Tutorial
Traffic Engineering Essentials for Automation Orientation

NANOG90 - FEB-2024


Diego Achaval


David Watkinson
bio.site/dwatkinson


Mau Rojas
bio.site/pinrojas

Abstract

- In this presentation you will be shown the essentials of Traffic Engineering to effectively automate and optimize network traffic management.
- Why Traffic Engineering is essential for network efficiency, detailing the role of MPLS-TE for reliable and customizable network performance, and the importance of core components such as Constraint-Based Routing, Admission Control, and Policy Enforcement.
- PCE/PCEP, BGP-LS, and Telemetry: Discusses the tools for optimal path calculation, network topology understanding, and real-time data gathering to maintain and improve Traffic Engineering.
- Introduces the use of standardized YANG APIs that help in maintaining uniform Traffic Engineering strategies across different network platforms.

Agenda

- Traffic Engineering Overview [Diego]
- PCE/PCEP Overview [David]
- BGP-LS Overview [David]
- IETF APIs for TE (standardized YANG) [David]
- Telemetry and PCE [David/Mau]

Introduction to Traffic Engineering

Diego Achaval

Understanding Traffic Engineering

What?

Process of **optimizing** the **performance** of operational networks to **ensure efficient data flow** and improve overall network resource **utilization**.

NANOG™

Understanding Traffic Engineering

- **Why is Traffic Engineering Crucial for MPLS?**
  - Predictability
  - Customization
  - Resilience

- **Key Components of Traffic Engineering:**
  - Traffic Trunks
  - Constraint-Based Routing
  - Admission Control
  - Policy Enforcement

NANOG™

Traffic Engineering Technologies

- **IETF**
  - **RSVP-TE** (Resource Reservation Protocol - Traffic Engineering)
  - **SR-TE** (Segment Routing - Traffic Engineering)
    - **SR-MPLS**
    - **SRv6**
    - **SR Flexible Algorithms**
  - **TE and Centralized Path Computation and Optimization**
    - **BGP-LS** (Border Gateway Protocol - Link-State)
    - **PCEP** (Path Computation Element Protocol)

- **Non-IETF or Proprietary**
  - **SDN** (Software Defined Networking)
  - **Closed Loop**: PCEP + Telemetry
  - **Automation Driven Traffic Steering (more details at Tutorial - part2)**
    - **Via APIs (i.e. RESTCONF, Python)**
    - **Demo lab in Tutorial – Part 2**

Traffic Engineering IETF

Diego Achaval

Resource ReSerVation Protocol – Traffic Engineering

- Traffic Engineering: find the lowest cost path that also meets specific constraints:
  - Shortest path based on different optimization objectives:
    - IGP cost, TE metric, e2e delay, including upper bounds
  - Constraints:
    - Bandwidth, local protection (link/node), link color, link/node/SRLG disjointness
  - Adapting to dynamic bandwidth demands using Auto Bandwidth
    - Periodically measure per LSP BW utilization
    - Re-signal BW values using RSVP
    - Instantiate new LSPs using MBB procedures

Resource ReSerVation Protocol – Adapting to topology changes and traffic demands

- Adapting to topology changes and traffic demands with Multi-path LSPs as defined in draft-kompella-mpls-rsvp-ecmp:
  - Dynamic creation/deletion of LSPs (aka as splitting/merging) based on upper/lower BW utilization thresholds
  - Dynamic LSP bandwidth resizing using Auto Bandwidth
  - Efficient load balancing
  - Resiliency with local protection and restoration
  - Automated tunnel management
  - Enhanced network adaptability

NANOG™

Segment Routing – Traffic Engineering

- Source Routing
  - The source/ingress node computes a path to an endpoint and encodes the forwarding instructions in the packet header as an ordered list of segments
  - Nodes in the network execute these segments/instructions to forward packet accordingly
- Source Routing with MPLS data plane:
  - With an MPLS data plane, and with either IPv4 or an IPv6 control plane, a segment or a list of segments, is represented by a label or a label stack

Segment Routing – Traffic Engineering

- Source Routing with IPv6 data plane:
  - With an IPv6 data plane, a segment or a list of segments, can be encoded in both the IPv6 DA field and in the IPv6 Segment Routing Header (SRH) within the IPv6 header.

NANOG™

Segment Routing – Traffic Engineering – SR-MPLS

- SR-TE LSP and/or SR-Policy:
  - Waypoints do not need to hold per-LSP/SR-policy state as there is no per-LSP/SR-Policy signaling nor per-LSP/SR-Policy BW reservations
  - Support for on-demand or intent-based instantiation
  - Uses Node-SIDs, Adj-SIDs, AdjSet-SIDs, Peer Node-SID/Adj-SIDs (SR-EPE) and Binding-SIDs.
  - Base MPLS Imposition may become a challenge when a large set of forwarding instructions (labels) is required. The use of Binding-SIDs, or other label-stack compression techniques may help.
  - SR-TE:
    - Explicit or dynamic paths supporting distributed or centralized path computation
    - Support for multiple optimization objectives: igp-cost, te-metric, e2e-delay
    - Traffic steering based on well known mechanisms like, shortcuts, auto-bind, etc

NANOG™

Segment Routing – Traffic Engineering – SR-MPLS

- SR-TE LSP and/or SR-Policy (continued):
  - SR-TE LSP (continued):
    - Local protection (FRR) and end-to-end protection (primary/secondary-standby)
    - ECMP/W-ECMP with multiple LSPs
  - SR-Policy:
    - Explicit or dynamic segment-lists supporting distributed or dynamic path computation
    - Support for multiple optimization objectives: igp-cost, te-metric, e2e-delay
    - Native color-based traffic steering
    - Local protection (FRR) and end-to-end protection with multiple candidate-paths
    - ECMP/W-ECMP with multiple segment-lists per candidate-path

