Evolution of MPLSDN for Open Networking Architectures

Bojian Wang bojwang@cisco.com
Customer Solution Architect, China SP

June, 2013

Agenda

1

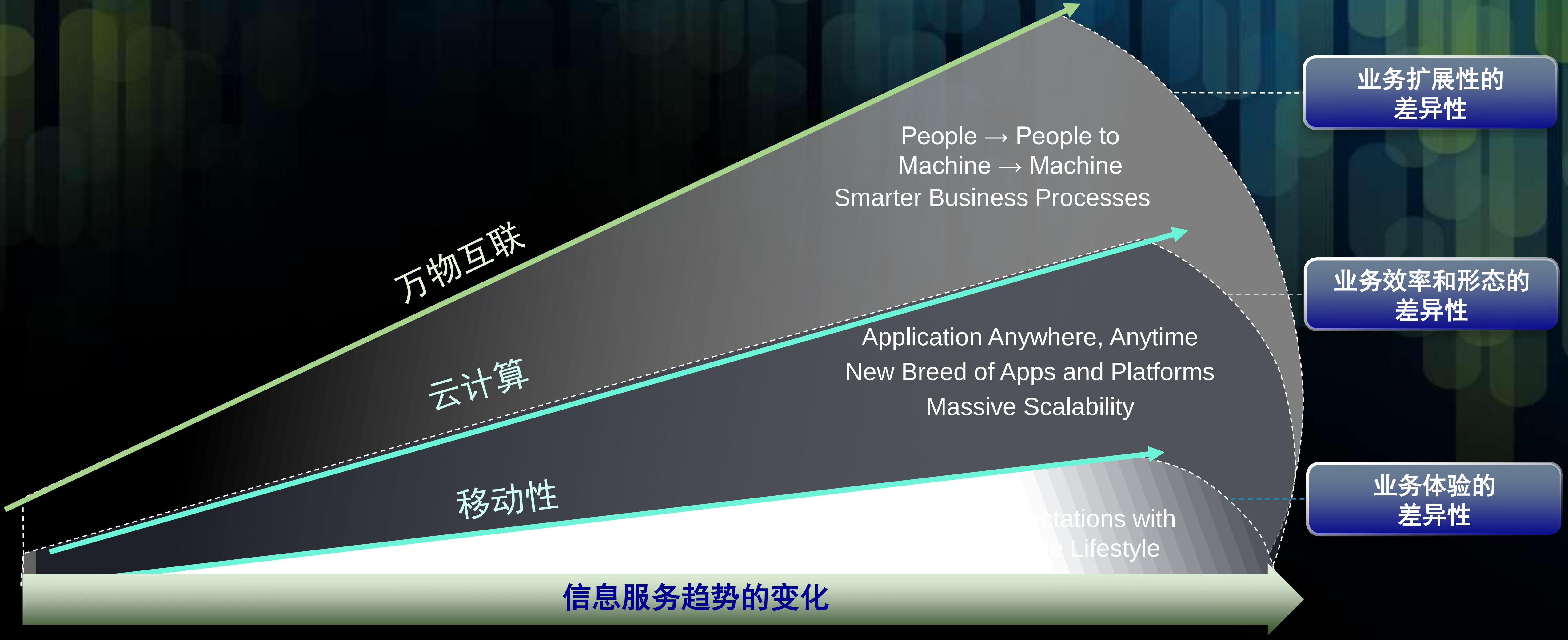
1 MPLSDN 技术展望

2 多层Stateful PCE的实现和应用

3 Segment Routing 介绍

4 总结

移动、云和万物互联时代对技术的影响



当前的MPLS技术

优点

Service Rich and Mature
Operational Experience
Widely Deployed
Standardized

存在的问题

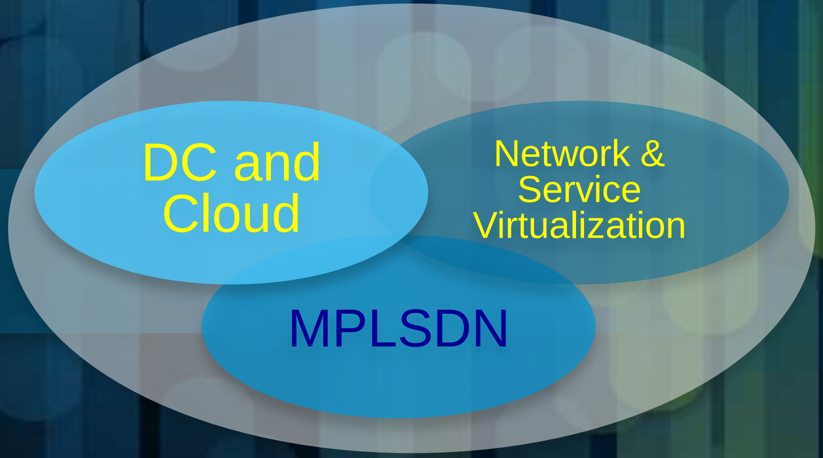
Network complexity
Expensive and Inefficient to Deploy/Operate
Non-Deterministic
Service velocity

新兴的技术趋势

SDN
Cloud Based Service Delivery and Network Function Virtualisation (NfV)
IP + Optical Integration

MPLSDN的技术演进

MPLSDN – 运营商网络承载技术演进的未来
简单、快捷、虚拟化、高效



Evolving Technology Landscape

- Embraces SDN and NFV Concepts
- Centralized Control for Network Optimization
- Dynamic Cloud Based Service Delivery over MPLS Networks
- Programmatic Interfaces and Orchestration

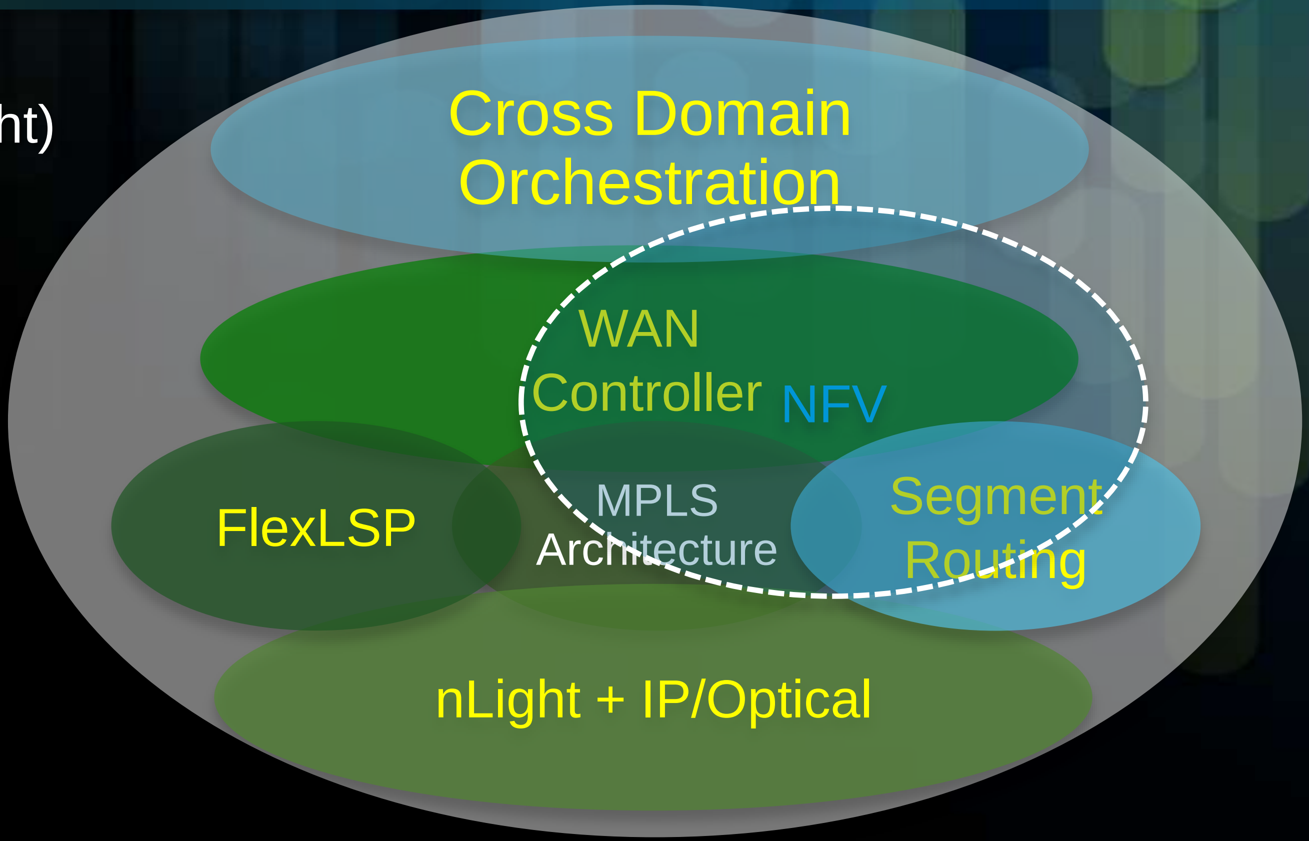
通过技术的继承和平滑演进解决存在的问题

- Leverages proven concepts in IP/MPLS and Optical Transport
- Operates with existing hardware platforms
- Simplified control plane – more scalable data plane
- Centralised control for Admission Control and Policy Engine for network optimisation and dynamic service placement

