Application Engineered Network
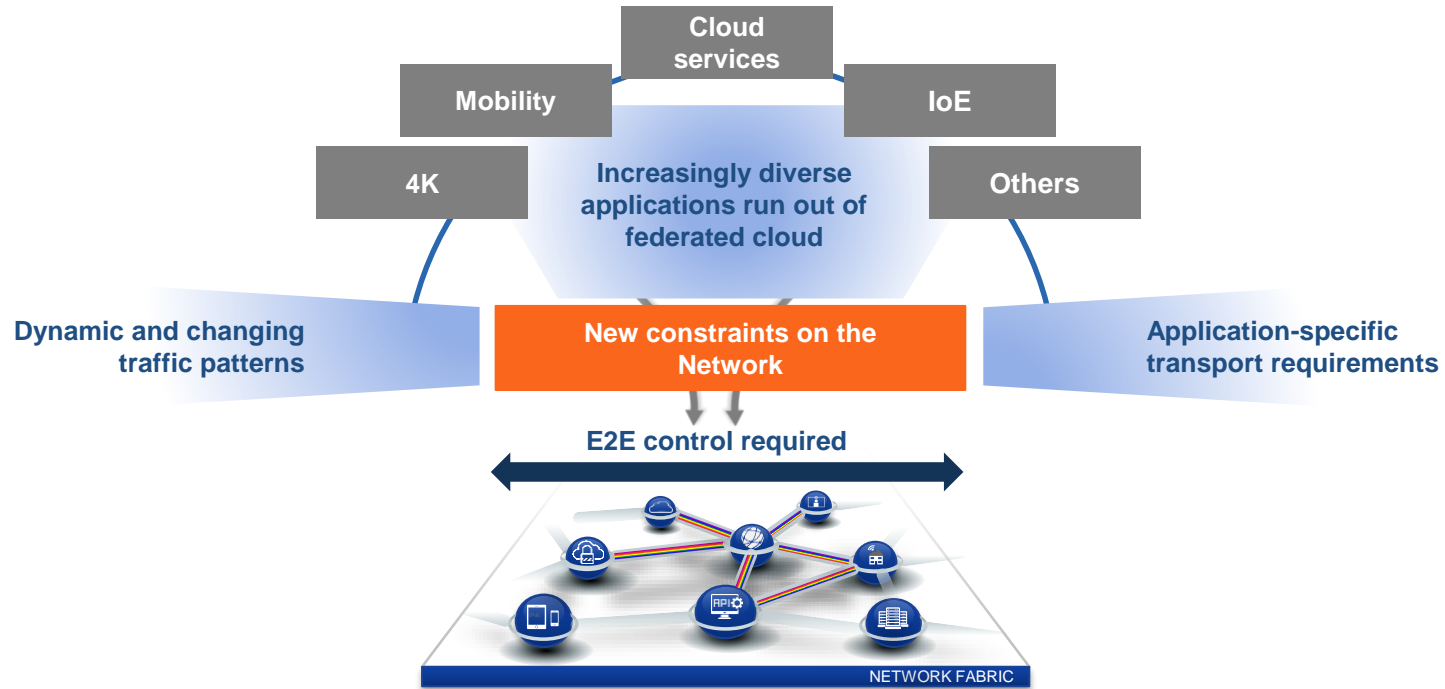
Eric Lam

Consulting Systems Engineer, Services Provider Business

ericla@cisco.com

The Problem

Cloud services

Mobility

IoE

4K

Others

Increasingly diverse applications run out of federated cloud

Dynamic and changing traffic patterns

New constraints on the Network

Application-specific transport requirements
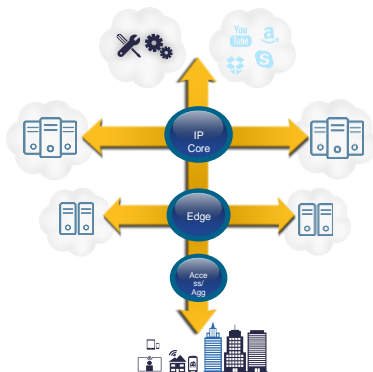
E2E control required

NETWORK FABRIC

SDN, NfV, Orchestration, Cloud, Big Data, Analytics, IoE, …
Getting all the buzzwords out of the way:
Applications and Network interaction is key
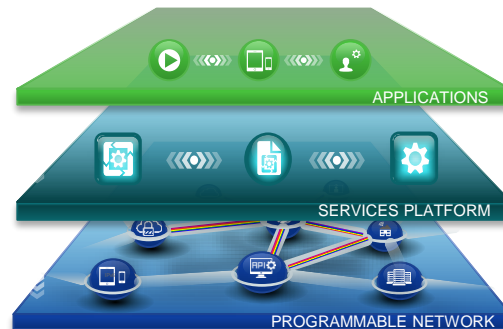
IP NGN Era	Programmable Network Era
Designed to support a set of services	Designed to support any kind of services
Static traffic patterns	Dynamic traffic patterns
Manual configuration (CLI)	Automation (APIs, Controllers, …)
Apps Independent of Network	App & Network Interaction

Applications & Network Interaction
Implications for the Network Fabric

Many applications with
dynamic and changing traffic patterns

IP Networks	IP Networks & Traffic Engineering

Shortest path with QoS

Traffic-engineered tunneling

Limitations

- Limited to a single network domain
- Scalability
- Configuration & troubleshooting complexity
- States to be maintained in each network node
- No RT programmability

✓ **Impediment to service creation**

✓ **Major scalability issues**

✓ **Operational challenges**

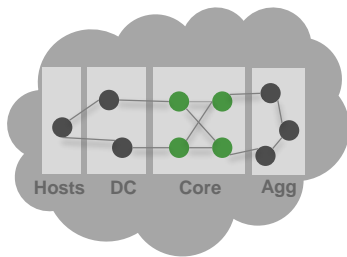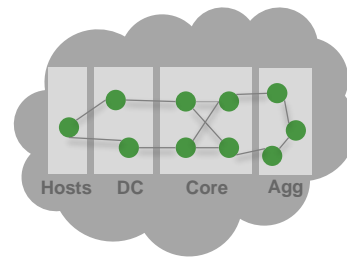✓ **Application desire != Network servicing those desires**

IP Networks Evolution

Specific approaches to the problem
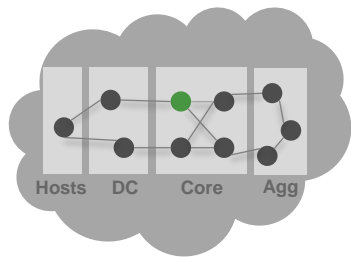
One device, single domain   **Many devices, single domain**
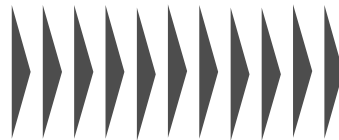
Many devices, across domains

IP NGN Era	
Policy-Based routing	**MPLS TE**

Effective solutions with some caveats:

- Little or no application / network interaction
- Scalability
- Configuration & troubleshooting complexity
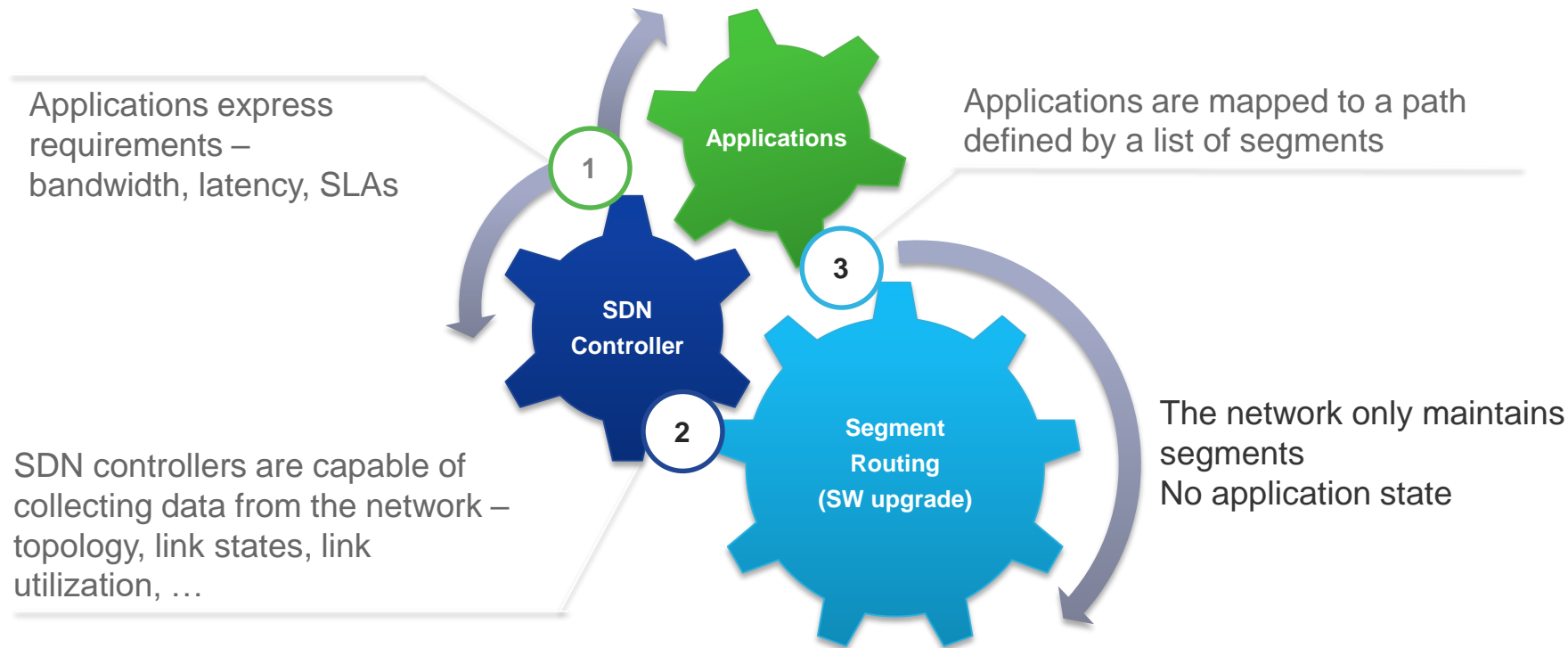- States to be maintained in each network node

