

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC



# Rethink your Edge Routing Architecture

BRKSP-2275

Takahiko Kodama  
Principal Solutions Engineer  
BRKSP-2275

CISCO *Live!*

#CiscoLiveAPJC

# Acknowledgements

- Thanks to Emerson Moura, Guillermo Trueba, Johan Gustawsson, Marty Fierbaugh, Phil Bedard, Rob Piasecki, LJ Wobker and other Cisco colleagues who kindly reviewed and provided feedback on this content.

# Cisco Webex App

## Questions?

Use Cisco Webex App to chat with the speaker after the session

## How

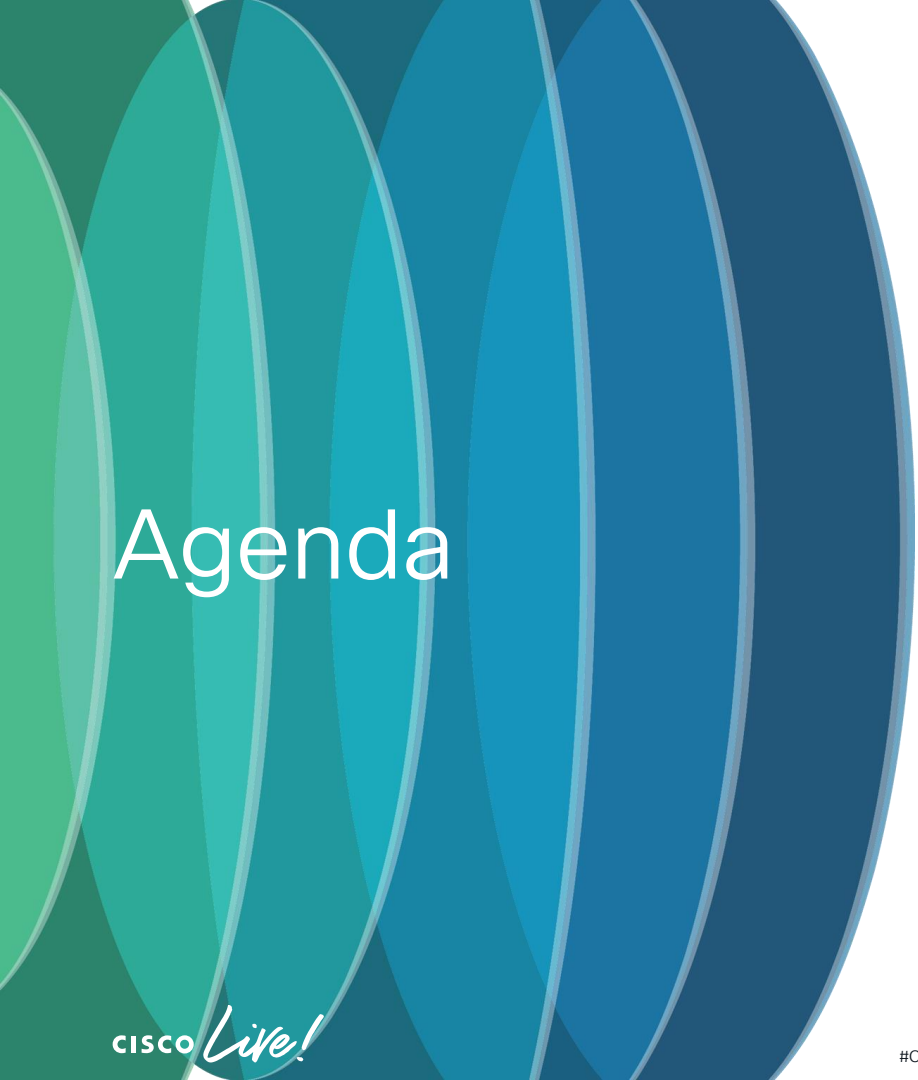
- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until November 15, 2024.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKSP-2275>

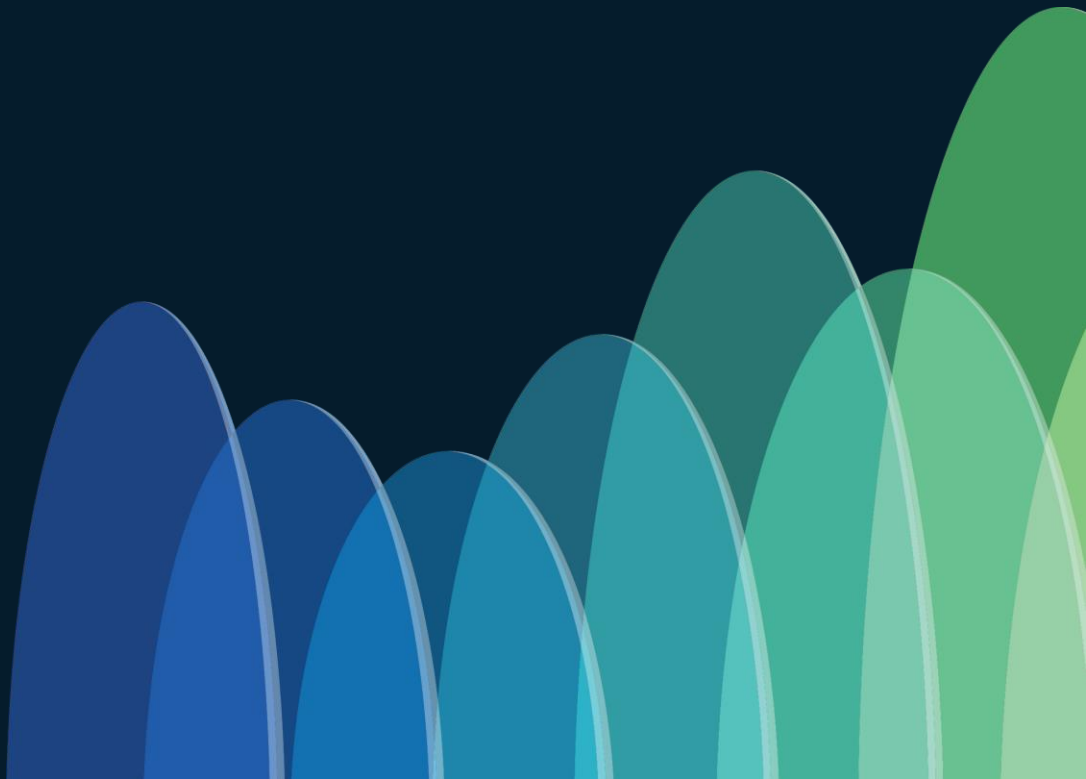




# Agenda

- Introduction – what are we trying to improve?
- Evolving Edge Routing architectures
- Design considerations
- Conclusion

# Introduction



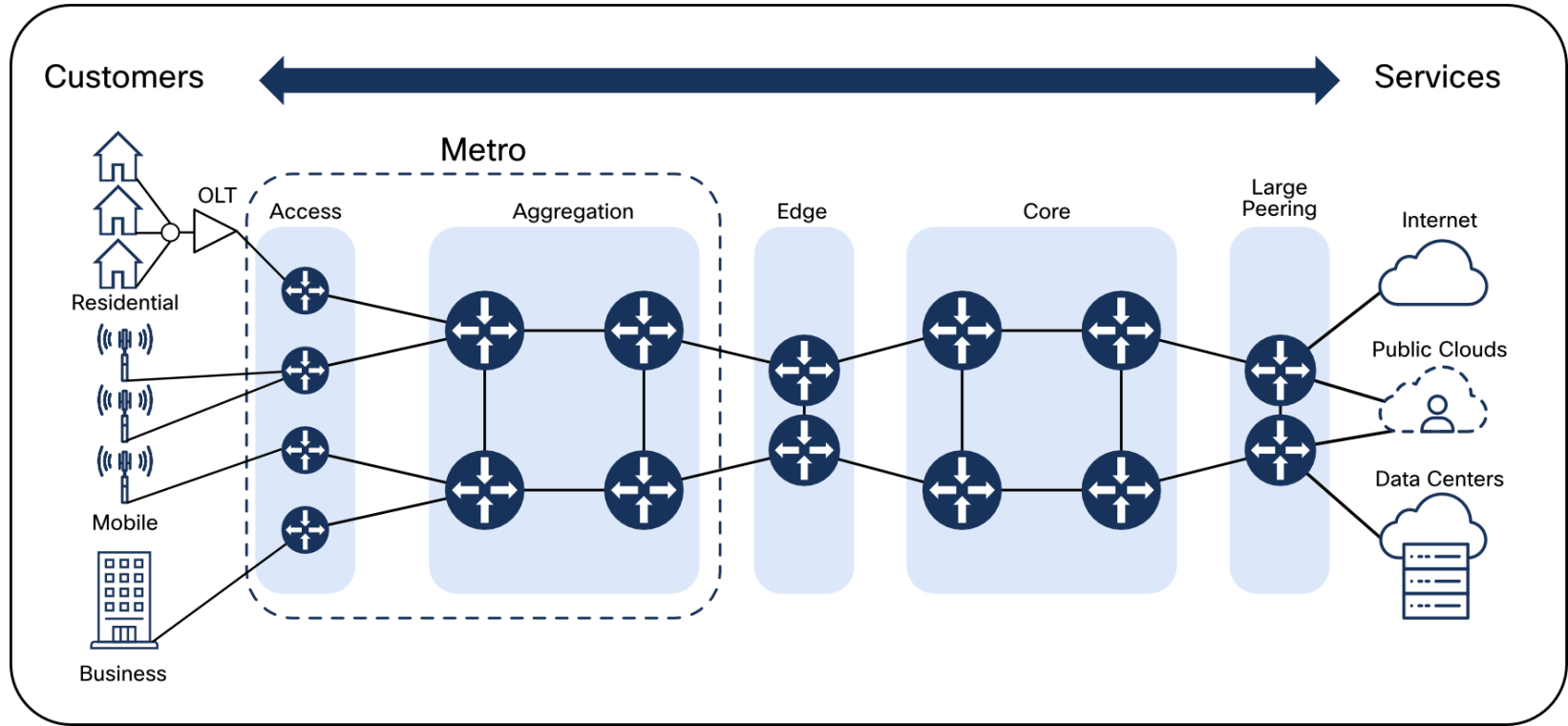
# What are we going to (do not) talk about?