新兴的技术引擎

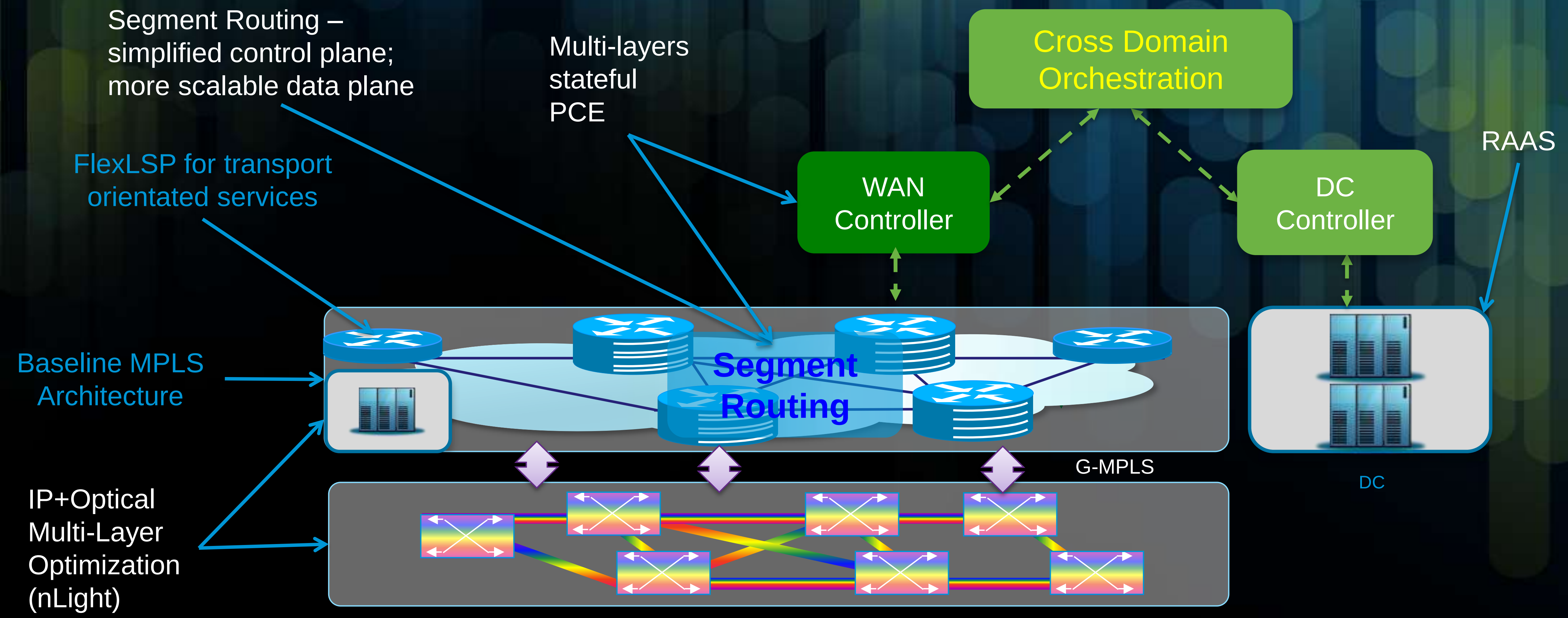
综合多个技术领域的创新成果

- Multi-layer stateful PCE
  - IP+Optical Multi-Layer Optimization (nLight)
  - WAN Controller
- Segment Routing
- FlexLSP
- IP/MPLS and Cloud Integration (RaaS)
- Cross Domain Orchestration

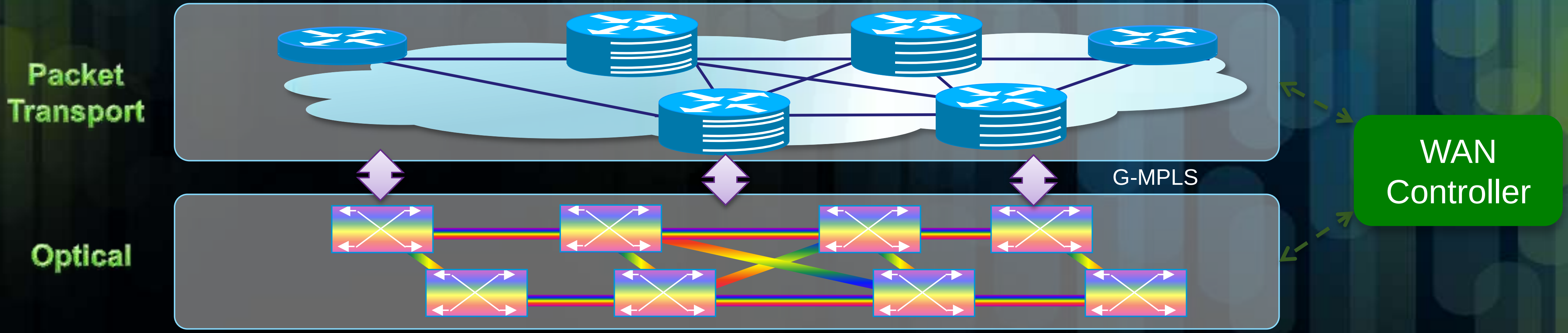


MPLSDN 技术引擎

MPLSDN的技术体系架构



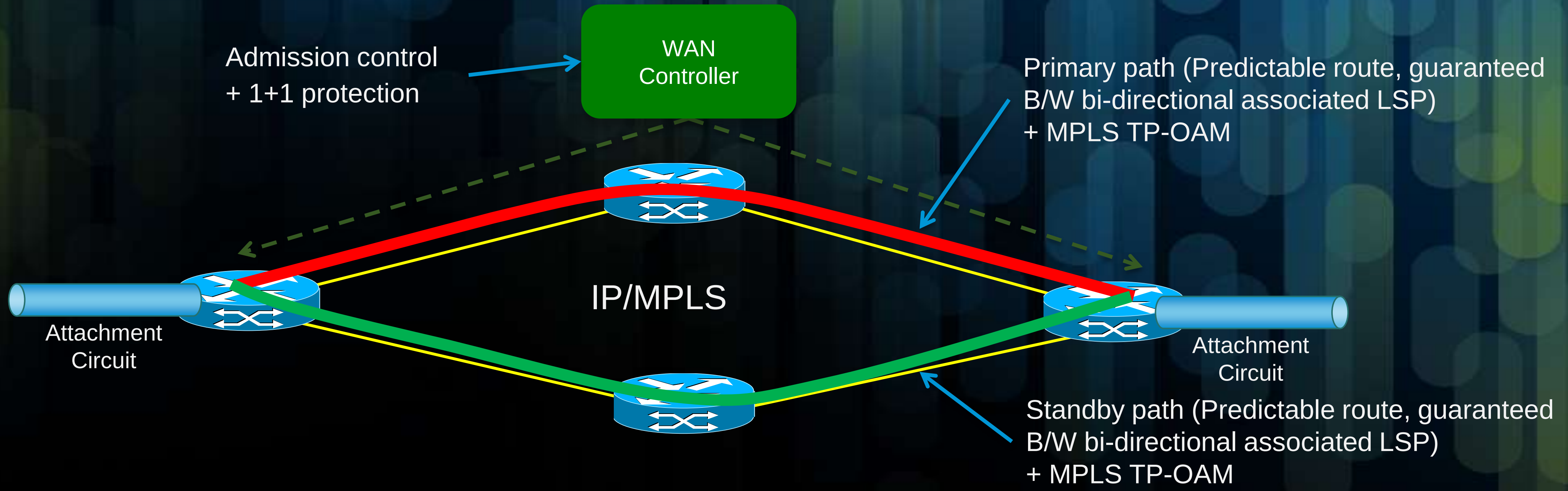
nLight与WAN Controller的结合



- Management and information exchange between optical and IP layers
- Dynamic optical control plane
- G-MPLS UNI between optical and IP domains
- Multi-layer optimization using WAN controller

Benefit: 60% saving in ports over 5 years via nLight+multi-layer WAN controller

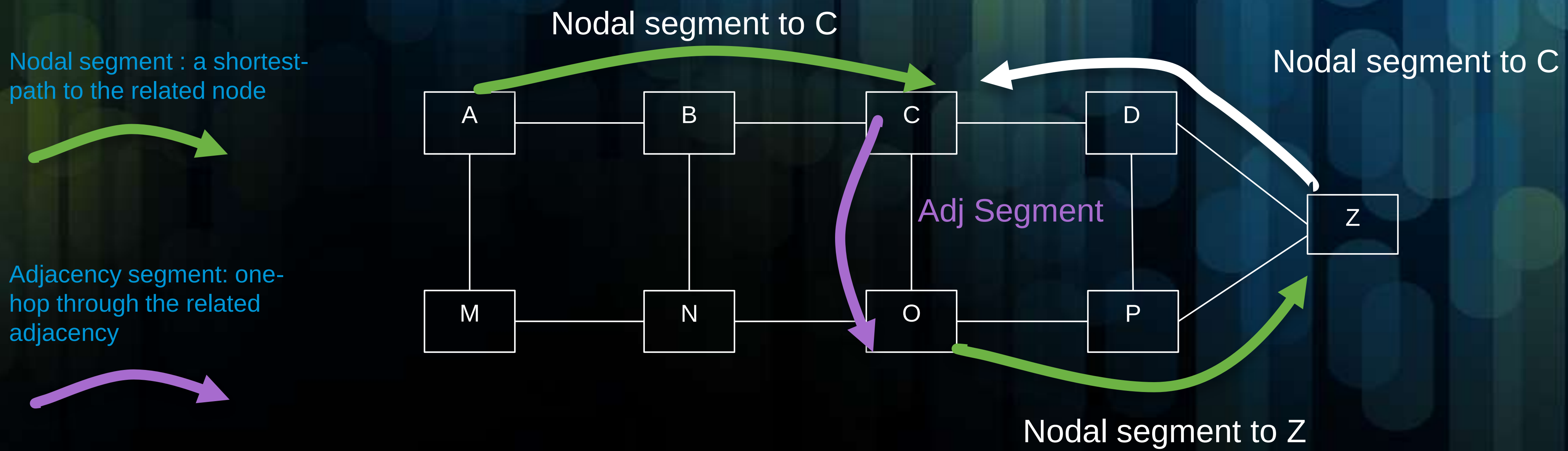
FlexLSP：MPLS-TP的演进



- FlexLSP brings transport orientated services to IP/MPLS environments
- Bi-directional transport orientated tunnels supporting pseudo-wires
  - Predictable route, guaranteed B/W bi-directional associated LSP
  - MPLS-TP OAM monitoring LSP status and driving protection
- Programmatic VPN services enabling NfV