Evolution required to address the new reality

Programmable Network Era

Application Engineered Routing

- Scalable

- Stateless

- Programmable

- Ease of configuration & troubleshooting

CISCO

The Solution - Application Engineered Routing

Applications

SDN Controller

Segment Routing (SW upgrade)

1

2

3

Applications express requirements – bandwidth, latency, SLAs

Applications are mapped to a path defined by a list of segments

SDN controllers are capable of collecting data from the network – topology, link states, link utilization, …

The network only maintains segments
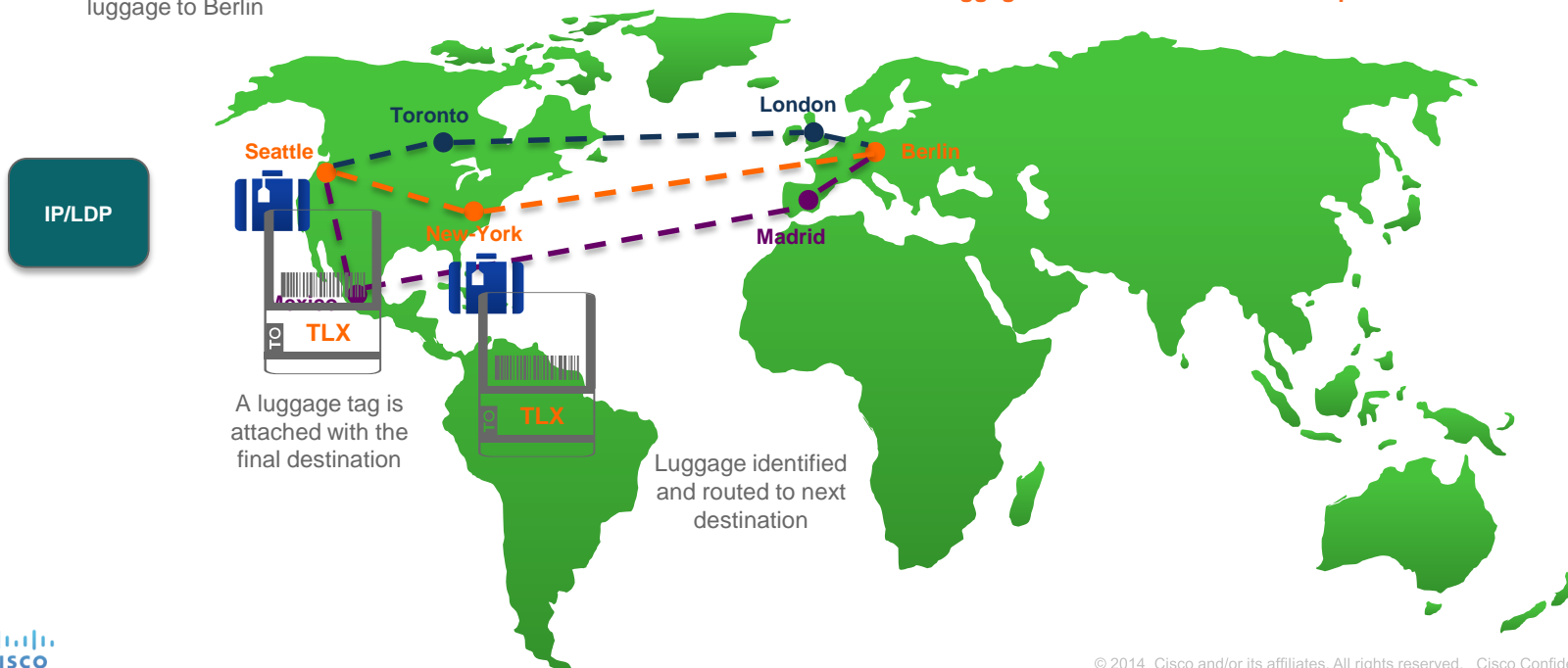No application state

Application Engineered Routing
Evolve MPLS with Segment Routing

Mission – Route the luggage to Berlin

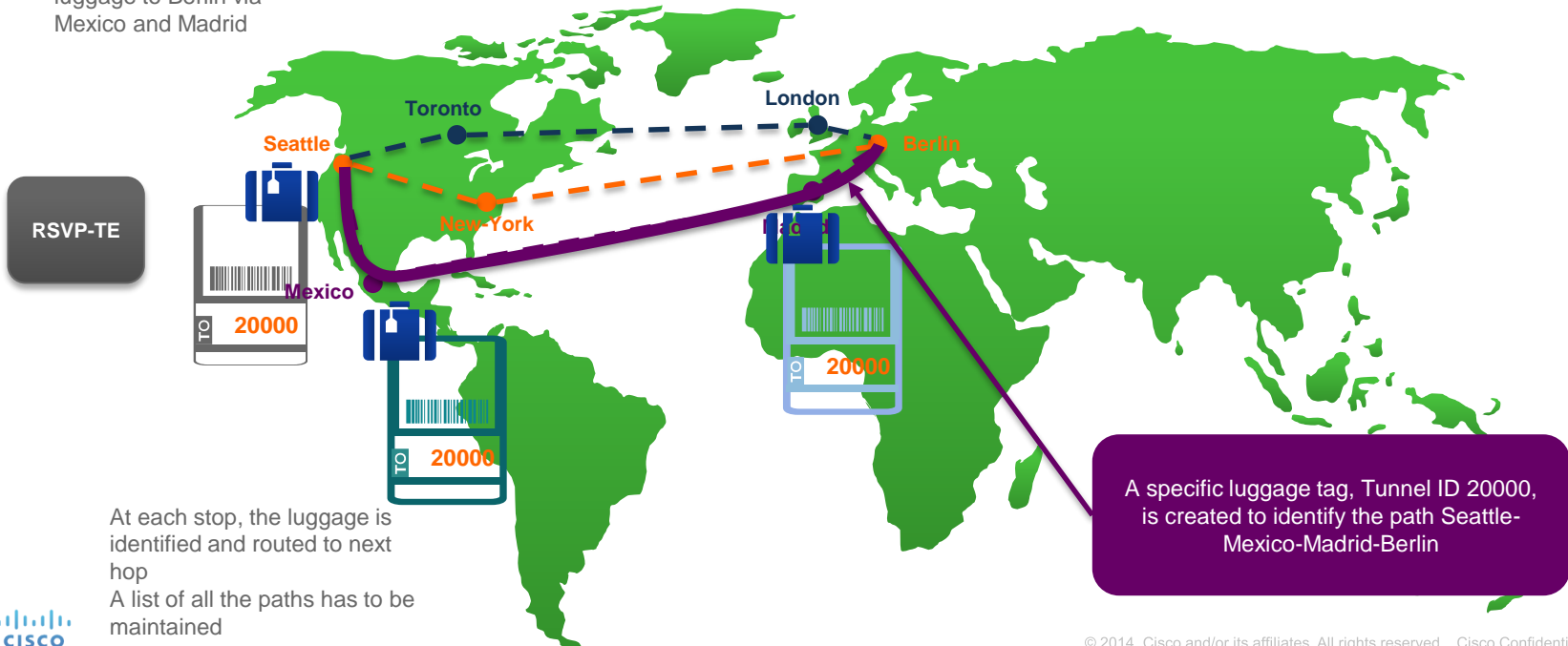No control over the path – Luggage is routed over the shortest path

Toronto

London

Seattle

Berlin

IP/LDP

New-York

Madrid

TLX

A luggage tag is attached with the final destination

TLX

Luggage identified and routed to next destination

CISCO

draft-ietf-spring

Application Engineered Routing
Evolve MPLS with Segment Routing

**Path can be controlled
Complexity and scalability
issues**

Mission – Route the
luggage to Berlin via
Mexico and Madrid

Toronto

London

Seattle

Berlin

RSVP-TE

New York

Madrid

Mexico

20000

20000

20000

At each stop, the luggage is
identified and routed to next
hop
A list of all the paths has to be
maintained

A specific luggage tag, Tunnel ID 20000,
is created to identify the path Seattle-
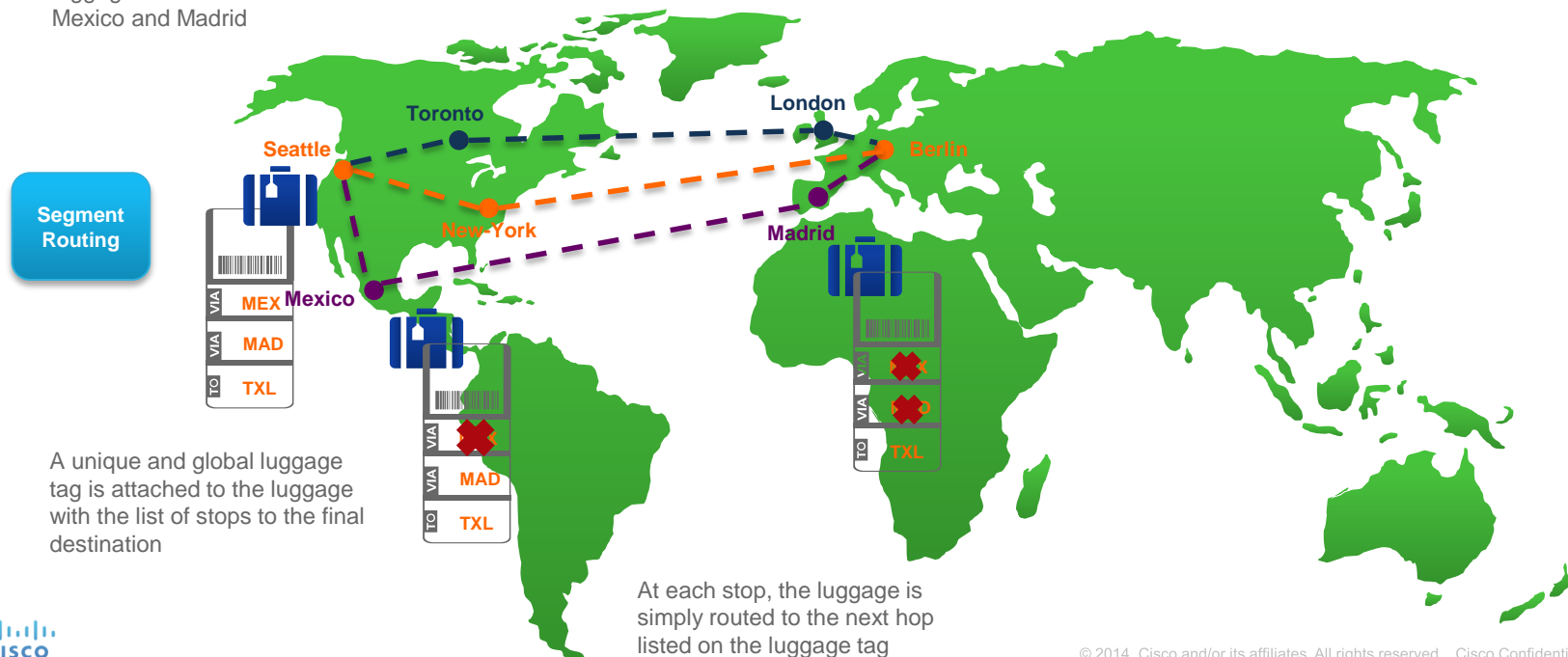Mexico-Madrid-Berlin