## Traditional Services Edge routers:

- Subscriber services
- Business VPN services (L2, L3)
- Internet peering
- Data Center Interconnect
- Cloud gateways

# Classic network architecture



# Unique attributes of services edge routers

## Services rich

- L2 VPN
- L3 VPN
- Internet
- Subscriber
- Business
- Multicast

## High scale

- Subscribers
- Services
- Queues
- Routes IGP, BGP
- Counters
- Policies

## Carrier-class

- Everything must have redundancy
- Dual homing routers

# Cost and complexity example: Edge Routers

Practical implications of edge routing requirements



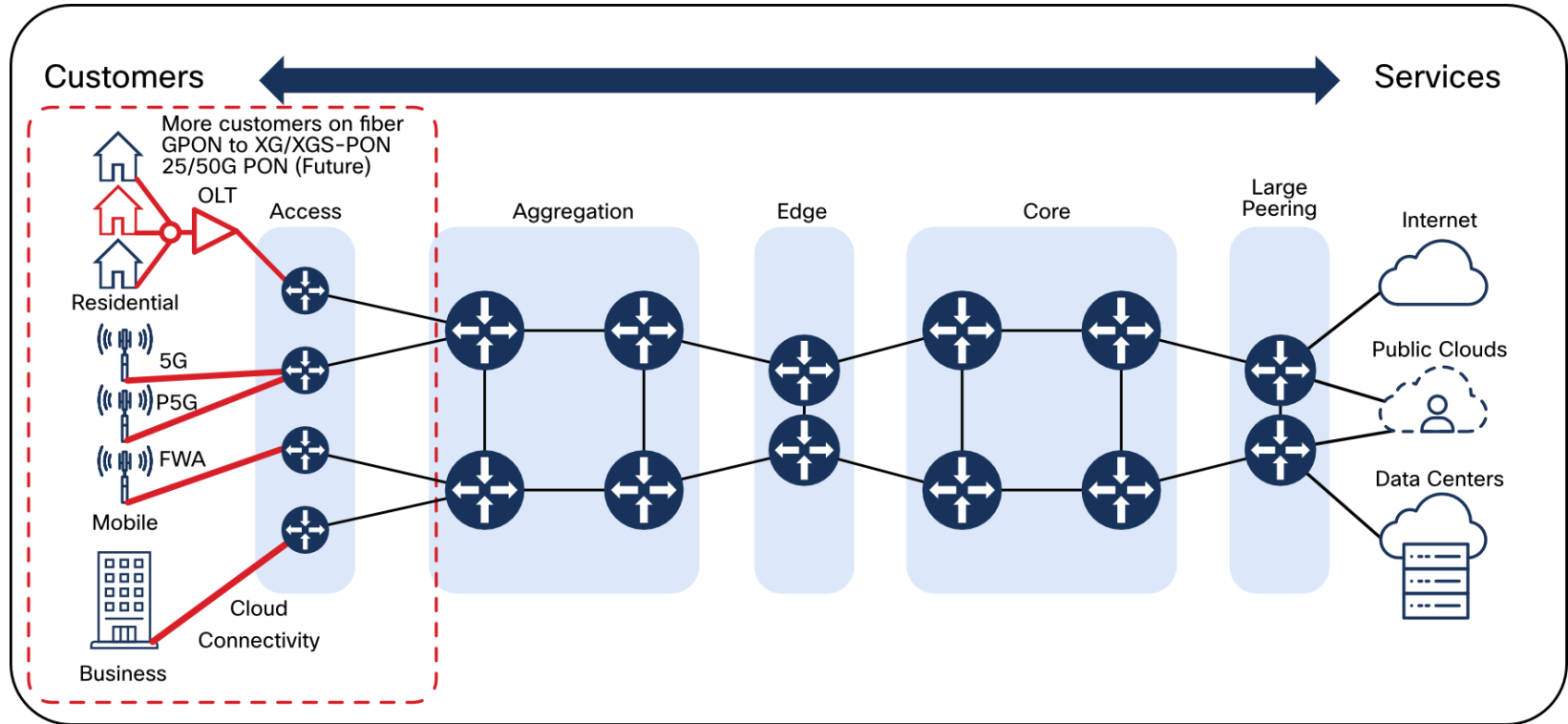
Design philosophy: Centralize it!

- Large modular chassis in as few edge locations as possible
- As many customers/services per router as possible
- Low-touch – avoid or delay major changes/upgrades.

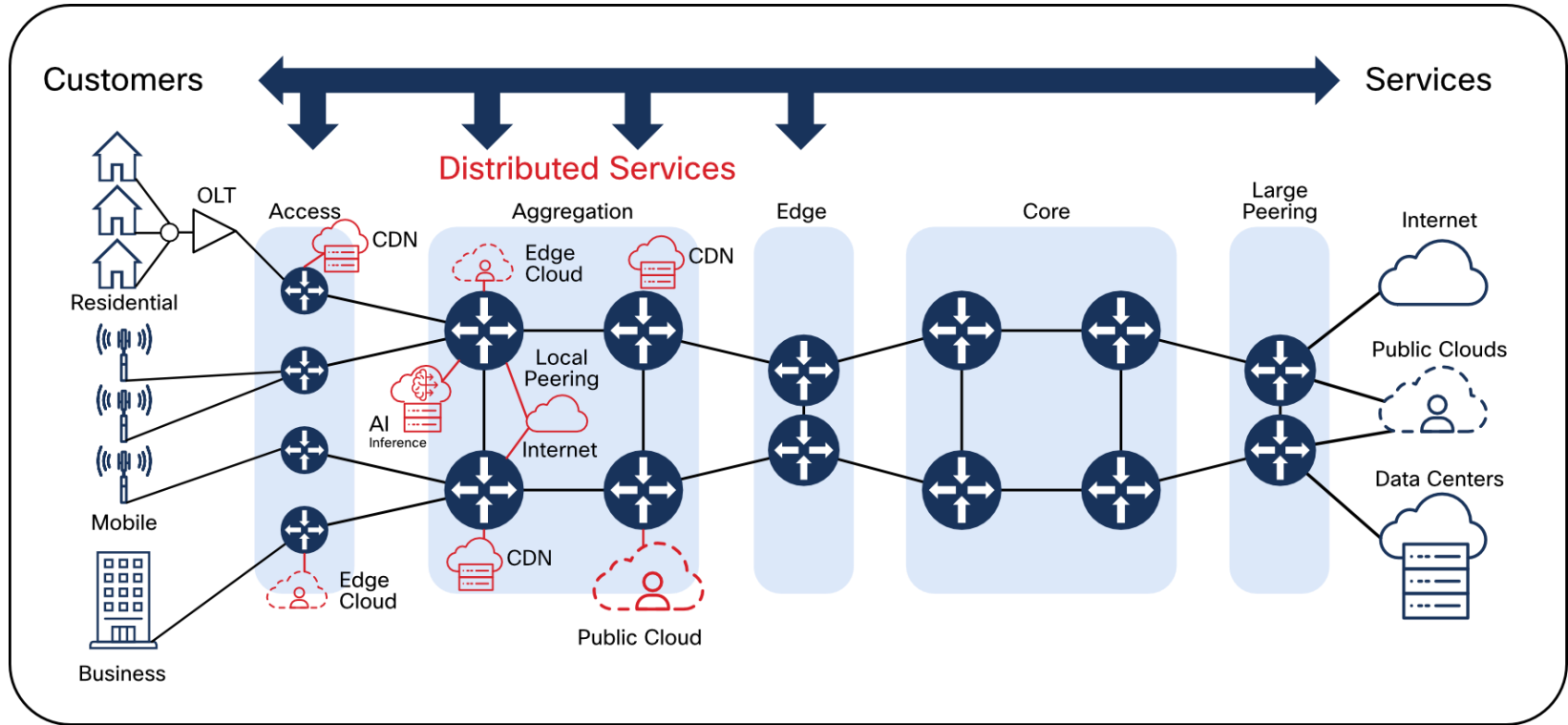
*However forklift upgrades are the norm due to compatibility challenges between technology generations. Low touch is not the reality.*

