cisco Live!

Let's go

The bridge to possible

Rethink your Edge Routing Architecture

Emerson Moura, Distinguished Solutions Engineer

Acknowledgements

- Thanks to 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.

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?

Services Edge routers:

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

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

Practical implications of edge routing requirements

$$\underbrace{\text{High cost} \quad + \quad \text{Complexity}}$$

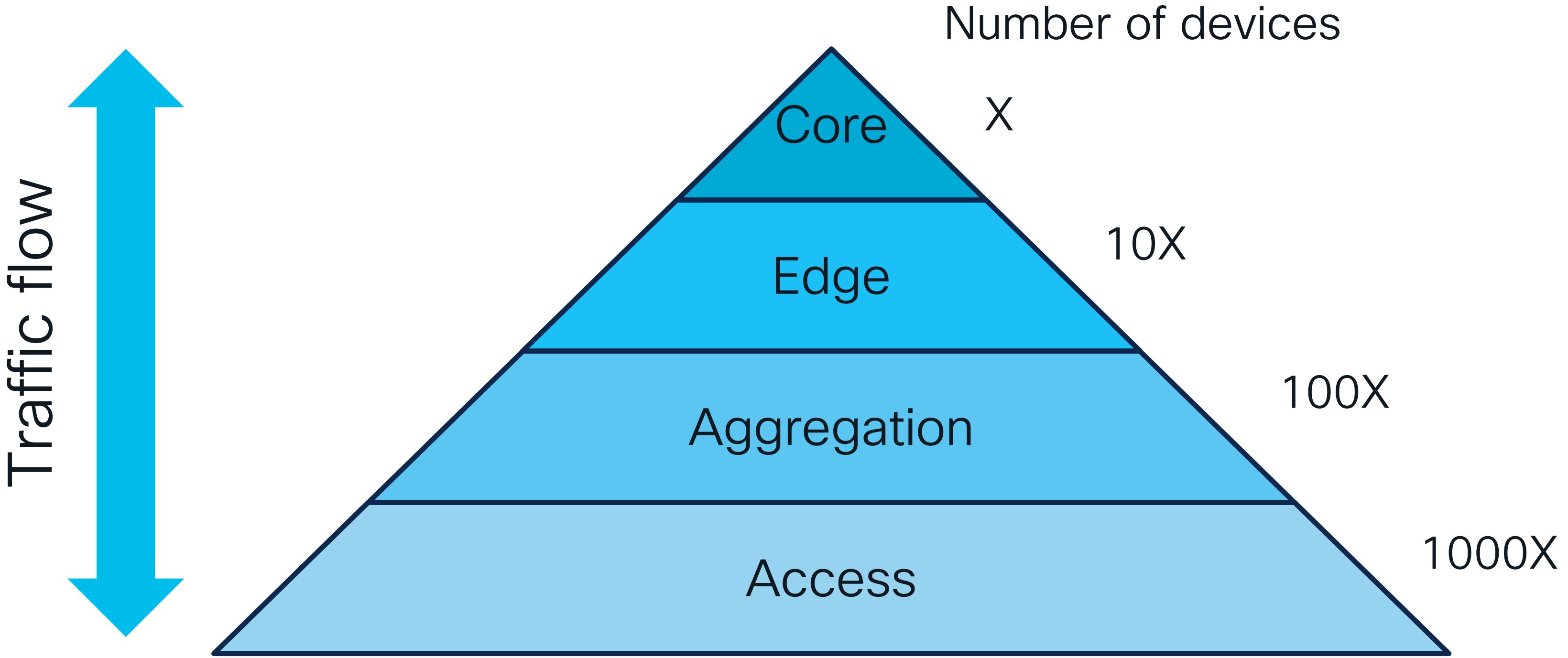
Design philosophy: Centralize it!

- 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.

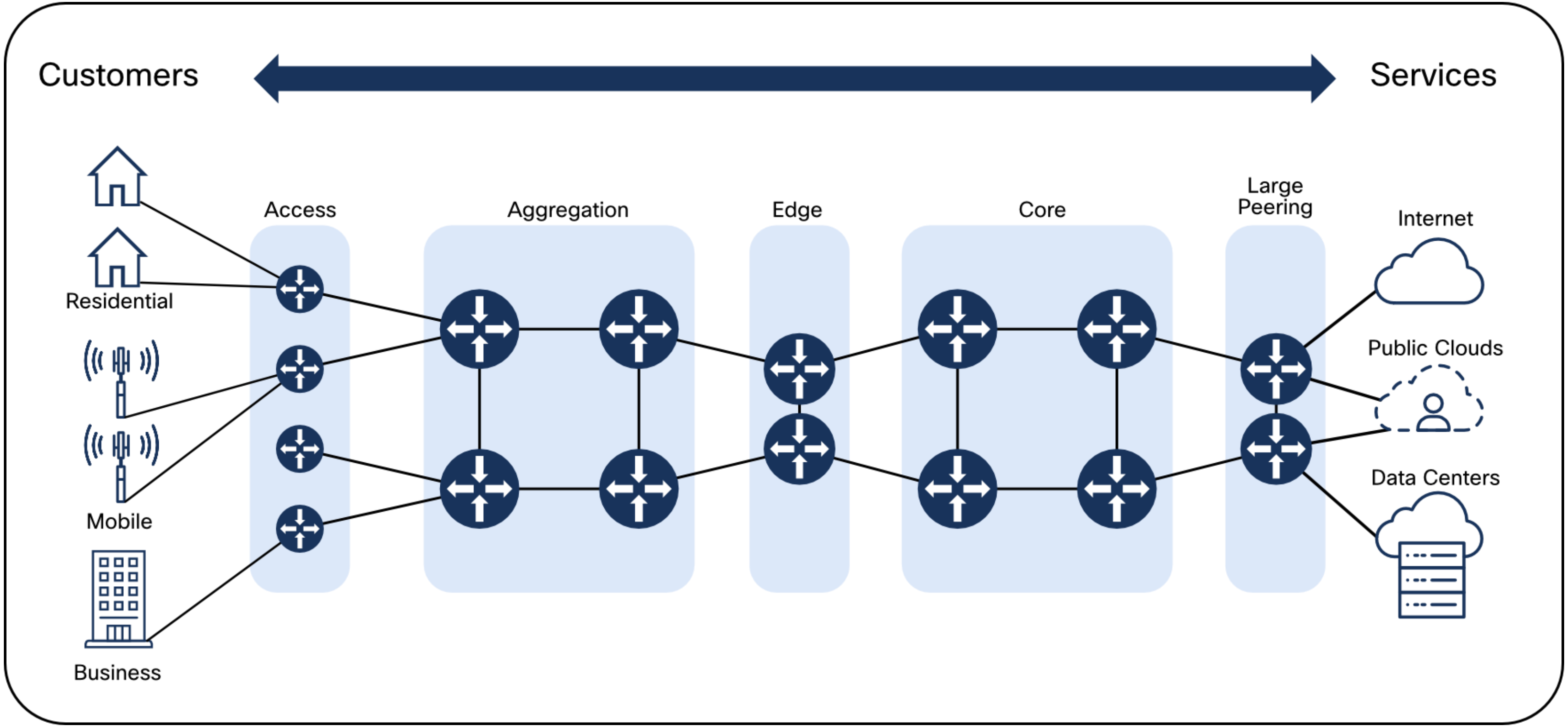
Slow innovation cycles, increased risks, late to realize technology gains