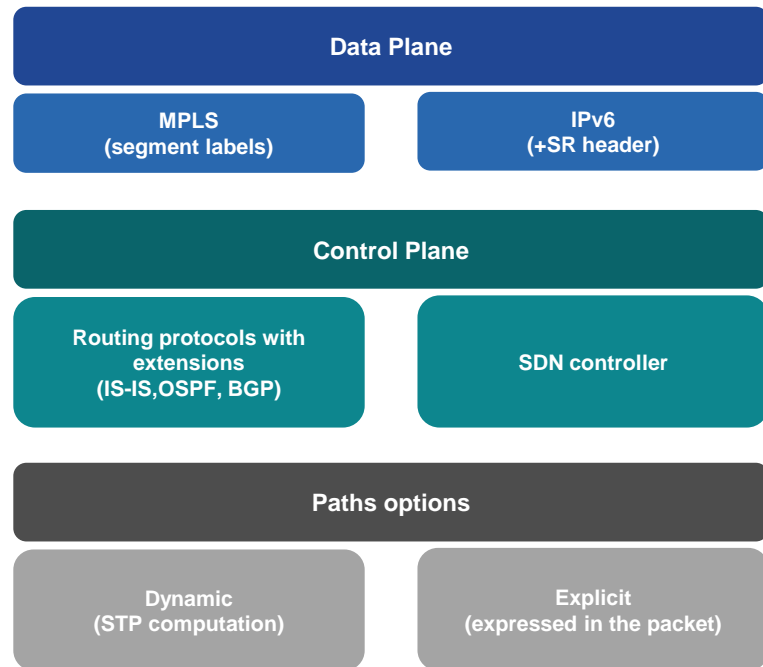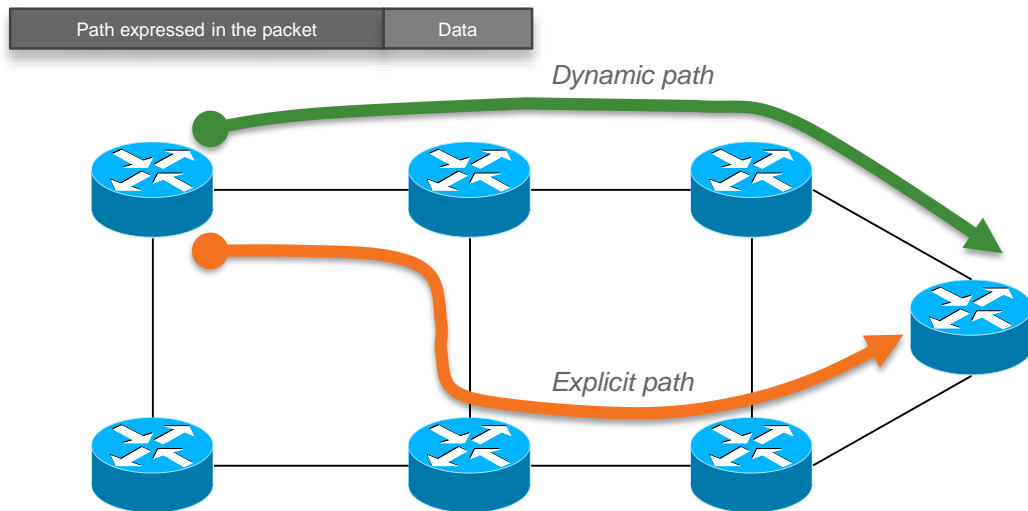# Application Engineered Routing
Evolve MPLS with Segment Routing

Mission – Route the luggage to Berlin via Mexico and Madrid

Path can be controlled
It is simple and scalable

Segment Routing

Toronto

London

Seattle

Berlin

New-York

Madrid

VIA MEX
VIA MAD
TO TXL

Mexico

VIA ✖
VIA MAD
TO TXL

VIA ✖
VIA ✖
TO TXL

A unique and global luggage tag is attached to the luggage with the list of stops to the final destination

At each stop, the luggage is simply routed to the next hop listed on the luggage tag

Application Engineered Routing
Segment Routing – Technical view

| Path expressed in the packet | Data |

Dynamic path

Explicit path

Data Plane

| MPLS (segment labels) | IPv6 (+SR header) |

Control Plane

| Routing protocols with extensions (IS-IS, OSPF, BGP) | SDN controller |

Paths options

| Dynamic (STP computation) | Explicit (expressed in the packet) |

Segment Routing

- Source Routing
  - the source chooses a path and encodes it in the packet header as an ordered list of segments
  - the rest of the network executes the encoded instructions without any further per-flow state

- Segment: an identifier for any type of instruction
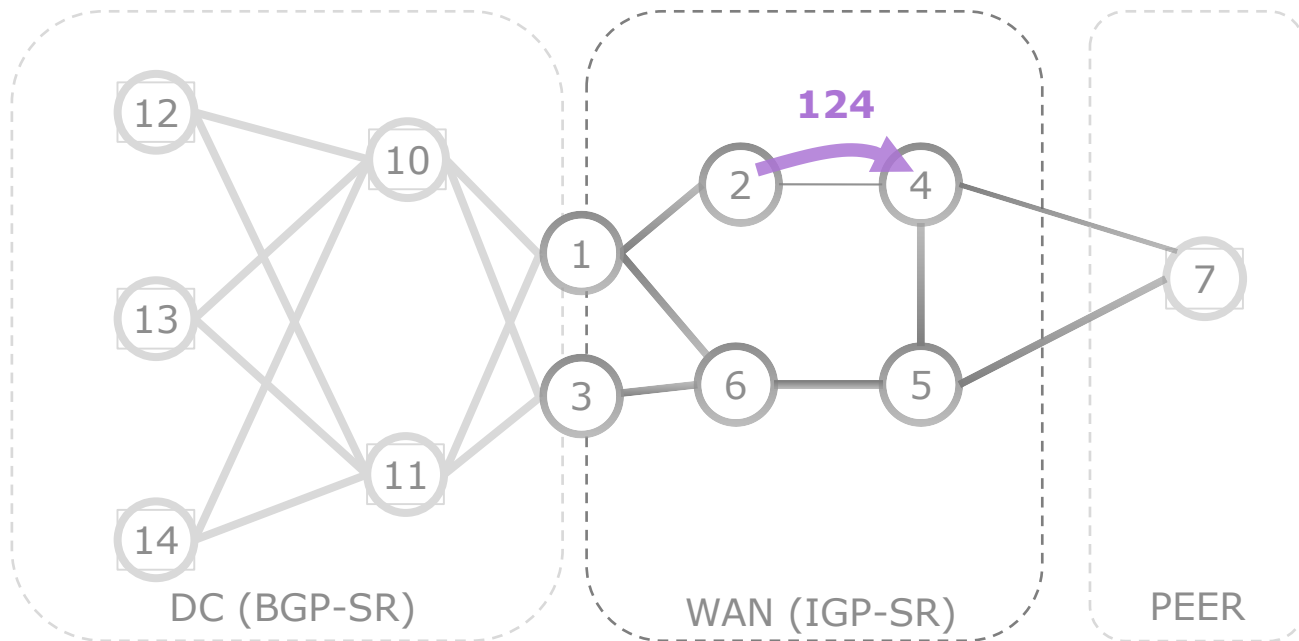 - forwarding or service

IGP Prefix Segment

- Shortest-path to the IGP prefix
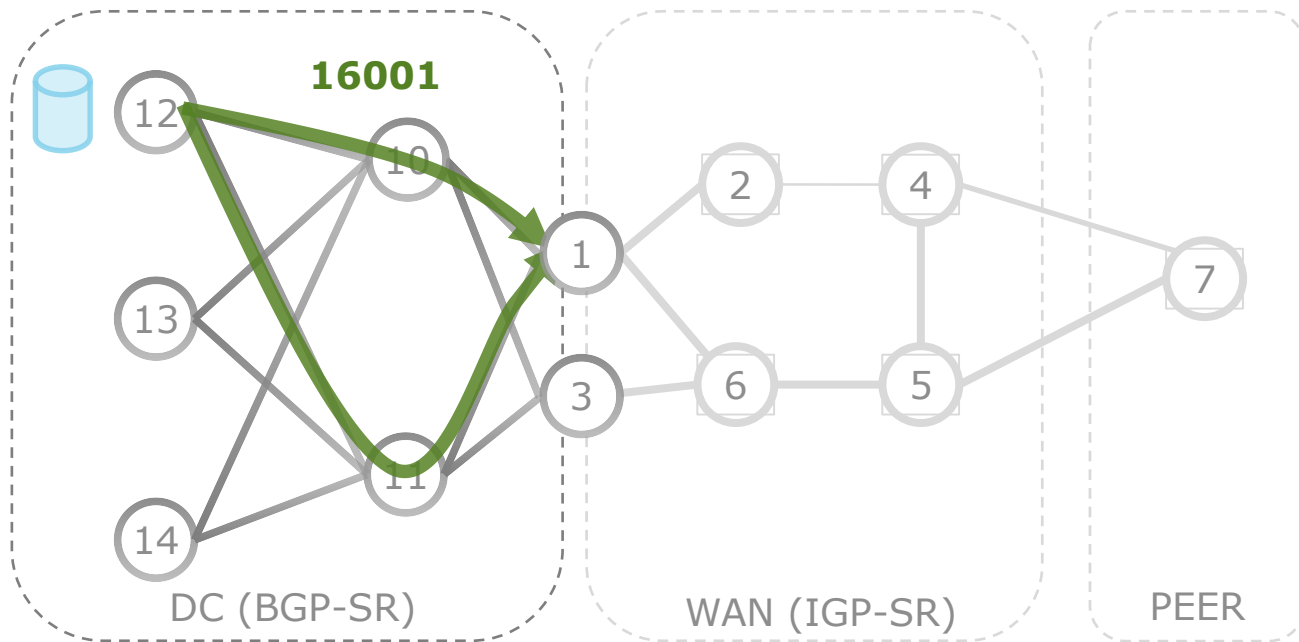
- Global