Benefit: 20-60% saving for transport services with FlexLSP vs. OTN

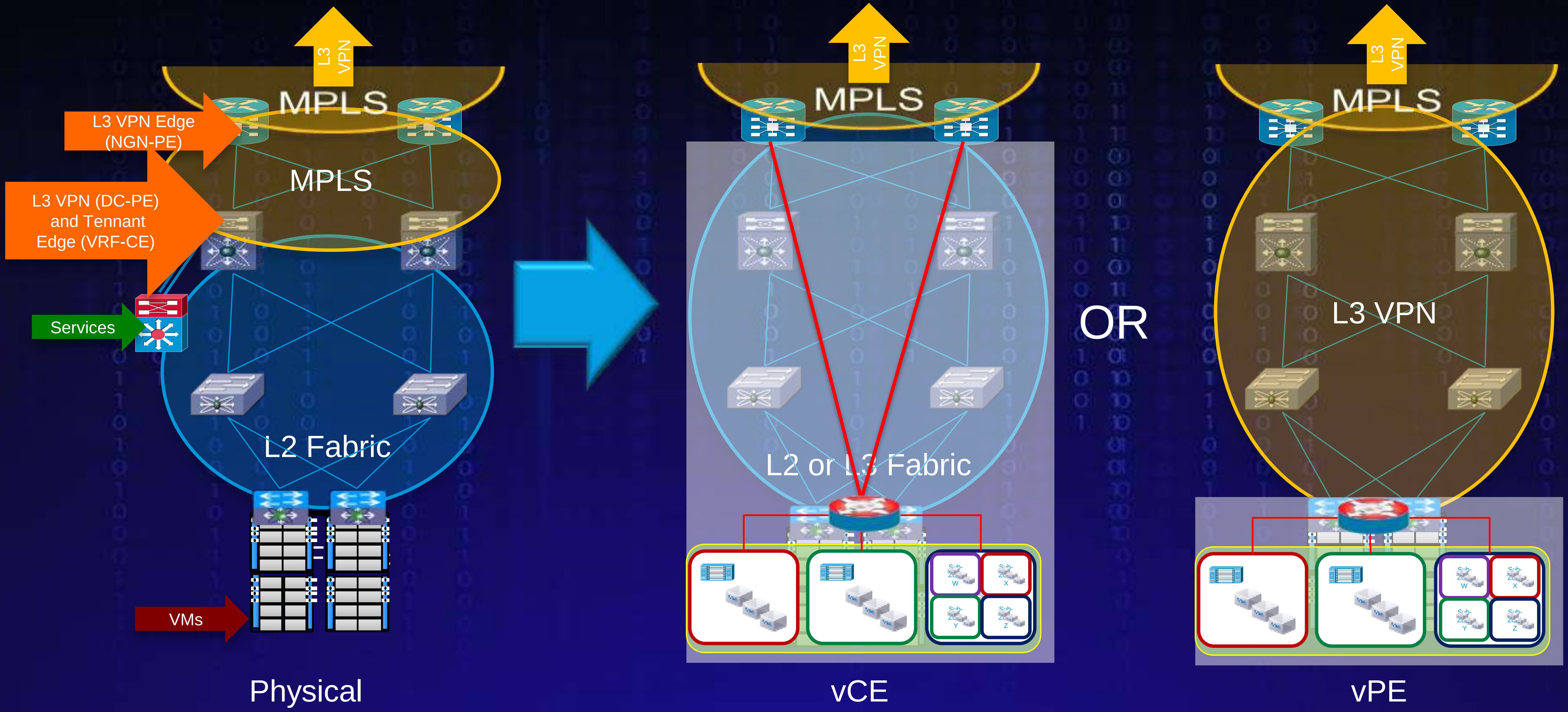
Segment Routing: 更加简化高效的MPLS



- Emergence of Stateless MPLS
- Simplification – label distribution via IGP; no need for LDP and RSVP
- Scale – less state for routers to maintain to maintain
- Combined with WAN controller for Admission Control and optimal path determination
- Backward compatible with existing networks

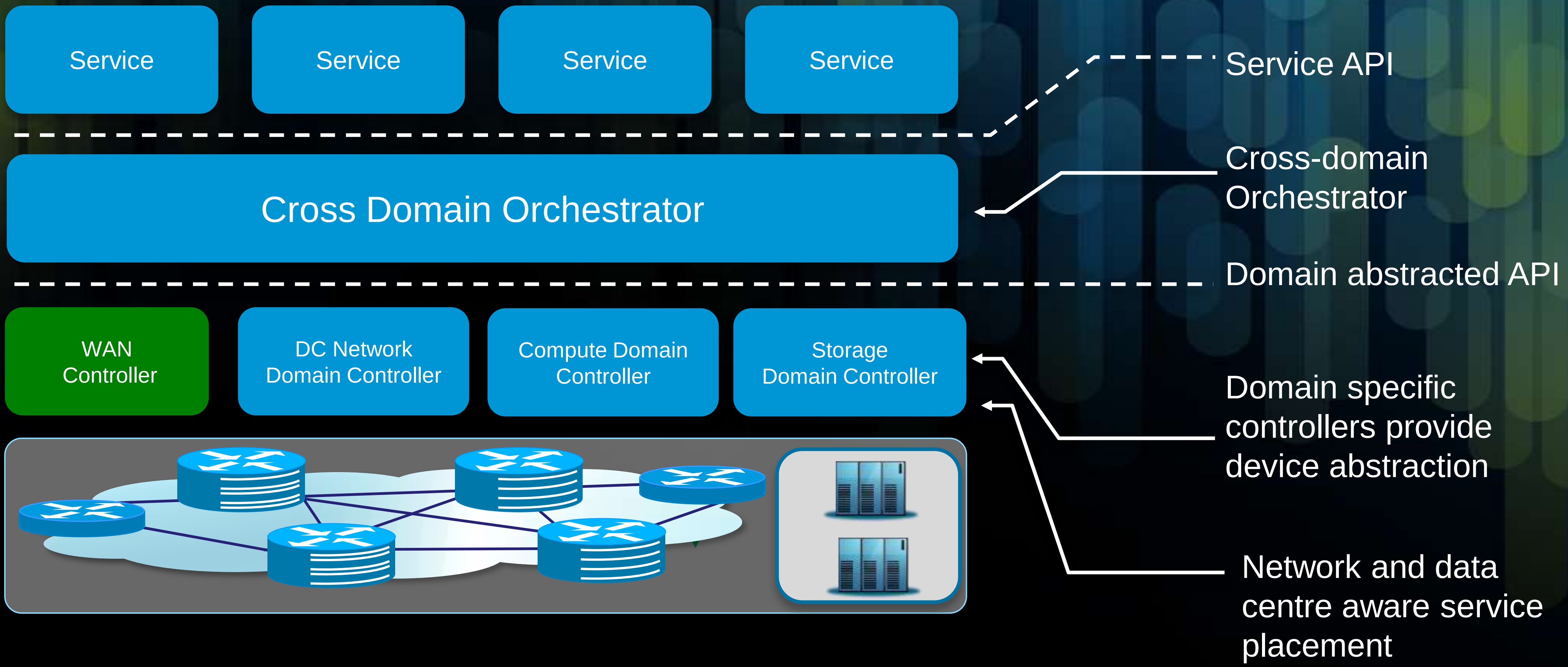
Benefit: Simple, scalable, programmable MPLS

MPLS技术向云端的延伸: RAAS



Benefit: Interconnectivity between IP/MPLS and cloud-based services

跨域的统一业务调度编排



Benefit: Cloud based service delivery + dynamic, deterministic, optimized network

Agenda

2

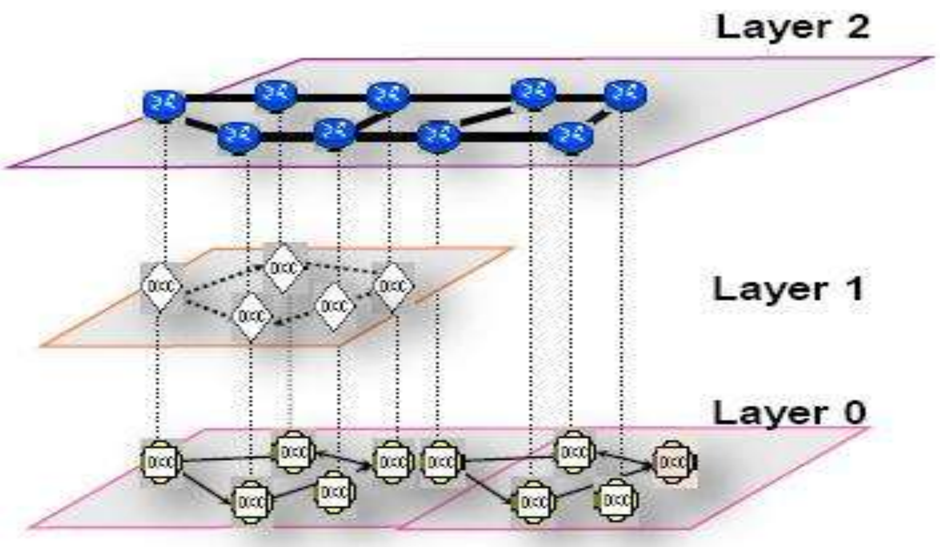
1 MPLSDN 技术展望

2 多层Stateful PCE的实现和应用

3 Segment Routing 介绍

4 总结

美国有线电视运营商：为何需要多层协作优化



Benefits of M-Layer optimization

- Cost-optimize network **dimensioning** and **capacity planning**
  - Maximize the available capacity while reducing transport cost
- Evaluate and plan for **service requirements**
  - Delay: Propagate connections/links delays to the packet layer
  - Availability: Constrain service routing based on availability needs
- Validate **optical network failures on the SLAs**
  - Assessing the impact of failures in the Optical network such as fiber cut, optical amplifier, transponder, regenerator, OXC failures on SLA
- Designing networks to achieve **multi-layer resilience**
  - Propagate SRLG to the packet layer to serve as routing constraints