Typical network scenario

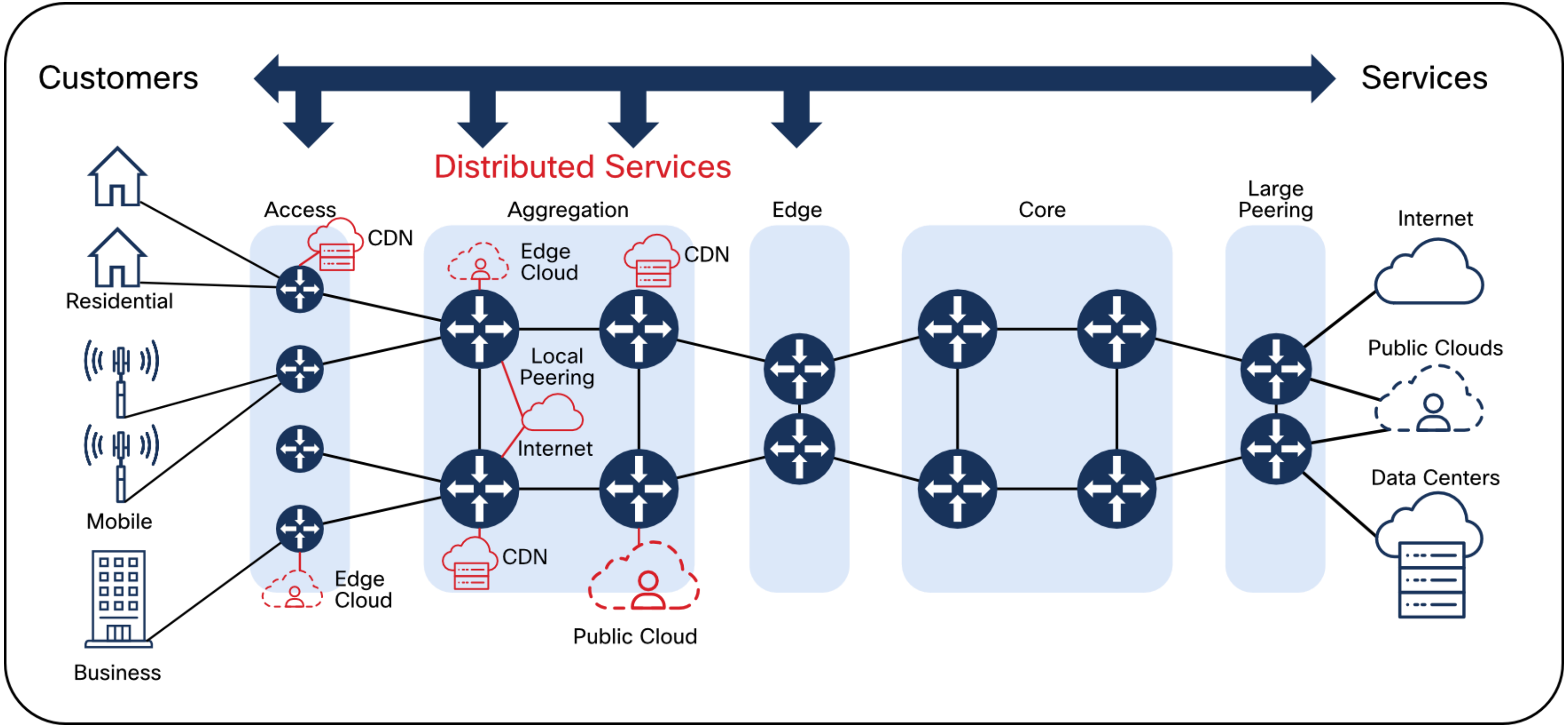
Built for North-South traffic patterns



Challenge #1 – Services are being distributed

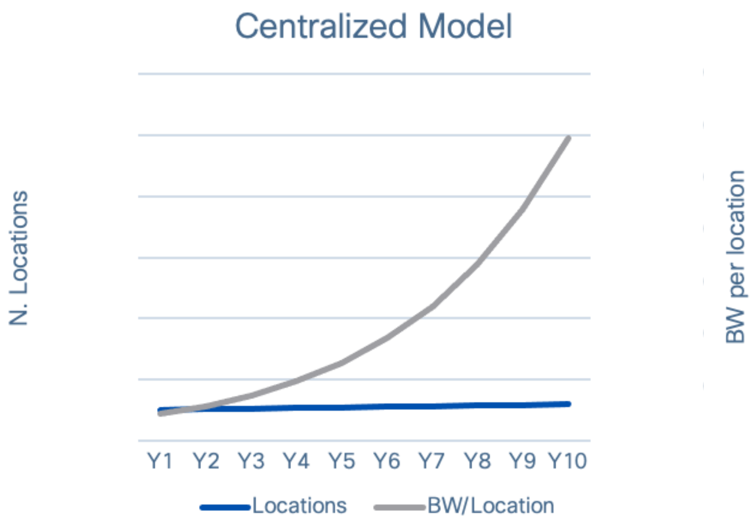


Challenge #1 – Services are being distributed



Implications of fully distributed services

Hypothetical model analysis



Note: the model above is based on an arbitrary network

Implication of fully distributed services

Hypothetical model analysis

Modeling result observations*:

- Flat to slow subscriber growth
- Continued 30% y/y BW growth
- 10x increase in service locations

BW profile and subscriber density for a distributed edge changes significantly:

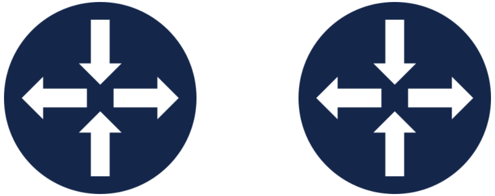
- BW at distributed service locations grows linearly (vs exponentialy)
- ~9x decrease in number of subscribers per distributed edge