- 16000 + Index

- Signaled by ISIS/OSPF
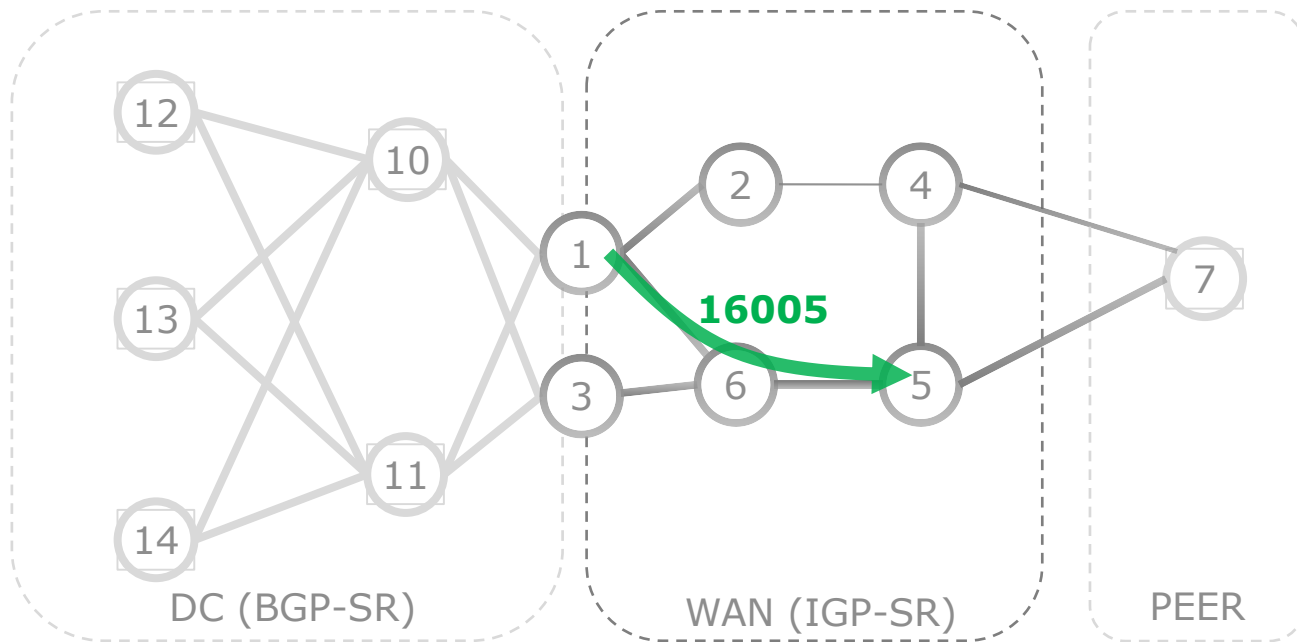
IGP Adjacency Segment

- Forward on the IGP adjacency

- Local

- 1XY
  - X is the "from"
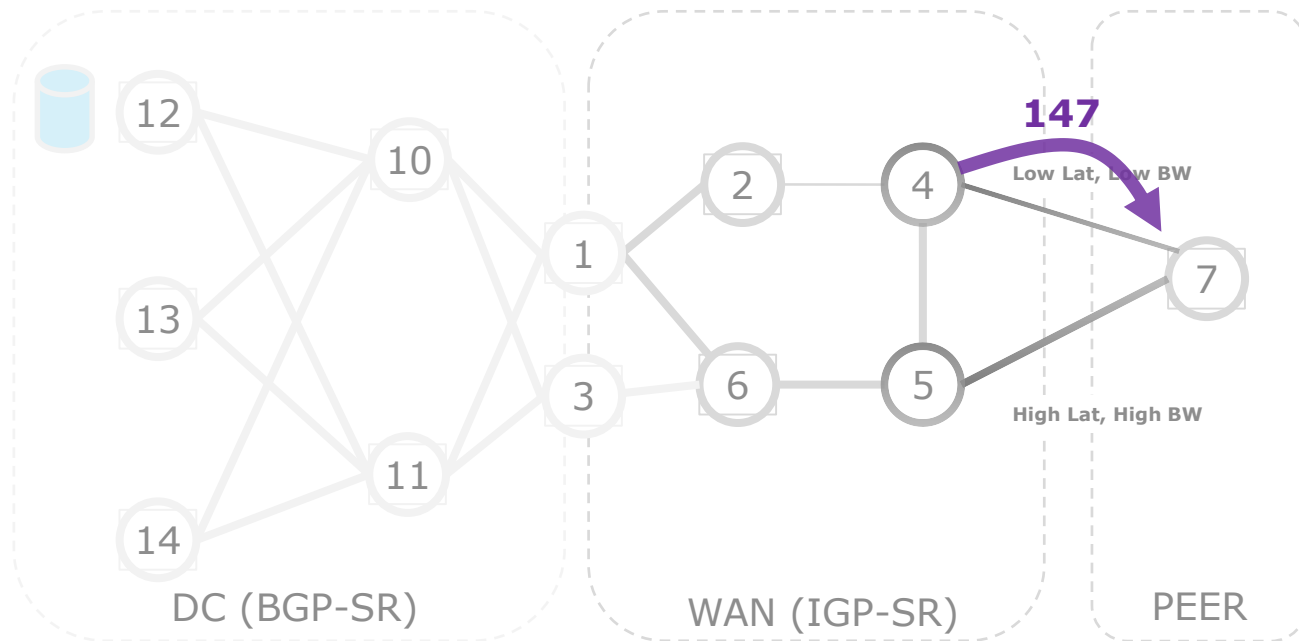 - Y is the "to"

- Signaled by ISIS/OSPF

BGP Prefix Segment

- Shortest-path to the BGP prefix

- Global

- 16000 + Index

- Signaled by BGP

16001

12   10

13   1

14   3

11   2   4

6   5   7

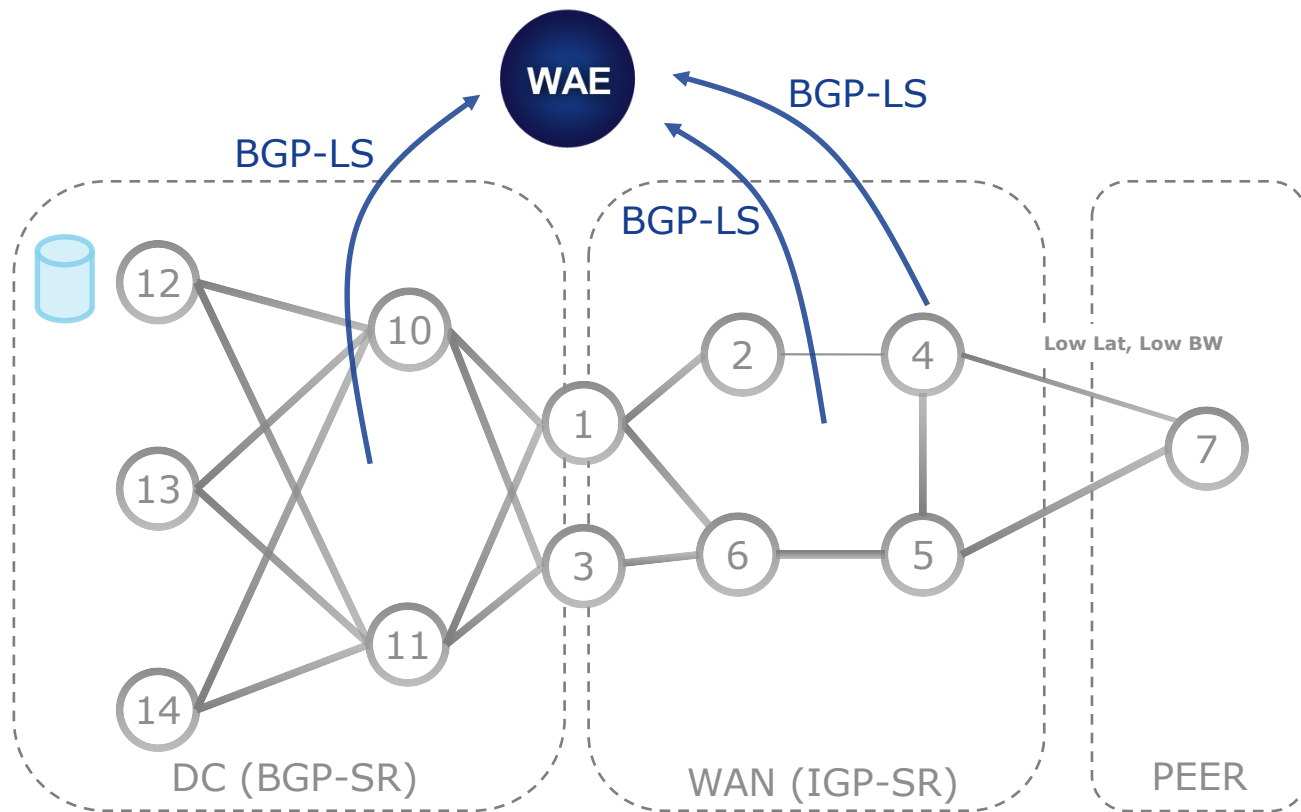DC (BGP-SR)   WAN (IGP-SR)   PEER

BGP Peering Segment

- Forward to the BGP peer

- Local

- 1XY
  - X is the "from"
  - Y is the "to"