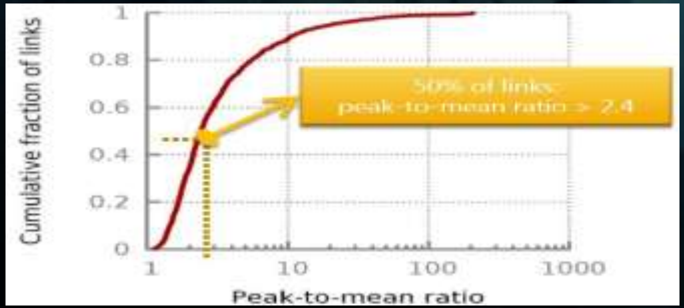
<Multilayer planning techniques optimizes expensive capacity while addressing the SLA and resiliency requirements>

MPLS 2012

Source: Designing Multi-Layer Carrier Networks for Capacity and Survivability, OPNET Technologies, Inc., OPNETWORK 2012.

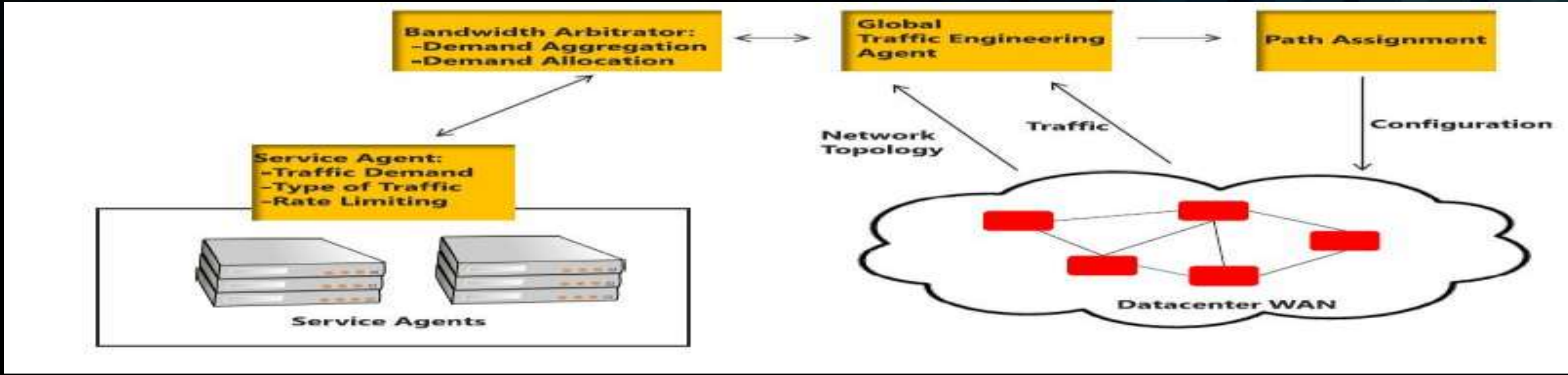
COX

微软：广域网带宽调度优化



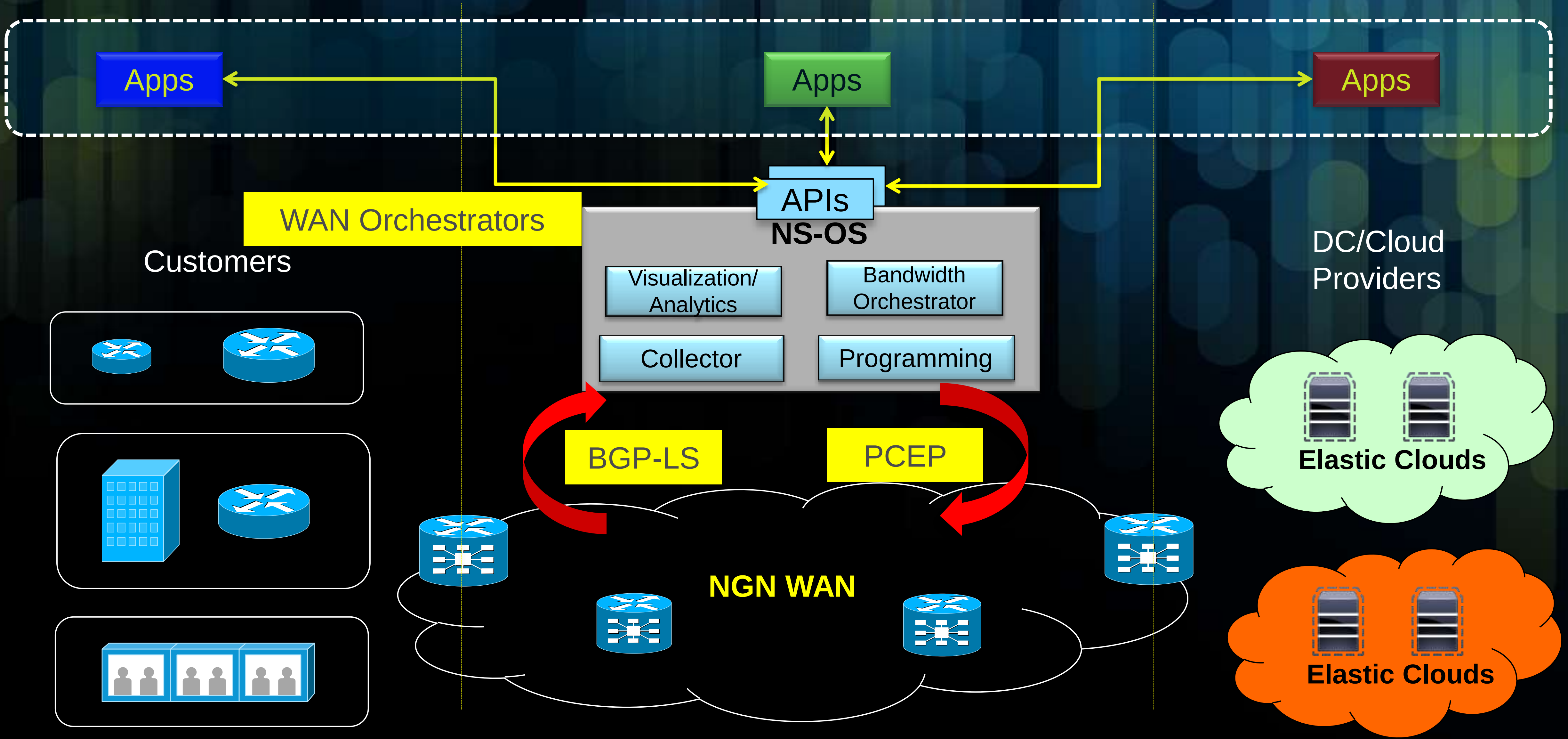
Network wide Link utilization

Link utilization



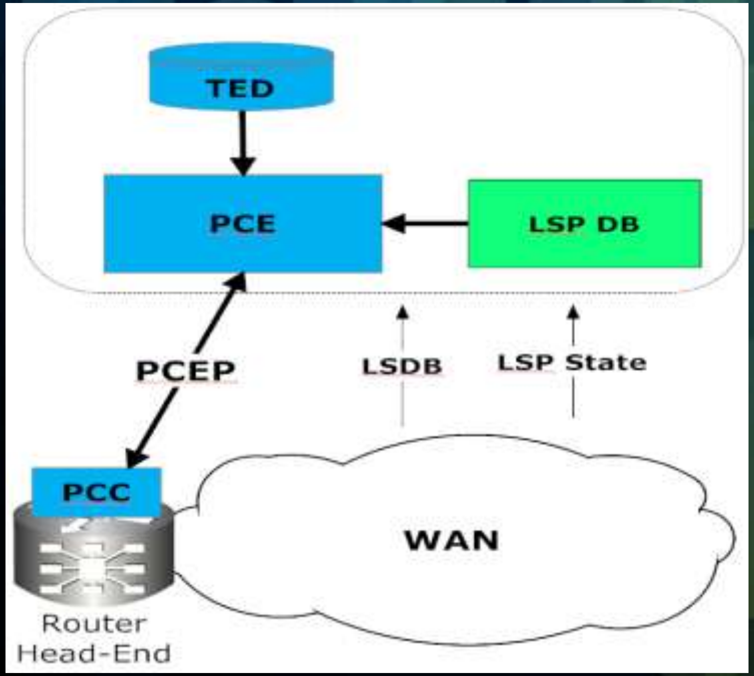
SDN based centralized TE

多层Stateful PCE



PCE技术背景以及有状态PCE协议

- RFC4655 defined PCE Architecture (2006)
  – Section 6.8 discussed Stateful vs Stateless PC
- RFC5440 defined PCEP (2009)
  – PCC-PCE; PCE-PCE communications
  – Router asks PCE for TE path in PCRreq message
  – PCE responds to router with PCRep message
  – PCE assumed to be stateless