Segment Routing – Traffic Engineering – SRv6

- SR-Policy:
  - Waypoints do not need to hold per-SR-Policy state as there is no per-SR-Policy signaling nor per-SR-Policy BW reservations
  - Support for on-demand or intent-based instantiation
  - Uses SRv6 SIDs with different behaviors and flavors. Behaviors like End/End.X/uN/uA/End-B6 and flavors like PSP, USP or USD or a combination.
  - MSD may become a challenge when a large set of forwarding instructions is required, particularly the Maximum H.Encaps supported by ingress nodes
  - SID compression is likely to be required for TE with SRv6 SIDs
  - Explicit or dynamic segment-lists supporting distributed or dynamic path computation
  - Support for multiple optimization objectives: igp-cost, te-metric, e2e-delay
  - Native color-based traffic steering
  - Local protection (FRR) and end-to-end protection with multiple candidate-paths
  - ECMP/W-ECMP with multiple segment-lists per candidate-path

NANOG™

Segment Routing – Traffic Engineering – Flexible Algorithms

- Flex-Algo allows the IGP to compute a constraint-based paths in the network
  - Support for multiple optimization objectives: igp-cost, te-metric, or e2e-delay
  - Support for SRLG and/or include/exclude amin-groups
- Segment Routing Flexible algorithms can be used with an MPLS or an IPv6 (SRv6) data plane.
  - Per Algo Prefix SIDs with MPLS
  - Per Algo Locators with SRv6
  - No per-LSP/SR-policy state is held in ingress nodes or waypoints.

Traffic Engineering – Centralized Path Computation and Optimization

- BGP Link State – Overview
  - Originally defined in RFC7752, now obsoleted by RFC9552. And further extended to support SR-MPLS, SRv6, Flex-Algo, etc
  - It has been developed to carry Link State and TE information to be shared with external applications, like a Path Computation Element
  - New BGP-LS NLRI types have been defined to help describe a Node, Link, an IPv4/IPv6 Prefix or a SRv6 SID
  - New BGP-LS Path Attribute has been defined to carry LS or TE information associated to links, nodes, prefixes and SRv6 SIDs.

Traffic Engineering – Centralized Path Computation and Optimization

- Path Computation Element Protocol – Overview
  - Originally defined in RFC 5440 and further extended in RFC8231, RFC8281, RFC8664, among others, to support RSVP-TE, SR-TE and SR-Policy with both MPLS and IPv6 data planes
  - PCEP: is a protocol used for communications between a PCC and a PCE. The protocol defines a set of messages and objects used to manage sessions, requests, replies, updates, and reports associated to LSPs and SR-Policies.
  - Path Computation Element (PCE): an entity that can compute paths that meet specific requirements/constraints based on local-policy or topology information received from the network
  - Path Computation Client (PCC): a client application, usually running on a router, that requests the PCE run path computations using relevant optimization objectives and constraints. PCC may also report and/or delegate control of LSPs and SR-Policies to a PCE

NANOG™

Traffic Engineering Non-IETF

David Watkinson
bio.site/dwatkinson

Non-IETF or Proprietary techs: why?

- Advanced Features
- Vendor Integration
- Performance and Optimization
- Customization
- Competitive Advantages
- Rapid Innovation

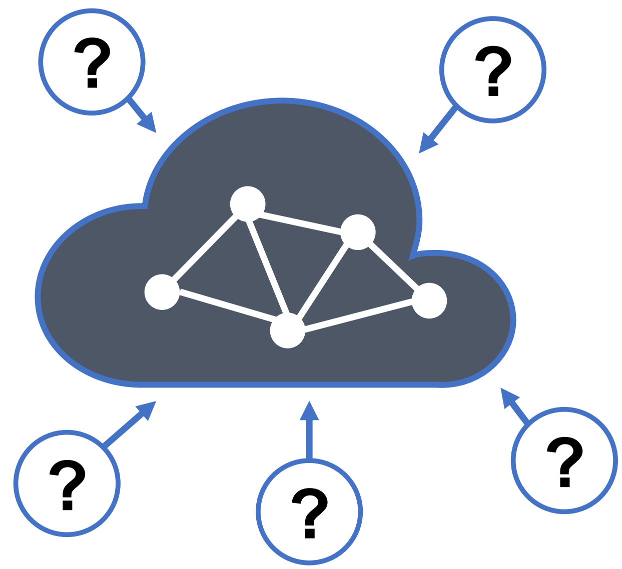
Why would we need this?

MPLS (RSVP-TE/SR-TE) is a great solution for **static or even stationary traffic patterns**

- However, traffic patterns become more **unpredictable and bursty** with changes due to factors such as 5G, streaming, cloud services, IoT, and remote work

The use of multiple LSPs, balanced via ECMP, brings better **network utilization and resilience**.

- However, LSPs are not elements that can be **added or removed on-demand**."



Closed Loop

This use of Closed Loop Automation involves

1. Measuring information about the network then
2. Reacting to those changes by changing the paths taken by LSPs in the network
3. With controls to ensure no instabilities arise

Closed Loop: Types of Telemetry

Utilization (Bandwidth)

- Use when trying to run networks at high capacities
- Reduces the probability of congestion related discards

Latency

- Use when Slices or Services need to be routed over a low latency path
- May need to use separate LSPs when both Utilization and Latency Optimization

Packet Loss

- Use when links have variable packet loss such as Microwave

Closed Loop – Use Cases

Real-world use cases illustrating how these technologies are applied:

Global Internet Service Provider (ISP) Network Optimization:

Use Case: A global ISP utilizes PCE along with telemetry data to optimize their MPLS-TE network. The PCE is responsible for calculating efficient routing paths based on real-time network conditions, including LSP bandwidth utilization.
Application: This approach allows the ISP to dynamically adjust and reroute traffic to manage congestion, enhance the quality of service, and ensure optimal utilization of network resources. For example, during peak **usage** times, the network can adapt to changing traffic patterns, avoiding bottlenecks.

Enhanced Mobile Backhaul Network Management:

Use Case: Mobile network operators implement MPLS-TE with PCE and telemetry to manage their backhaul networks, which are crucial for connecting cellular towers to the core network
Application: The integration of PCE and telemetry allows for real-time monitoring and dynamic adjustment of LSPs based on current network load and bandwidth utilization. This is particularly beneficial for handling the variable and high-volume traffic characteristic of mobile networks, especially with the advent of 5G.

