SDN for SP Networks
2022 Edition

Phil Bedard – Principal Engineer

NANOG84

Agenda

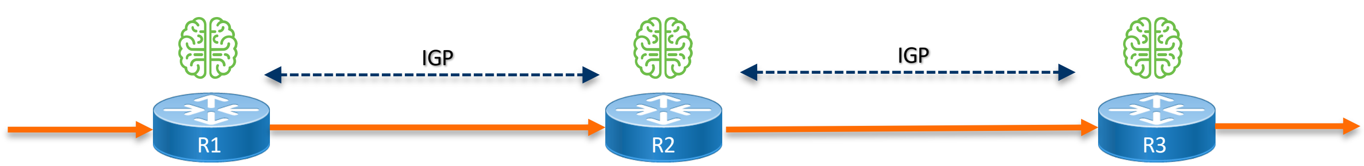
- What is “SDN” and what is an “SDN Controller”
- Current industry standards and implementation initiatives
- Network use cases

🚧 Caution! Industry buzzword alert!

SDN Background

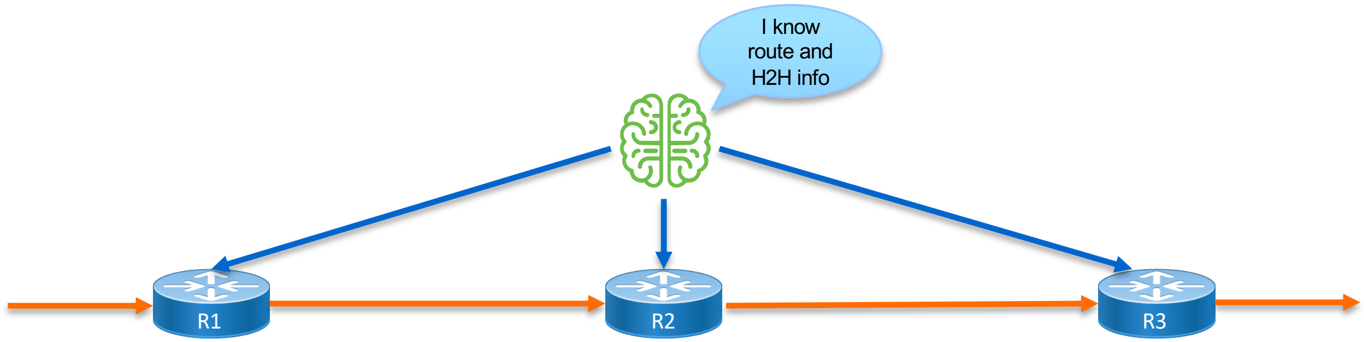
What is a ”non-SDN” network?

- **What most networks are today and have been for the last several decades**
- Network has a distributed control plane using protocols to communicate routing and forwarding information across the network
- This does not mean it is not intelligent, we’ve added all kinds of protocols and extensions to enhance distributed control planes and influence end-to-end forwarding
  - RSVP, RSVP-TE, Segment Routing, BGP-Flow Spec, onboard applications like Auto-Bandwidth
  - In some cases, the head-end node influences forwarding across the network, but the intelligence is still embedded within the device



What is Software Defined Networking?

- **Purist view**: Physical separation of control and data plane, external control plane is responsible for all management and routing functions
- Centralized (off-box) control plane makes all routing and forwarding decisions
- Simplified hardware was one of the drivers
  - Standard interface for programming routing/forwarding tables at each hop
  - Standard languages to define forwarding mechanics: Openflow, P4



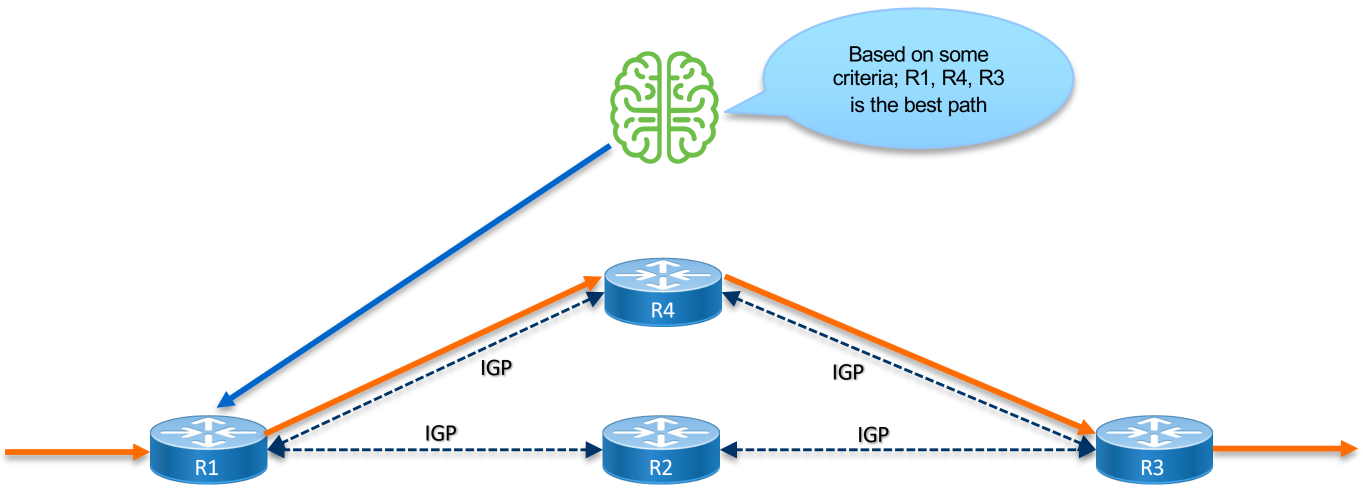
R1 Instructionc	
SrcIP	1.1.1.1
DstIP	2.2.2.2
OIF	Eth1/1
DMAC	57:ff:de:cd:00:01