Slow innovation cycles, increased risks, late to realize technology gains

# Challenge #1 – Access bandwidth growth



# Challenge #2 – Services are being distributed



# Challenge #3 – Optimization problem

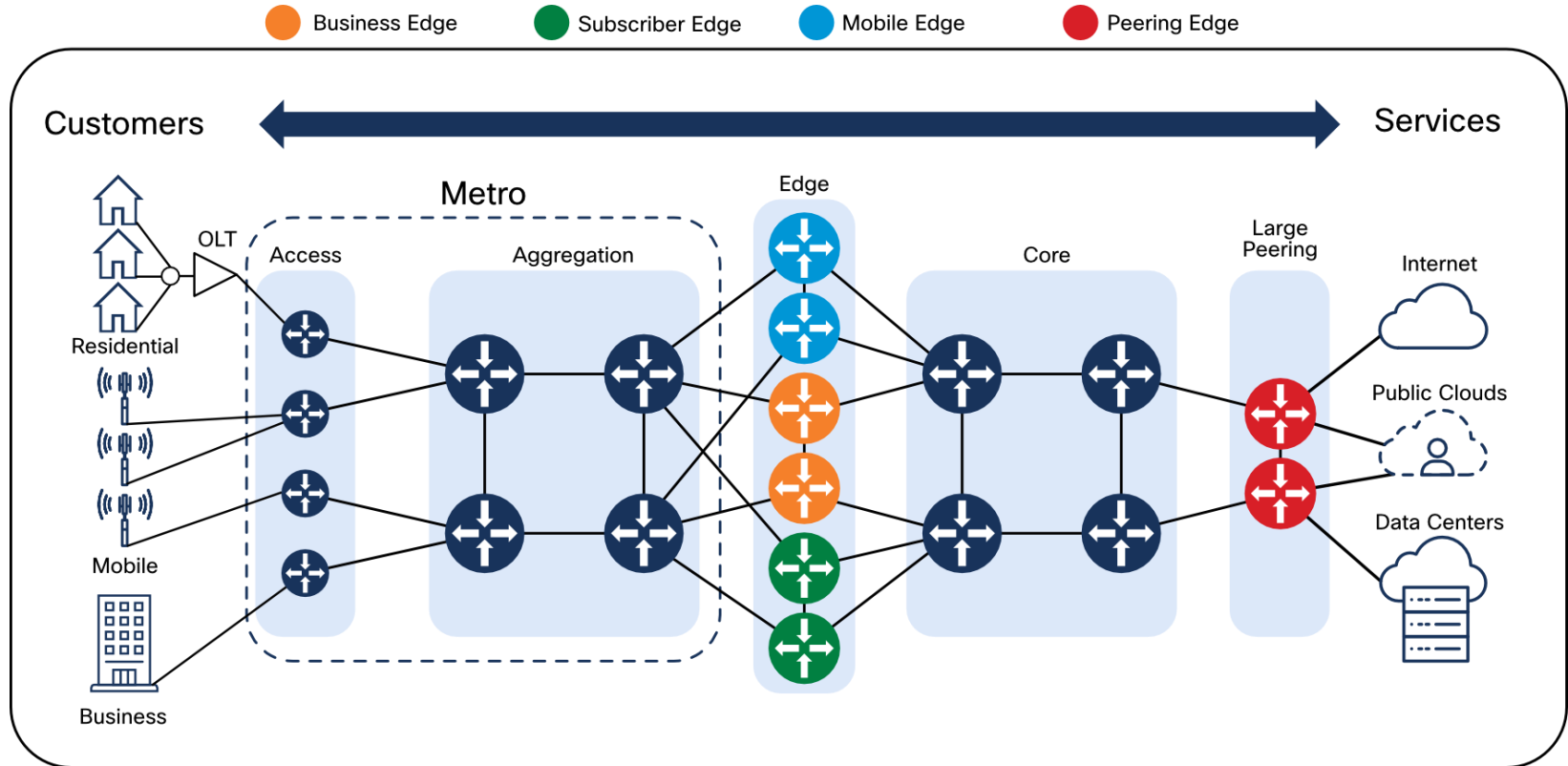
| Service            | Main Features*                                                                                                                                                                                                                     | Main scale Attributes*                                                                                                                   |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Business L2 VPN    | <ul style="list-style-type: none"><li>• VLAN tag processing</li><li>• Traffic Policers (2R3C)</li><li>• BPDU tunnelling, processing</li><li>• Service multiplexing</li><li>• Connectivity models (E-Line, E-LAN, E-Tree)</li></ul> | <ul style="list-style-type: none"><li>• Bandwidth</li><li>• MAC tables</li><li>• N. of Bridge groups</li><li>• N. PW EVPN/VPWS</li></ul> |
| Business L3 VPN    | <ul style="list-style-type: none"><li>• Import/Export route filters</li><li>• Traffic policers, Shapers</li><li>• ACLs</li><li>• Accounting</li></ul>                                                                              | <ul style="list-style-type: none"><li>• Bandwidth</li><li>• VRFs</li><li>• BGP neighbors</li><li>• Routing tables</li></ul>              |
| Subscriber (BNG)   | <ul style="list-style-type: none"><li>• Traffic policers, shapers</li><li>• Interfaces for OSS/BSS</li></ul>                                                                                                                       | <ul style="list-style-type: none"><li>• Subscriber sessions</li><li>• Queues</li><li>• Counters</li><li>• ARP entries</li></ul>          |
| Internet (Peering) | <ul style="list-style-type: none"><li>• ...</li></ul>                                                                                                                                                                              | <ul style="list-style-type: none"><li>• ...</li></ul>                                                                                    |

\* Note: Examples only. Not an all inclusive list.

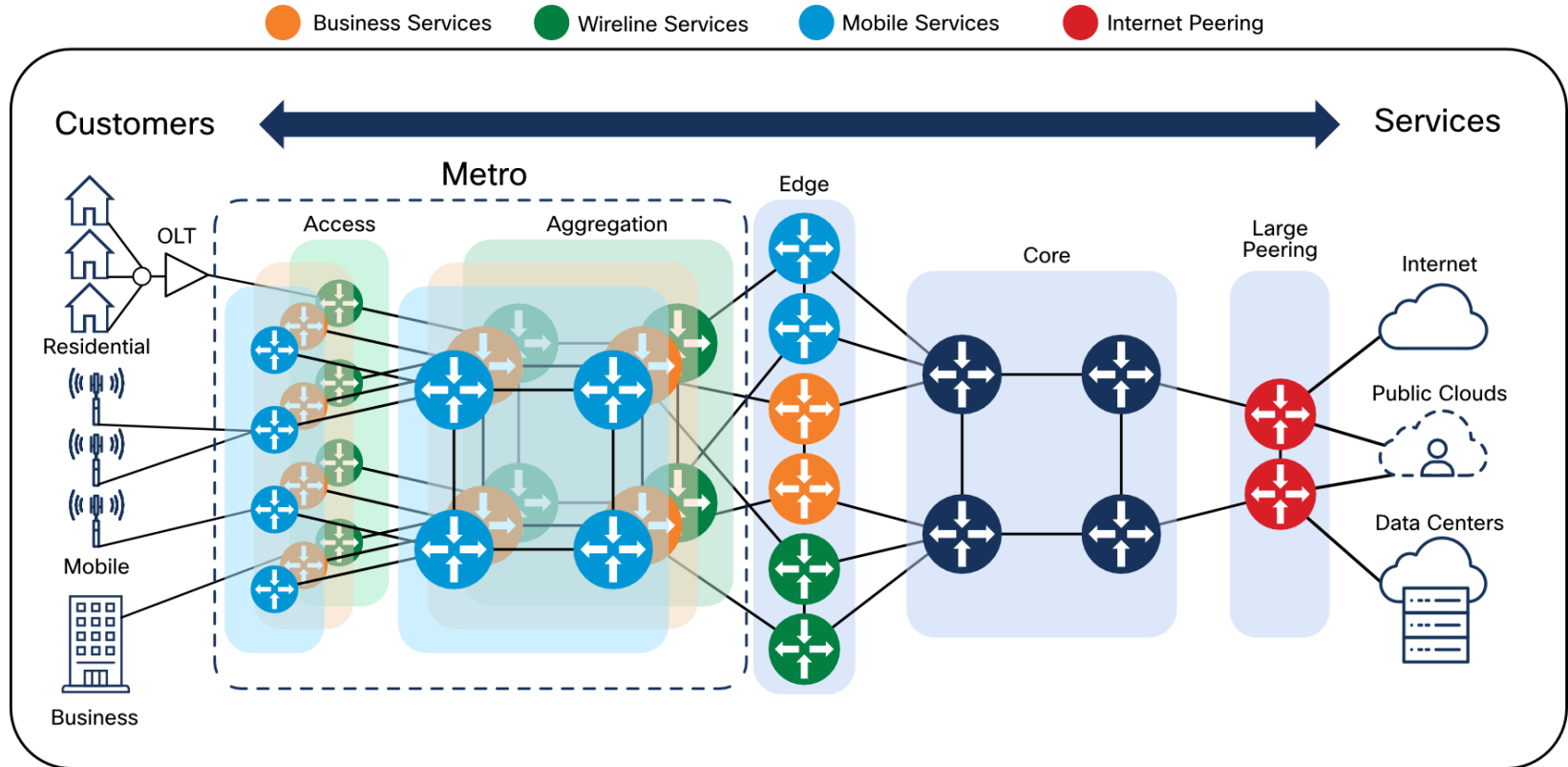
# Challenge #3 – Optimization problem

- Cost of building a solution that meets all the service requirements at scale
- Highest common denominator problem – most demanding services sets the price for all others
- Common solution: multiple routers

# Common solution to deploying multiple services at scale: Multiple parallel networks



# Common solution to deploying multiple services at scale: Multiple parallel networks



# Challenge #4 – Gigantic blast radius for failures



Solution



Geo redundant  
Dual-homing



Cost + Complexity

Failure impact:

- Redundant RPs, Fabric, LC doesn't cover for site failures
- 10's of thousands of customers
- Critical business services

Do you still need all of these?

- ISSU, NSF, NSR
- Redundant processors, fabrics, line cards, power supplies, ports, plus chassis
- **Overkill?**

# Summary

## Traditional edge routing challenges



Platform lock-in - large systems remain in place for many years, slowing down innovation



Costly and complex to optimize for different services



Limits operators who want simplicity and efficiency



Forklift upgrades - delayed as much as possible

Moving forward, to continue meeting business requirements we'll need to **consider other deployment models** for edge routing beyond the traditional centralized model.