Path Computation Element Intro

David Watkinson
bio.site/dwatkinson

Path Computation Element (PCE)

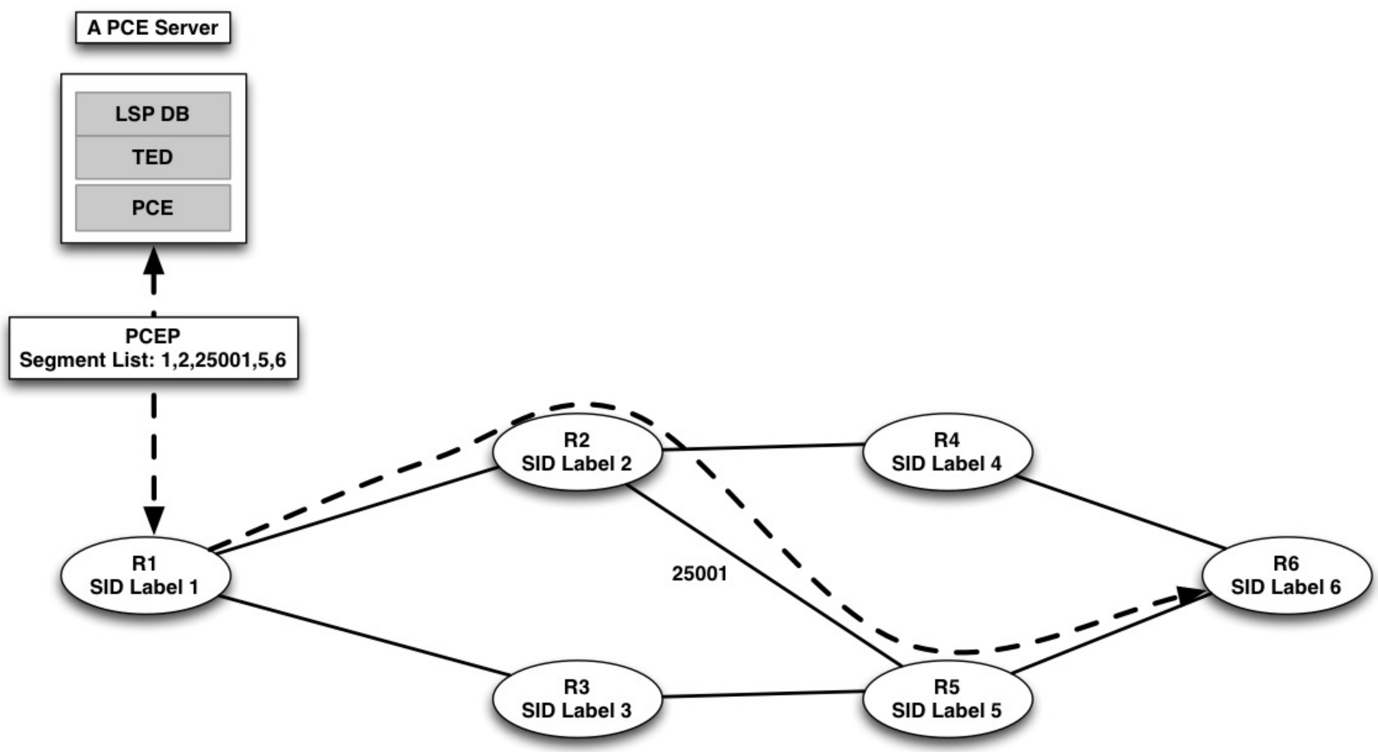
- PCE and the associated Protocol for Communication between PCE and PCC (PCEP) are IETF standards.
- PCE is responsible for computing paths for traffic engineering purposes.
- PCE is defined in RFC 4655
- PCEP is the protocol used for communication between a (PCC) and a PCE
- PCEP allows a PCC to request path computations from a PCE and receive computed paths
- PCEP is defined in RFC 5440
- Stateful extensions to PCEP are defined in RFC 8231
- **Yes, you can consider a PCE (Path Computation Element) controller as a traffic engineering technique**

NANOG™

PCE Benefits

- **Reduced operational complexity**
- **Unified network view**
- Centralized path computation
- Optimized network routing
- Real-time adaptability
- Enhanced scalability
- Efficient resource utilization