R2 Instruction	
SrcIP	1.1.1.1
DstIP	2.2.2.2
OIF	Eth1/3
DMAC	57:cd:93:ce:ac:27

R3 Instruction	
SrcIP	1.1.1.1
DstIP	2.2.2.2
OIF	Eth1/27
DMAC	57:ce:af:73:21:32

What is SDN? – "Hybrid" SDN

- **Pragmatic view:** Common distributed networking with additional intelligence
- Network programmability through standard protocols and open interfaces
- Can augment different layers of the network: Service endpoints, overlay paths, device-level forwarding
- Delegated PCE is a very simple example



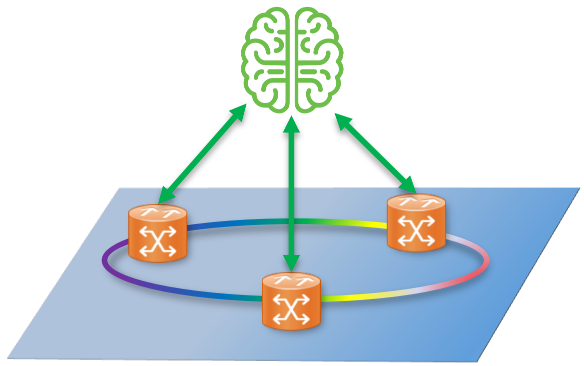
R1 Instruction	
SrcIP	1.1.1.1
DstIP	2.2.2.2
Path	R1, R4, R3

Intent Based Networking (IBN)

- Often goes hand in hand with “SDN”
- IBN drives network configuration from desired end goal
- Declarative configuration driven by higher layer requirements
- ”I need L2 connectivity between applications A and Z”
  - Software decomposes ask into network services and configuration
  - Each component is only responsible for the components it knows about
- Key is to create abstract representations of services and networks higher layer elements can consume

What is an "SDN controller"?

- A "controller" directs the actions or function of (something), in this case network devices or other network controllers
- Southbound adapters to interact with network or other controllers
- Aggregates standard and proprietary network data
  - Network data is the foundation of more intelligent networks
  - Consumed by controller applications



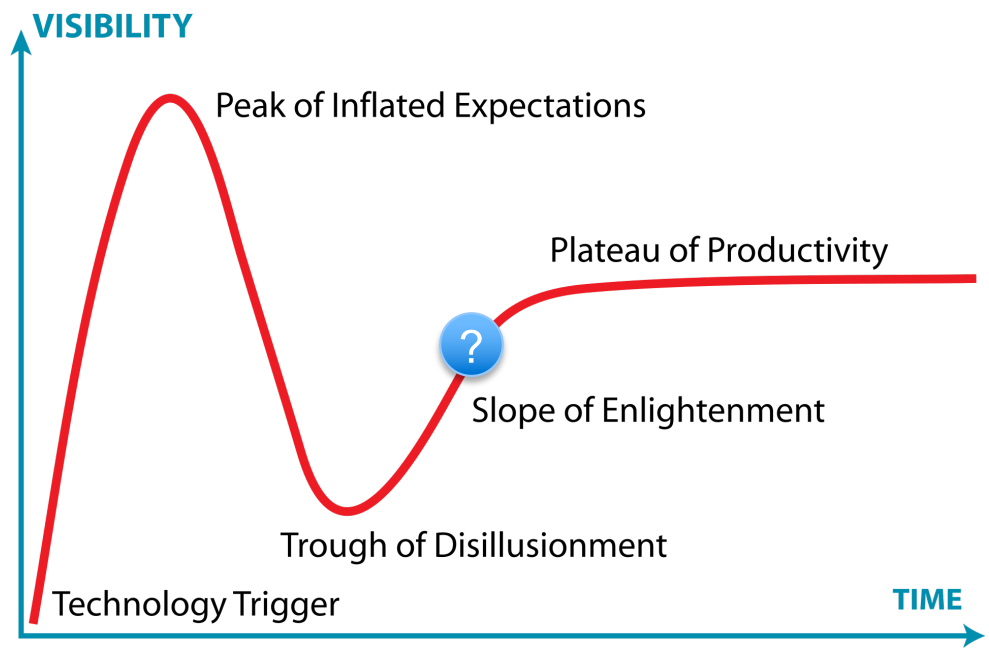
- Render north bound API interfaces
  - Allows other controllers to consume data or perform application operations on its domain elements
- Contains intelligent network applications
  - Translate Intent into network configuration

Why use SDN?

- Enhance network decision making with data that's not feasible to use as input into a distributed on-box routing protocol
  - Embedding complex functions into device level software is difficult
  - Upgrading device software is still relatively painful
- Intelligent multi-layer networks
  - GMPLS is a distributed IP+Optical control plane but has not seen adoption
  - Is SDN the savior of failed device-level control plane interop?
- Optimize network resources
- Ancillary benefits include network visibility and fault identification/correlation

Who is using SDN today?

- Some have been using “SDN” for some time, EG: Content providers with more advanced automation
- Majority of SPs are not using what I would call “SDN” but many are looking at it