* Note: looking at model sensivity, not aiming at precise figures

Challenge #2 – 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?

Challenge #3 – Optimization problem

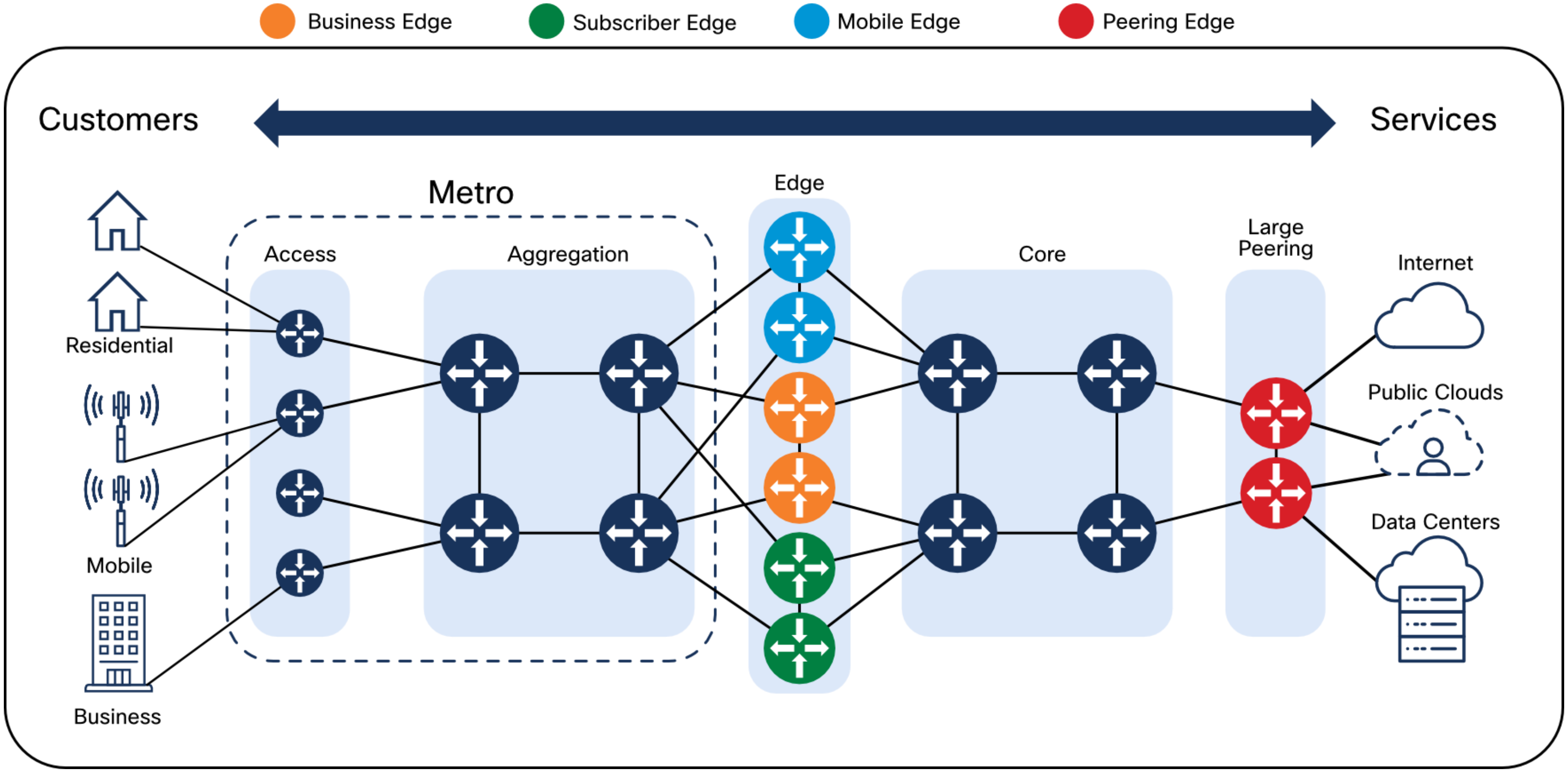
Service	Main Features*	Main scale Attributes*
Business L2 VPN	• VLAN tag processing • Traffic Policers (2R3C) • BPDU tunelling, processing • Service multiplexing • Connectivity models (E-Line, E-LAN, E-Tree)	• Bandwidth • MAC tables • N. of Bridge groups • N. PW EVPN/VPWS
Business L3 VPN	• Import/Export route filters • Traffic policers, Shapers • ACLs • Accounting	• VRFs • BGP neighbors • Routing tables
Subscriber (BNG)	• Traffic policers, shapers • Interfaces for OSS/BSS	• Subscriber sessions • Queues • Counters • ARP entries
Internet (Peering)	• …	• …