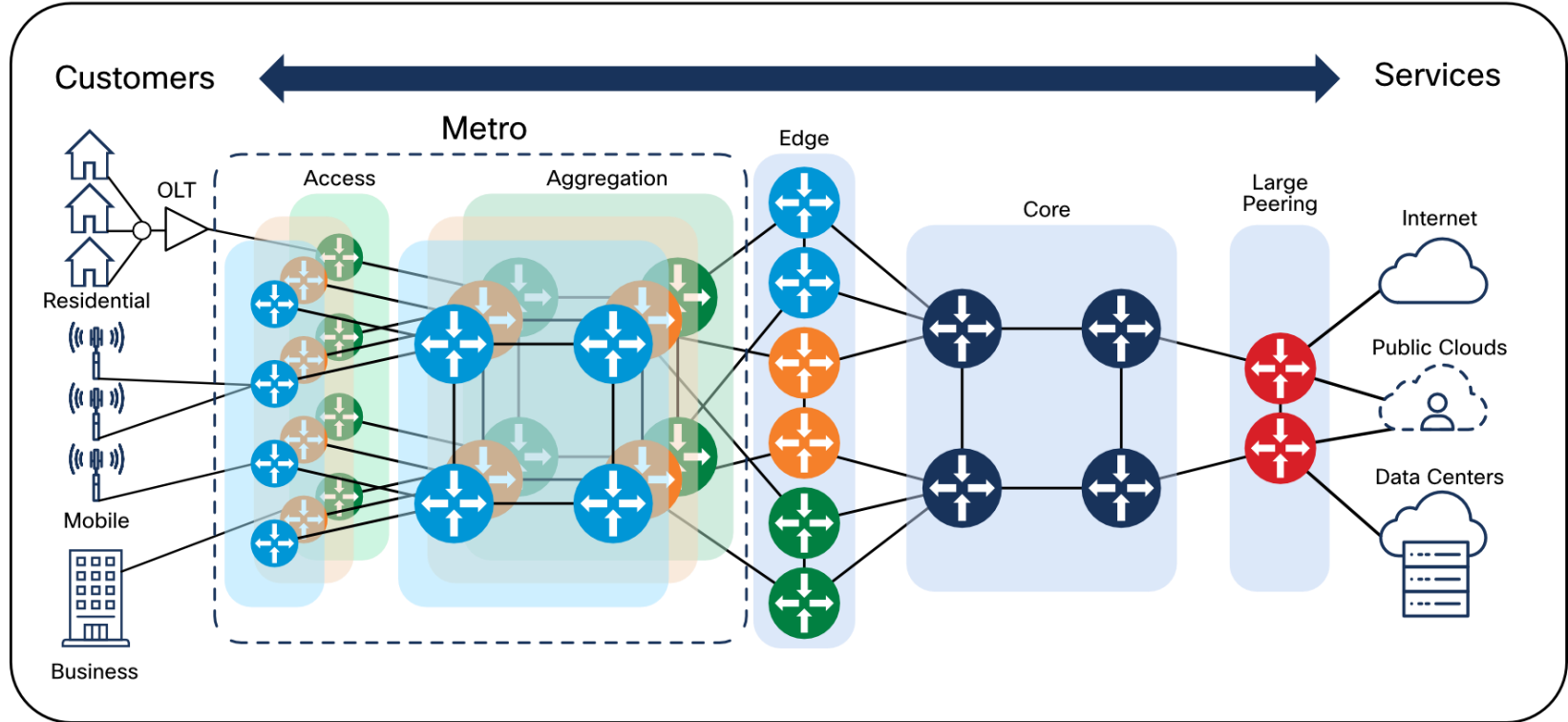
How can we build Metro networks such that the edge routing architectures align with new bandwidth demands services delivery requirements?

# Evolving Metro Network Architectures



# Today's Metro networks

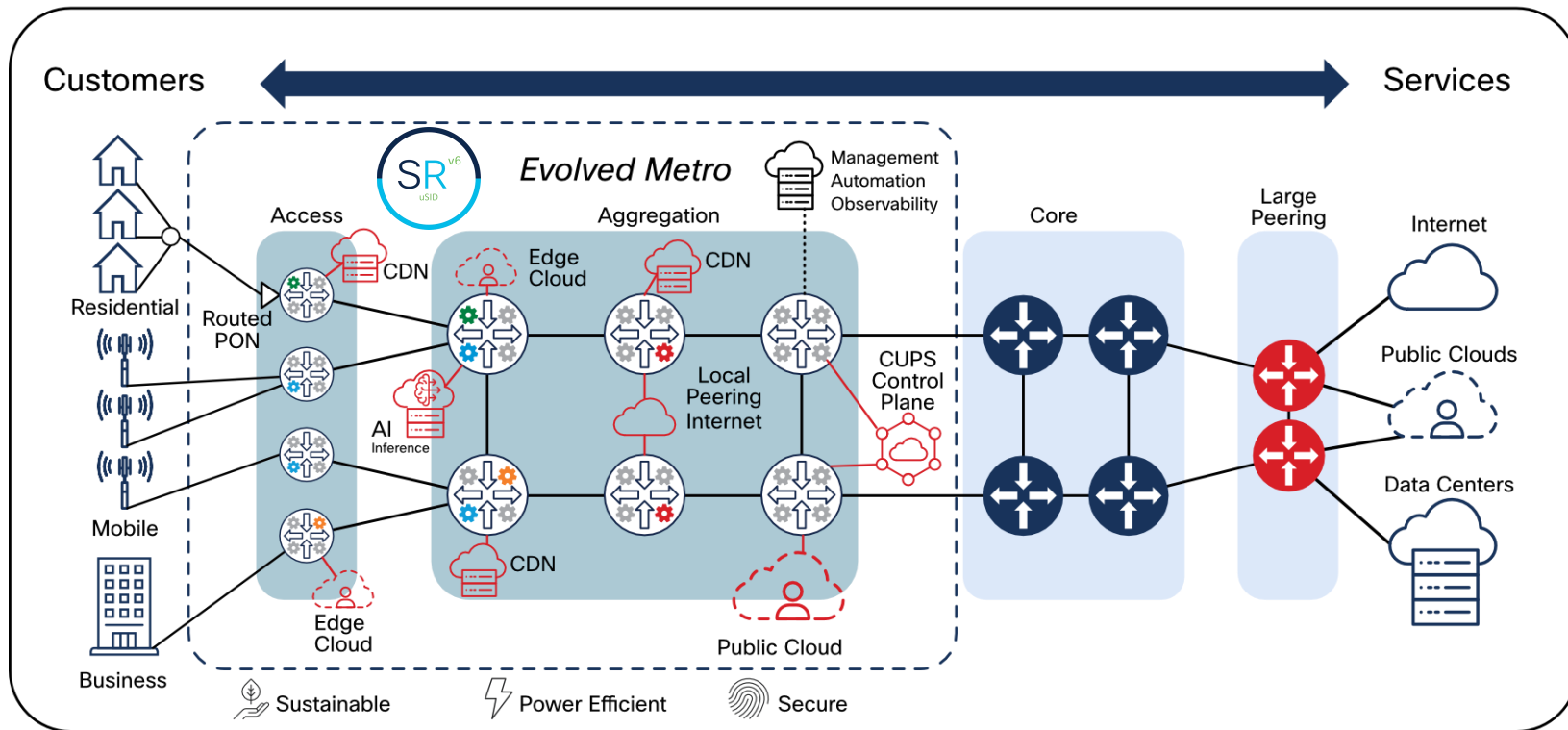
Business Services Wireline Services Mobile Services Internet Peering



# Cisco's Vision – Evolving Metro networks

## Fully converged, distributed network architecture

Floating edge functions:  Business Edge  Subscriber Edge - BNG CUPS UPF  Mobile Edge - 5G CUPS UPF  Peering



# Metro network evolution drivers

Triggered by new points of service delivery closer to consumers

## Technology drivers

- High-capacity edge silicon
- Convergence of functions
- Smaller systems and form factors

## Business drivers

- Deliver services closer to user/apps
- Cost savings
- Sustainability

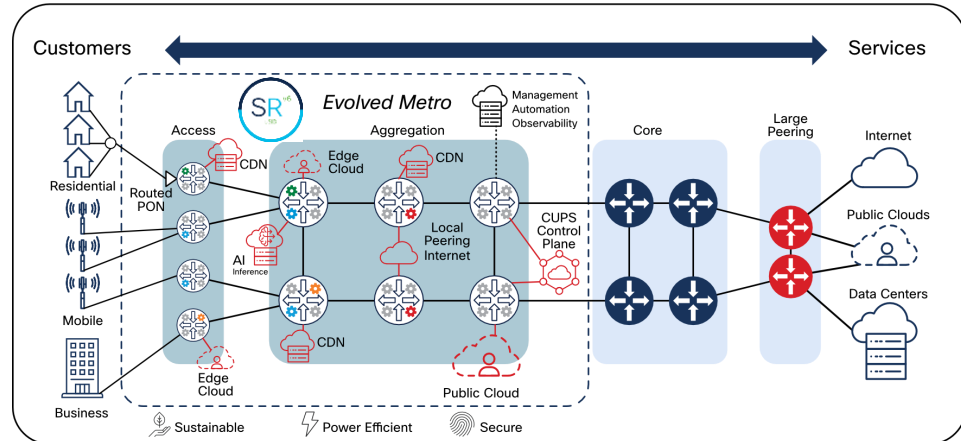
## Operational drivers

- Improved services resiliency
- Network efficiency (ex. avoid hair-pinning)
- Automation and orchestration to avoid operational burden

### Technology Building blocks



Floating edge functions: Business Edge Subscriber Edge - BNG CUPS UPF Mobile Edge - 5G CUPS UPF Peering





# Router platform evolution

New and emerging technologies can be applied to new architectures



## Networking processor silicon:

- Feature richness across the network with IOS-XR
- Tbps of capacity, high-speed at very low power consumption.  
All you need to build a spine-leaf architecture.
- Feature rich and “on-demand”



## Simpler fixed form factor platforms:

- Deliver services at right scale and cost points required for distributed architectures
- Can be used in-line for full featured router with edge capabilities on-a-stick or collapsed, or as edge fabric leaves