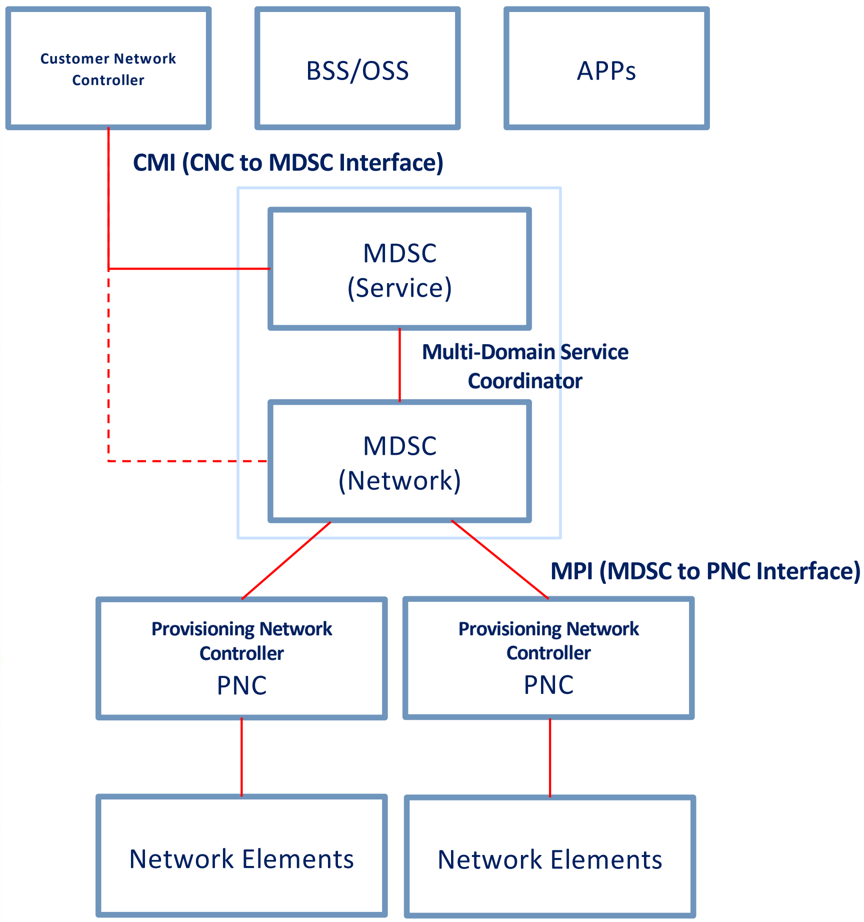
SDN Standards and Initiatives

What allows SDN controller interop?

- Standard architecture definition for how controllers interact
- Standard data models
  - YANG is the de facto standard language today
  - Standards covering different domains and technologies
  - Device-level, controller-level, service-level
  - Necessary to normalize proprietary data into vendor-agnostic data
- Standard interface protocols
  - NETCONF to devices is widely implemented and deployed
  - gNMI, gNOI are additional standards for device-level interaction
  - RESTCONF has become the de facto standard for controller to controller interaction

IETF ACTN

Abstraction and Control of Traffic Engineered Networks



- RFC8453 defines hierarchical framework of controllers
- Assigns roles and responsibilities to different elements
- Framework defines loose coupling between components but not interfaces and encoding

HCO or "Hierarchical Controller" is a common name for MDSC

"Domain controller" is a common name for PNC

IETF ACTN – Controller roles

- CNC – Customer Network Controller
  - Highest level orchestration / workflow system defining intent
- MDSC (Multi-Domain Service Controller) or HCO
  - Aggregate information from downstream controllers to perform multi-domain tasks, for example IP+Optical provisioning or stitching a single service across two ASNs
- PNC – “Provisioning Network” aka ”Domain” controller
  - Communicates using SBI to network elements for provisioning, fault, and performance data collection
  - May be proprietary or standard SBI to devices

IETF – Additional Topics

- Service level standard models
  - IETF L2NM for L2VPN, L3NM for L3VPN
- Network models
  - Network topology and TE models
- Continued work on enabling protocols such as PCEP, IS-IS, BGP-LS extensions, Network Telemetry Framework