* 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

Multiple edge routers in networks



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

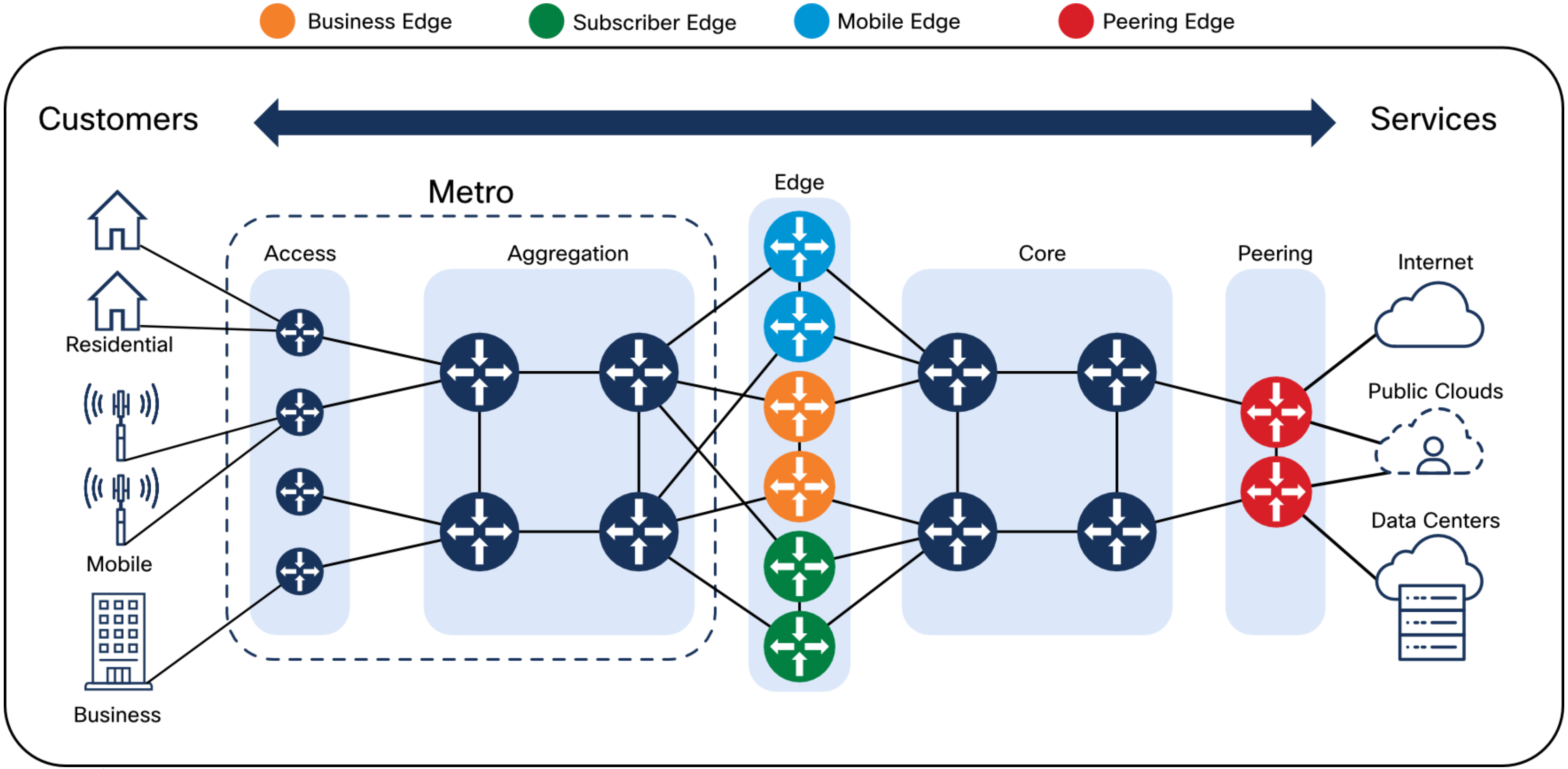
Forklift upgrades - delayed as much as possible

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 networks such that the edge routing architectures align with services architectures - living anywhere in the network?