## Disaggregated models (Control/User Plane Separation – CUPS)

- Further simplifies edge routers, moving control plane load to compute nodes



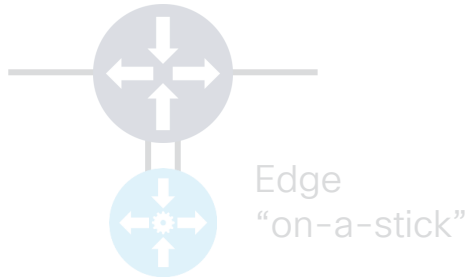
# Edge as a Function

New technology enables new consumption models

- Granular feature enablement, potentially to the port level
- Feature set aligned with use case (ex. subscriber or business edge)
- Paired with other technology advances, e.g. CUPS model
- Better cost/performance

# Edge routing architecture evolution options

Aggregation Router



Edge *on-a-stick*

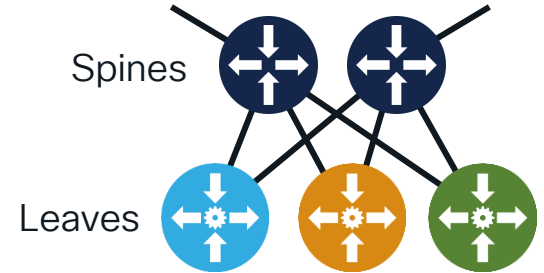
- Edge router connected to an aggregation router
- Distributed model
- Edge routing just “pushed” closer to the user

Fully distributed Edge



Collapsed Edge function

- Aggregation router (eventually access) “absorbs” Edge routing functions
  - Services at the appropriate scale for a distributed model
- No additional devices
- Fully distributed model



Edge *Fabric*

- Spine-leaf architecture
- Multi-service or role specific leaves
- Operated as a fabric vs individual routers
- Local (centralized) or remote (distributed) leaves

# Qualitative analysis

- Impact on router scale and redundancy requirements
- Network scale implications and how to address them
- Qualitative impact on services architecture
- Qualitative changes in cost structure
- Operational overhead (and how they can be overcome)

# Edge “on-a-stick”

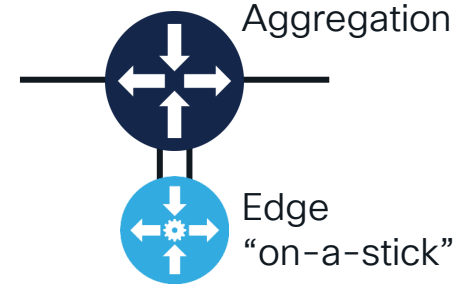
Based on traditional routers

- Pros:

- Smaller blast radius and smaller scale requirements means use of simpler edge routers
- Reduced transport costs
- Latency can be reduced with shorter fiber distances to reach the edge

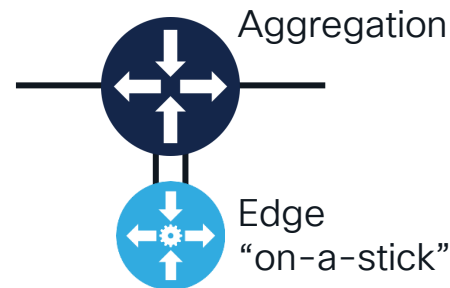
- Cons:

- Increased number of network devices
- Impact on routing protocol scale
- Higher cost if all edge requirements are kept intact
- Higher number of devices to manage



# Edge “on-a-stick”

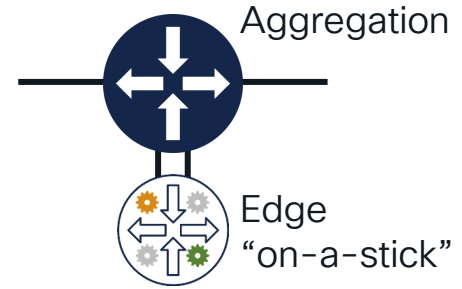
Based on traditional routers



| Router Attribute | Model Impact     | Notes                                                                                                                                                                                                                                                       |
|------------------|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive         | Scale requirements are reduced significantly compared to the centralized model                                                                                                                                                                              |
| Redundancy       | Positive         | Smaller blast radius can relax platform redundancy requirements                                                                                                                                                                                             |
| Network scale    | Negative         | Edge routers added across the network. Potentially > X increase in number of edge nodes (value of X is debatable).                                                                                                                                          |
| Services impact  | Positive         | Traffic reaches end-user and applications faster, crossing fewer network devices and using less transport capacity, reducing competition for network resources                                                                                              |
| Cost impact      | Neutral/Negative | If router requirements are not relaxed to couple with new architecture, overall cost can be higher. For the impact to be neutral, the cost of the new distributed edge routers has to be the same as the centralized option minus the savings on transport. |

# Edge “on-a-stick”

Based on state-of-art silicon and router platforms



- Pros:

- All the previous plus...
- Optimized hardware, all the way to silicon... Lower cost
- Feature consistency across the network

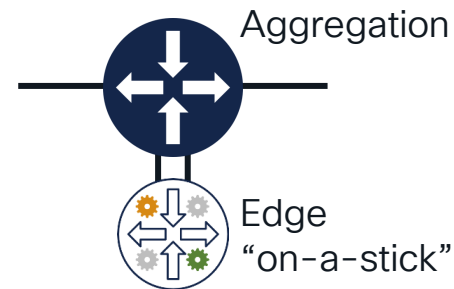
- Cons:

- No additional cons...\*

\*Just think about this as new routers optimized for distributed architectures instead of moving an existing one far into the network

# Edge “on-a-stick”

Based on state-of-art silicon and router platforms

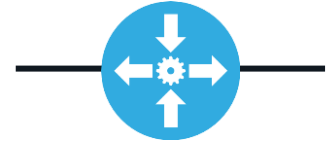


| Router Attribute | Impact   | Note                                                                                                                                                                                               |
|------------------|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive | Scale requirements are reduced compared to the centralized model                                                                                                                                   |
| Redundancy       | Positive | Smaller blast radius can relax platform redundancy requirements                                                                                                                                    |
| Network scale    | Negative | Edge routers added across the network. Potentially X increase in number of edge nodes (value of X is debatable)                                                                                    |
| Services impact  | Positive | Traffic reaches end-user and applications faster, less transport requirements and competition for network resources                                                                                |
| Cost impact      | Neutral  | Potential cost of adding more devices in the network is balanced by reduced transport requirements and use of simpler edge routing platforms with lower cost/bit compared to traditional platforms |

# Collapsed Edge function

Based on traditional routers

Collapsed Router  
Aggregation/Edge



## Pros:

- Lower cost – one device instead of two
- Smaller blast radius and smaller scale requirements means use of simpler edge routers
- Reduced transport costs
- Reduced latency

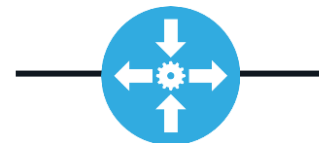
## Cons:

- Potentially higher cost for transit traffic
- Impact on routing protocol scale, with an increase in number of edge devices that BGP control planes for services

# Collapsed Edge function

Based on traditional routers

Collapsed Router  
Aggregation/Edge



| Router Attribute | Impact   | Note                                                                                                                                                                                                                                         |
|------------------|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive | Scale requirements are reduced compared to the centralized model                                                                                                                                                                             |
| Redundancy       | Positive | Smaller blast radius can relax platform redundancy requirements                                                                                                                                                                              |
| Network scale    | Negative | Even though the number of routers is nearly the same, the scale of the services control plane, ie. BGP will be higher.                                                                                                                       |
| Services impact  | Positive | Traffic reaches end-user and applications faster, less transport requirements and competition for network resources                                                                                                                          |
| Cost impact      | Neutral  | Potential cost of adding a more sophisticated router in many locations in the network to provide services edge capabilities is balanced by collapsing the transport or aggregation function as well as removing the centralized edge router. |

# Collapsed Edge function

Based on state-of-art silicon and router platforms

Collapsed Router  
Aggregation/Edge



## Pros:

- Same as previous, plus...
- Router is optimized to collapse transport and services functions at right feature/cost

## Cons:

- No additional cons...

# Collapsed Edge function

Based on state-of-art silicon and router platforms

Collapsed Router  
Aggregation/Edge

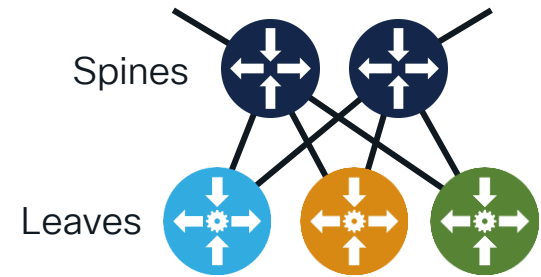


| Router Attribute | Impact   | Note                                                                                                                                                                                                                                  |
|------------------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive | Scale requirements are reduced compared to the centralized model                                                                                                                                                                      |
| Redundancy       | Positive | Smaller blast radius can relax platform redundancy requirements                                                                                                                                                                       |
| Network scale    | Negative | Even though the number of routers is nearly the same, the scale of the services control plane, ie. BGP will be higher.                                                                                                                |
| Services impact  | Positive | Traffic reaches end-user and applications faster, less transport requirements and competition for network resources                                                                                                                   |
| Cost impact      | Positive | Simpler routers are used and Edge functions can be enabled as required, without necessarily penalizing the cost model of the routers by enabling service capabilities based on the most demanding services everywhere in the network. |

# Edge Fabric

- What makes a Fabric different:
  - Set of routers are managed as a group
  - Routers have specific roles – spines, leaves
  - Scale-out model supporting multiple layers
  - Oversubscription as part of the design
    - You can match bandwidth to requirements, unlike in a chassis where you have to take whatever oversubscription the system gives you
  - Focus on device simplicity. Improved availability/reliability achieved at system level
    - “Cattle” versus “Pet” design philosophy

Fabric can be seen as a “distributed chassis” without the backplane limitations. (Do NOT confuse with a “virtualized chassis”)