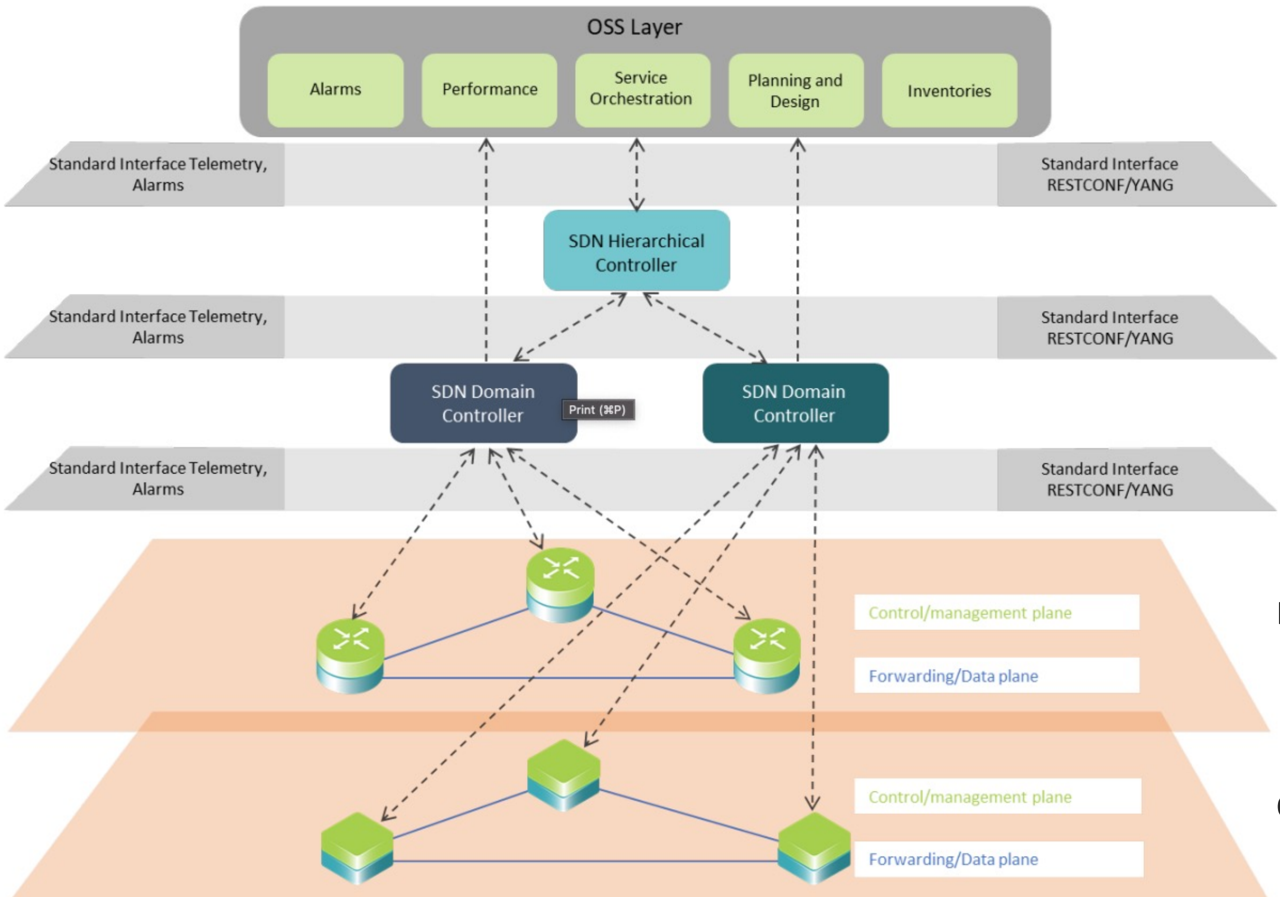
ONF Telecom Infra Project

TELECOM INFRA PROJECT

- OOPT – Open Optical & Packet Transport Project Group
- TIP OOPT MUST - Mandatory Use Case Requirements for Transport SDN
  - Focused on defining a SDN controller-based architecture like ACTN, but adds transport protocols, encoding, and model requirements based on standards or de-facto standards
  - Covering IP/MPLS and Optical transport networks
  - MUST has generated controller requirements documents for the NBI/SBI models, interfaces, along with the use cases which must be supported
- Driven primarily by operators, not vendors

https://telecominfraproject.com/oopt/

TIP MUST Open Transport SDN



Similar structure to ACTN

Communication is not always hierarchical, only where it makes sense

IP/MPLS/SRv6 Network

Optical DWDM Network

ONF OTCC Project and Transport API (T-API)

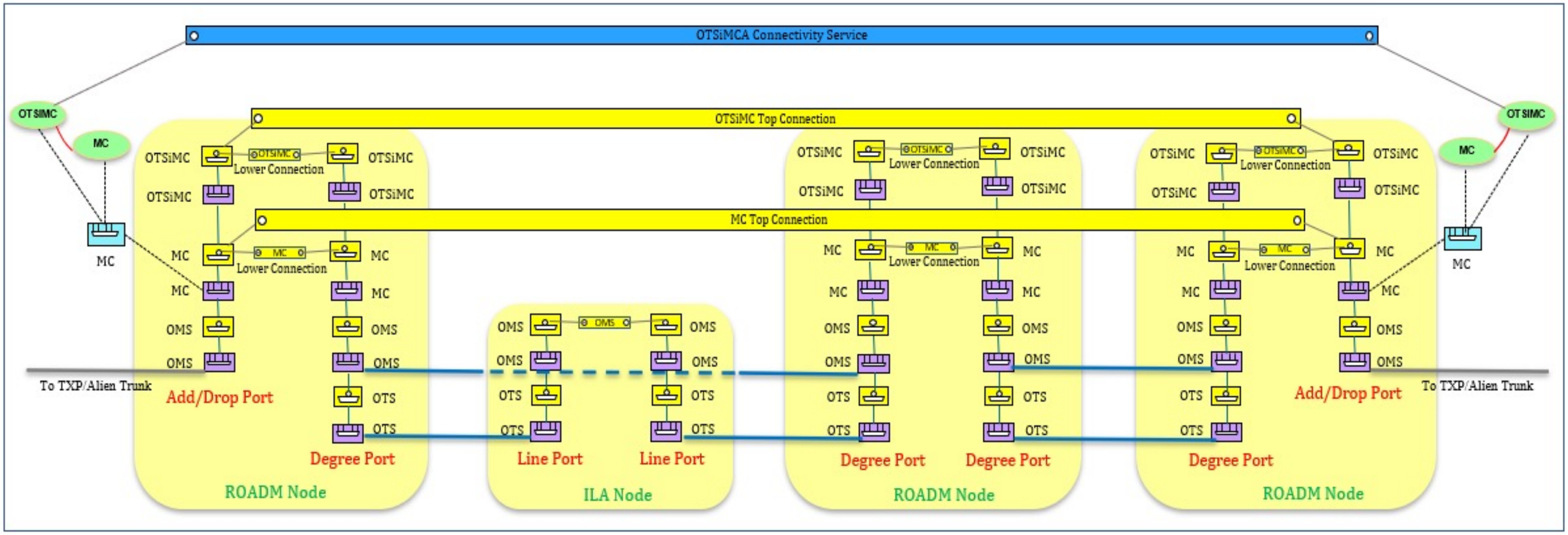
Open Transport Configuration and Control

- Defines a set of models for managing optical transport networks
  - Equipment inventory, topology, connectivity-service, fault, and performance covered in defined models
  - Used by external systems, T-API does not define SBI from controller to device

- Has become the de-facto standard for optical domain controllers
  - Still in early stages of implementation, many differences between implementations still
  - OIF (Optical Interworking Forum) public interop events
  - Used by TIP MUST and other groups defining architectural standards