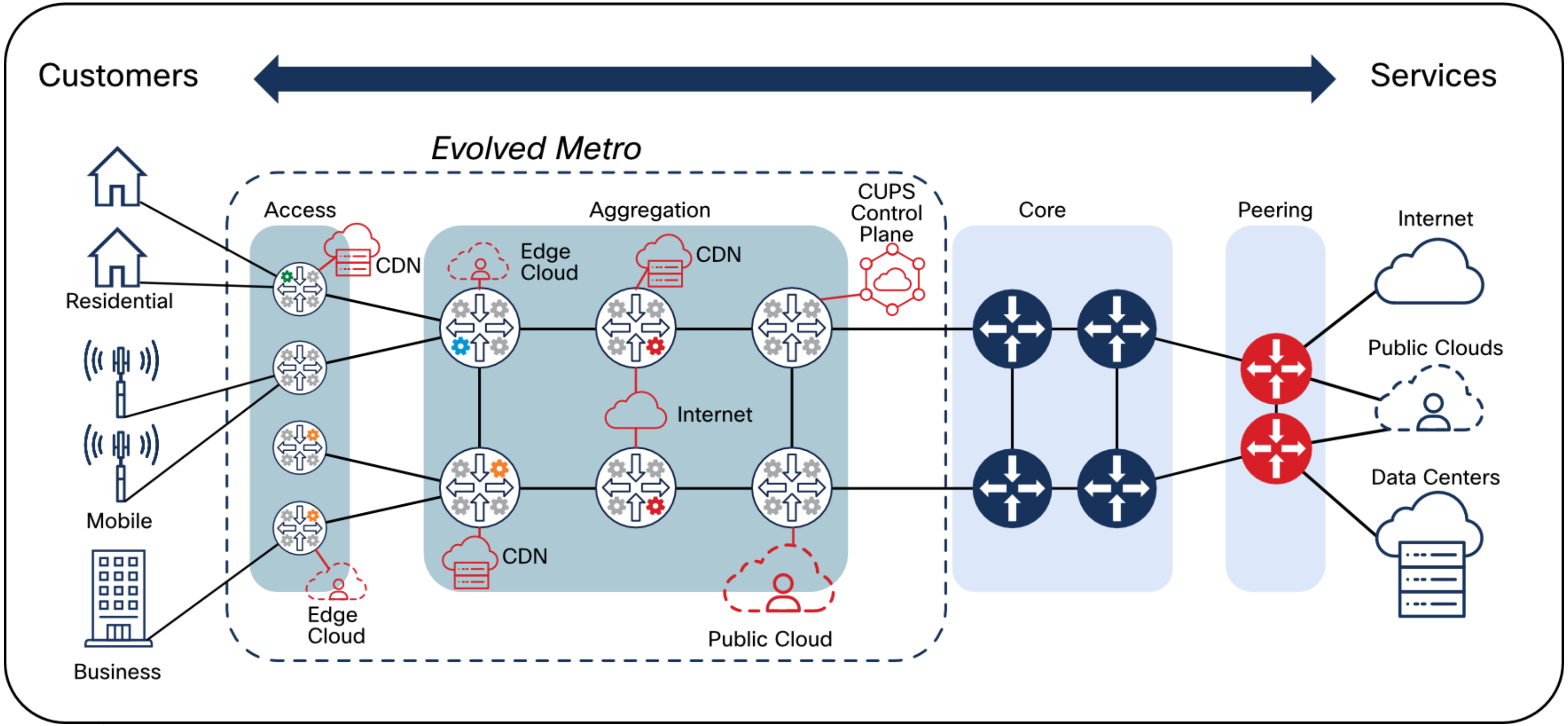
Evolving Edge Routing Architectures

Traditional network architecture



Cisco's Vision – Evolving Metro networks

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



Edge routing evolution

Triggered by new points of service delivery closer to consumers

Technology drivers

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

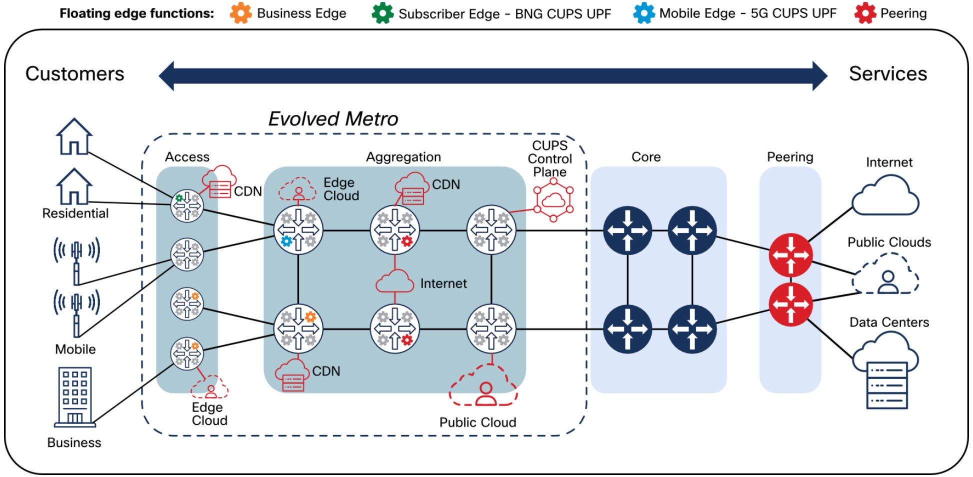
Business drivers

- 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



Big network changes take time.
How do we get there?
What are other options?

Edge routing architecture evolution options

Aggregation Router

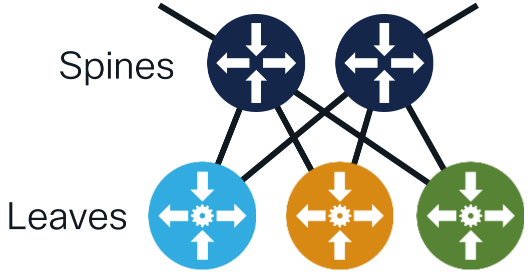
Services *on-a-stick*

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

Fully distributed Edge

Edge *as-a-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 (distribtued) leaves

Edge router platform evolution

New and emerging technologies can be applied to new architectures

Networking processor silicon:

- Consistent features across the network. Ex. Cisco Silicon One
- 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 *as-a-function, on-a-stick* 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 busines edge)
- Paired with other technology advances, e.g. CUPS model
- Better cost/performance

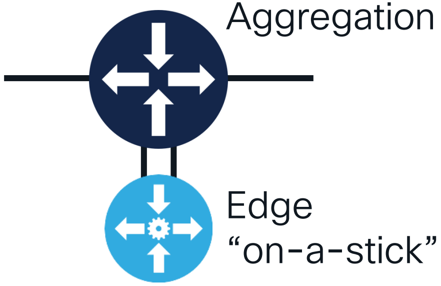
How do these options compare?