# Fabric model redundancy\*

## Chassis-level availability



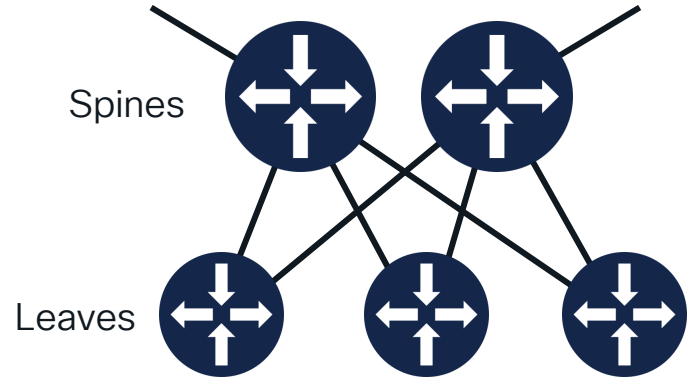
Dual Route Processors

Dual Line Cards

Downtime: < 5.3min

Availability: > 99.998%

## System-level availability



Customers dual homed to leaves

Leaves dual homed to spines

Downtime: < 6s

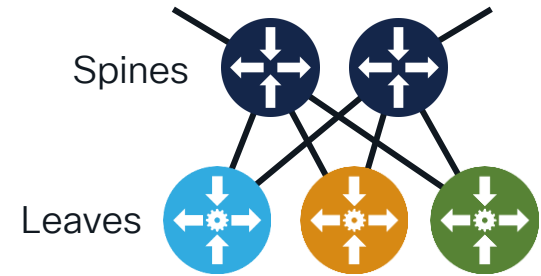
Availability: > 99.9999%

\*Notes:

- Based on actual MTBF numbers
- Doesn't consider planned operations, ex. software upgrades

# Edge Fabric

Based on traditional routers



## Pros:

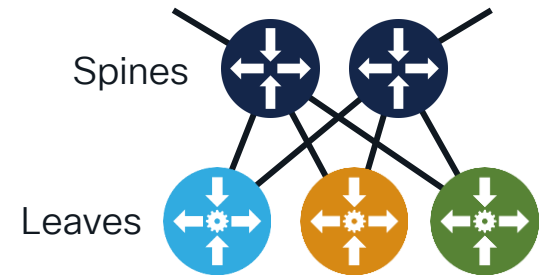
- Unmatched scale
- Easy software upgrade without “ISSU” with complex rules
- Improved solution level availability
- Add capacity or scale by adding new leaves (when spines are sized accordingly)
- Less-disruptive upgrades/refreshes
- Choice of platforms for specific roles
- Flexible oversubscription

## Cons:

- Impact on network scale with increased number of edge routers
- More devices to manage, more interconnections between devices (but simpler devices though)
- Engineering to right-size the fabric may be new to some people
- Changes in your protocol stack may be required (for the better though)

# Edge Fabric

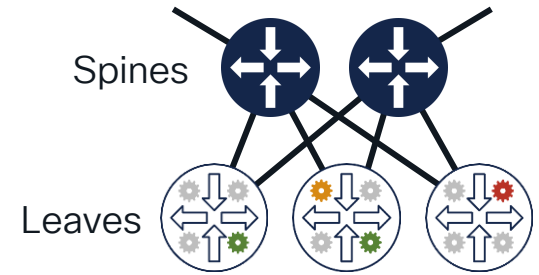
Based on traditional routers



| Router Attribute | Model Impact     | Notes                                                                                                                                                                                                                                                                                                                     |
|------------------|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive         | Horizontal scaling is unmatched by modular chassis. Each leaf scales independently.                                                                                                                                                                                                                                       |
| Redundancy       | Positive         | Smaller blast radius for each leaf, with redundant spines and load-balancing.                                                                                                                                                                                                                                             |
| Network scale    | Negative         | With more routers performing edge roles, the scale of the services control plane can increase significantly, ie. BGP will be higher.                                                                                                                                                                                      |
| Services impact  | Neutral/Positive | If the spine-leaf architecture is deployed in a centralized model, there should be no changes in the services architecture. If leaves are deployed remotely, for instance as a collapsed aggregation and edge node, for locally switched traffic the impact on the services is positive.                                  |
| Cost impact      | Neutral/Negative | When simpler, more cost effective routers are deployed and sizing considers oversubscription/multiplexing, the cost of having additional devices could be balanced out with the more efficient design. However, if the edge requirements remain exactly same, and a non-blocking model is chosen, the cost can be higher. |

# Edge Fabric

Based on state-of-art silicon and router platforms



## Pros:

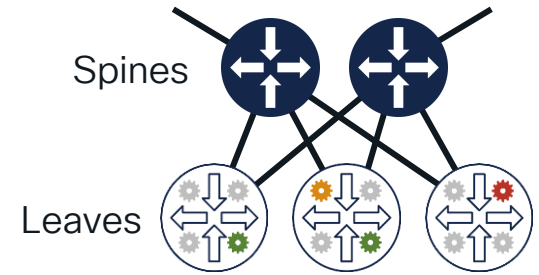
- Same as previous, plus...
- Additional cost benefits resulting from the use of optimized silicon and simpler, more optimized routers for spine and leaf roles

## Cons:

- No additional cons...

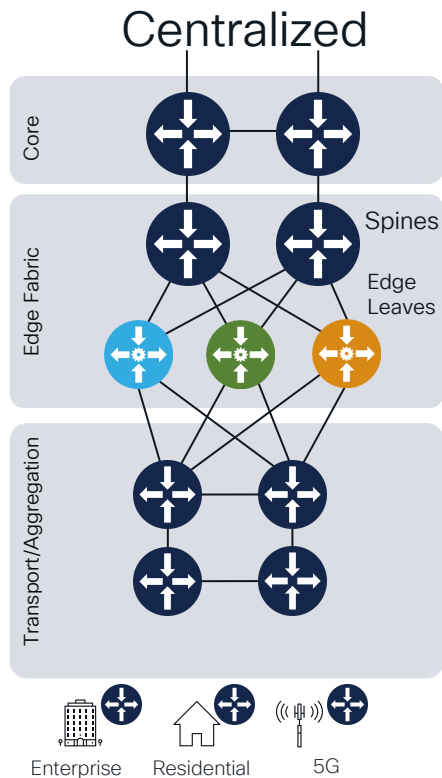
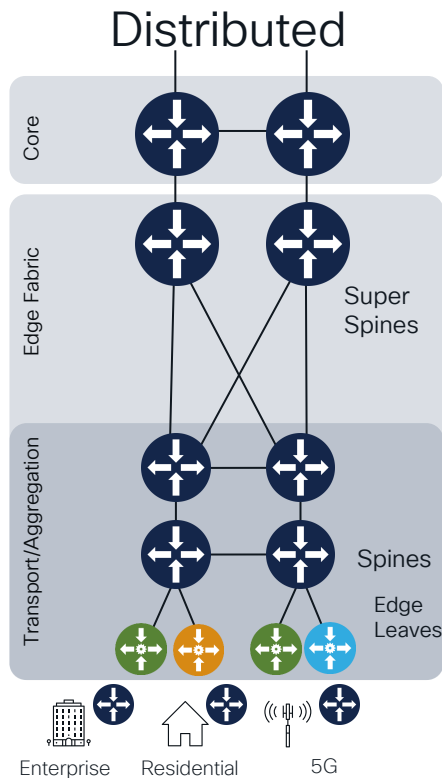
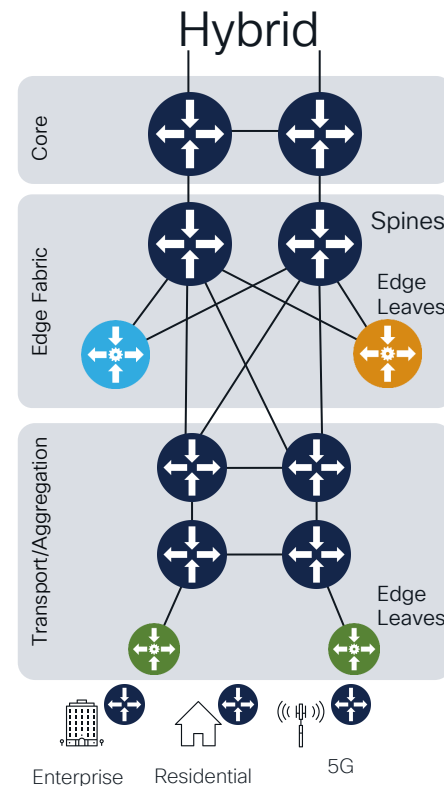
# Edge Fabric

Based on state-of-art silicon and router platforms



| Router Attribute | Model Impact     | Notes                                                                                                                                                                                                                                                                                    |
|------------------|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Router scale     | Positive         | Horizontal scaling is unmatched by modular chassis. Each leaf scales independently.                                                                                                                                                                                                      |
| Redundancy       | Positive         | Smaller blast radius for each leaf, with redundant spines and load-balancing.                                                                                                                                                                                                            |
| Network scale    | Negative         | With more routers performing edge roles, the scale of the services control plane can increase significantly, ie. BGP will be higher.                                                                                                                                                     |
| Services impact  | Neutral/Positive | If the spine-leaf architecture is deployed in a centralized model, there should be no changes in the services architecture. If leaves are deployed remotely, for instance as a collapsed aggregation and edge node, for locally switched traffic the impact on the services is positive. |
| Cost impact      | Neutral/Positive | The use of highly optimized network processors like Silicon One bring additional cost benefits that can overcome the extra cost of deploying more devices in the network. Likewise, the use of simpler, role optimized routing platforms also contribute to additional cost savings.     |

# Edge Fabric - deployment options



● L2VPN Business Edge Leaf
 ● L3VPN Business Edge Leaf
 ● Subscriber Edge (BNG) Leaf
 ● Peering Edge