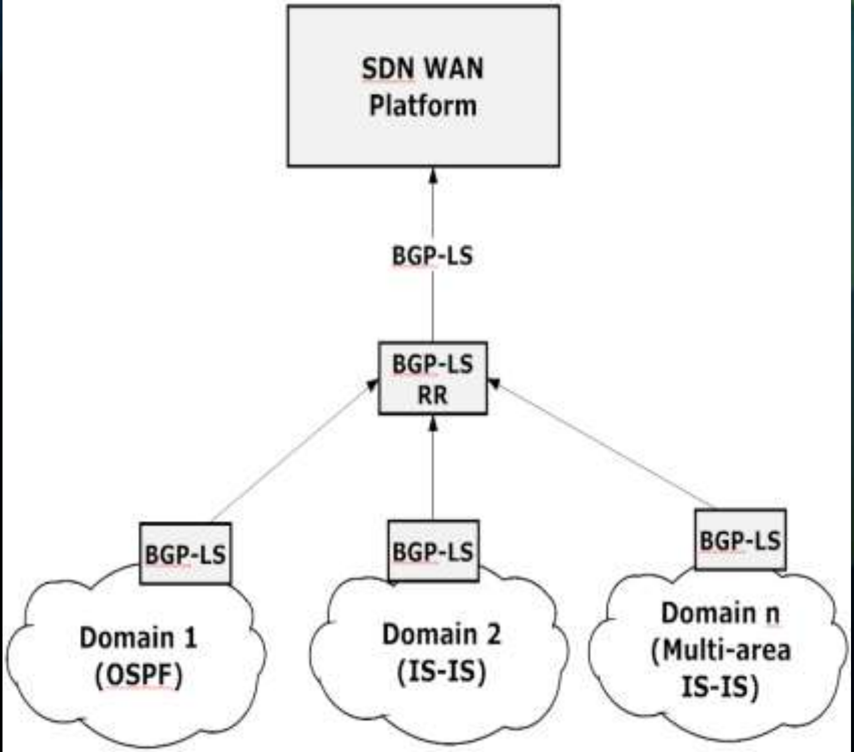
- Stateful
  – Topology, resource and TE state synced to PCE
  – Considers topology, resources and TE state when computing paths
- Stateless
  – PCE does not know or or remember computed TE paths, thus they are not considered when new paths are computed

有状态PCE的协议扩展

- Need PCE to program demand and path placement in network
  - Bandwidth scheduling
  - Demand Engineering
- draft-ietf-pce-stateful-pce
  - Defined new functions in support of stateful PCE including capabilities, LSP state sync, LSP update request, LSP state report and LSP control delegation
- draft-crabbe-pce-pce-initiated-lsp
  - PCCreate message
  - PCupd message with remove bit deletes LSP

BGP-LS

- Introduces the ability to "redistribute" an IGP topology into a BGP
- Redistribution takes the IGP LSDB as the input but…
- Redistribution is NOT limited to the content of an LSDB
  - Ability to extend/enrich topology data
  - Ability to aggregate/hide/abstract topology data
- Allows over-the-top topology export
  - No need to access IGPs from external topology consumers
  - Topology Servers/Systems are BGP-LS consumers: PCE, NPS/ALTO, …
- BGP policy, security and reliable TCP transport
- Control is kept by network operator

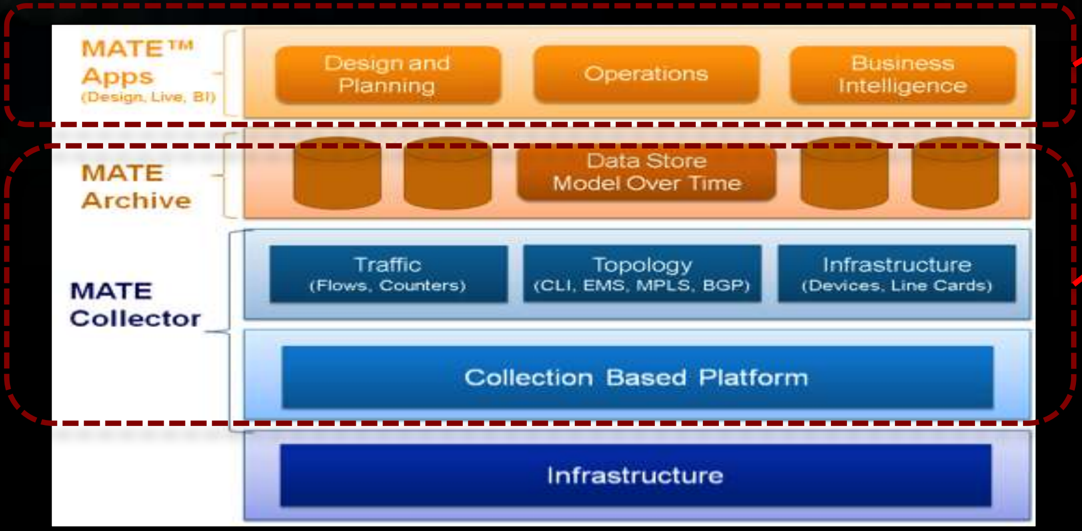


BGP-LS 和 IGP 扩展

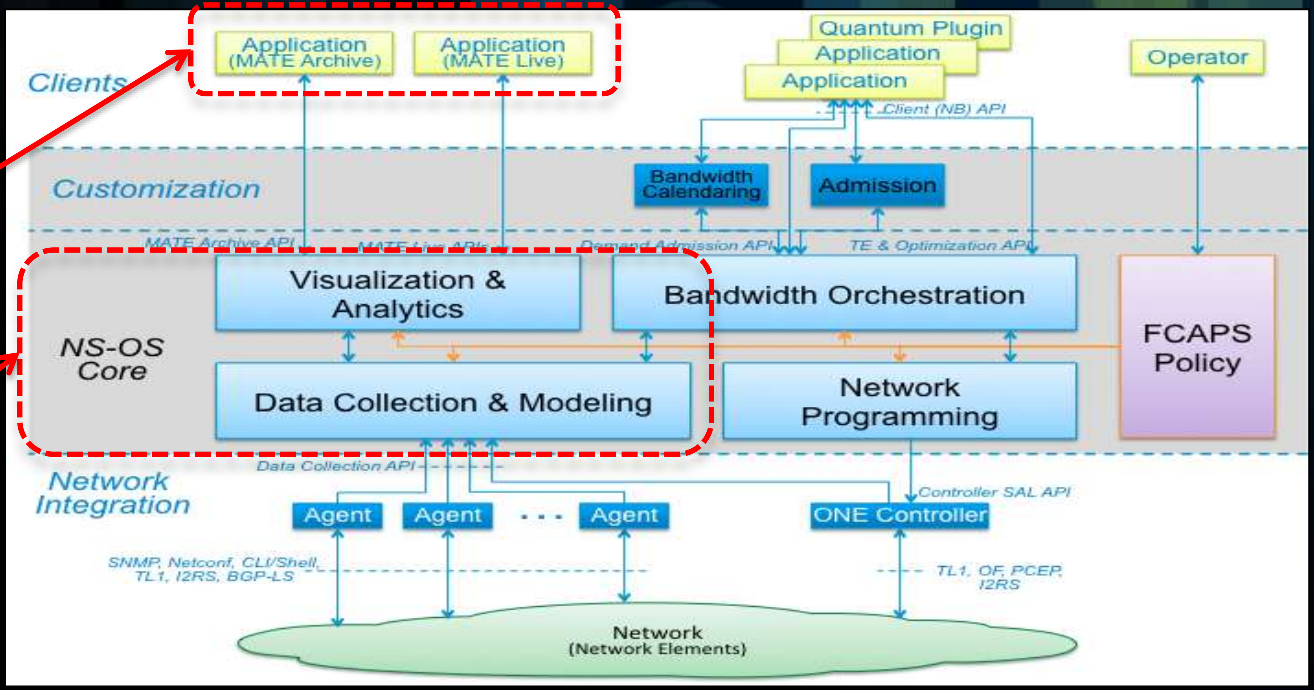
- Latest extensions to ISIS/OSPF allow the advertisement of new subTLVs
  - draft-previdi-isis-te-metric-extensions
  - draft-giacalone-ospf-te-metric-extensions
- Delay, BW and Loss information
- Allow IGP to carry resources utilization/availability from a “real” use perspective
  - Vs. TE-provisioning info
- Goal: enhance SPF/CSPF/xSPF tree computation with additional metrics
  - Natural extension to the 4 metrics of ISIS (Default, Delay, Expense, Error)
- BGP-LS is agnostic regarding IGP data
  - Transparently advertise IGP TLVs
  - Extensions to IGPs are de facto integrated into BGP-LS

思科WAN Orchestrator — NS-OS

- NS-OS Platform integrates:
  - Existing Cariden software components
  - Northbound Restful API and southbound API/Protocol (PCEP,BGP-LS, onePK, Openflow, I2RS, etc..)
  - New Cisco software developments