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 router “on-a-stick”

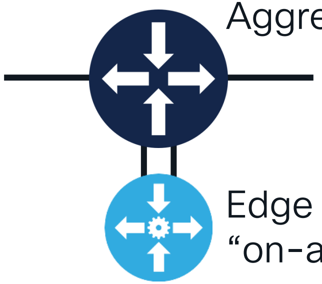
Based on traditional routers



- 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

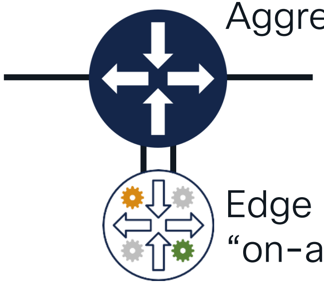
Edge router “on-a-stick”

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 router “on-a-stick”



Based on state-of-art silicon and router platforms

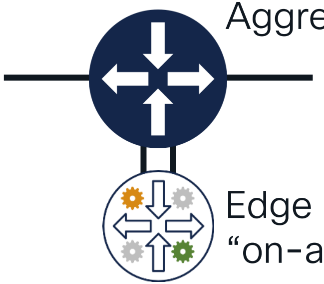
- Pros:
  - All the previous plus…
  - Optimized hardware, all the way to silicon… Lower cost
  - Feature consistency across the network

- 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 router "on-a-stick"

Based on state-of-art silicon and router platforms

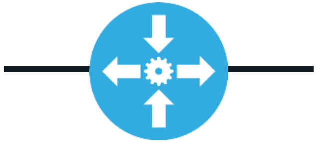


Router Attribute	Impact	Note
Router scale	Positive	Scale requirements are reduced compared to the centralized model
Redundancy	Positive	Smaller blast radius can relax platform redundancy requirements
Network scale	Negative	Edge routers added across the network. Potentially X increase in number of edge 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

Edge "as-a-function"

Based on traditional router platforms

Collapsed Router Aggregation/Edge

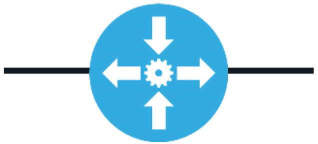


- Pros:
  - Lower cost – one device instead of two
  - Smaller blast radius and smaller scale requirements means use of simpler edge routers
  - Reduced transport costs
  - Reduced latency
- Cons:
  - Potentially higher cost for transit traffic
  - Impact on routing protocol scale, with an increase in number of edge devices that BGP control planes for services

Edge “as-a-function”

Based on traditional routers

Collapsed Router Aggregation/Edge

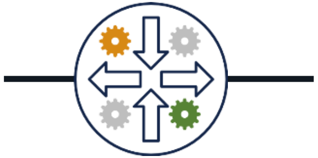


Router Attribute	Impact	Note
Router scale	Positive	Scale requirements are reduced compared to the centralized model
Redundancy	Positive	Smaller blast radius can relax platform redundancy requirements
Network scale	Negative	Even though the number of routers is nearly the same, the scale of the services 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.

Edge “as-a-function”

Based on state-of-art silicon and router platforms

Collapsed Router Aggregation/Edge

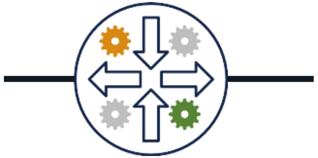


- Pros:
  - Same as previous, plus…
  - Router is optimized to collapse transport and services functions at right feature/cost
- Cons:
  - No additional cons…

Edge “as-a-function”

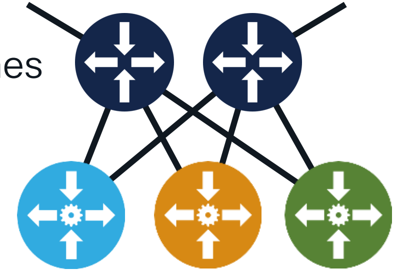
Based on state-of-art silicon and router platforms

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*

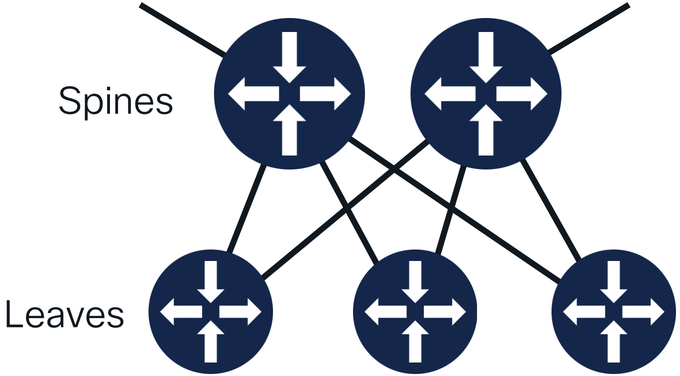
Chassis-level availability

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

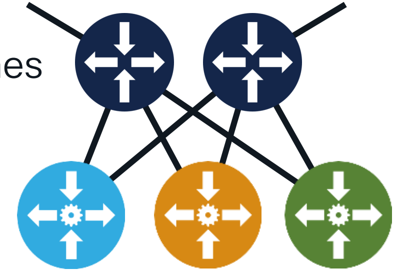
Downtime: < 6s

Availability: > 99.9999%

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

Edge Fabric

Based on traditional routers

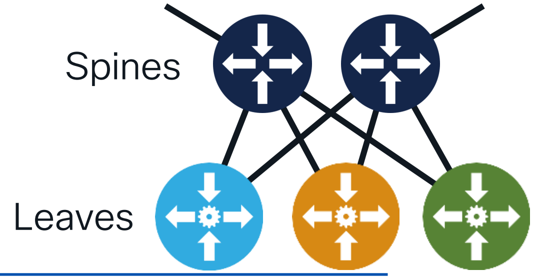


- Pros:
  - Unmatched scale
  - Easy software upgrade without "ISSU" with complex rules
  - Improved solution level availability
  - Add capacity or scale by adding new leaves (when spines are sized accordingly)
  - Less-disruptive upgrades/refreshes
  - Choice of platforms for specific roles
  - Flexible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