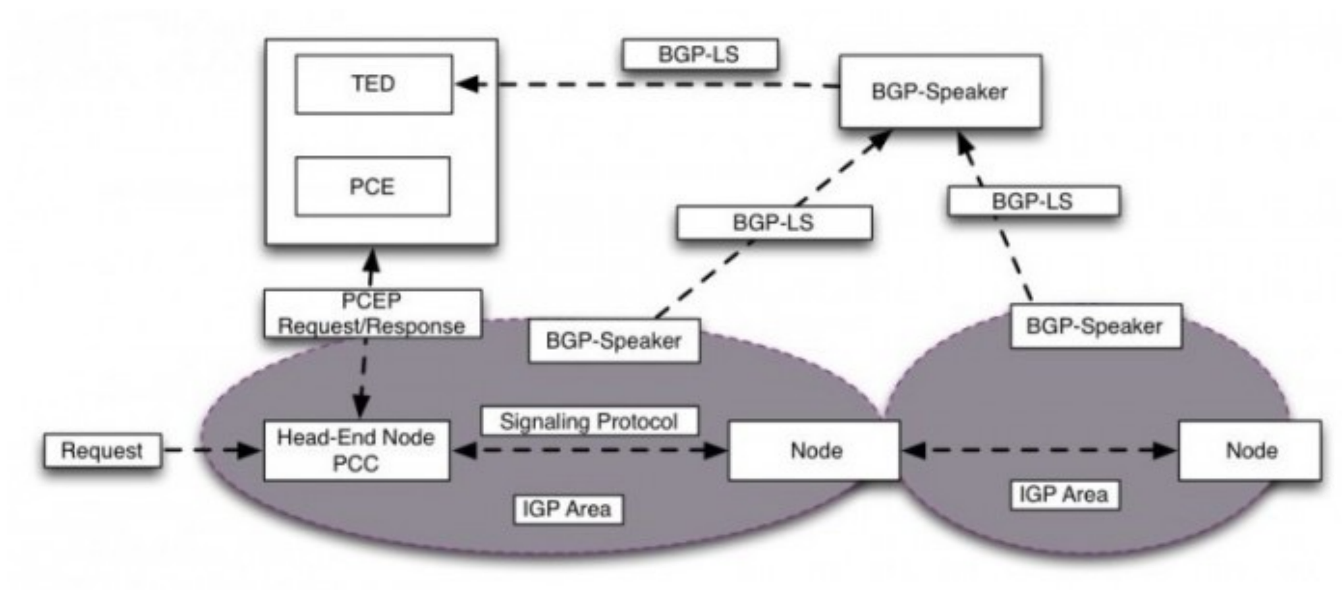
PCE Architecture



NANOG™

Image Source: Packet Pushers

PCE Architecture



NANOG™

Image Source: Packet Pushers

PCEP Key Messages and Objects

- Session
  - PCEP Open
  - PCEP Keepalive
  - PCEP Close
- Stateless Messages
  - PCEP Request PCC->PCE
  - PCEP Reply PCE->PCC
- Stateful Messages
  - PCEP Report PCC->PCE
  - PCEP Update PCE->PCC
  - PCEP Initiate PCE->PCC
- Stateless Objects
  - Request Parameters (RP)
  - Endpoints (SRC + DST)
  - Bandwidth
  - Metric
  - LSPA (Admin Groups, Priority)
  - ERO PCE->PCC
- Stateful Objects
  - LSP (for identification)

NANOG™

PCEP use Cases

Traffic Engineering Path Setup	Network Optimization	Resource Allocation	Path Re-optimization
• Setting up traffic-engineered paths when a new LSP needs to be established. • Compute a path that minimizes latency while avoiding congested links, ensuring optimal performance.	• Optimizing network resources. • Compute and recommend changes to existing LSPs • Suggest the creation of new LSPs to better utilize available network resources • Can dynamically reroute LSPs to balance the traffic load	• Determining how to allocate bandwidth, labels, and other resources for LSPs based on their requirements • Preventing one tenant from monopolizing resources	• Dynamic path re-optimization in response to network changes, such as link failures or changes in traffic demand. • A link failure occurs due to maintenance work. The PCE detects this event and quickly computes new paths for affected LSPs to bypass the failed link, minimizing service disruption

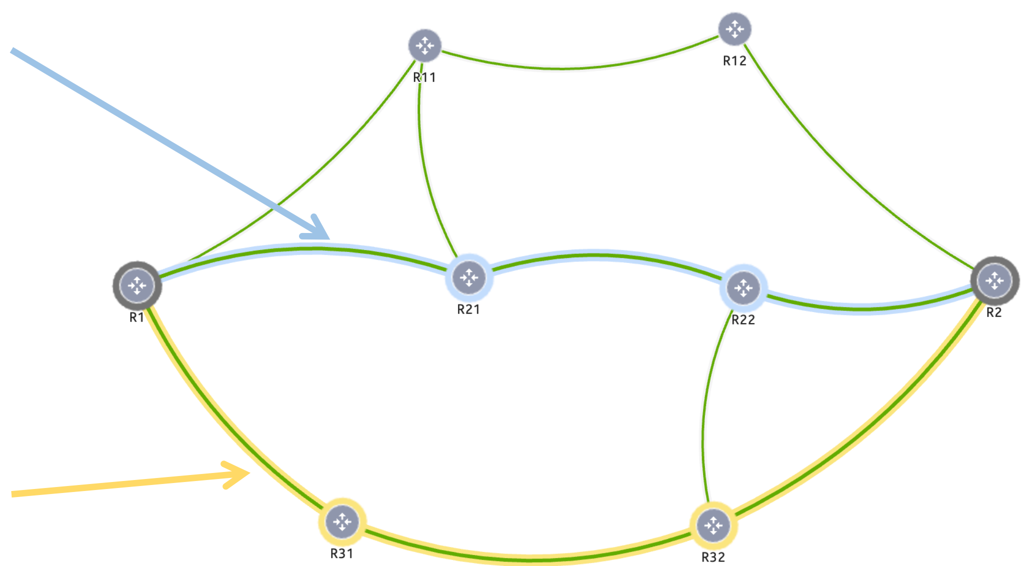
Real-World Examples

- **Telecom Service Provider:**
  - Use Case: A telecom service provider uses PCEP to **optimize its core network**. The PCE continuously **evaluates traffic patterns, rerouting LSPs** as needed to maximize network efficiency.
  - Example: During a **sudden spike in video streaming traffic** due to a live sporting event, the PCE detects congestion on certain links. It **quickly re-routes traffic** by rerouting existing LSPs or establishing new LSPs along less congested paths, ensuring high-quality streaming for customers.

Example: Disjoint Link_Strict

```
"name": "LSP-1",
... # this is an extract
"association-objects": {
    "association-object-extended": [
    {
        "association-key" : "nokia-path-profile-1",
        "id": "10102",
        "extended-id" : "3E"
```

```
"name": "LSP-2",
... # this is an extract
"association-objects": {
    "association-object-extended": [
    {
        "association-key" : "nokia-path-profile-1",
        "id": "10102",
        "extended-id" : "3E"
```



NANOG

Summary

- Importance of PCE and PCEP in Traffic Engineering.
- Additional Resources
    - https://packetpushers.net/pce-pcep-overview/

IETF APIs for TE (YANG)

David Watkinson
bio.site/dwatkinson

Yet Another Next Generation

- YANG (Yet Another Next Generation) is crucial in network technologies for several reasons
  - Standardization
  - Interoperability
  - Automation
  - Extensibility
  - Human Readability
  - Integration with GRPC, NETCONF and RESTCONF

IETF Tunnel (LSP) Model

- Link: https://datatracker.ietf.org/doc/draft-ietf-teas-yang-te/
- YANG Data Model
  - Describes a YANG data model for Traffic Engineering (TE) tunnels, Label Switched Paths (LSPs), and interfaces.
  - Divided into two modules: 'ietf-te.yang' (generic and device-independent) and 'ietf-te-device.yang' (device-specific).