Current MATE Platform



NS-OS Platform

Agenda

3

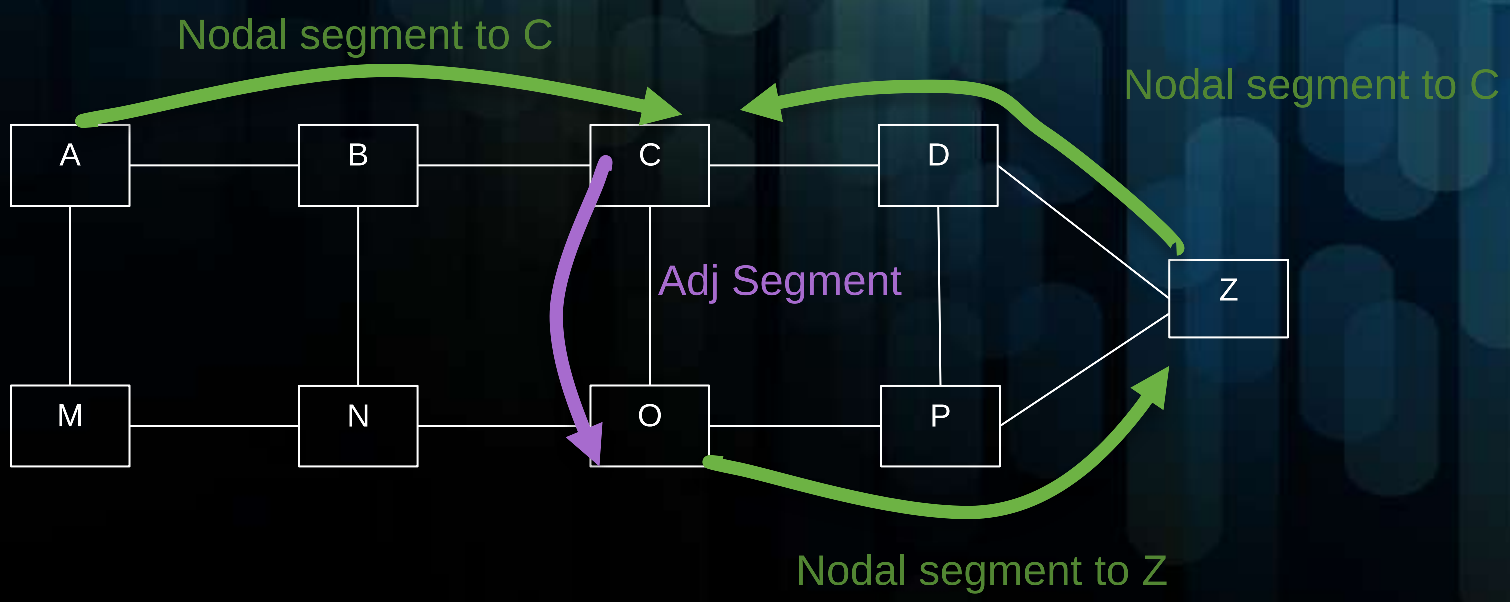
1 MPLSDN 技术展望

2 多层Stateful PCE的实现和应用

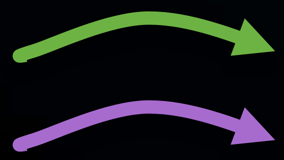
3 Segment Routing 介绍

4 总结

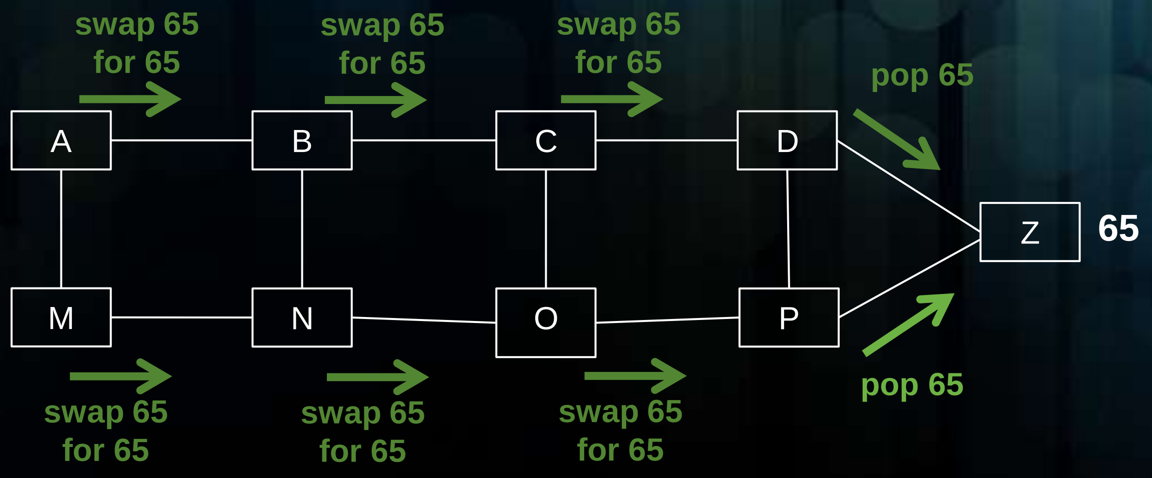
ISIS/OSPF 协议扩展



- ISIS automatically builds and maintains segments
  - Nodal: a shortest-path to the related node
  - Adjacency: one-hop through the related adjacency



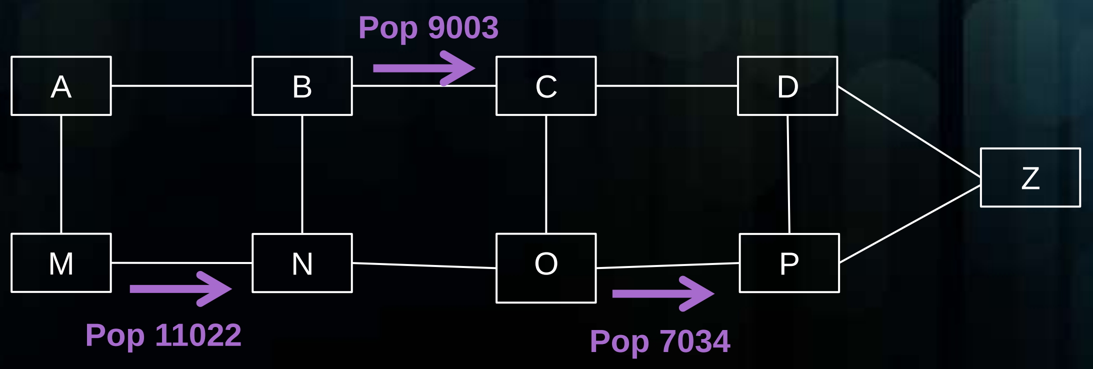
Nodal Segment 是全局标签



A packet injected anywhere with top label 65 will reach Z

- Only two IP-alike configuration steps
- Operator allocates a label block for Segment Routing
  - [64, 5000] is the “SR registry”
- Operator allocates a label from the SR registry to each node
  - Z is given label 65

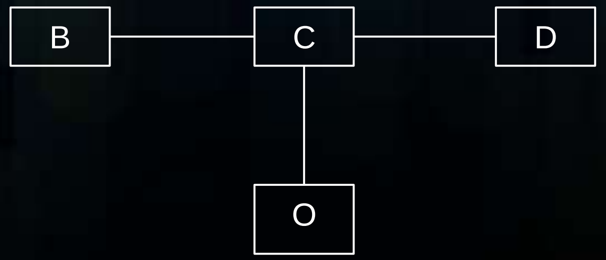
Adj Segment 是本地标签



A packet injected at node B with label 9003 is forced through datalink BC

- Node automatically allocates a local label for each adjacency
  - label taken outside the SR registry

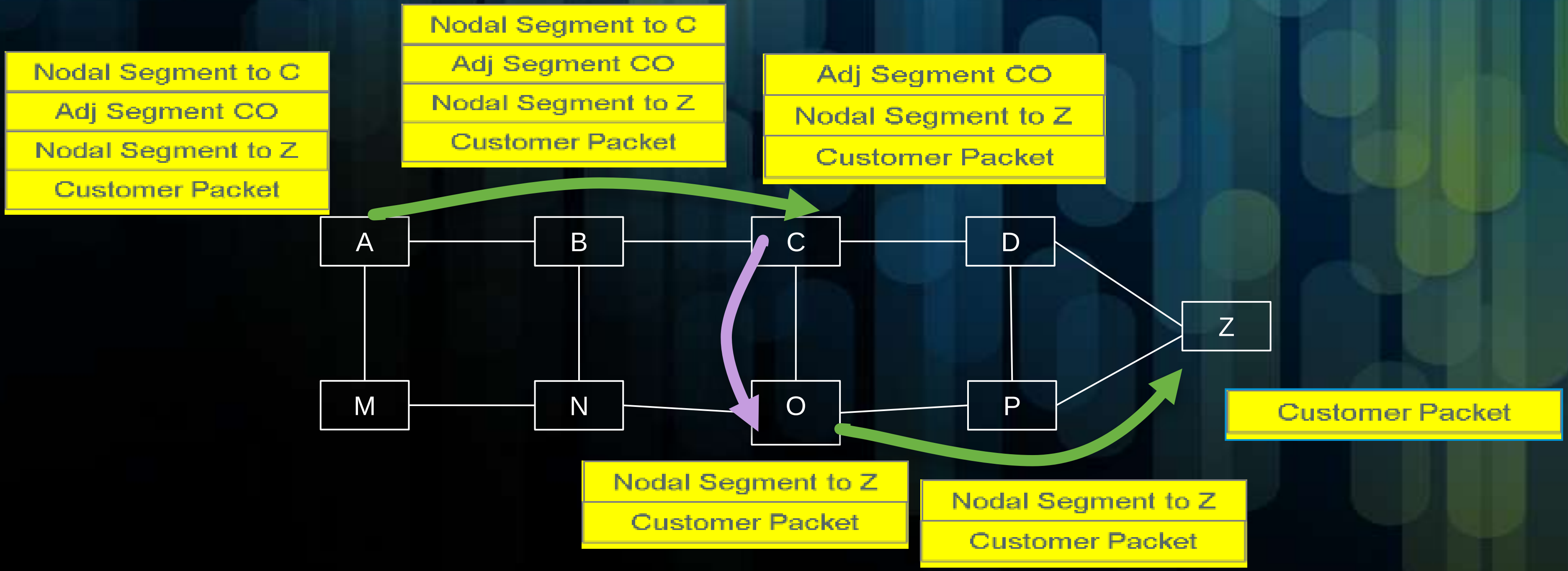
IGP协议通告SR的标签信息



C's linkstate LSP advertises
Leaf C/C with global nodal label 67
Adjacency CB with local label 9001
Adjacency CD with local label 9002
Adjacency CP with local label 9003

- Simple extension
  - One single 4-byte Segment sub-TLV
- Nodal segment: sub-TLV attached to leaf TLV
  - leaf is loopback
- Adjacency segment: sub-TLV attached to adjacency TLV

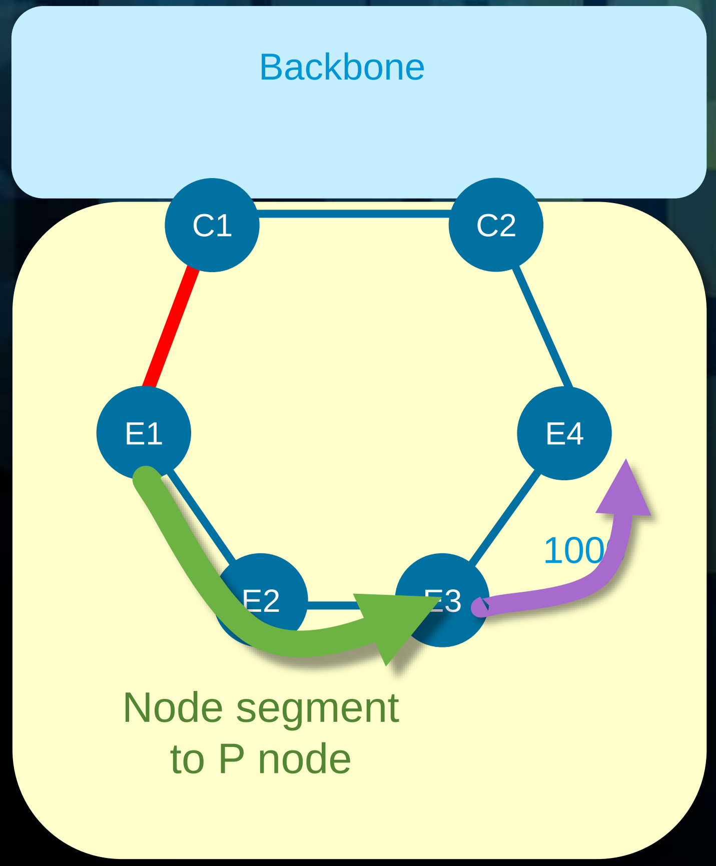
SR实现源路由



- Path is computed by source
  - potentially with help of central optimization
- Path is encoded by the source in the packet header as a label stack
  - a path is an ordered list of segment
- Each hop along the path forwards according to classic MPLS dataplane

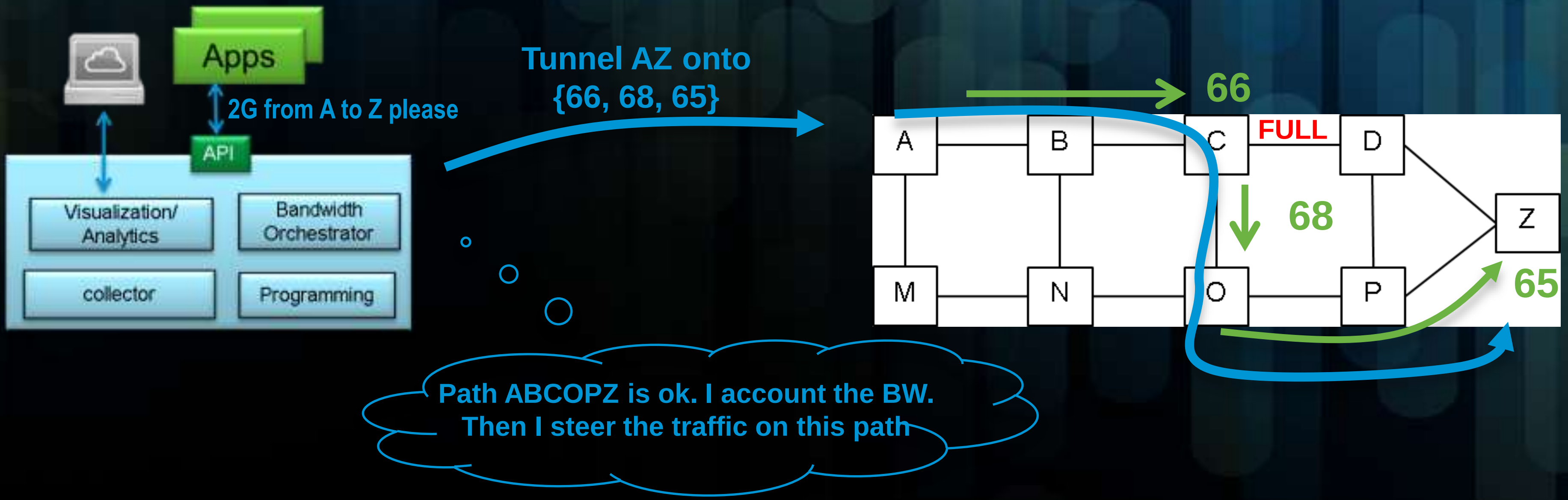
自动和通用化的FRR实现

- IP-based FRR is guaranted in any topology
  - 2002, LFA FRR project at Cisco
  - draft-bryant-ipfrr-tunnels-03.txt
- Directed LFA (DLFA) is guaranteed when metrics are symetric
- No extra computation (RLFA) — Directed LDP
- Simple repair stack
  - node segment to P node
  - adjacency segment from P to Q



Default metric: 10

SDN控制下的SR路由选择



The network is simple, highly programmable and responsive to rapid changes

简化的多路径多平面设计

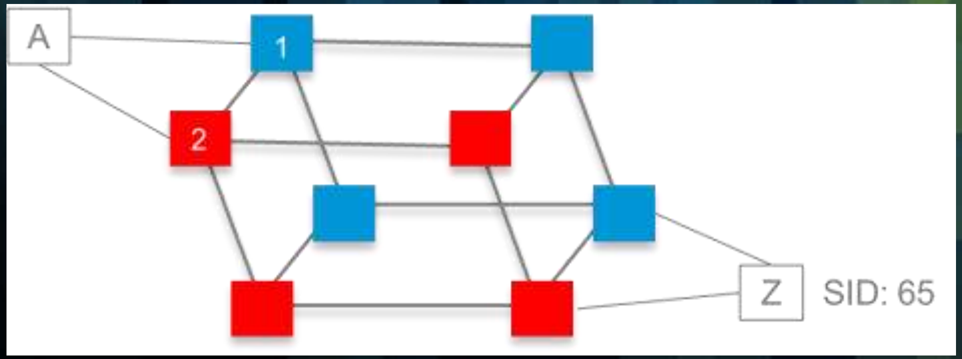
SR avoids state in the core

SR avoids enumerating RSVP-TE tunnels for each ECMP paths

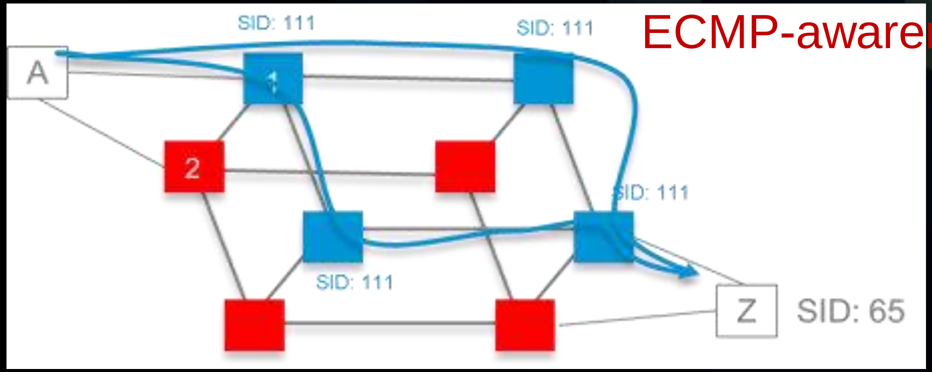
- 无特定路径、平面选择
  A sends traffic with [65]
  Classic ecmp “a la IP”

- 选择特定平面
  A sends traffic with [111, 65]
  Packet gets attracted in blue plane and then uses classic ecmp “a la IP”

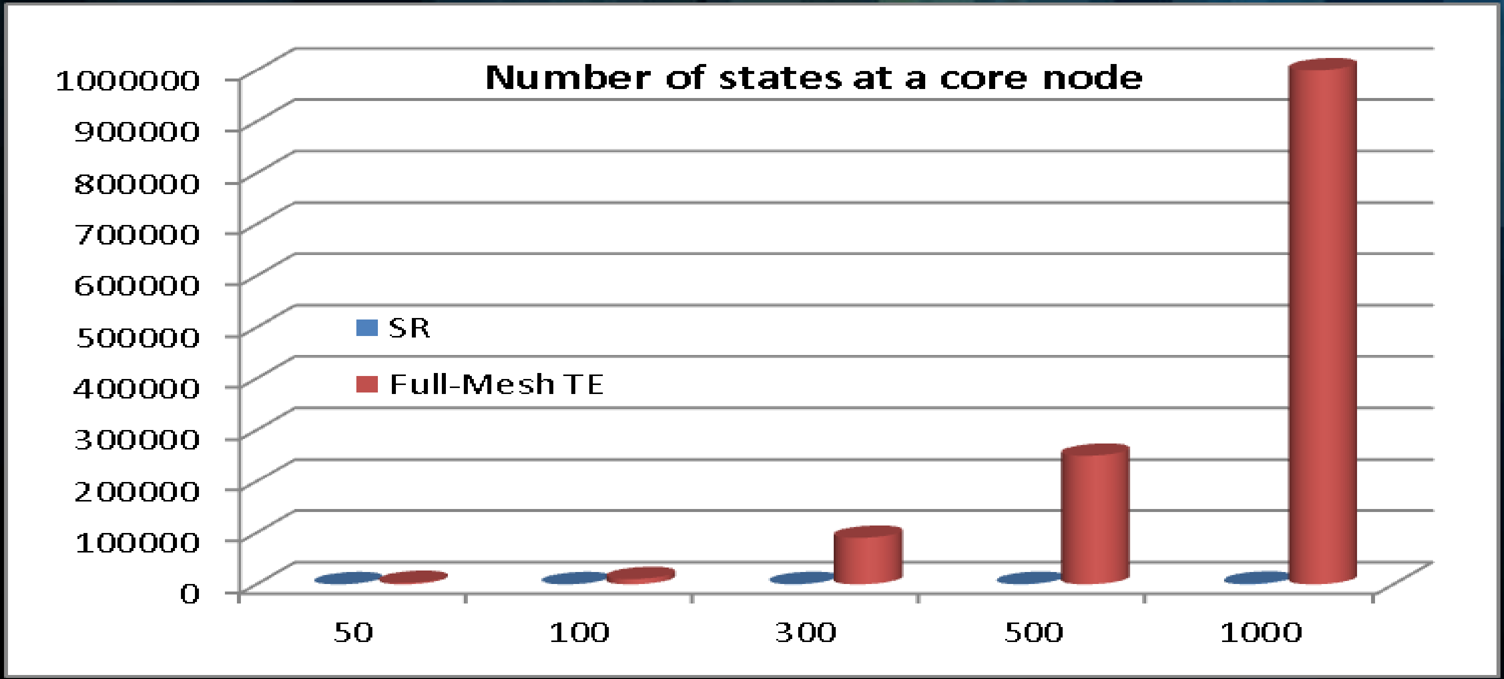
- 选择双平面
  A sends traffic with [111,65]
  via Blue Plane
  A sends traffic with [112,65]
  via RedPlane



ECMP-awareness!



传统流量工程技术与SR在扩展性上的比较

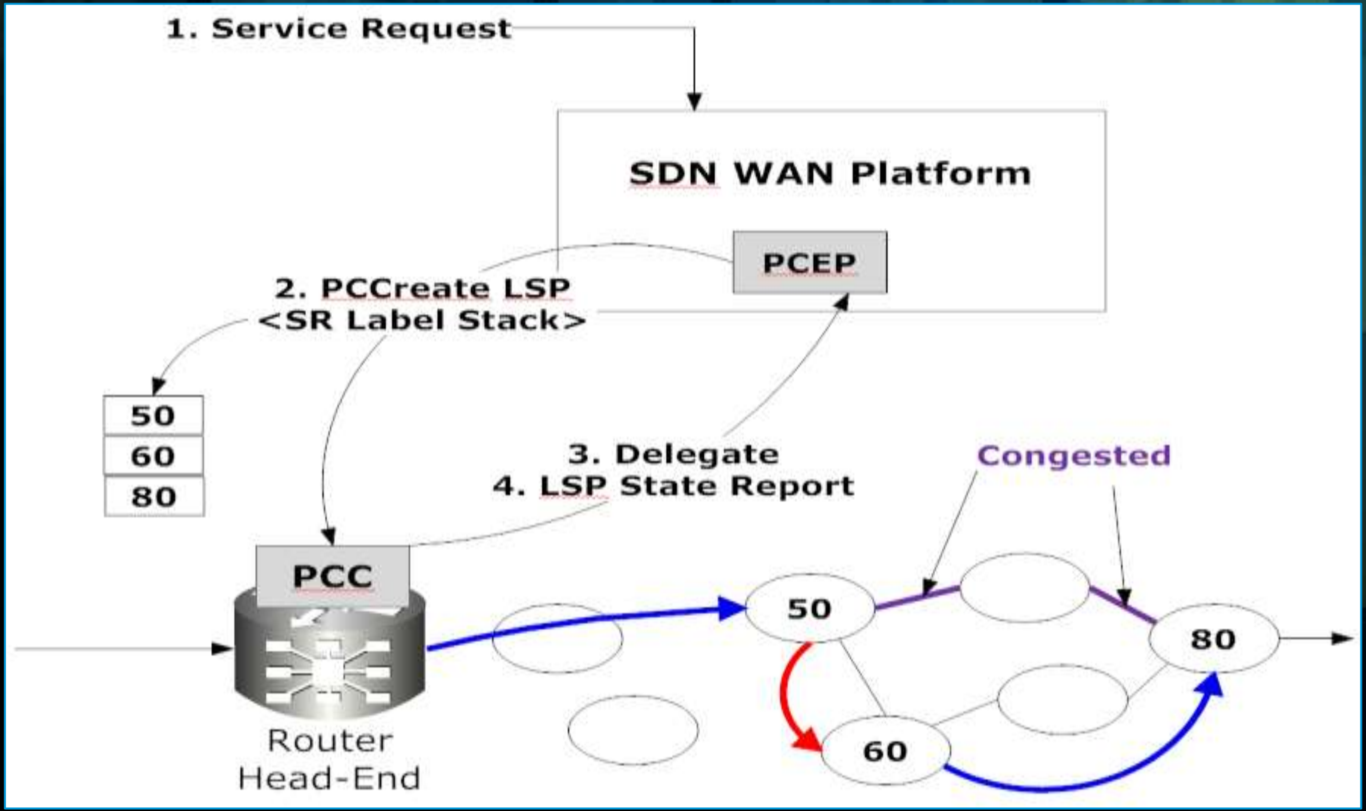


- An SR core router scales much than with RSVP-TE
  - The state is not in the router but in the packet
  - N+A vs N^2

N: # of nodes in the network
A: # of adjacencies per node

SR技术在多层PCE架构网络中的演进

- PCE knows topology and node/adj segment IDs via BGP-LS
- Computes path that avoids congested links (based on service request constraints)
- PCEP extensions needed to program SR path (label stack) in router
  - SR path (label stack prepended to each packet)
- No RSVP-TE signaling needed
- draft-sivabalan-pce-segment-routing



SR技术的优势

- No MPLS signaling or per-explicit path state in network
- Guaranteed FRR
- Scalable MPLS TE since only edges need to be programmed with explicit path label stack
- Topology flexibility and scale
  - E.g. create redundant, disjoint routing topologies, arbitrary detours away from the SPT, distinct service topologies, predictable load-balancing schemes, etc
- Coexistence with existing MPLS data-plane and signaling protocols
- Leverages mature link-state IGP; straightforward to push SR labels to SDN WAN Platform (Stateful PCE) via BGP-LS

Agenda

4

1 MPLSDN 技术展望

2 多层Stateful PCE的实现和应用

3 Segment Routing 介绍

4 总结

MPLSDN – MPLS与SDN技术融合的解决方案

更快速地业务部署

- ➔Real-time WAN Orchestration + Cross Domain Orchestration
- ➔Network Function Virtualization with RaaS (vCE / vPE)

更高的效率

- ➔IP / Optical Integration and Optimization (nLight)
- ➔Unified MPLS Transport – FlexLSP

更紧密的应用可编程、可调度的能力

- ➔Paths: Segment Routing + WAN Orchestration
- ➔Demand Admission / Placement: WAN Orchestration

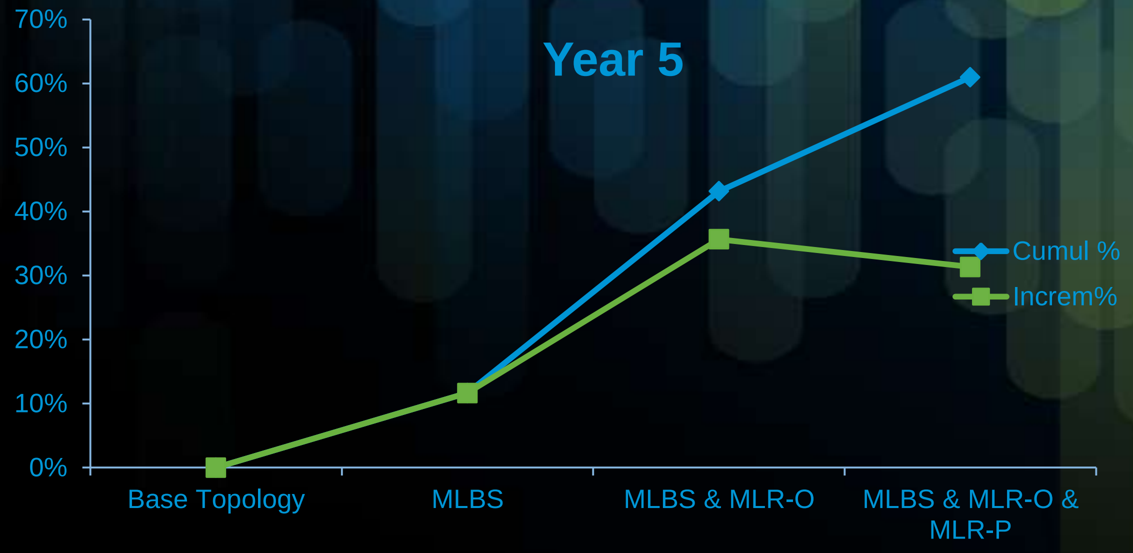
更加简化和更高的扩展性

- ➔Fewer protocols: Segment Routing
- ➔Reduced data plane state: Segment Routing

MPLSDN – 提高TCO效率

Estimated TCO Savings of >50% through operational efficiency

~30% bandwidth gain through centralised explicit path control and effective service placement
Estimated~30%+ Savings with FlexLSP



~60% saving in IPoDWDM ports over 5 years through nLight + multi-layer WAN controller

MPLSDN向开放的网络架构演进

MPLSDN – MPLS与SDN技术的紧密结合
简单、快速、虚拟化、效率

- ➔Centralized WAN orchestration for path optimization and dynamic service placement
- ➔Adds Admission Control and Real-time Policy
- ➔RaaS – Simplifies and Accelerates Cloud Delivery over MPLS Networks
- ➔Programmatic Interfaces and Orchestration

The next step in the evolution of SP core networks – providing a platform for the next 10 years