https://wiki.opennetworking.org/display/OTCC/TAPI+Documentation

T-API Connectivity Service



- Hierarchical services definitions
- Hierarchical topology

T-API Connectivity Service


```json
{
"tapi-connectivity:connectivity-context": {
"connectivity-service": [
{
"uuid": "0695d528-ad1c-4ba5-b4af-8a55e9ce1a65",
"end-point": [
{
"local-id": "EndPoint11",
"layer-protocol-name": "PHOTONIC_MEDIA",
"layer-protocol-qualifier": "tapi-photonic-media:PHOTONIC_LAYER_QUALIFIER_MC",
"service-interface-point": {
"service-interface-point-uuid": "12256896-f41c-30e7-846a-a3cc48f929d8"
},
"connection-end-point": [
{
"topology-uuid": "4b1b5fac-a97f-32bc-af8a-7fd5cec82ad7",
"node-uuid": "97ae548c-0632-3d23-8fb4-a4614c1e50b1",
"node-edge-point-uuid": "500680c6-ce80-398b-bb78-580c020d396a",
"connection-end-point-uuid": "b24c67c4-d036-34ba-94fd-649828e68465"
}
],
"direction": "BIDIRECTIONAL",
"role": "SYMMETRIC",
"protection-role": "NA",
"name": [
{
"value-name": "CONN_SERVICE_END_POINT_NAME",
"value": "EndPoint11_mc"
}
],
"administrative-state": "UNLOCKED",
"operational-state": "ENABLED",
"lifecycle-state": "INSTALLED",
"tapi-photonic-media:mca-connectivity-service-end-point-spec": {
"mc-config": [
{
"local-id": "EndPoint11_mc",
"name": [
{
"value-name": "CSEP_MC_NAME",
"value": "EndPoint11_mca_csep"
}
]
}
],
"number-of-mc": "1"
}
},
{
"local-id": "EndPoint12",
"layer-protocol-name": "PHOTONIC_MEDIA",
"layer-protocol-qualifier": "tapi-photonic-media:PHOTONIC_LAYER_QUALIFIER_OTSiMC",
"service-interface-point": {
"service-interface-point-uuid": "12256896-f41c-30e7-846a-a3cc48f929d8"
},
"connection-end-point": [
{
"topology-uuid": "4b1b5fac-a97f-32bc-af8a-7fd5cec82ad7",
"node-uuid": "97ae548c-0632-3d23-8fb4-a4614c1e50b1",
"node-edge-point-uuid": "ab81e7cf-bd39-3332-9c8b-f72c9c1690b8",
"connection-end-point-uuid": "15fa01b2-da13-32b5-b7bc-34565221014e"
}
],
"direction": "BIDIRECTIONAL",
"role": "SYMMETRIC",
"protection-role": "NA",
"server-connectivity-service-end-point": {
"connectivity-service-uuid": "0695d528-ad1c-4ba5-b4af-8a55e9ce1a65",
"connectivity-service-end-point-local-id": "EndPoint11"
{
"name": [
{
"value-name": "CONN_SERVICE_END_POINT_NAME",
"value": "EndPoint12_otsimc"
}
],
"administrative-state": "UNLOCKED",
"operational-state": "ENABLED",
"lifecycle-state": "INSTALLED",
"tapi-photonic-media:otsia-connectivity-service-end-point-spec": {
"otsi-config": [
{
"local-id": "EndPoint12_otsi",
"central-frequency": {
"central-frequency": "192350000"
},
"application-identifier": {
"application-identifier-type": "PROPRIETARY",
"application-code": "00B08E#QDD-400G-ZRP-S#1#5005"
},
"transmit-power": {
"total-power": "-11.5"
}
}
],
"number-of-otsi": "1"
},
"tapi-photonic-media:mca-connectivity-service-end-point-spec": {
"mc-config": [
{
"local-id": "EndPoint12_mc",
"name": [
{
"value-name": "CSEP_OTSI_MC_NAME",
"value": "EndPoint12_otsi_mc"
}
]
}
],
"number-of-mc": "1"
}
},
{
"local-id": "EndPoint21",
"layer-protocol-name": "PHOTONIC_MEDIA",
"layer-protocol-qualifier": "tapi-photonic-media:PHOTONIC_LAYER_QUALIFIER_MC",
"service-interface-point": {
"service-interface-point-uuid": "3c3ccf81-4e9b-381a-9a6d-a93a45641dc9"
},
"connection-end-point": [
{
"topology-uuid": "4b1b5fac-a97f-32bc-af8a-7fd5cec82ad7",
"node-uuid": "0fbc938b-4362-3eca-95b3-02293ec7dc1b",
"node-edge-point-uuid": "56025e0b-256e-32bb-a17f-1a2dfdf3dcef",
"connection-end-point-uuid": "b64f4832-ac3f-3f74-a9b0-6e4045c3b36e"
}
],
"direction": "BIDIRECTIONAL",
"role": "SYMMETRIC",
"protection-role": "NA",
"name": [
{
"value-name": "CONN_SERVICE_END_POINT_NAME",
"value": "EndPoint21_mc"
}
],
"administrative-state": "UNLOCKED",
"operational-state": "ENABLED",
"lifecycle-state": "INSTALLED",
"tapi-photonic-media:mca-connectivity-service-end-point-spec": {
"mc-config": [
{
"local-id": "EndPoint21_mc",
"name": [
{
"value-name": "CSEP_MC_NAME",
"value": "EndPoint21_mca_csep"
}
]
}
],
"number-of-mc": "1"
}
},
"local-id": "EndPoint22",
"layer-protocol-name": "PHOTONIC_MEDIA",
"layer-protocol-qualifier": "tapi-photonic-media:PHOTONIC_LAYER_QUALIFIER_OTSiMC",
"service-interface-point": {
"service-interface-point-uuid": "3c3ccf81-4e9b-381a-9a6d-a93a45641dc9"
},
"connection-end-point": [
{
"topology-uuid": "4b1b5fac-a97f-32bc-af8a-7fd5cec82ad7",
"node-uuid": "0fbc938b-4362-3eca-95b3-02293ec7dc1b",
"node-edge-point-uuid": "01fc37b0-61c4-3b1c-bb07-ec79783b9156",
"connection-end-point-uuid": "5230d167-c187-343f-9de2-fe6db5d9f60e"
}
],
"direction": "BIDIRECTIONAL",
"role": "SYMMETRIC",
"protection-role": "NA",
"server-connectivity-service-end-point": {
"connectivity-service-uuid": "0695d528-ad1c-4ba5-b4af-8a55e9ce1a65",
"connectivity-service-end-point-local-id": "EndPoint21"
},
"name": [
{
"value-name": "CONN_SERVICE_END_POINT_NAME",
"value": "EndPoint22_otsimc"
}
],
"administrative-state": "UNLOCKED",
"operational-state": "ENABLED",
"lifecycle-state": "INSTALLED",
"tapi-photonic-media:otsia-connectivity-service-end-point-spec": {
"otsi-config": [
{
"local-id": "EndPoint22_otsi",
"central-frequency": {
"central-frequency": "192350000"
},
"application-identifier": {
"application-identifier-type": "PROPRIETARY",
"application-code": "00B08E#QDD-400G-ZRP-S#1#5005"
},
"transmit-power": {
"total-power": "-11.5"
}
}
],
"number-of-otsi": "1"
},
"tapi-photonic-media:mca-connectivity-service-end-point-spec": {
"mc-config": [
{
"local-id": "EndPoint22_mc",
"name": [
{
"value-name": "CSEP_OTSI_MC_NAME",
"value": "EndPoint22_otsi_mc"
}
]
}
],
"number-of-mc": "1"
}
}
],
"connection": [
{
"connection-uuid": "04e632ed-2ed2-3825-8b33-6b05c134cf86"
},
{
"connection-uuid": "224abed2-664a-3380-a89c-0ce0cadf9ecf"
}
],
"name": [
{
"value-name": "SERVICE_NAME",
"value": "ac7c171f_ddac_45f0_8e3b_55f1aee4956d"
}
],
"service-layer": "PHOTONIC_MEDIA",
"service-type": "POINT_TO_POINT_CONNECTIVITY",
"connectivity-direction": "BIDIRECTIONAL",
"is-exclusive": false,
"administrative-state": "UNLOCKED",
"operational-state": "ENABLED",
"lifecycle-state": "INSTALLED"
}
```