- Signaled by BGP-LS (topology information) to the controller

147

Low Lat, Low BW

High Lat, High BW
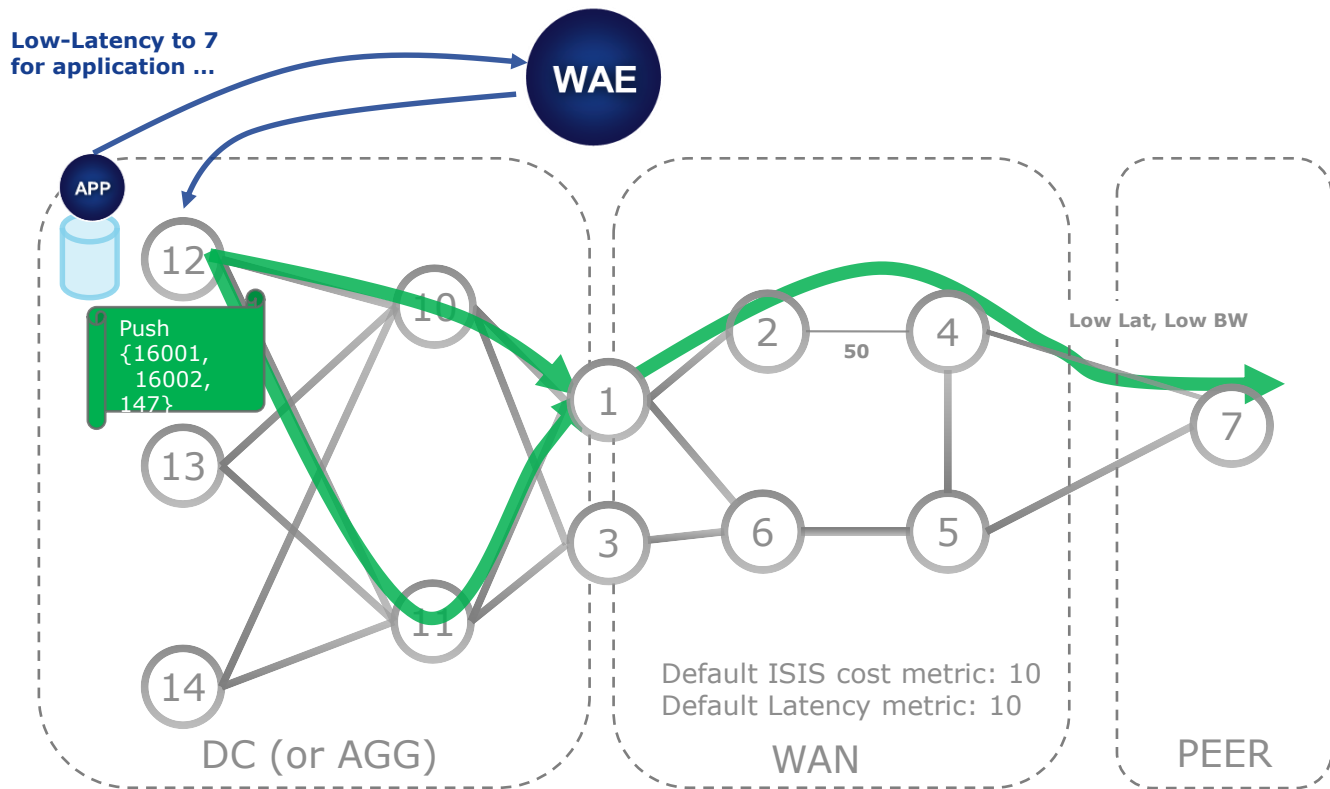
DC (BGP-SR)

WAN (IGP-SR)

PEER

WAN Controller

- WAE collects via BGP-LS
  - IGP segments
  - BGP segments
  - Topology

An end-to-end path with binding segment

- WAE computes that the green path can be encoded as
  - 16001
  - **16002**
  - 147

- WAE programs a single per-flow state to create an application-engineered end-to-end policy

Low-Latency to 7 for application ...

WAE

APP

Push {16001, 16002, 147}

12 10 1 2 50 4 7

13 3 6 5

14 11

Low Lat, Low BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG) WAN PEER

Application Engineered Routing

- Applications program the network on a per-flow basis

- End-to-End policy
  - DC, WAN, AGG, PEER

- Millions of flows
  - No per-flow midpoint state
  - No reclassification at boundaries

- Simple
  - BGP and ISIS/OSPF

High-BW to 7 for application ...

WAE

APP

Push {16001, 16005}

12
10
13
2   50   4
1
3   6   5
11
14

Low Lat, Low BW
High Lat, High BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)   WAN   PEER

7

Application Engineered Routing

High-BW to 7
for application ...

WAE

APP

Push
{16001,
16005}

12

10

13

11

14

1

3

6

5

2

4

50

7

Low Lat, Low BW

High Lat, High BW

• Automated 50msec FRR

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)

WAN

PEER

CISCO

Application Engineered Routing

High-BW to 7
Load-share across DCedges
for application ...

WAE

APP

Push
{16008,
16005}

12

10

13

14

11

8

8

2

4

6

5

2

50

7

Low Lat, Low BW

High Lat, High BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)

WAN

PEER

- Any policy can be programmed by the application

- The network scaling and simplicity is preserved
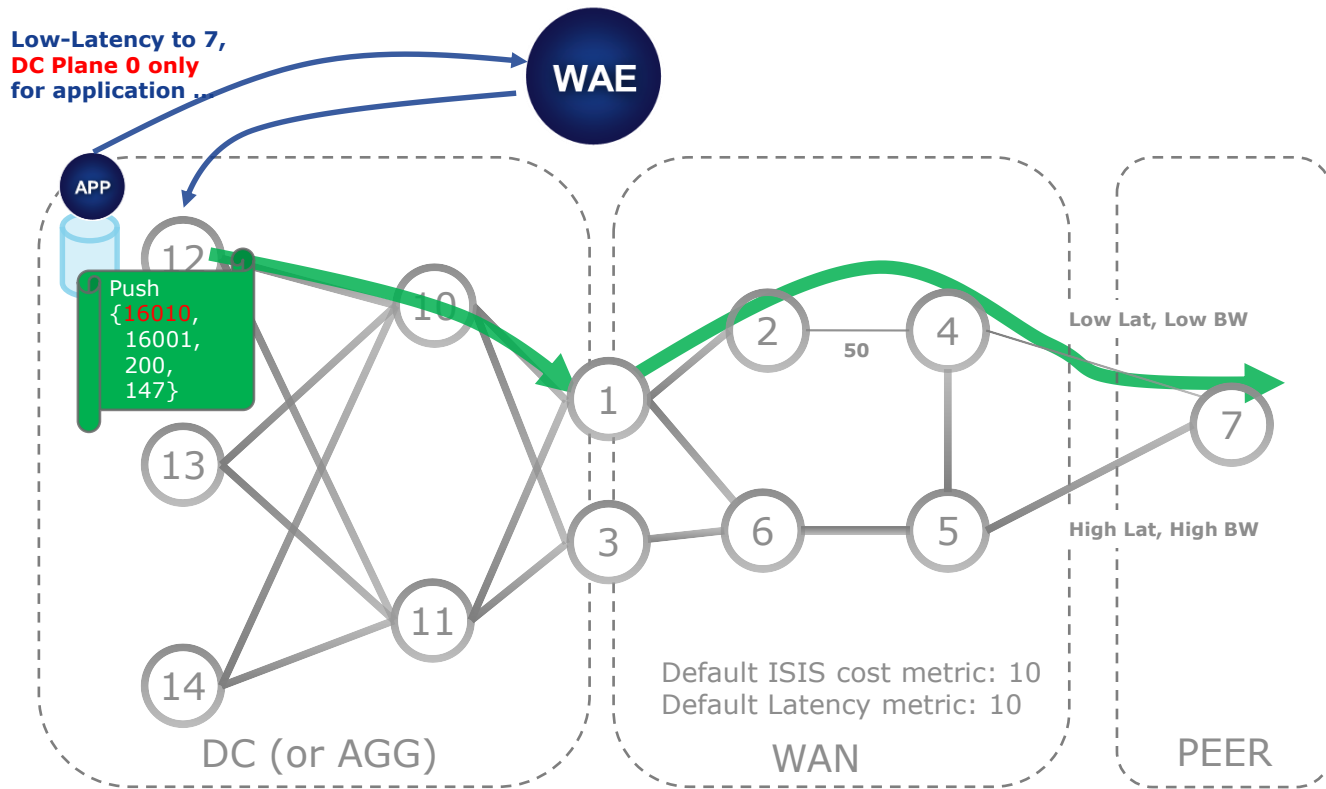
CISCO
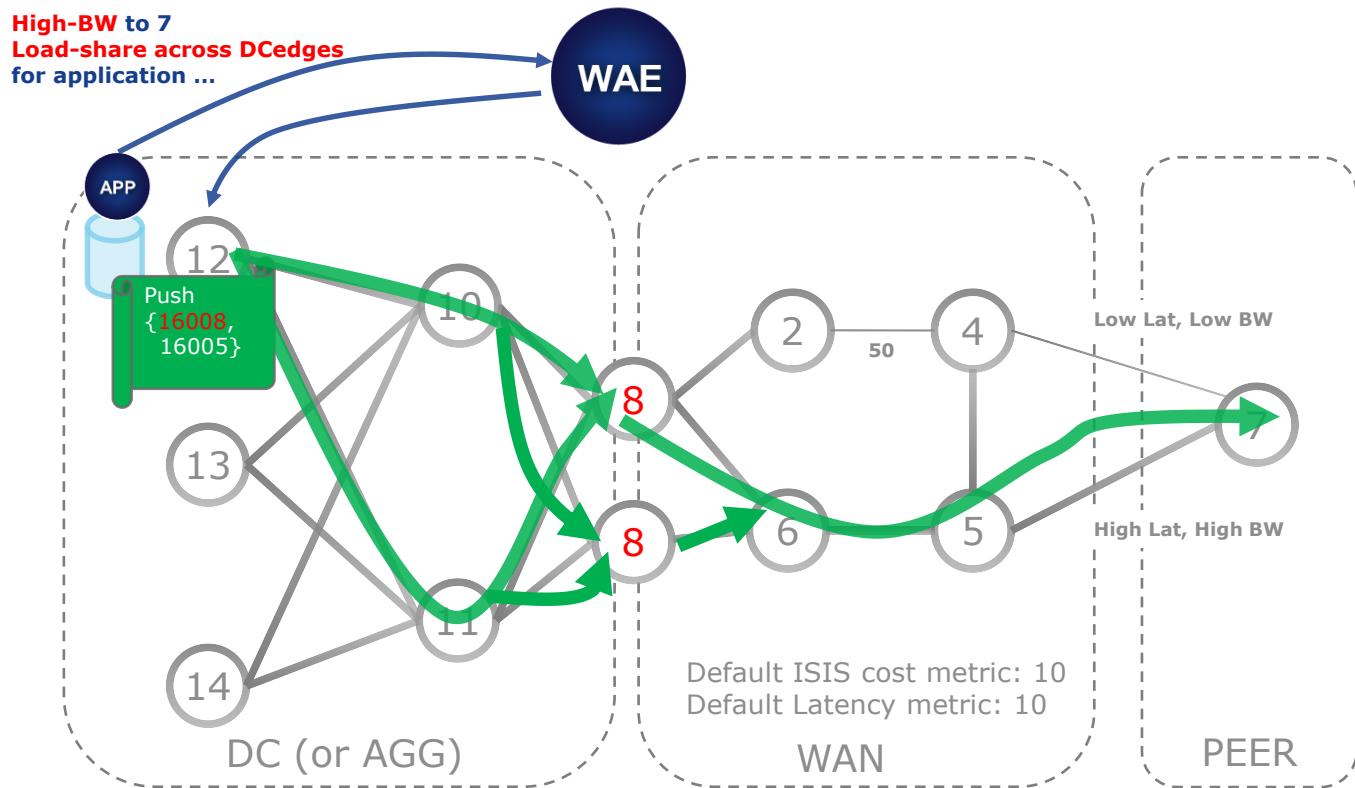
Application Engineered Routing

- Any policy can be programmed by the application

- The network scaling and simplicity is preserved

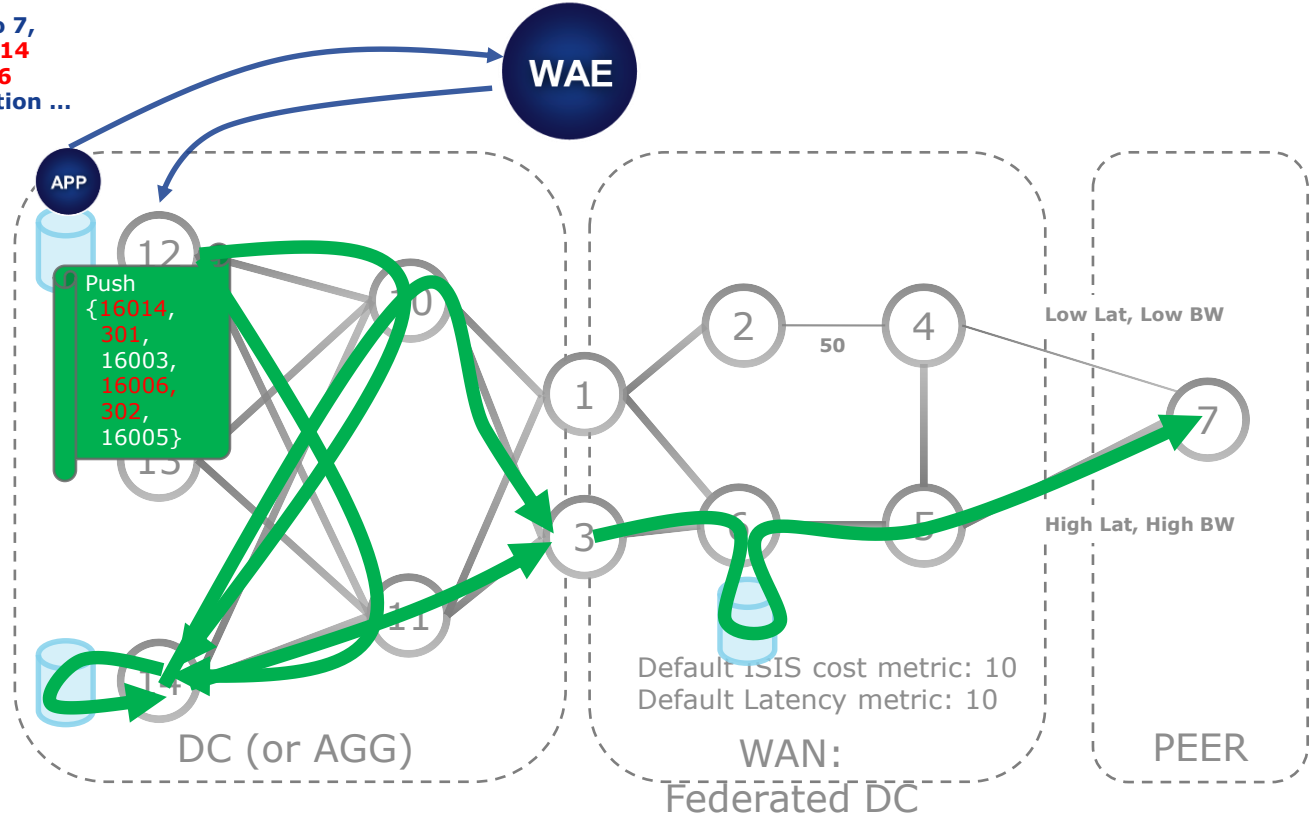Low-Latency to 7,
DC Plane 0 only
for application ...

WAE

APP

Push
{16010,
16001,
200,
147}

Low Lat, Low BW

High Lat, High BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)

WAN

PEER

Application Engineered Routing

High-BW to 7,
1st VNF at 14
2nd VNF at 6
for application ...

WAE

APP

Push
{16014,
301,
16003,
16006,
302,
16005}

2   50   4

1

3   6   5   7

Low Lat, Low BW

High Lat, High BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)   WAN:
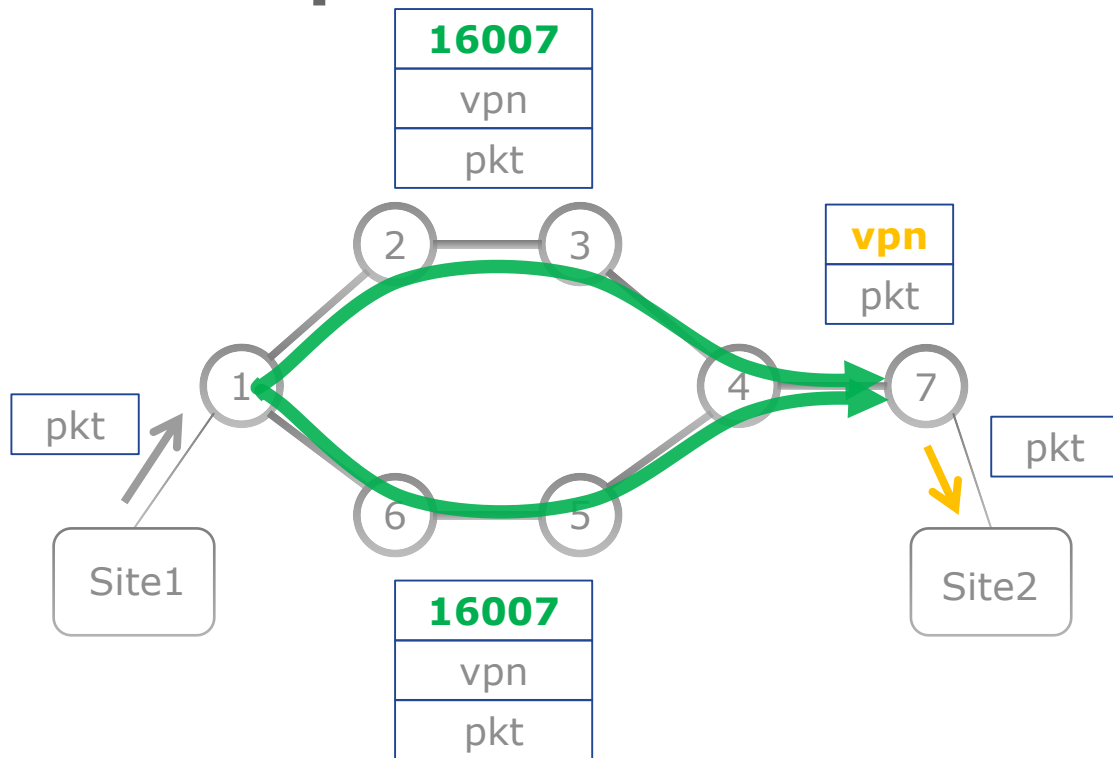   Federated DC   PEER

- Any policy can be programmed by the application

- The network scaling and simplicity is preserved

# Incremental Deployment Use-Cases

IPv4/6 VPN/Service transport

- IGP only
  - No LDP, no RSVP-TE

- ECMP

TI-LFA FRR

- 50msec FRR in any topology

- IGP Automated
  - No LDP, no RSVP-TE

- Optimum
  - Post-convergence path

- No midpoint backup state

- Detailed operator report
  - S. Litkowski, B. Decraene, Orange

- Mate Design
  - How many backup segments
  - Capacity analysis

16007
pkt

16007
pkt

16005
16007
pkt
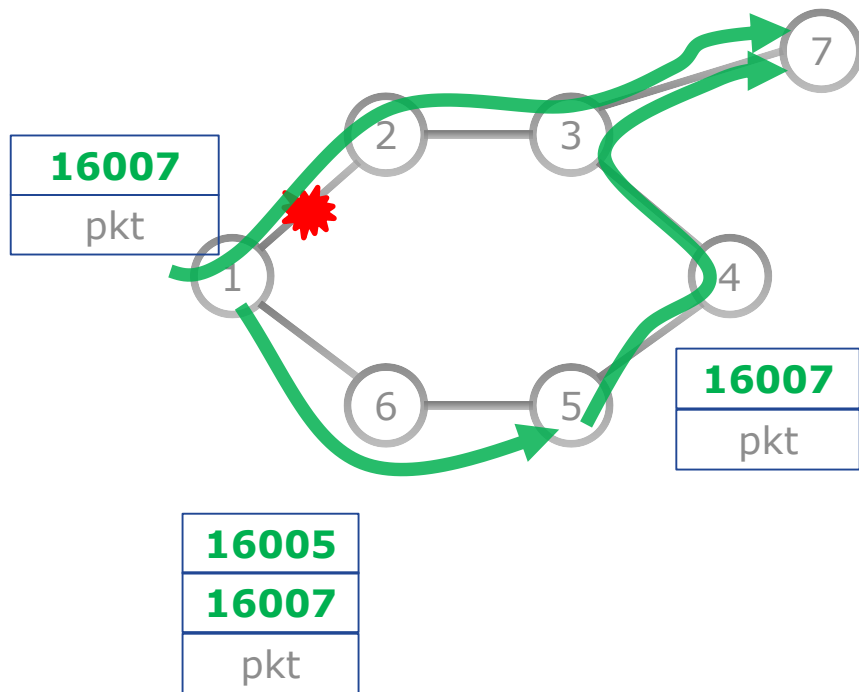
Automated Traffic Matrix Collection

