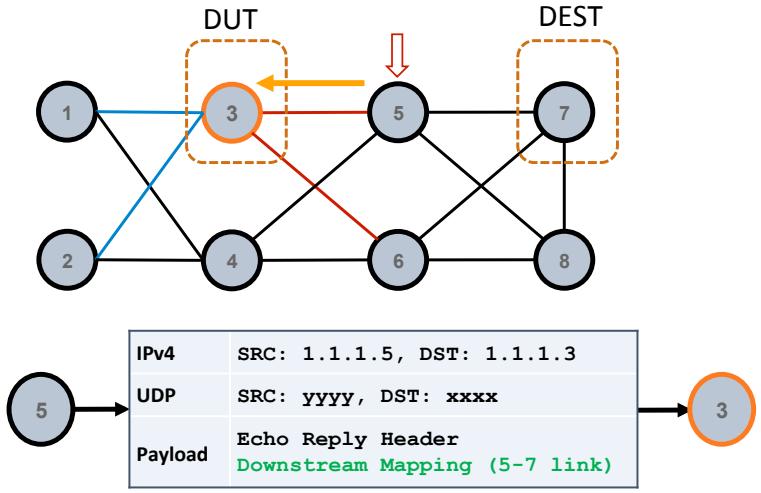
# Phase 2 – Prefix Validation DUT Processing

- For each upstream and downstream link combination for any destination Prefix SID, DUT (**3**) initiates following process:
  - a) Builds an MPLS echo request packet identifying destination and targeted downstream link.
  - b) Switches the packet to the upstream neighbor (**1**) from a particular link.
  - c) Upstream (**1**) pops and label switches the packet back to DUT **3** on the same link.
  - d) 3 hashes incoming packet with prefix SID label towards downstream (**5**).
  - e) TTL of the packet expires and it's processed at downstream (**5**).



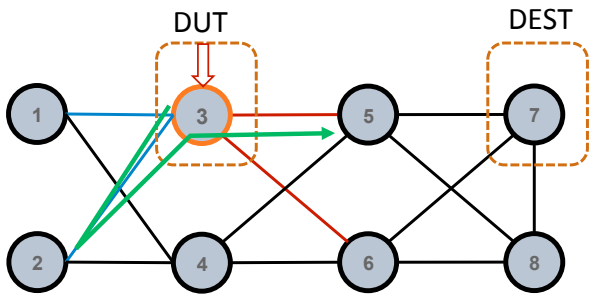
# Phase 2 – Prefix Validation Downstream Proc.

- Downstream node (5) processes the incoming echo request:
  a. Verifies that it has a path to reach the destination Prefix SID (**7**).
  b. Verifies that packet was received on the targeted node and link.
  c. Returns the result of the verification back to the originator (**3**).
- Downstream node does not require any software change, or even be DPM-aware.



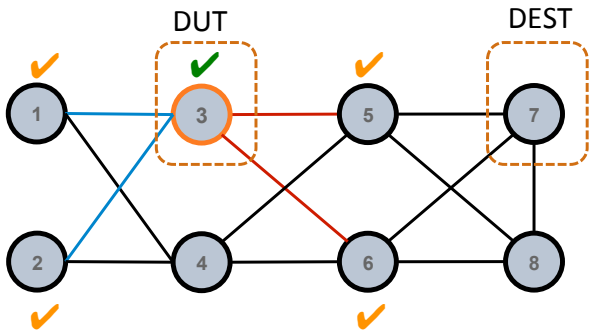
# Phase 2 – Prefix Validation Response Proc.

1. Originator (3) upon receiving the echo reply from the downstream node (5):
   a. Validates the content of the response against the request sent.
   b. Returns the returned info to **dpm** process.
2. **dpm** processes the response and if a fault is identified, generates a message to operator for further review.
   a. Continues to process the echo request for subsequent upstream-downstream neighbor combinations and remaining prefix SIDs.



# Prefix Validation Reach

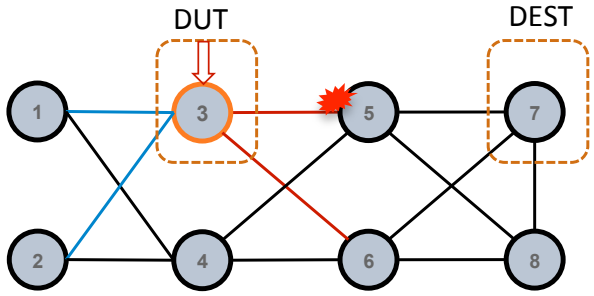
- SR-DPM can detect any forwarding issue at DUT. ✔
  - SR-DPM works in distributed fashion where devices across the network run SR-DPM.
- SR-DPM can even detect most forwarding issues at downstream device. ✔
- For optimal coverage of your topology, deploy SR-DPM at strategic nodes.



# Fault Notification

- Upon detecting a fault, **dpm** process generates a syslog notification for operator review.
- Contains the impacted nodes/links and a brief error description.
- Also has the ability to notify the operator using streaming telemetry messages.
- **dpm** process can also perform automatic fault remediation and reduce operator overhead.

```
Aug 10 11:28:37.358 : DPM_ERR prefix:[1.1.1.7], upstream link: Gi0/0/0/1
(99.1.3.3), downstream link: Gi0/0/0/3 (99.3.5.3). Output code:[N] - no
RX label 16997
```



# On-demand and Event Driven Monitoring

- For most networks, SR-DPM can actively run in the background for fast proactive detection.
- SR-DPM can be invoked on-demand for a particular prefix to validate all upstream and downstream data plane path combination.
- SR-DPM can also be invoked when certain network events are triggered. Triggers may include:
  - Interface state change
  - IGP convergence after a network change.
  - Reception of a trigger from a controller

# Conclusion

- A unique and innovative approach to tackle data plane consistency verification and traffic black hole detection challenges.
- Overcomes scale challenges by distributing the detection process, while still achieving validation of entire customer traffic path.
- Interoperable by design – no special processing needed at neighbors. No standardization required at IETF.
- Complements existing OAM solutions instead of replacing them.
- Ready for adoption by customers.