# Edge Fabric Management and Automation

Unique requirements to avoid operational overhead

## Management\*

- Device grouping
- Resource monitoring
- Inventory
- Fault/Performance management
- Alarm filtering

## Lifecycle automation\*

- Device on-boarding
  - Automatic safety checks
  - Zero-touch provisioning
  - Auto-discovery of interconnections
- Fleet upgrades

Many of these capabilities are supported by  
Cisco Crosswork software suite

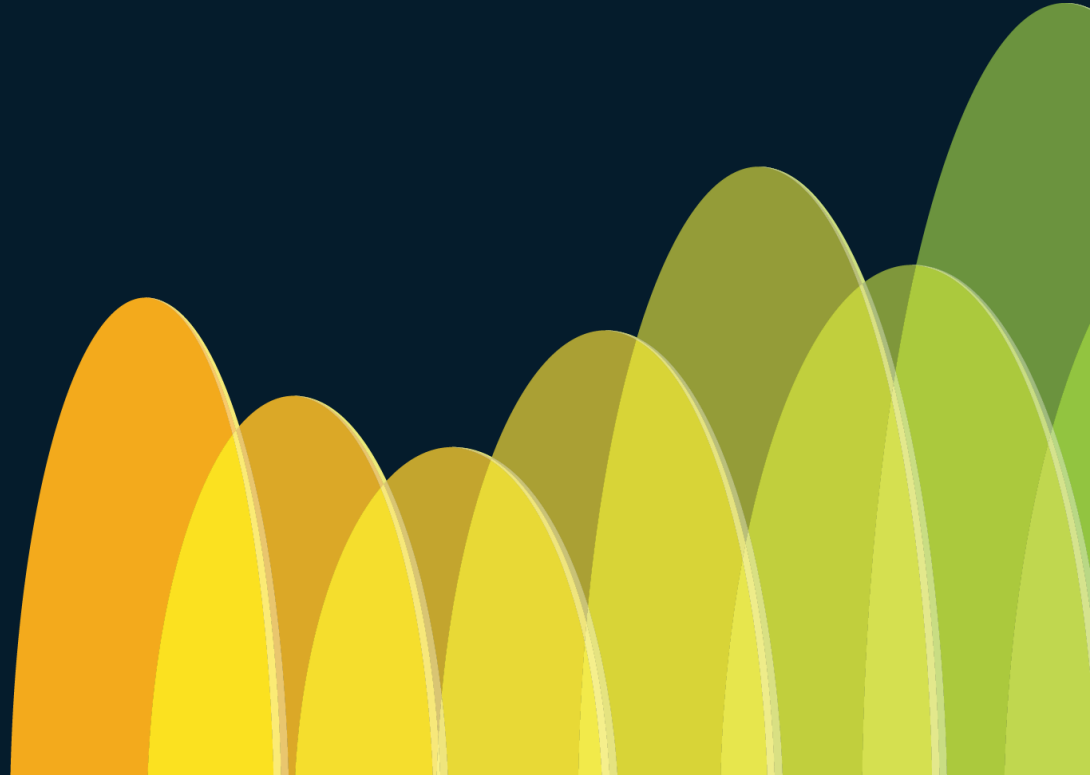
\*Note: Examples only, not an all-inclusive list

# Edge Fabric

## Additional considerations

- Fabric is one more option, not the ultimate one
- Clear benefits when:
  - Very large scale is required
  - Freedom to select different leaves is required (ex. role optimized leaves)
  - Oversubscription is part of the design
  - New, more cost-effective technologies can be used (ex. Silicon One spines)

# Design considerations



# Impact on routing protocol scale

- Distributed design and Fabric model can lead to major increase on number of routers in the IGP control plane
- This problem has been addressed before:
  - IP/MPLS – RFC 3107:
    - widely adopted for mobile x-haul and converged networks, some with over 100k routers
    - Segmentation via IGP areas and BGP AS
    - In-line route reflectors
  - Segment Routing provides a simpler solution
    - Controller based (PCE) model
    - Automated Steering
    - On-Demand Next-hop

# Segment Routing IPv6 Benefits

Segment Routing IPv6 makes networking simpler. **It's just IP routing!**



Maximum scale

Simple and efficient

Services rich

Operational simplicity

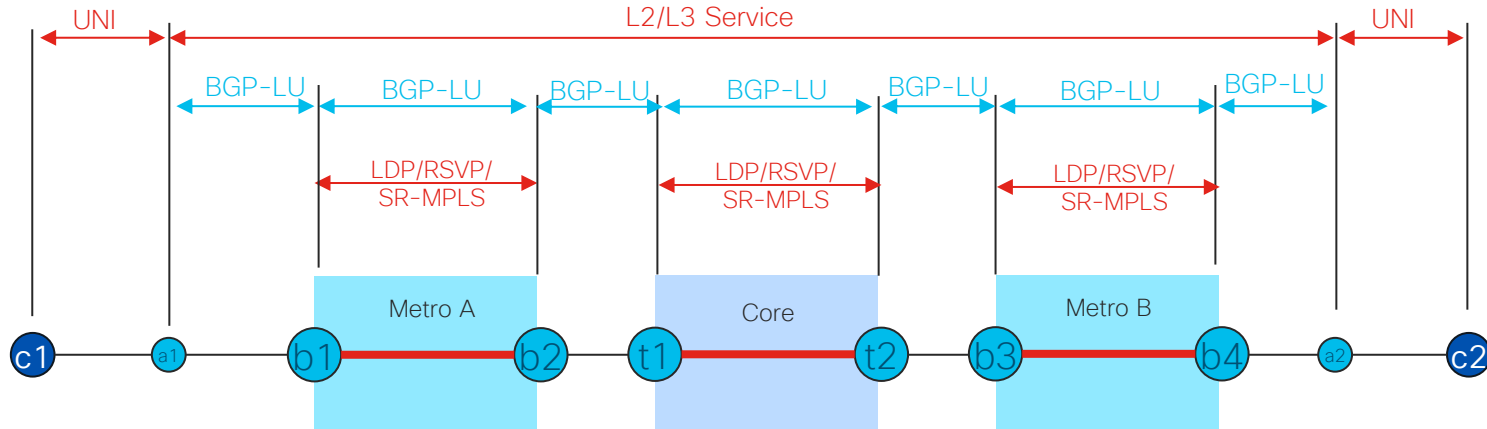
Seamless deployment



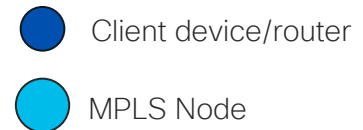
# SRv6 uSID Unique Benefits

- **Eliminates** numerous **legacy protocols**, including LDP, RSVP-TE, BGP-LU, and MPLS OAM
- **Maximizes** network **scale** – enables IP route summarization between areas/domains, unlike MPLS
- Optimal **ECMP** load balancing (no sub-optimal MPLS label hashing)
- **Better FIB scaling** – no ip2mpls, mpls2mpls, and mpls2ip entries
- Optimal **MTU efficiency** with micro-segment (uSID)
- Solution **universality**: WAN, xHaul, DC, Metro, Far/Near Edge, IoT, Host, etc.
  - Elimination of complexity and cost at silo boundaries (i.e., DPI at VxLAN/MPLS boundaries)
  - Native Host and Cloud support: SRv6 is IP (MPLS is neither in the IP socket nor in the cloud)
- Enables tight **application interaction** with the network, i.e., application-driven network programmability as well as service function chaining
- **Seamless brownfield deployment** with classic IPv6 nodes
  - For example, if no TE, Flex Algo or FRR, then core routers only need to support IPv6, and SRv6 is not required
- Provides more **QoS** markings for packet classification and drop profiles (6-bits versus 3-bits)

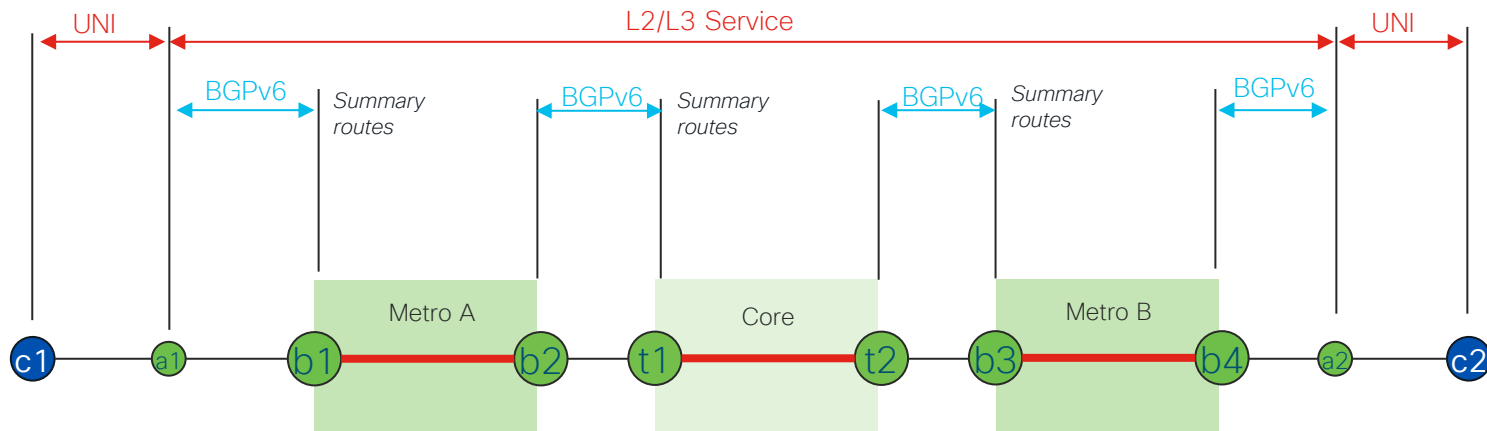
# Inter-Domain using BGP-LU



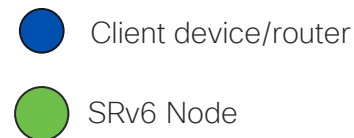
- This common design requires BGP-LU to stitch between domains within the SP
  - Also commonly used to stitch between sub-domains within a domain
- BGP-LU adds a substantial tax of complexity and limits scalability