- Traffic Matrix is fundamental for
  - capacity planning
  - centralized traffic engineering
  - IP/Optical optimization

- Most operators do not have an accurate traffic matrix
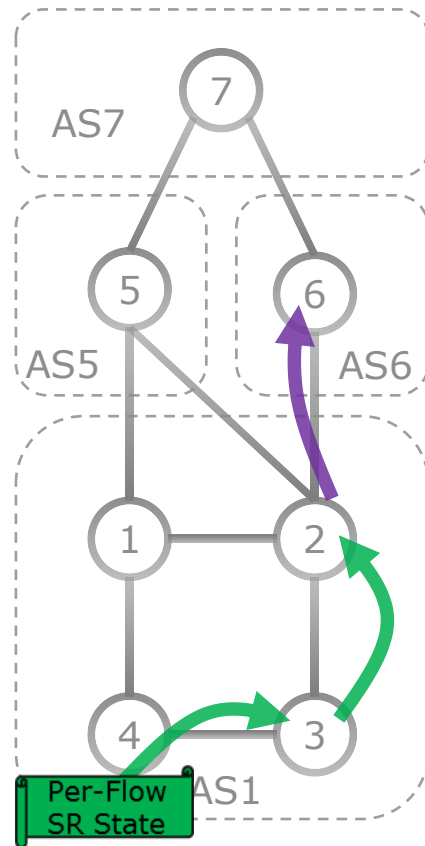
- With SR, the traffic matrix collection is automated

	1	2	3	4
1				
2				
3				
4				
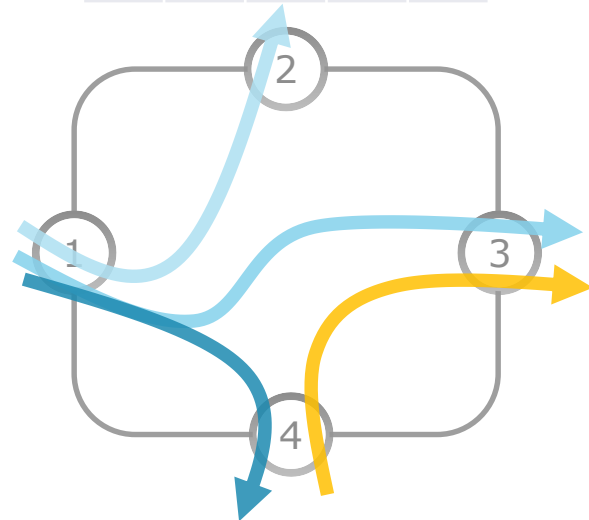
Optimized Content Delivery

- On a per-content, per-user basis, the content delivery application can engineer
  - the path within the AS
  - the selected border router
  - the selected peer

- Also applicable for engineering egress traffic from DC to peer
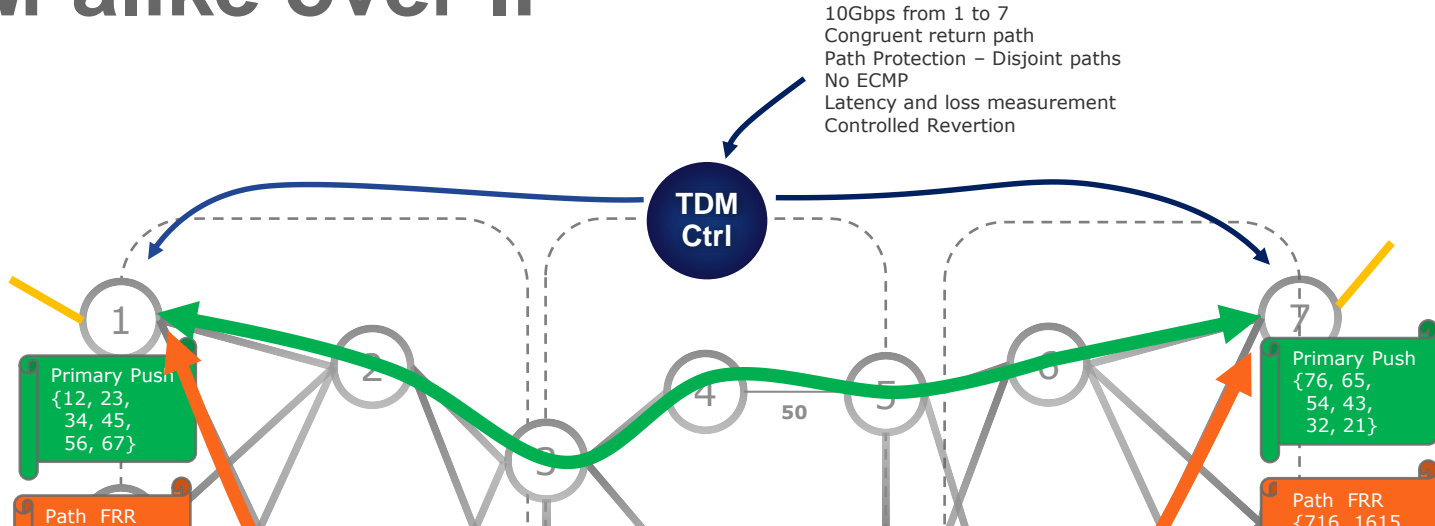 - BGP Prefix and Peering Segments

MPLS services across 100k nodes

PW Service 10 to 20

SDN

Push
{16001,
16002,
16020}

Default ISIS cost metric: 10
Default Latency metric: 10

METRO1

WAN

METRO2

TDM-alike over IP

10Gbps from 1 to 7
Congruent return path
Path Protection – Disjoint paths
No ECMP
Latency and loss measurement
Controlled Revertion