IETF Network Topology

- General
  - Composed of nodes, links, and termination points.
  - Layered model of maps
  - https://datatracker.ietf.org/doc/rfc8795/
- L2 is Ethernet view or layer
  - https://datatracker.ietf.org/doc/rfc8944/
- L3 is IP and IGP view or layer
  - https://datatracker.ietf.org/doc/rfc8346/
- L4 is Service view or layer
  - Service Attachment Point (SAP) : point of interconnection with customer

RSVP-TE/SR-TE and bidirectional (POST)


```
{
    "tunnel": [
        {
            "name": "LSP1",
            "encoding": "ietf-te-types:lsp-encoding-packet",
            "admin-state": "ietf-te-types:tunnel-admin-state-up",
            "bidirectional": "true",
            "signaling-type": "ietf-te-types:path-setup-sr",
            "source": "10.30.1.1",
            "destination": "10.30.1.2"
        }
    ]
}
```

BGP Link State

David Watkinson
bio.site/dwatkinson

BGP Link State

- BGP-LS allows for the aggregation of routing information from multiple IGP domains
- BGP-LS is defined in
  - RFC 7752 using a new address family of 16388
  - RFC 8571 for performance metric extensions
  - RFC 9086 for adding eBGP links for Egress Peer Engineering
- IGP Area information can be imported into BGP
- A PCE only needs one connection to get all TE info

NANOG™

Telemetry (GRPC)

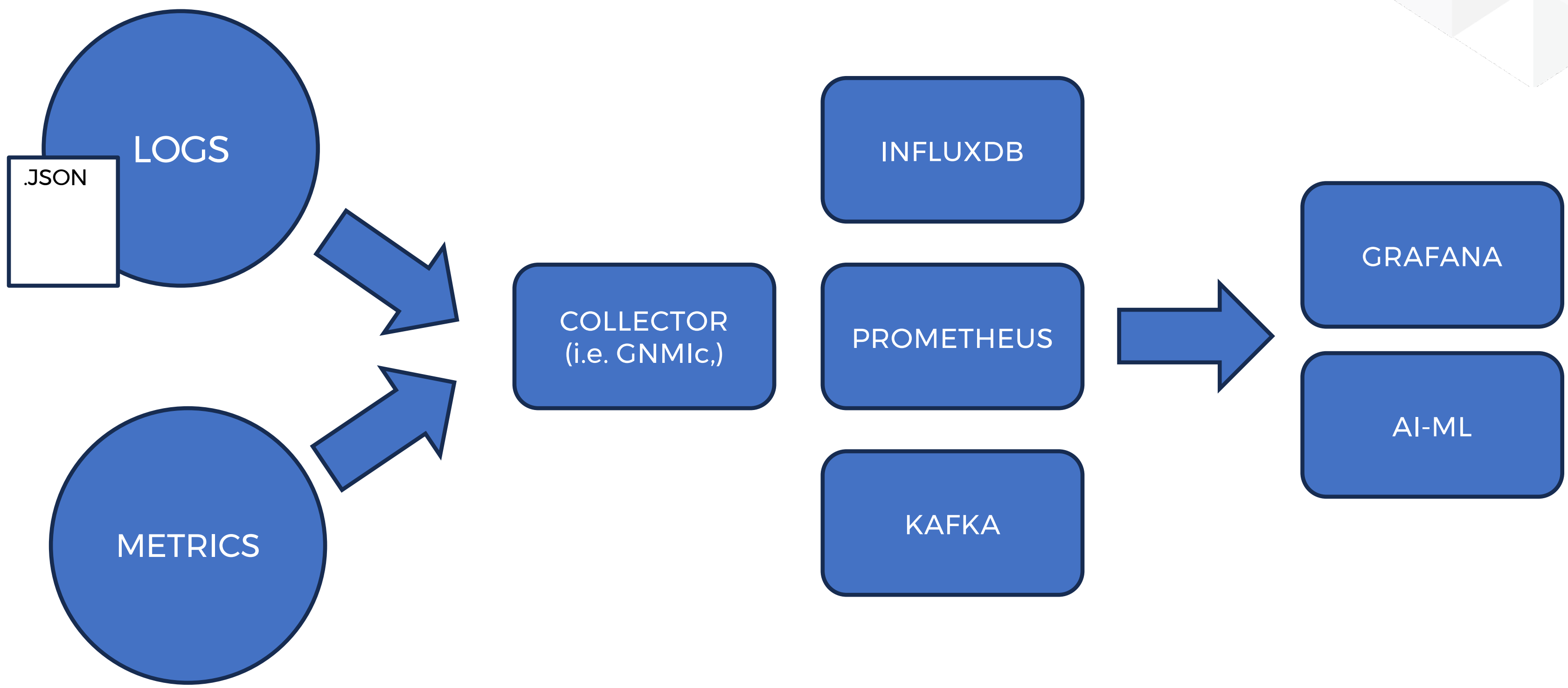
Mau Rojas

bio.site/pinrojas

Why Telemetry?

- Proactive Issue Detection
  - Early detection of network anomalies or issues, such as increased traffic, packet loss, or device failures, before they impact user experience.
- Performance Optimization
  - Real-time data collection helps in optimizing network performance by identifying bottlenecks, optimizing routing, and ensuring efficient resource utilization.
- Capacity Planning
  - Telemetry data aids in capacity planning by providing insights into long-term trends, helping organizations scale their network infrastructure as needed.
- Security
  - Crucial role in network security by detecting and responding to suspicious or unauthorized activities promptly.

Telemetry Architecture



NANOG

Telemetry Protocols

- SNMP

  Protocol for network device management.

  Use Case: Monitor router performance metrics (CPU usage, bandwidth) to ensure network reliability.

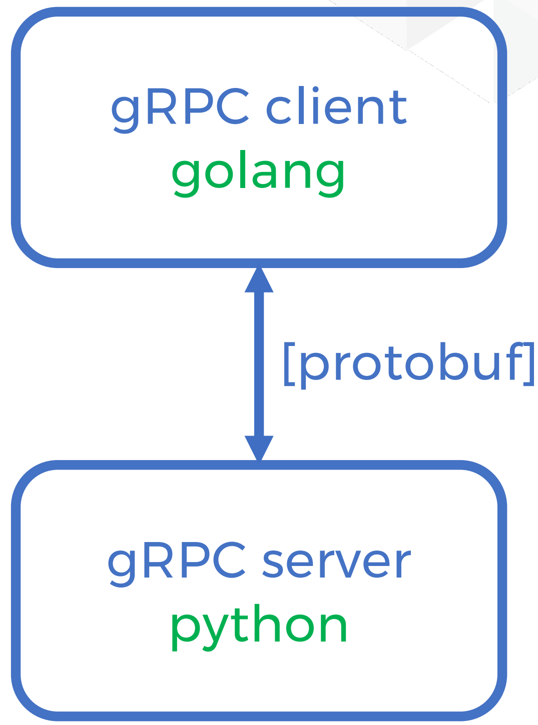
- gRPC

  High-performance RPC framework for distributed systems.

  Use Case: Efficiently exchange real-time telemetry data among microservices in a cloud-based network.

What is GNMI?

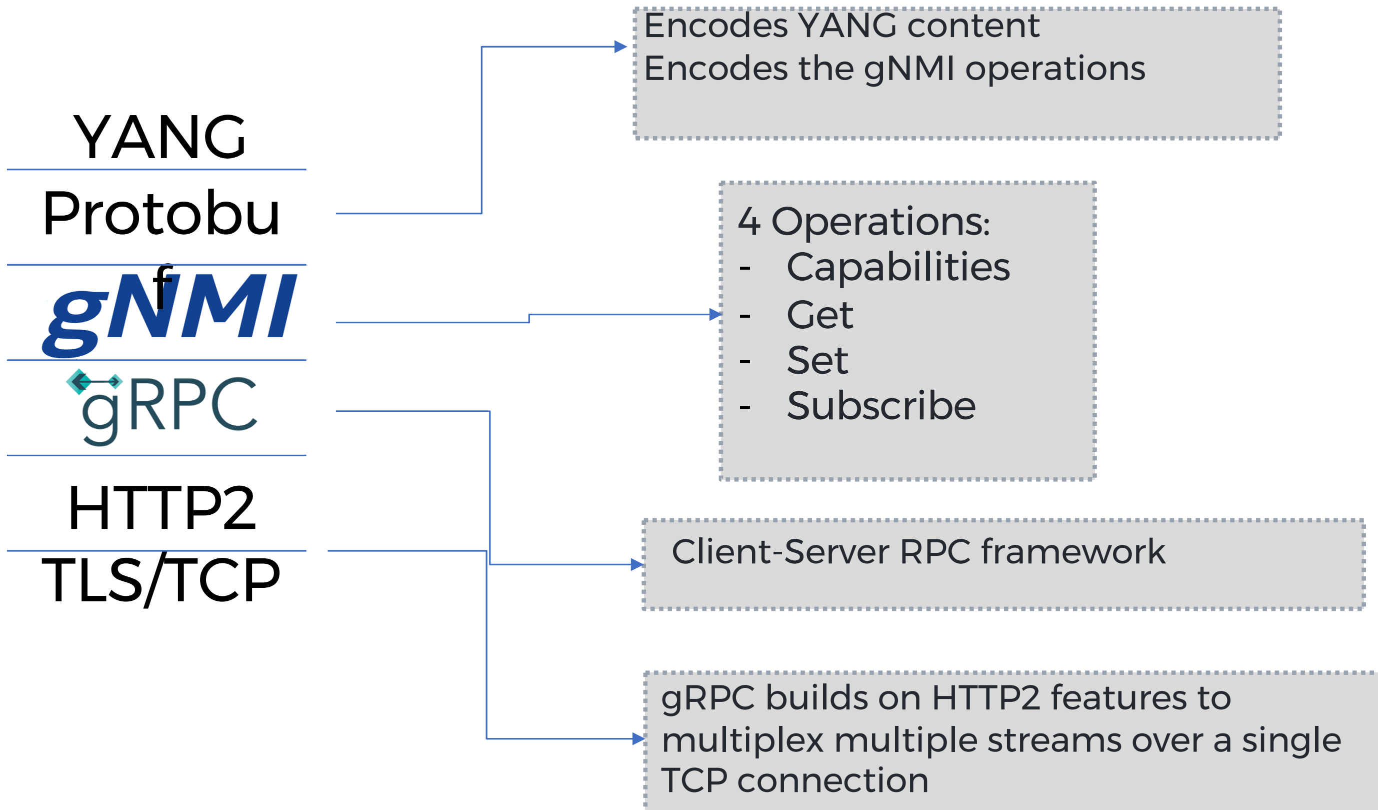
- gRPC Network Management Interface
- Open-source protocol developed by Google
- Versatile, efficient, and scalable
  - Protocol buffers
- Retrieve and configure network state information