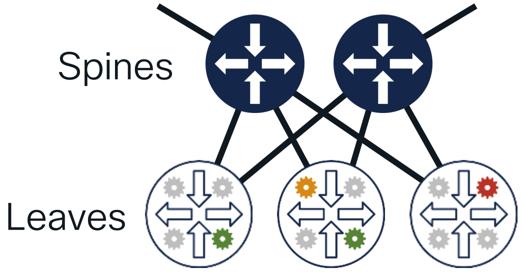
Edge Fabric

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-balacing.
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

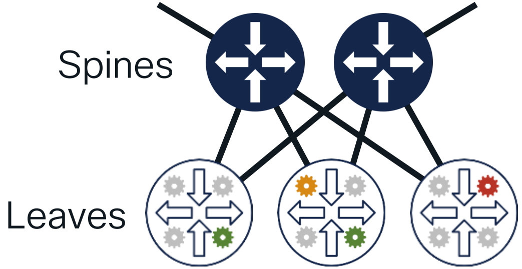


Based on state-of-art silicon and router platforms

- 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…

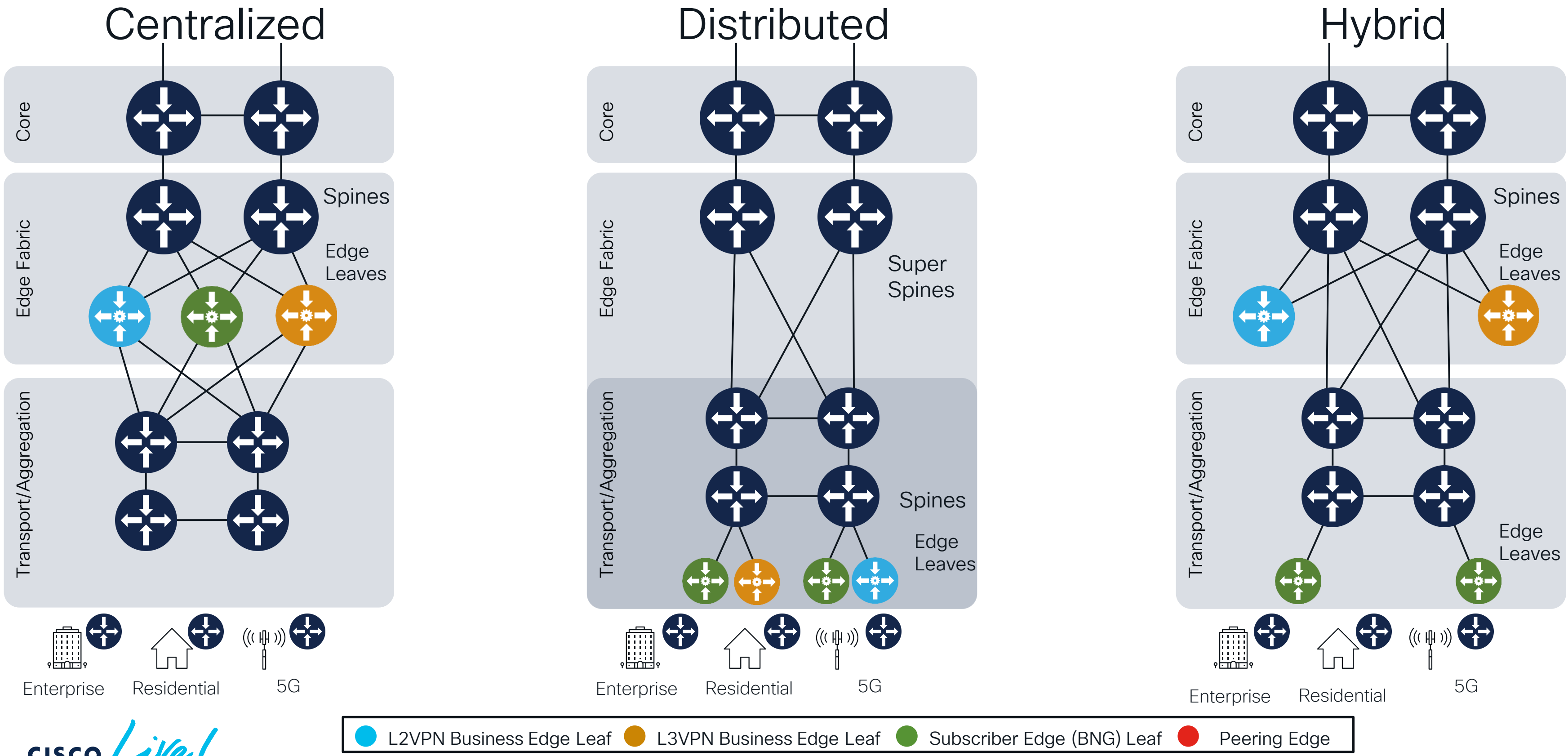
Edge Fabric

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-balacing.
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



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
    - ODN

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

Edge Fabric Management and Automation

Unique requirements to avoid operational overhead

Management*

- Device grouping
- Resource monitoring
- Inventory
- Fault/Peformance 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 sizing

Examples

Number of ports in Spine	32
Spine-Leaf port speed (Gbps)	400
Number of customer ports in Leaf	60
Leaf to customers port speed (Gbps)	100

Key take away: oversubscription ratio has a big impact on fabric size
In reality, there will be no more than 2-4 spine routers.