TDM Ctrl

Primary Push
{12, 23,
34, 45,
56, 67}

Primary Push
{76, 65,
54, 43,
32, 21}

Path FRR

Path FRR
{716, 1615

50
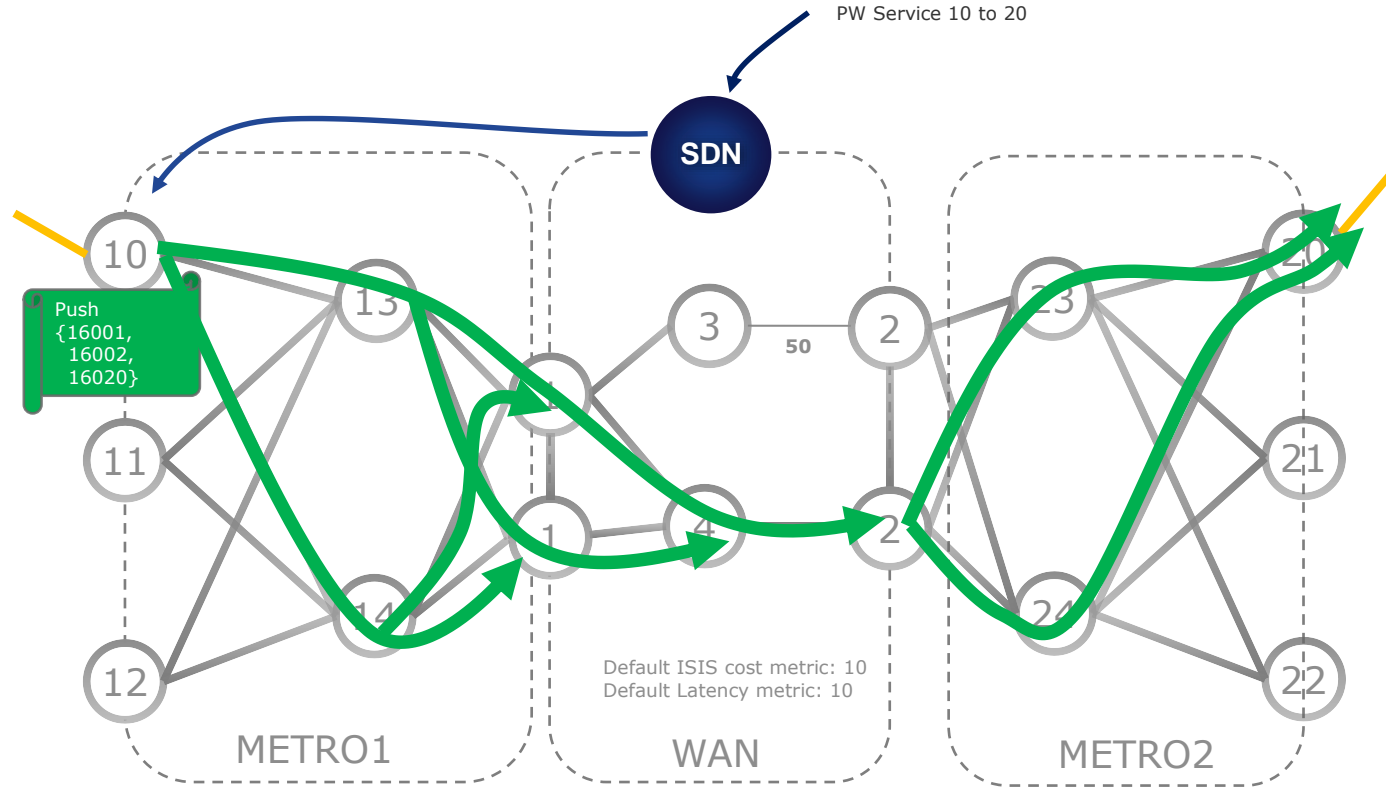
The centralized nature of SR lands very well to provide TDM-alike services over an IP-optimized infrastructure
Controller completely controls the path of the service (BW and latency guarantees)
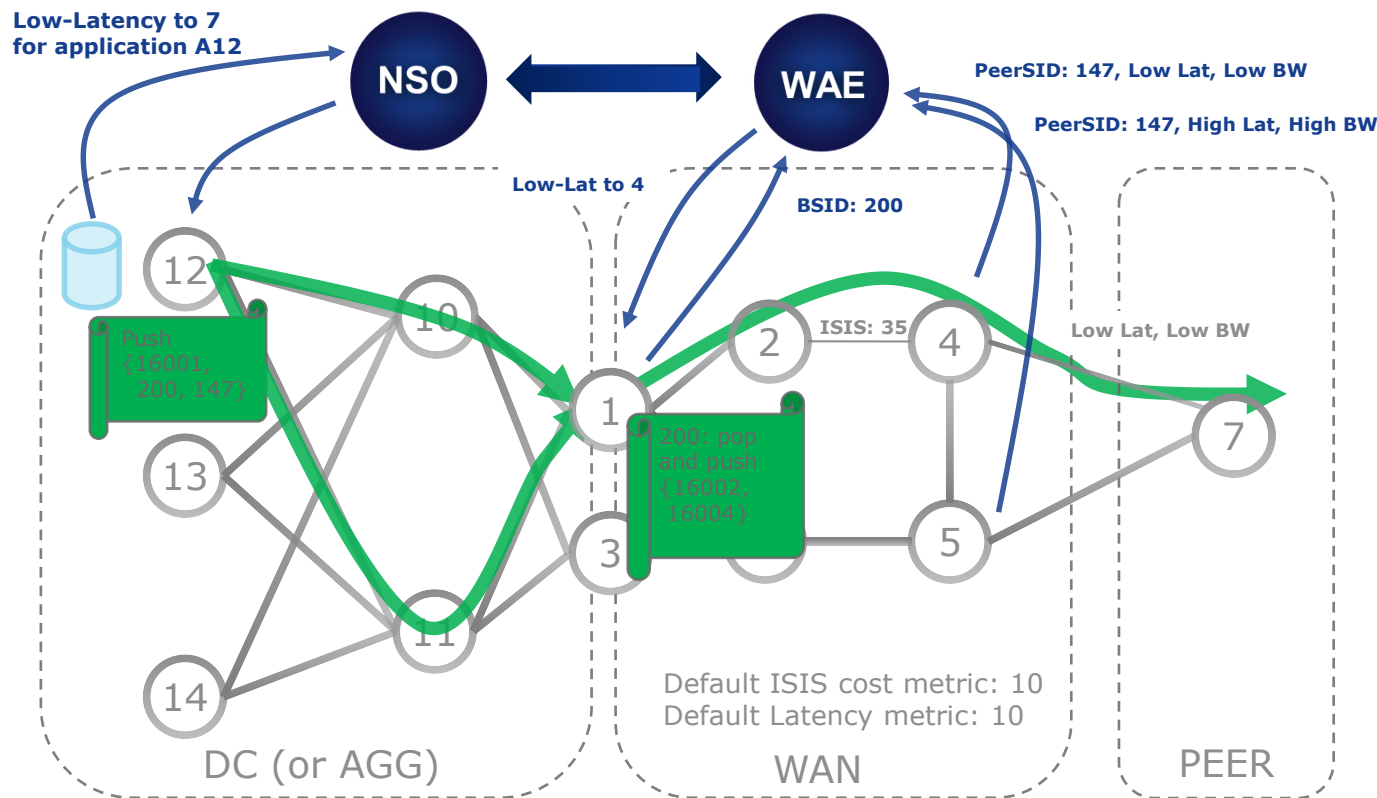No dependency on any IP signalling.
No ECMP
End-to-end service through metro's and WAN
TDM-alike service (path protection over disjoint paths, performance monitoring, no ECMP)

METRO1   WAN   METRO2
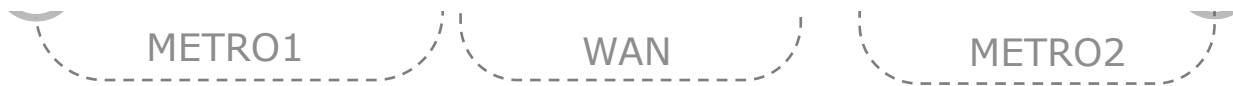
Application Engineered Routing

- Per-application flow engineering
- End-to-End
  - DC, WAN, AGG, PEER
- Millions of flows
  - No signaling
  - No midpoint state
  - No reclassification at boundaries

Low-Latency to 7 for application A12

NSO ⬌ WAE

PeerSID: 147, Low Lat, Low BW
PeerSID: 147, High Lat, High BW

Low-Lat to 4

BSID: 200

Push {16001, 200, 147}

12
10
2   ISIS: 35   4
1
13
200: pop and push {16002, 16004}
3
5
11
14
7

Low Lat, Low BW

Default ISIS cost metric: 10
Default Latency metric: 10

DC (or AGG)   WAN   PEER

Application Engineered Routing Journey
Adding value at your own pace

Enable Segment Routing on the network (Software only)

Insert Orchestration, SDN controller Cisco WAE, Cisco NSO (Tail-F)

Connect with Cisco's and third party VNFs

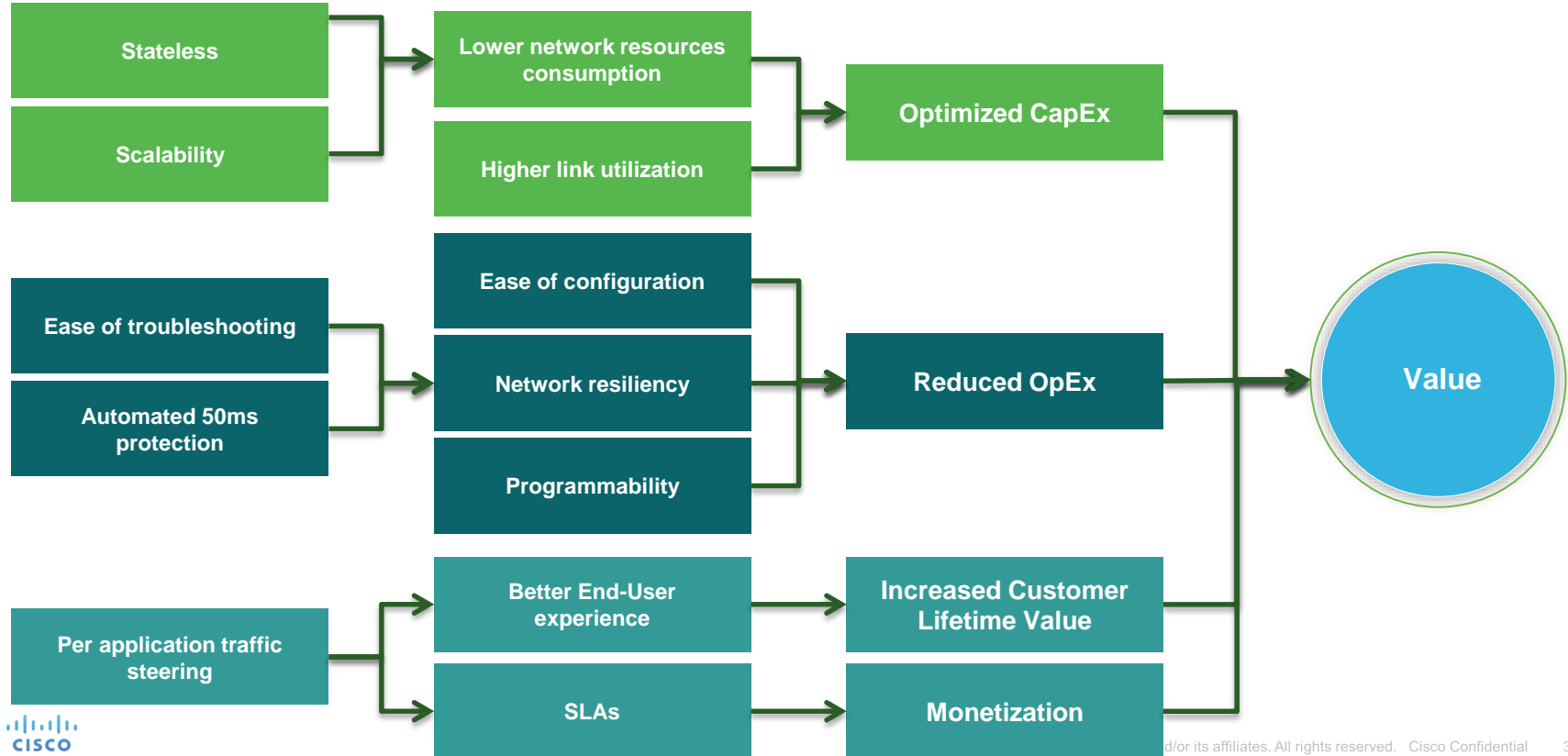| Benefits			
Network Simplification	●	●	●
Network Resiliency	●	●	●
End-User Experience	●	●	●
Network Optimization		●	●
Service Velocity			●
E2E Application Control			●

CISCO

Conclusion

Traffic Engineering with SR

- No midpoint state (n^2 scale In RSVP-TE)

- No extra protocol (RSVP-TE)

- Native ECMP

- Few segments are required
  - Apply Cisco WAE and NSO (Tail-f) on your data

- Distributed computation or Centralized
  - Optimize on Cost, Latency or BW
  - Include/exclude Address, Affinity, SRLG
  - Disjointness

- Integration with IP/Optical optimization

Technology innovation driving business outcomes

Stateless	Lower network resources consumption	
Scalability	Higher link utilization	Optimized CapEx
Ease of troubleshooting	Ease of configuration	
Automated 50ms protection	Network resiliency	Reduced OpEx
	Programmability	
Per application traffic steering	Better End-User experience	Increased Customer Lifetime Value
	SLAs	Monetization

Value

Thank you

CISCO

TOMORROW *starts here.*