NANOG™

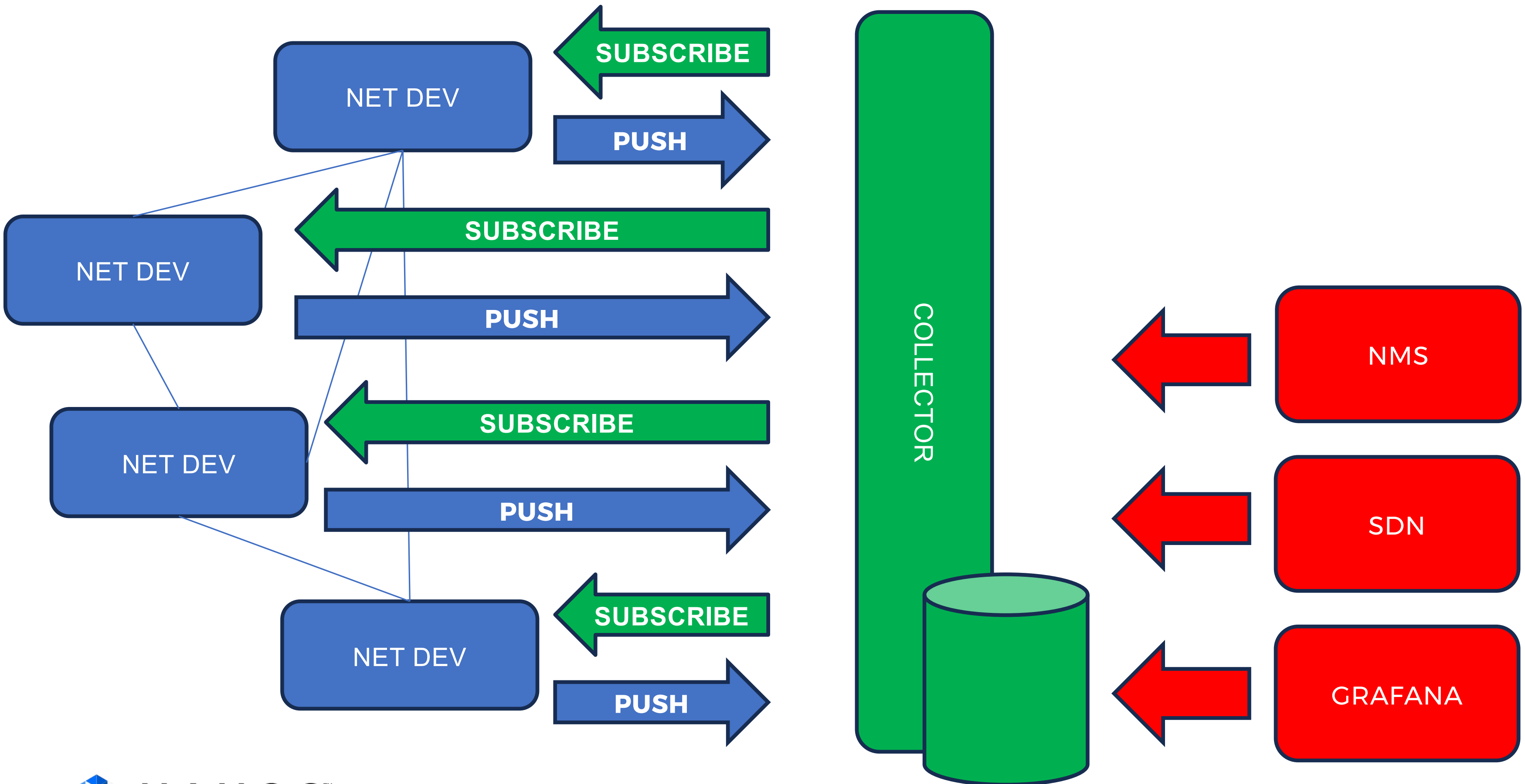
Source: https://gnmic.kmrd.dev/

gRPC Network Management Interface



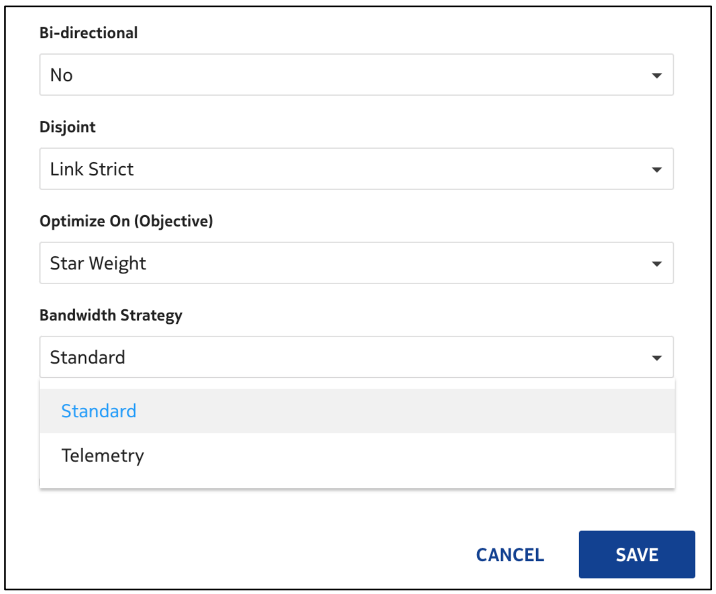
NANOG

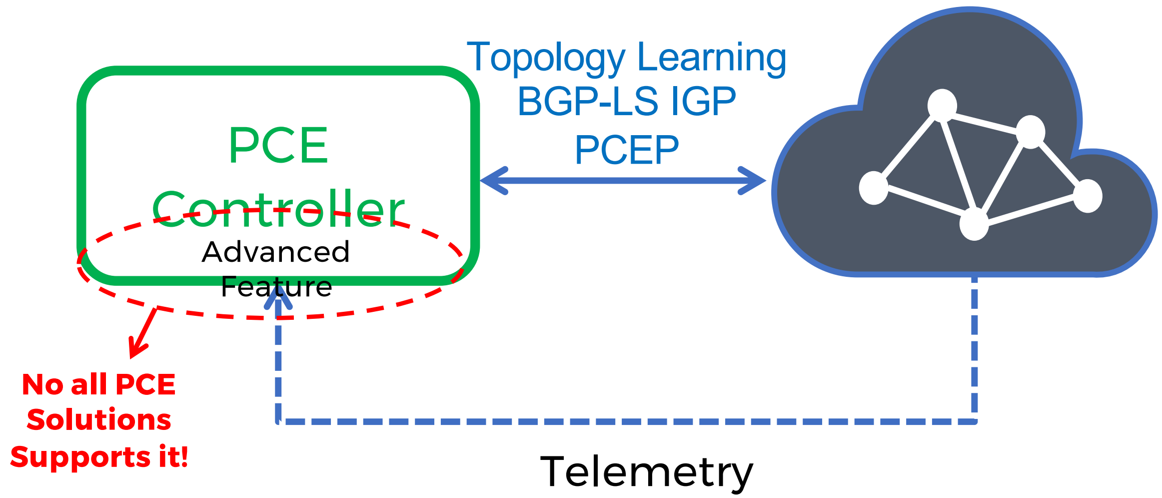
GRPC Subscription



NANOG™

Closed Loop: PCE + Telemetry





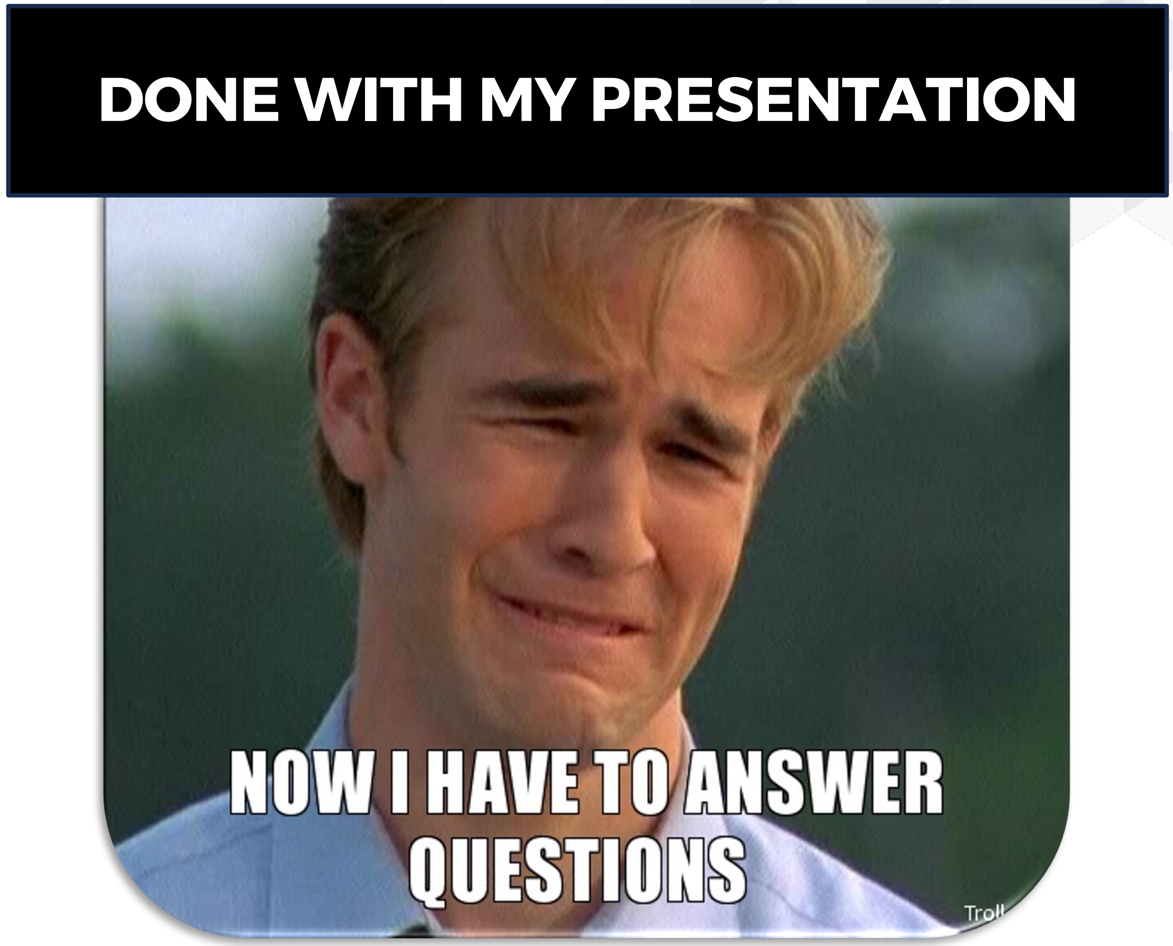
Almost done!

Additional resources

Needs more work

- Containerlab – NANOG talk - Running networking labs with Docker UX – Roman and Karim
  - https://youtu.be/qigCla1qY3k
- gNMIc – NANOG Talk - An intuitive gNMI CLI and a feature-rich telemetry collector - Karim
  - https://youtu.be/v3CL2vrGD_8
- Importance of PCE and PCEP in Traffic Engineering:
  - https://packetpushers.net/pce-pcep-overview/

Thanks!



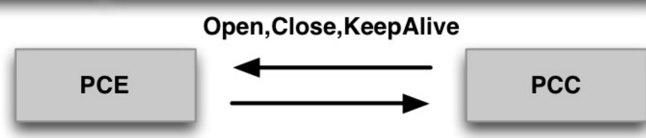
NANOG™

Backup Material

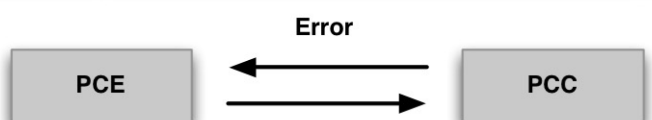
NANOG™

PCEP Messages

Messages	Purpose	PCE Type
Open, Close, Keep-alive	These are basically housekeeping messages. Open Message is also use for capability negotiation between PCE and PCC.	Stateless and Stateful



Messages	Purpose	PCE Type
Error Messages	Used for Error handling between PCE and PCC in case something goes wrong.	Stateless and Stateful



NANOG™

Image Source: Packet Pushers

PCEP Messages

Messages	Purpose	PCE Type
PCC Report	PCC Report messages from PCC are used for multiple functions: 1) State Synchronization After the session between PCC and a stateful PCE is initialized, PCE must synchronize the state of a PCC‘s LSPs before it performs path computations or update LSP attributes in a PCC. 2) LSP State Report A PCC sends an LSP state report to a PCE whenever the state of an LSP changes. 3) LSP Control Delegation A PCC grants to a PCE the right to update LSP attributes on one or more LSPs; the PCE becomes the authoritative source of the LSP‘s attributes as long as the delegation is in effect; the PCC may withdraw the delegation or the PCE may give up the delegation.(active stateful PCE only)	Stateful PCE

NANOG™

Image Source: Packet Pushers

PCEP Messages

Messages	Purpose	PCE Type
Path Computation Request and Reply	PCE Request is sent by PCC to PCE asking for Path computation. PCE Reply is sent by PCE to PCC in response of PCE Request.	Stateless and Stateful

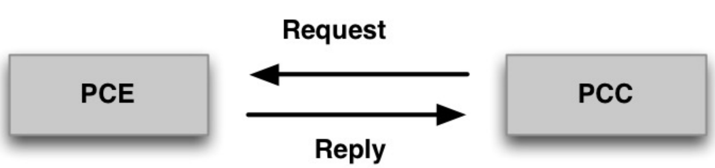


Image Source: Packet Pushers

PCEP Messages

Messages	Purpose	PCE Type
PCC Update	A PCE requests updates of attributes on a LSP. (active stateful PCE only). PCC reports back with the status report messages to the PCE.	Stateful PCE

Messages	Purpose	PCE Type
PCE Initiate	Used by PCE to Initiate LSPs on PCC. A PCC reports back on the status to the PCE.	Stateful PCE

NANOG™

Image Source: Packet Pushers