Number of Leaf user ports	256	256	256	512	512	512	1024	1024	1024
Oversubscription ratio	2	4	6	2	4	6	2	4	6
Leaf routers	5	5	5	9	9	9	18	18	18
Spine routers (2:1 Redundancy)	2	2	2	3	2	2	6	3	3
Spine-Leaf links	32	16	11	64	32	22	128	64	43

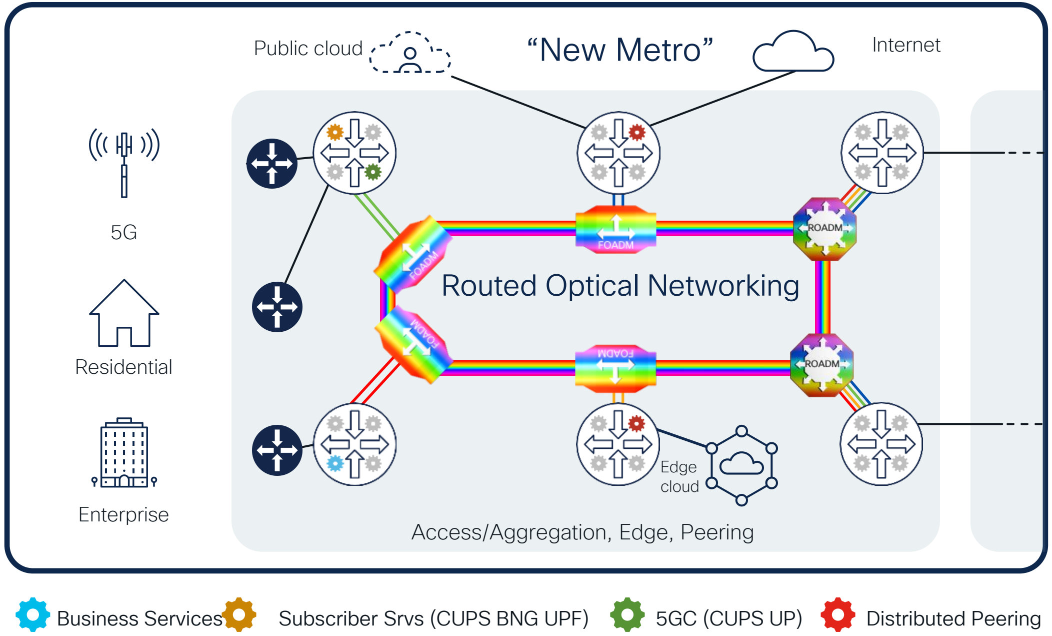
Edge Fabric benefits – real-life example Recent APJ customer case study

	Configuration 1	Configuration 2	Configuration 3
Rack Space reduction	57%	57%	67%
Power reduction (DC Power vs chassis)	63%	46%	52%
Power Cost savings	64%	46%	52%
Bandwidth Provided		150% more	

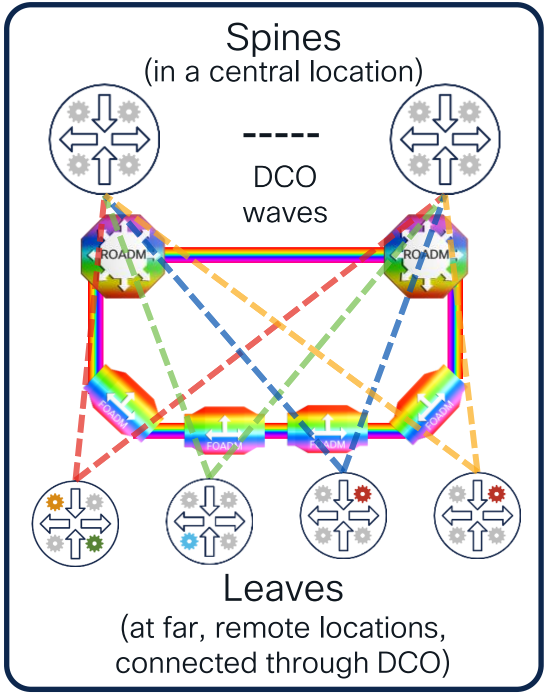
Fabric architecture could provide 150% more bandwidth at lower CAPEX plus other significant saving in power and rackspace for high-capacity Edge locations

Synergies with other solutions
Routed Optical Networking example

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



Distributed
Fabric Architecture



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)

You should look for a Fabric model when:

- You have a real need for:
  - Horizontal scale
  - Splitting services in different leaves to balance cost
  - System level oversubscription
  - Split large blast radius (modular chassis has become big and too risky)
- You can operationalize it:
  - Fabric managers have not been designed for Edge use cases before (Work in Progress)
  - For larger systems, overhead will balance out by managing more devices that are lower touch
  - Use an automation first approach
  - Adjust IGP/BGP (or overlay/underlay) architecture to address scale and operations

Conclusion

Summary

- Traditional Edge Routing designs are challenged by new traffic and business realities
  - Ever growing bandwidth with flat ARPU
  - Increasing cost and complexity of scaling centralized architectures
  - Counter to content and application evolution towards distributed architectures
- New and evolving router and network architectures should be considered as design options moving forward
  - Fully distributed systems
  - Fabric models
  - Each with unique benefits and trade-offs, but overall better alternatives that can be combined
- Modular chassis will still play an important role for some applications
  - Ex. High-capacity peering, Core

The bridge to possible

Thank you

Let's go