SDN Controller Data Collection

Network Data – Packet

- Realtime topology data is needed to make accurate traffic routing decisions

- Network to IP domain controller
  - BGP-Link State (BGP-LS) is de facto standard today for IGP data
  - PCEP is de facto standard for Traffic Engineering tunnel information (RSVP-TE/Segment Routing)
  - Direct device interrogation using CLI (screen-scraping) or SNMP, pushed from device using telemetry

- Controller to controller
  - RFC 8345 (Base network model) is used today for topology data
    - RFC 8346 covers L3 topology, 8944 L2 topology, 8795 Traffic Engineering
    - Drafts for Segment Routing and other attributes
  - Additional IP Traffic Engineering (RSVP-TE LSP, SR Policy) data is conveyed via IETF TEAS models
  - RESTCONF notification subscription over SSE (Server-side events) or Websockets

Network Data – IETF RFC8345 topo example


```
{
"node-id": ”router-r1",
"ietf-network-topology-state:termination-point": [
{
"tp-id": "FourHundredGigE0/0/1/4",
"cisco-crosswork-topology-state:termination-point-attributes": {
"l2-termination-point-attributes": {
"mac-address": "34:ed:1b:35:93:28",
"unnumbered-id": [
132
],
"encapsulation-type": "ietf-l2-topology:ethernet"
},
"l3-termination-point-attributes": {
"ip-address": [
"100.8.1.5"
]
}
}
},
{
"tp-id": "HundredGigE0/0/0/0",
"cisco-crosswork-topology-state:termination-point-attributes": {
"l2-termination-point-attributes": {
"mac-address": "6c:03:09:6a:35:40",
"unnumbered-id": [
111
],
"encapsulation-type": "ietf-l2-topology:ethernet"
},
"l3-termination-point-attributes": {
"ip-address": [
"100.1.1.15"
]
}
}
},
{
"tp-id": "FourHundredGigE0/0/1/8",
"cisco-crosswork-topology-state:termination-point-attributes": {
"l2-termination-point-attributes": {
"mac-address": "34:ed:1b:35:93:38",
"unnumbered-id": [
135
],
"encapsulation-type": "ietf-l2-topology:ethernet"
},
"l3-termination-point-attributes": {
"ip-address": [
"100.20.1.9"
]
}
}
}
],
"ietf-l3-unicast-topology-state:l3-node-attributes": {
"name": ”router-r1",
"prefix": [
{
"prefix": "100.0.0.27/32",
"ietf-sr-mpls-topology-state:sr-mpls": [
{
"algorithm-value": 0,
"last-hop-behavior": "php",
"is-node": true,
"start-sid": 16127,
"algorithm": "ietf-segment-routing-common:prefix-sid-algorithm-shortest-path",
"value-type": "absolute",
"is-local": false,
"range": 1
}
]
}
],
"ietf-sr-mpls-topology-state:sr-mpls": {
"srlb": [
{
"lower-bound": 15000,
"upper-bound": 15999
}
],
"srgb": [
{
"lower-bound": 16000,
"upper-bound": 23999
}
],
"msd": 10,
"node-capabilities": {
"transport-planes": [
{
"transport-plane": "ietf-segment-routing-common:segment-routing-transport-mpls"
}
]
}
},
"router-id": [
"100.0.0.27"
],
"cisco-crosswork-isis-topology:isis-node-attributes": [
{
"system-id": "1000.0100.0027",
"level": "level-2"
}
]
}
},
```

Network Data – Packet metric data

- Performance data
  - Typical network statistics, IE: Interface stats, resource utilization

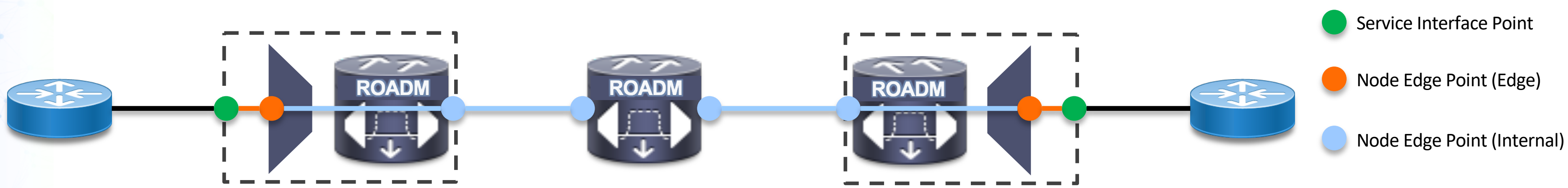
- Flow data
  - If available Netflow / IPFIX can add additional detail to network level flows
  - Applications steering traffic to/from external destinations relyon flow information to make routing decisions

- Metric data is overlayed onto topology data to optimize routing and forwarding decisions

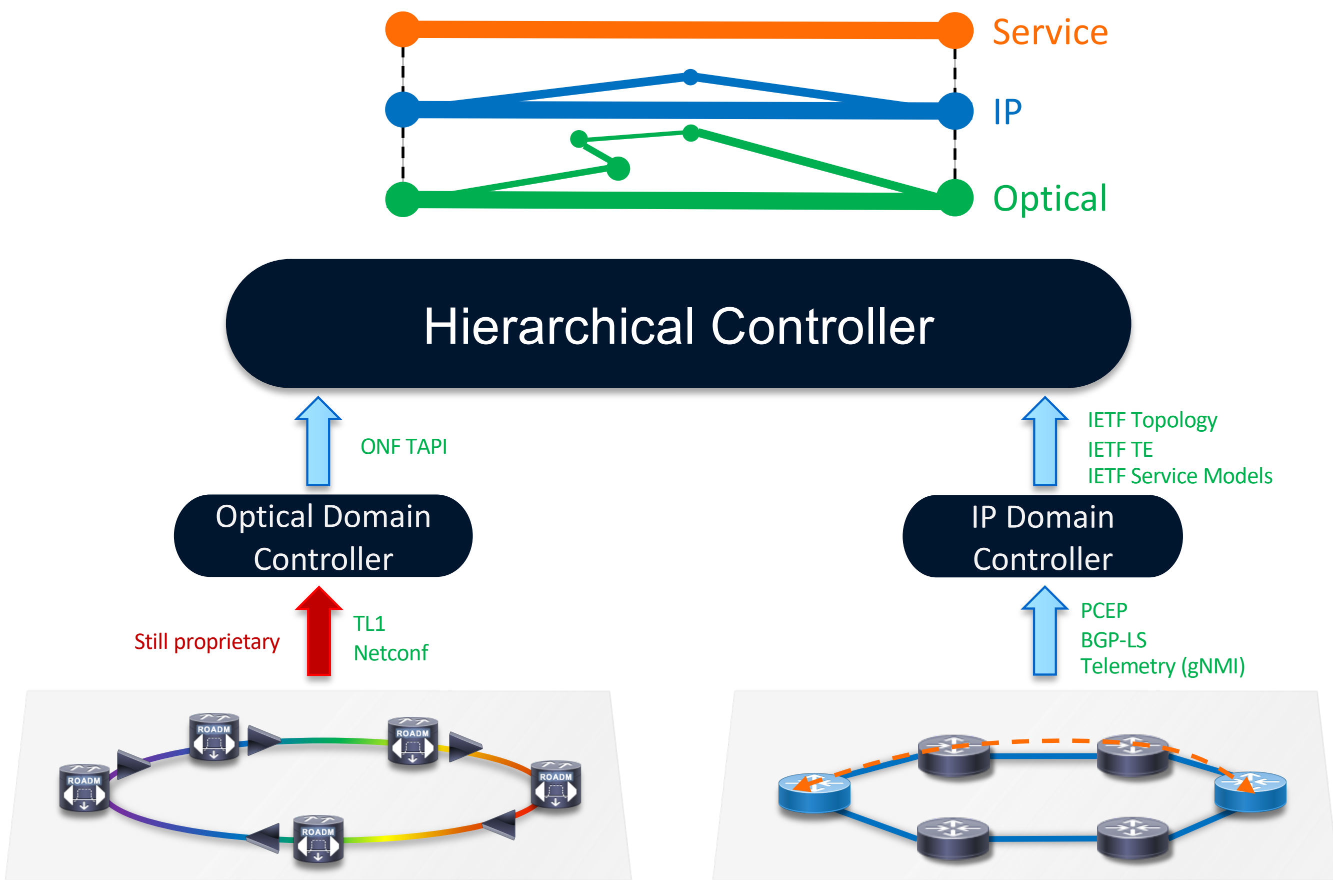
Network Data – Optical

- Controller to network
  - Typically, standard interfaces like TL1, SNMP, or NETCONF are used to the device but encoding and models are proprietary

- Controller to controller
  - T-API is the current de-facto standard
    - topology-context includes full internal optical topology with nodes, network-edge-points (NEP)
    - service-interface-point model contains all client facing interfaces that are a possible termination point for service creation
    - Topologies in TAPI can be recursive



Network Data



Standards gaps

Packet Controller to HCO gaps

- Physical inventory
  - draft-yg3bp-ccamp-optical-inventory-yang-01 (ietf.org) is a recent proposal for a NBI controller model to carry aggregated inventory information for a domain it controls
- Fault
- Performance metric data
- In the IP world we typically get these direct from a device or proxy device-level data through an intermediate application. Proxy is still an option but not ideal

Popular open source controllers

- Open Daylight
  - Linux Foundation project initiated in 2013 by commercial networking vendors
  - Java OSGi based framework; applications for data collection, TE, orchestration, provisioning for various technologies
  - Base for various open source and commercial vendor controllers

- ONOS (Open Network Operating System)
  - Work initiated by ON.Lab in 2012, now driven by ONF (Open Networking Foundation)
  - Java OSGi based distributed framework using network and data abstractions to facilitate network application development
  - uONOS is the newer microservices based ONOS
  - Wide range of use cases: BNG, WAN, BGP EPE

Network Use Cases

SDN Controllers – Real world view

- Controller can be considered a part of the network, managing network resources like an embedded device control plane

- Operational visibility is a good starting place for deployment
  - Provisioning and more advanced control functions can be added later

- Fully closed loop automation is rare

Use Case – Shared Risk Link Groups

- IP links over optical DWDM circuits often converge on a single fiber, conduit, node
- Traffic engineering and protection on the IP network requires it be seeded with information about shared resources, otherwise a single failure could take down working and protect paths
- Optical restoration requires dynamic updates of SRG information on the IP network
- GMPLS does solve some of this, but was complicated and not interoperable between vendors

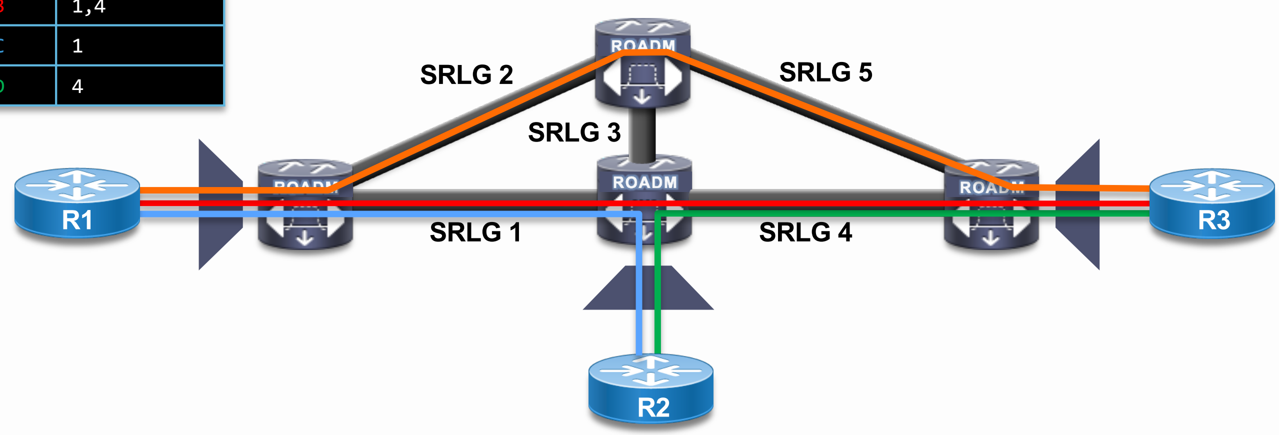
Use Case – Shared Risk Link Groups

- IP layer needs SRLG information for service and protection diversity
- Easier to solve by using intelligent software at hierarchical controller level
- Source of truth for SRLG information is optical network

HCO	
R1 Eth0	Circuit A
R1 Eth1	Circuit B
R1 Eth2	Circuit C

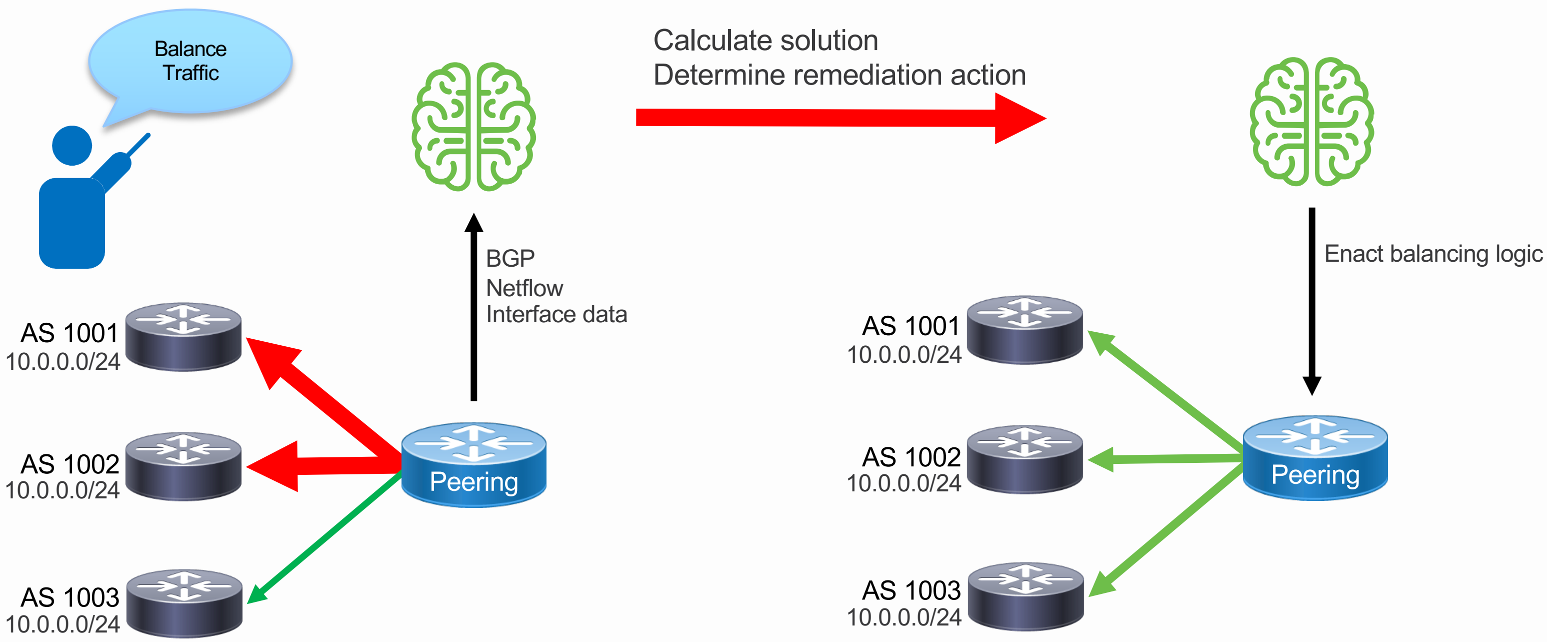
Optical Controller	
Circuit A	1,2,5
Circuit B	1,4
Circuit C	1
Circuit D	4

- HCO has discovered or manually entered Inter-Layer Link
- HCO programs SRLG information on each router, R1 will augment Eth1 protection to use Eth0 and not Eth2



Use Case – Edge traffic balancing

- More of a true "SDN" use case of defining an end goal and using network data and intelligent software
- End state: Balance traffic to 10.0.0.0/24 across all possible peers



Q & A

cisco