# Inter-Domain simplification with SRv6

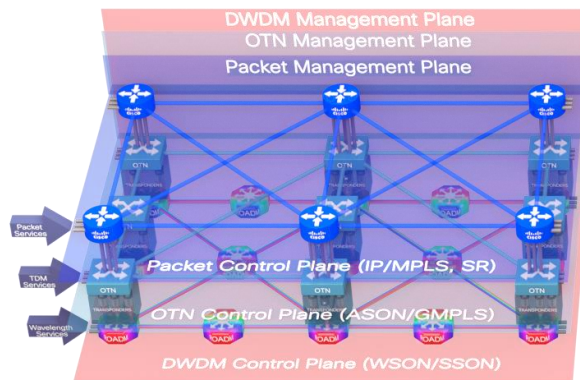


- SRv6 eliminates the BGP-LU shim layer and significantly improves scalability through summarization



# Synergies with other solutions

## Routed Optical Networking architecture



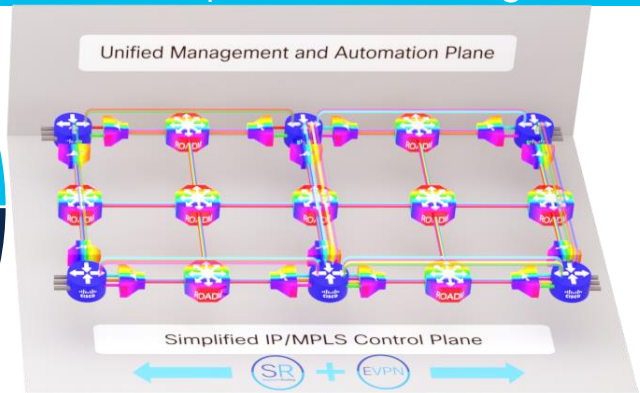
Today's Networks  
3-4 Layered Architecture

**Collapses 3 layers into 1**

- 1 Switching Element
- 1 Control plane
- 1 Services layer

**Single pane of glass**

## Internet for the Future Routed Optical Networking



Route when you can and  
ROADM when you need to

*RON is to Networking what EVs are to the Automotive industry as it redefines the way we build IP and Optical Networks by simplifying by providing a more sustainable way to build networks*

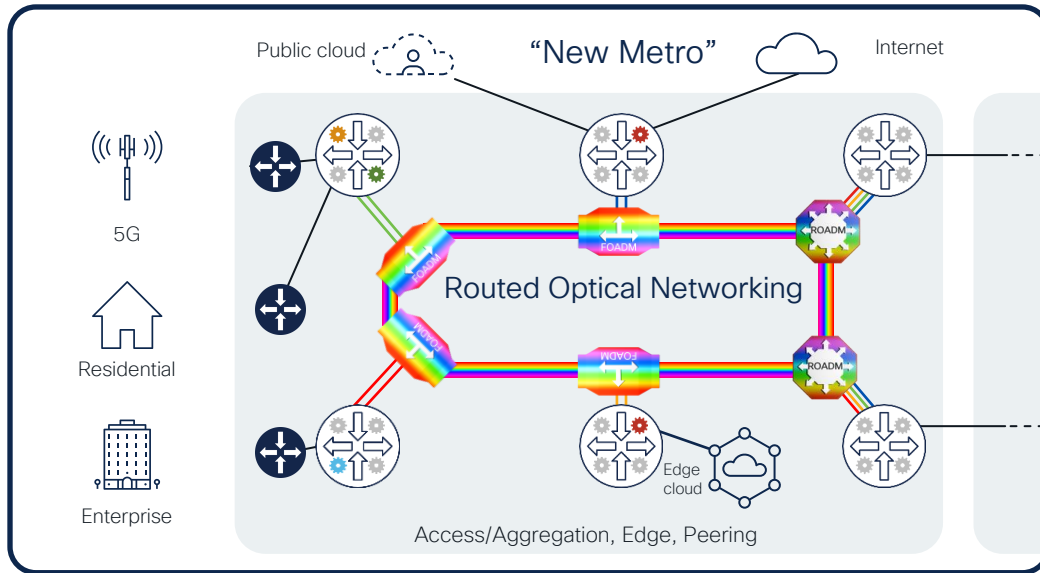
**3S' of RON :**

**S**implification, **S**avings, **S**ustainability

# Synergies with other solutions

## Routed Optical Networking example

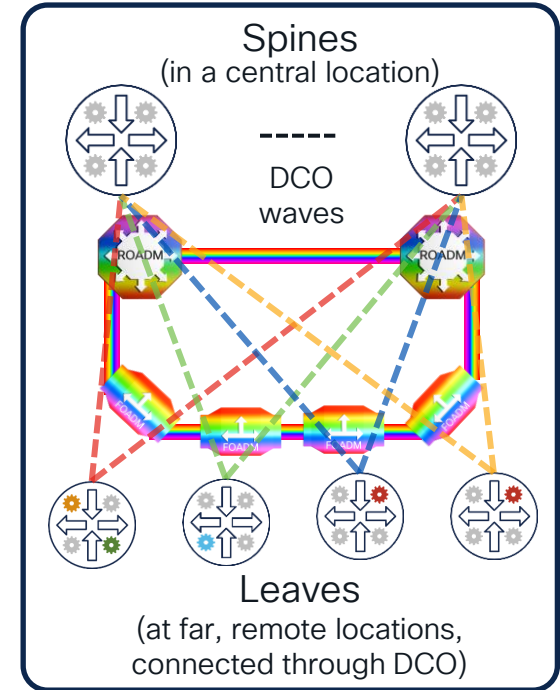
Routed Optical Networking flexible connectivity, scale and cost effective bandwidth enable the new Metro architecture



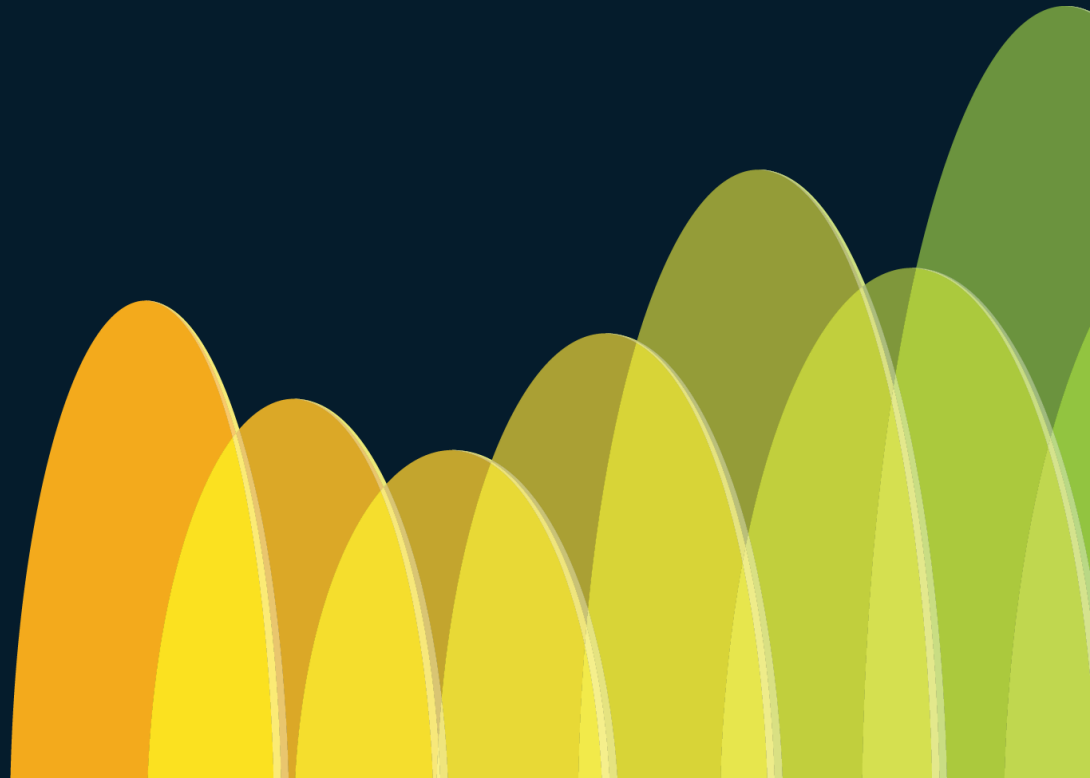
Business Services Subscriber Srvs (CUPS BNG UPF) 5GC (CUPS UP) Distributed Peering

**CISCO** Live!

## Distributed Fabric Architecture



# Summary



# Summary

- Traditional Metro network designs are challenged by new traffic and business realities
  - Ever growing bandwidth with flat ARPU
  - Increasing cost and complexity of running multiple networks and scaling centralized services architectures
  - Counter to content and application evolution towards distributed architectures
- New and evolving router and network architectures should be considered in network designs moving forward
  - Fully distributed systems using simpler routing architectures
  - Fabric models with horizontal, scale-out properties
  - Each with unique benefits and trade-offs.
  - Consider all of them and choose based on your business requirements

# Complete Your Session Evaluations



Complete a minimum of 4 session surveys and the Overall Event Survey to claim a **Cisco Live T-Shirt**.

---



Complete your surveys in the **Cisco Live mobile app**.

---



# Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at [www.CiscoLive.com/on-demand](https://www.CiscoLive.com/on-demand)



# Thank you

CISCO *Live!*

#CiscoLiveAPJC

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC