The bridge to possible

Best Practices for Ensuring High Availability

In Service Provider IP/MPLS Backbone Networks

David J. Smith, Distinguished Solutions Engineer
Lokesh Khanna, Solutions Engineer

BRKSPG-2080

cisco Live!

#CiscoLive

Cisco Webex App

https://ciscolive.ciscoevents.com/ciscolivebot/#BRKSPG-2080

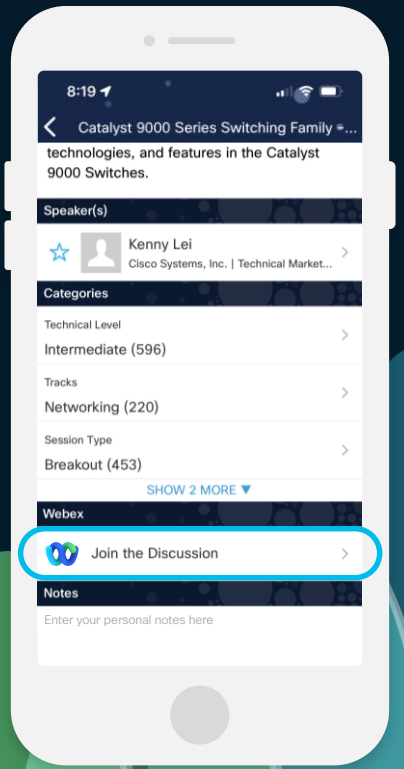
Questions?

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

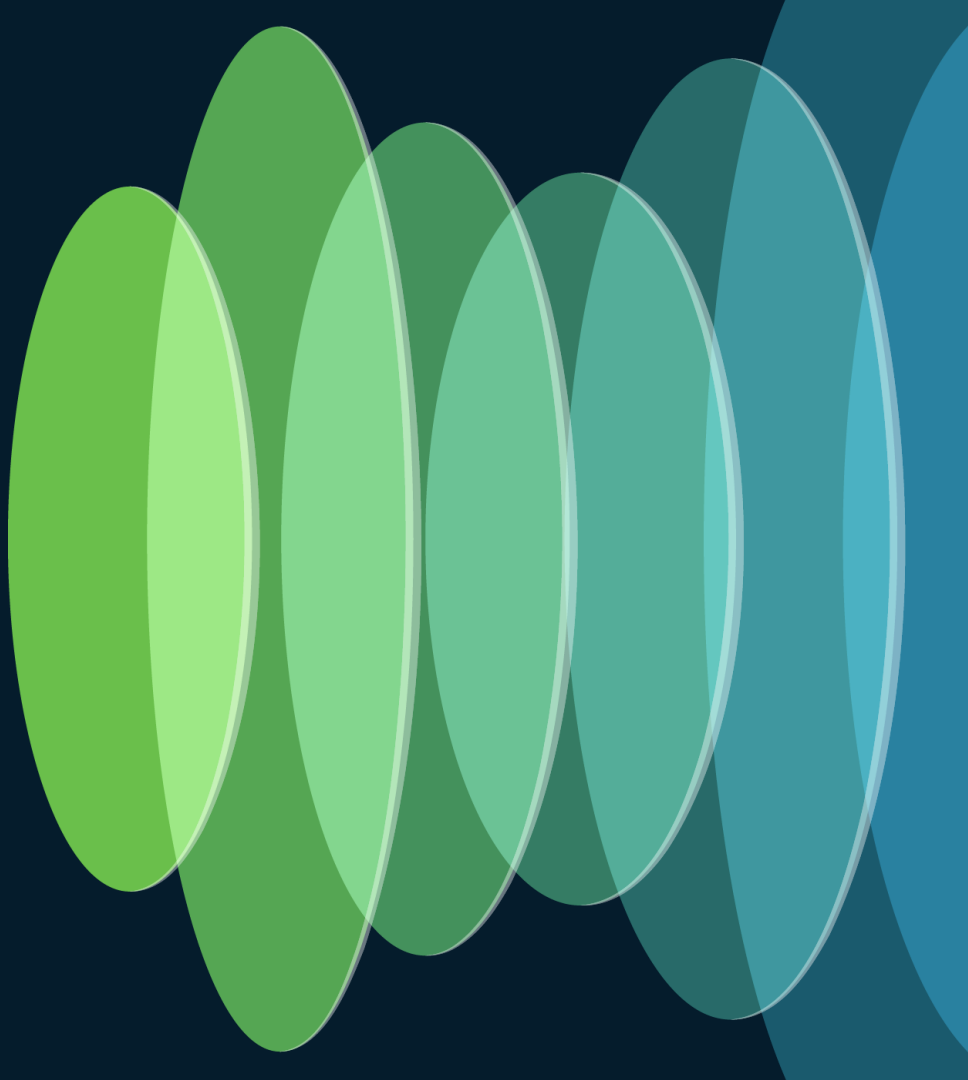
Webex spaces will be moderated by the speaker until June 7, 2024.



Agenda

- Introduction
- Architectural Considerations
- Operational Considerations
- Conclusion

Introduction



Goals for Network Availability

- Maintain the network with stable reliability above SLOs and SLAs
- Mitigate the impact of:
  - Failures
  - Errors
  - Security attacks
  - Maintenance activities
- Note, network redundancy is great when network elements fail hard and fast
- However, when something malfunctions (errors) but doesn't fail completely, the consequences can be devastating despite redundancy

Key Tenets of Highly Available Networks

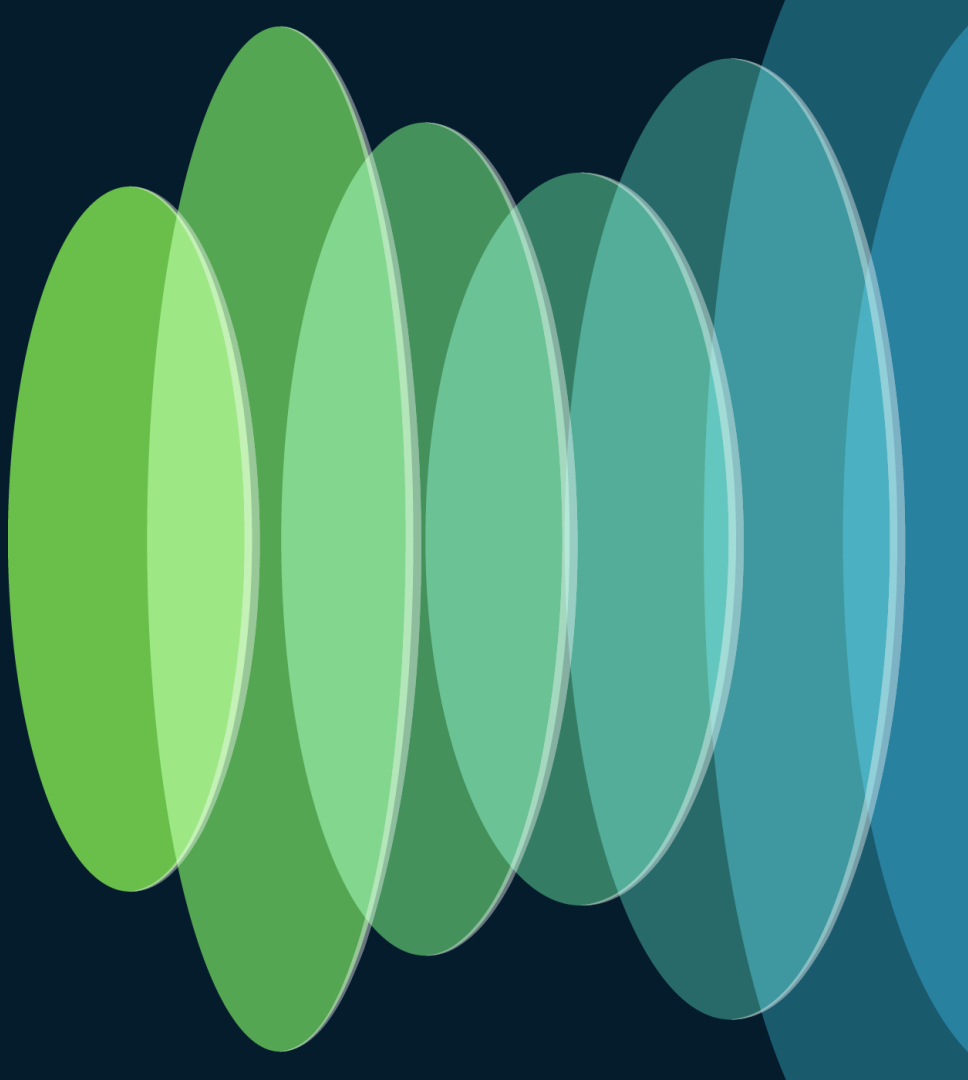
Architectural

- Node resilience
- Network redundancy
- BCPs for the IP control, data & management planes
- Network simplification

Operational

- Operational process rigor
- Network management & tooling
- Network automation
- Incident response

Architectural Considerations



Router Architectures and Resilience

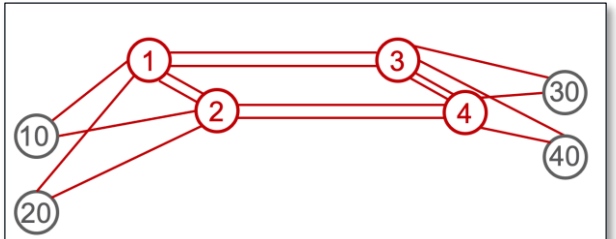
	Distributed / Modular	Centralized	Fixed
HW redundancy	Full redundancy (RP, fabric, power, fans)	Full redundancy	Partial redundancy (power/fans only, not RP)
Scaling	Scales vertically; Enables bandwidth scaling	Scales horizontally & vertically	Scales horizontally; Enables service scaling
Blast radius	Larger	Smaller	Smaller

- All support IP control, data and management plane best practices for network resiliency
- A hybrid deployment may be the most optimal in terms of availability

Network Redundancy

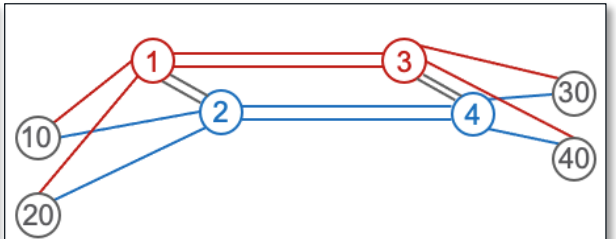
- Minimize single points of failure with redundancy and diversity:
  - Nodes (e.g., core routers, BGP RRs)
  - Router links (e.g., Ethernet Bundles, IP ECMP)
  - IP paths (e.g., BGP multi-homing, multi-plane architecture)
  - Optical paths (i.e., separate conduits) – if some fibers must use the same conduit, SRLGs may help
  - Geo-redundancy of a central office (CO), DC, Colo, MTSO, PoP, etc.

Multi-Plane Architectures - Examples



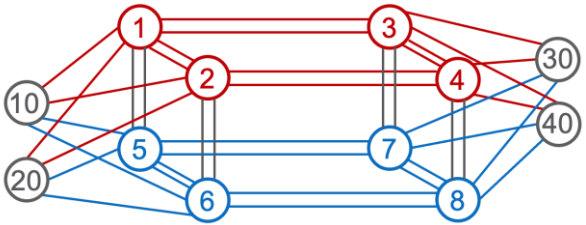
Single Physical Plane
Single Logical Plane
Single IGP

1



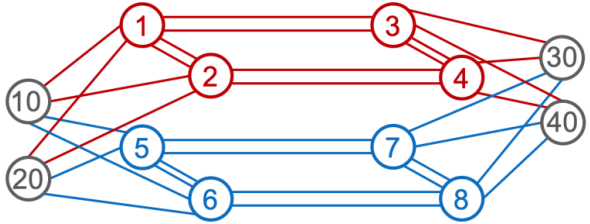
Single Physical Plane
Two Logical Planes
Single IGP

2



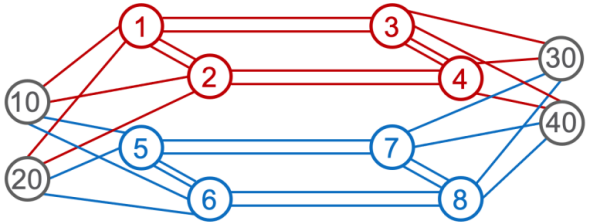
Two Physical Planes
Two Logical Planes
Single IGP

3



Two Physical Planes
Two Logical Planes
Single IGP

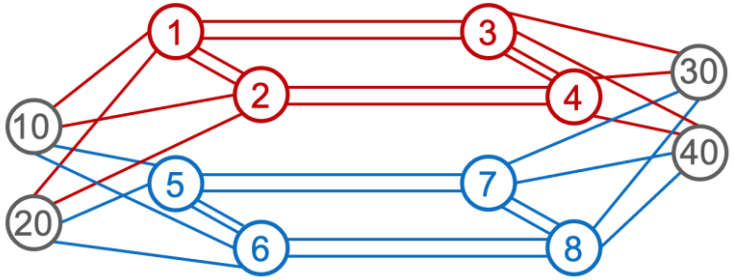
4



Two Physical Planes
Two Logical Planes
Two IGPs (one per plane)

5

Multi-Plane Traffic Steering



- Traffic steering options:
  - IGP costs/metrics
  - MP-BGP policies and service route attributes (e.g., local preference, next hops)
  - TE policies (e.g., RSVP-TE, SR-TE)
  - Segment Routing (SR) Flexible Algorithms
- Variety of multi-plane service models. Examples:
  - Active / Active load balancing
  - Active versus Backup
  - Path protection / disjointness (e.g., Live-Live contribution video)
  - Service-based routing (e.g., VPN versus Internet, Mobility versus Wireline)
  - Secure (e.g., MACsec encrypted) versus Unsecured circuits

IGP Hardening / Best Practices

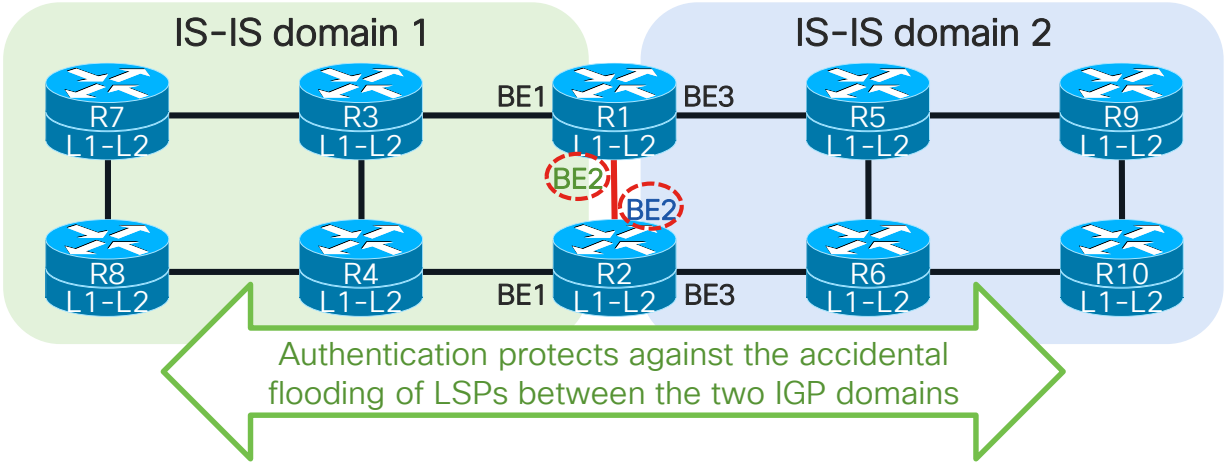
- Avoid BGP redistribution into IGP; If not done correctly, may cause IGP failure or routing loops
- Ensure an upper limit on the number of prefixes that can be redistributed into the IGP to protect against a misconfiguration

```
maximum redistributed-prefixes 10000 75 /* IOS XR default */
```

- Configure OSPF 'max-lsa' commands to limit the number of non-self-generated LSAs kept in the LSDB
  - Prior to IOS XR 7.9.1 'max-lsa' was disabled by default
  - IOS XR 7.9.1 added 'max-lsa' default of 500K
  - IOS XR 7.10.1 added 'max-external-lsa'
  - Not applicable to IS-IS; However, IS-IS restricts the max number of LSPs an IS-IS node can originate to 256

IGP Hardening / Best Practices

- Configure IGP cryptographic authentication to:
  1. Mitigate the risk of malicious IGP attacks
  2. Protect against the accidental collapsing of two IGP domains



R1 config snippet:

```
router is-is 1
  net 49.0001.0001.0001.0001.00
  interface Bundle-Ether 2
    address-family ipv4 unicast
    hello-password hmac-md5 encrypted [pwd1]
```

R2 config snippet:

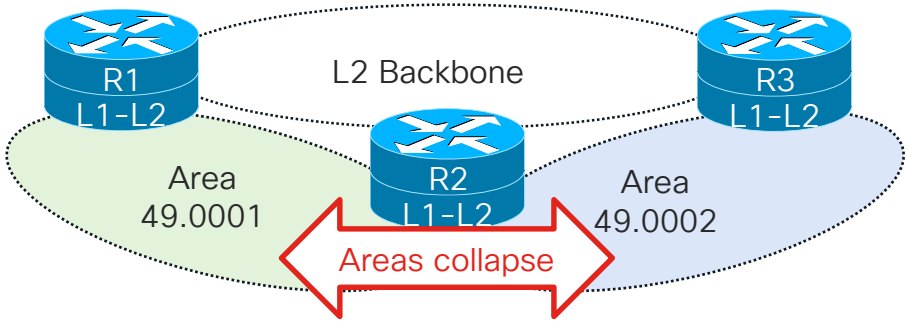
```
router is-is 2
  net 49.0002.0001.0001.0001.00
  interface Bundle-Ether 2
    address-family ipv4 unicast
    hello-password hmac-md5 encrypted [pwd2]
```

IGP Hardening / Best Practices

- Configure the IS-IS routing process type along with proper area addresses (or NETs) to ensure the proper level of adjacencies
  - **`is-type level-1`** specifies that a router only establish adjacencies with other routers in the same area
  - **`is-type level-2-only`** specifies that a backbone router cannot communicate with level-1 only routers
  - **`is-type level-1-2`** specifies that a router act as a gateway (e.g., ABR) to connect different areas (IOS XR default) – use this mode on ABRs only

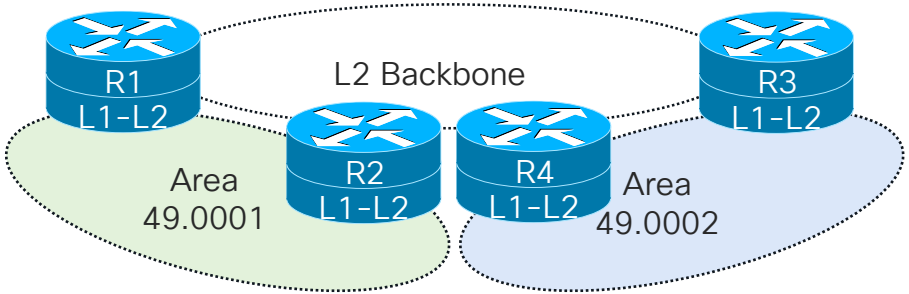
IGP Hardening / Best Practices

- Do not configure multiple area addresses for a single IS-IS instance
  - Only useful (temporarily) to merge multiple IS-IS areas or split up an area – use with caution!
  - Otherwise, risk of accidental collapse of / LSP flooding between areas



R2 config snippet:

```
router is-is 1
  net 49.0001.0001.0001.0001.00
  net 49.0002.0001.0001.0001.00
```



- IS-IS associates routing nodes (not interfaces) to an area
- Use separate IS-IS ABRs per area

IGP Hardening / Best Practices - Scaling

- Excessive IGP scale may affect network stability
- Manage IGP scale properly to reduce IGP churn and/or blast radius
- Multiple well-known techniques are available to protect and scale the IGP:
  - Hierarchical IGP (i.e., using multiple areas)
  - Multi-domain IGP with BGP-LU (aka Unified MPLS / Seamless MPLS)
  - Multi-domain IGP with IP route summarization (SRv6 only)
  - Converged SDN Transport using SR-PCE
  - Multi-plane architecture with each plane having its own distinct IGP

Other IGP Best Practices (1)

Reference

- Optimal IGP exponential backoff algorithm timers to rate limit LSA/LSP generation and SPF computation during network instability (IOS XR default)
- IGP prefix prioritization for IPv4 /32 and IPv6 /128 host prefixes (i.e., BGP next hops) during SPF run to minimize network convergence for transit traffic (IOS XR default)
- Configure "network point-to-point" for all router-to-router IGP links – otherwise they become LAN which is complicated (DR/BR, LSAs, features)
- Don't have too many ABRs between the backbone and a non-backbone area as it increases scale of inter-area/summary LSAs
- Control plane policing (e.g., IOS XR LPTS / CoPP) to protect router CPU and ensure control plane stability (IOS XR default)
- OSPF TTL security (RFC 3682 GTSM) to filter remote attacks against OSPF. IS-IS runs directly on Layer 2 so it is not exposed to remote attacks

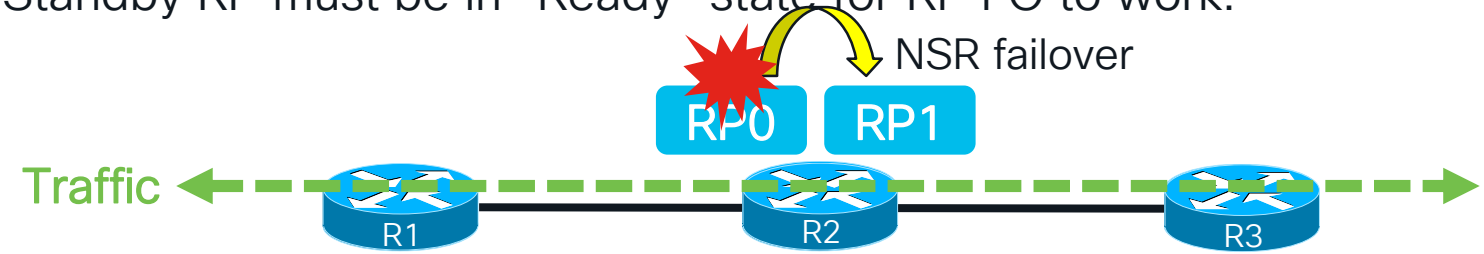
Reference

Other IGP Best Practices (2)

- Consider IGP prefix-suppression to avoid carrying the P2P prefixes of transit links in the LSDB, thereby, reducing IGP convergence time
- In multi-plane architectures with multiple IGPs, avoid redistributing IGP routes between planes
- Enable LDP/IGP synchronization to prevent IGP forwarding on a link when the associated LDP session is down
- Configure LDP label allocate for host-only and LDP label advertise to permit I-BGP next hops only (e.g., PE loopbacks) so that only transit traffic is MPLS LSP forwarded
- LDP session protection minimizes traffic loss and provides faster network convergence during link DOWN➔UP events

Non-Stop Routing (NSR)

- Enables non-stop traffic forwarding during an RP failover
- Backup RP synchronizes and preserves the routing state, which includes protocol sessions and routing process data (e.g., BGP table, IGP LSDB)
- During RP failover, the backup RP is used to maintain control plane sessions and traffic forwarding without interruption
- Peer routers are unaware of such events – no protocol signaling required with Peers, however, backup RP is required
- Standby RP must be in “Ready” state for RP FO to work.

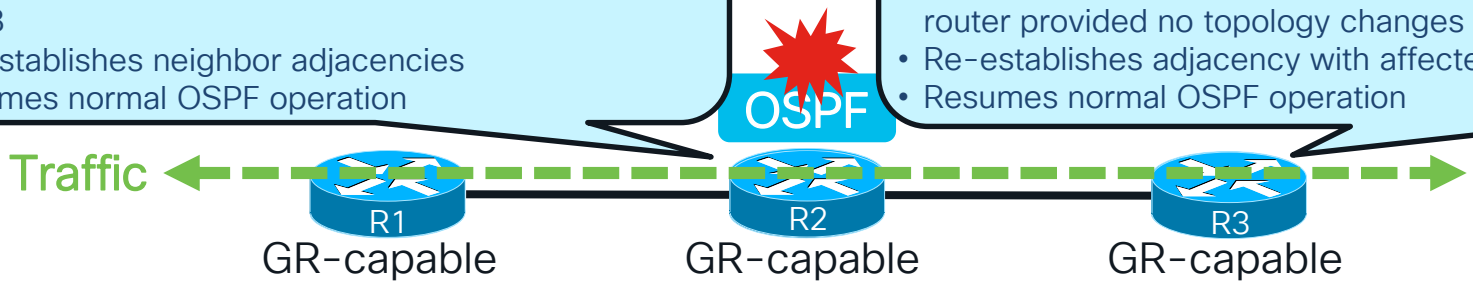


Non-Stop Forwarding (NSF) with Graceful Restart

- OSPF (RFC 3623), IS-IS (RFC 8706), LDP (RFC 3478), BGP (RFC 4724)
- Enables a routing process to restart without traffic loss
- Redundant RP is not required. However, neighbors must be GR-enabled

Graceful Restart procedures on affected router:
- Originates grace-LSAs (prior to restart)
- Starts grace period (lifetime) timer
- Restarts OSPF process while preserving OSPF routes in FIB
- Re-establishes neighbor adjacencies
- Resumes normal OSPF operation

Graceful Restart procedures on neighbors:
- Receives grace-LSAs from affected router
- Starts grace period (lifetime) timer
- Preserves OSPF routes and forwarding via affected router provided no topology changes
- Re-establishes adjacency with affected router
- Resumes normal OSPF operation



- If the topology changes or the NSF/GR timer expires before OSPF functions return to normal, the NSF/GR routes will be purged, potentially impacting traffic forwarding

Interface Fault Handling

- Carrier Delay to delay the processing of link-up notifications

```
R1(config-if))# carrier-delay down 0 up 10 /* IOS XR default */
```

  - Link-down 0 (default) improves convergence time for interface failures to reduce traffic loss
  - Link-up 10 (default) delays traffic forwarding onto a link until it is stable
  - Note, “up 3000” is recommended when a router is connected to L2 Ethernet switch which is often slower to detect the link up

- Event Dampening to limit the propagation of frequent interface state changes, thereby, improving network control plane stability

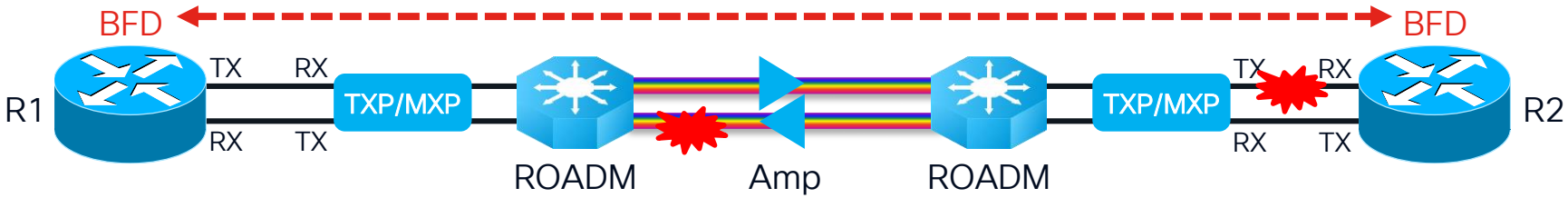
```
R1(config-if))# dampening /* disabled in IOS XR by default */
R1# show dampening interface - Displays dampened interfaces.
```

- Cisco recommends using Carrier Delay and Interface Dampening to avoid churn in the control plane caused by unstable interfaces

BFD (Bidirectional Forwarding Detection)

- Recommended for fast failure detection and, in turn, to rapidly trigger IGP/BGP control plane convergence, IP/MPLS FRR protection and BGP PIC Edge
- Also provides end-to-end optical path verification and advanced capabilities

Failure Detection Method	Detection Time	Applicability
Optical LoS/LoF	~10 ms	• Relies on optical network to propagate remote faults
Routing protocol timers	>30 secs	• Lower timers affect routing scale
Ethernet CFM w/ EFD	12 ms or more	• HW offload enables fast timers
BFD	12 ms or more	• HW offload enables fast timers • Supports advanced capabilities: BFD multi-hop, BFD strict mode, BFD dampening, BFD over Bundles



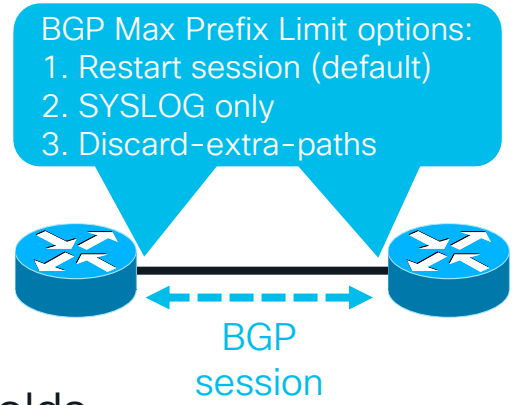
BFD (Bidirectional Forwarding Detection)

- BFD is recommended for IGPs, eBGP and Ethernet Bundles
  - Specifically, single-hop, async mode, strict-mode and with dampening
- Past multi-vendor interoperability issues related to BFD for Ethernet Bundles should no longer exist thanks to IETF standards
- BFD for Ethernet Bundles includes two (2) deployment modes:
  1. BFD over Bundles (aka "BoB") - defined in RFC 7130. "BoB" is where a (micro) BFD session is run on each member link within the bundle. Cisco supports two "BoB" modes: Cisco mode (pre-standard) and IETF mode. IETF mode is recommended and required for multi-vendor BoB interoperability
  2. BFD over Link Bundles (aka "BLB") - defined in RFC 5880. "BLB" is where a single BFD session is run for the entire bundle
- "BoB" is recommended and provides faster detection versus "BLB"

BGP Considerations

- Again, avoid BGP redistribution into IGP
  - If not done correctly, may cause IGP failure or routing loops
- IP/MPLS core routers should not carry MP-BGP service routes
  - No advantage carrying Internet routes / participating in the Internet control plane; BGP-free core recommended
  - Transit traffic should be forwarded via classic MPLS, SR-MPLS or SRv6

BGP Max Prefix Limit



- Be aware of any maximum prefix limit settings on BGP sessions and what could happen if exceeded; Optimize accordingly
  - Applicable to both eBGP and iBGP sessions
  - Different BGP address families have different thresholds
  - The default behavior is that if limit is hit, a BGP session would be taken down and remain down until cleared by an operator
  - It is highly recommended to define the action if limit is hit
  - Session resets and discarding prefixes may cause traffic disruptions which may persist until the neighbor(s) comply with locally configured max prefix limits

BGP Error Handling

- Basic error handling is enabled by default, and it prevents BGP session reset if a malformed update is received
- Treats a malformed update as withdraw, which removes the prefixes with malformed attribute from BGP table
- Cisco XR supports attribute filtering which allows removing the malformed attribute instead of withdrawing the prefixes

```
router bgp 65530
attribute-filter group BGP-ATTRIBUTE-FILTERING
attribute range 28 discard
!
neighbor 10.101.203.203
 update in filtering attribute-filter BGP-ATTRIBUTE-FILTERING
!
```

Extended BGP Error Handling

- Cisco added extended error handling capabilities to avoid session reset under following conditions (as per RFC 7606)
  - NEXT_HOP length is not 4
  - MP_REACH Nexthop length is 0
  - MP_REACH Nexthop length is invalid for the MP AFI/SAFI
  - Duplicate Non-OptionalTransitive attributes
- “Extended” revised error handling in IOS XR requires the following:

```
R1#configure
R1(config)#router bgp 65530
R1(config-bgp)#update in error-handling extended ebgp
R1(config-bgp)#update in error-handling extended ibgp
```

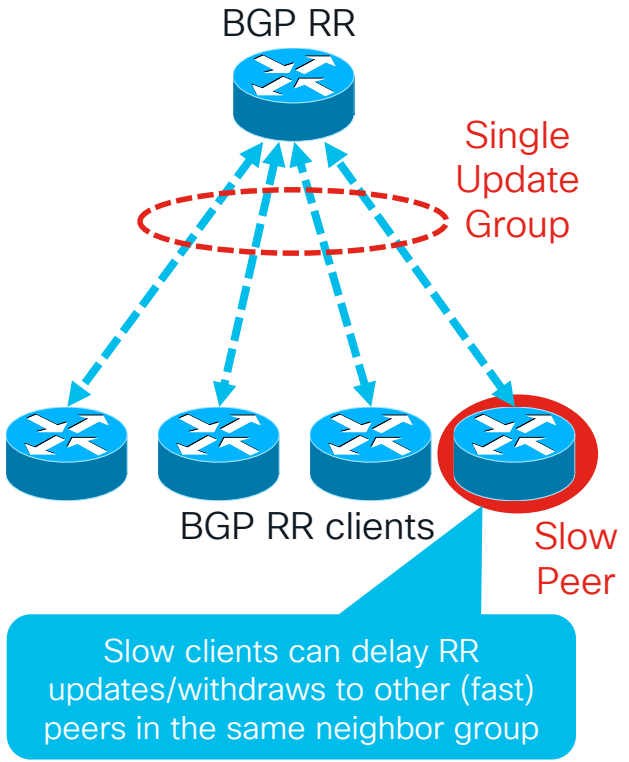
- Modern BGP routers should comply with RFC 7606

BGP Route Reflectors

- RRs simplify BGP control plane provisioning in large-scale BGP networks
- RR redundancy eliminates single points of failure, however, excessive redundancy can be detrimental
  - Increases BGP packet processing
  - Increases the number of BGP paths
  - Delays BGP convergence
- A fault within a RR cluster has the potential to impact the entire cluster. By maintaining smaller RR cluster sizes, the extent of traffic disruption caused by a fault can be contained

BGP RR Slow Peers

- BGP RR clients that are ‘permanently slow’ should be moved into their own distinct neighbor group separate from fast peers
  - To prevent slow peers from delaying BGP RR Updates/Withdraws to fast peers
  - IOS XR BGP slow peer detection and processing is not intended for peers that are permanently slow given its highly CPU intensive
  - Note, the actual RR route-policies for slow and fast peers don't have to differ; they simply require distinct names; This method ensures that slow peers are placed in a separate neighbor group from fast peers



BGP RR Considerations

- BGP RRs often hide paths which can cause suboptimal routing or route oscillations; Capabilities are available and recommended to prevent this:
  - BGP Add Paths
  - BGP ORR (Optimal Route Reflection)
- Consider the use of separate BGP RRs for increased fault isolation:
  - BGP Internet service routes (IPv4 unicast, IPv6 unicast, 6PE)
  - BGP VPN service routes (VPNv4, VPNv6, EVPN)
  - BGP transport routes (BGP-LU aka IPv4 labelled unicast)
  - BGP-LS for IGP topology export

Reference

Other BGP Best Practices (1)

- Static configuration of router ID using loopback address to prevent changes to the router ID and consequent flapping of BGP sessions
- Enable TCP Path MTU discovery to enable use of the largest packet size that does not require fragmentation anywhere along the path between two BGP speakers
- eBGP route policies to filter routes accepted from / advertised to eBGP neighbors (e.g., bogons, more specifics, infrastructure routes)
- Delete inbound communities, especially if doing VRF peering; some vendors may accept routes with an RT set from an eBGP neighbor
- Limit eBGP peering to only explicitly configured neighbors
  - Restrict the number of dynamic BGP sessions on devices under your administration

Reference

Other BGP Best Practices (2)

- BGP cryptographic authentication using TCP-AO with MD5, SHA or AES
- eBGP TTL security (i.e., RFC 3682 GTSM) to help protect against external remote BGP attacks
- AS-PATH limits to filter prefixes with an AS-PATH length greater than a specific value (e.g., 50)
- eBGP Route Flap Damping (RFD) is recommended to suppress Internet BGP churn (see http://rfd.rg.net/)
- Use IP routing advertisements to restrict external IP reachability to internal infrastructure and, thereby, make infrastructure attacks more difficult

Reference

Other BGP Best Practices (3)

- BGP RPKI prefix origin validation to drop invalid prefixes and help prevent prefix hijacking
- BGP flowspec for rapid, intra-domain, distributed attack mitigation
  - See RFC 9117
  - Take caution not to inadvertently filter eBGP sessions with flow specifications
- IETF BCP 194 (RFC 7454: BGP Operations and Security) if providing Internet service
- Be aware that different vendors use different RIB administrative distances and, therefore, have different preferences for IGP routes versus eBGP routes
  - In a multi-vendor environment, RIB admin distances should align

MPLS Label Scale Considerations

	ASR 9900	NCS 5700	8000
LDP label allocate for host-routes	Recommended	Recommended	Recommended
MP-BGP label allocation modes (L3VPN, 6PE)	Per-Prefix Per-VRF Per-CE	Per-VRF* (required)	Per-VRF* (required)
I-BGP next hop (IPv4, IPv6)	next-hop-unchanged next-hop-self	next-hop-self (Recommended)	next-hop-self (Recommended)

* In addition to configuring Per-VRF label allocation on a single node, it might also be necessary to configure it network wide
* This depends on the number of remote MP-IBGP L3VPN prefixes learned and installed
* Note, BGP per-prefix label allocation is the default IOS XR behavior

QoS Considerations

- DiffServ QoS plays a vital role in ensuring high network availability
  - To ensure control and management plane traffic is not discarded due to congestion
  - To isolate traffic classes and guarantee priority customer traffic
- Designate a traffic class with ample bandwidth exclusively for SP infrastructure control and management plane traffic
- On ingress to the network edge, mark (aka color) traffic to guarantee accurate classification of traffic flows downstream

Network Simplification

- A simplified network with fewer technology layers and protocols, helps to reduce complexity and potential failure scenarios as well as makes automation easier and more sustainable long-term
- Key tenets of a simplified IP/MPLS network:
  - MP-BGP as the unified service overlay control plane for all network services
  - Segment Routing (SR) as the unified forwarding plane including FRR protection
  - SR-TE (Traffic Engineering) for advanced control of traffic flows and bandwidth
  - Centralized (SDN) Controller for bandwidth reservation bookkeeping and network-wide orchestration of SR-TE policies
  - BGP-LS for export of topology link-state and TE information to Controller
  - BFD for fast failure detection and, in turn, to rapidly trigger network convergence
  - DiffServ QoS on every link to isolate traffic classes and guarantee priority traffic
  - YANG model-driven programmability & telemetry for management and automation
  - Routed Optical Networking for greater network efficiencies and economics

Segment Routing (SR)

- A programmatic IP source-routing architecture that provides the optimal balance between distributed intelligence and centralized control
- Mass simplification
  - Reduces control plane protocols (no LDP, no RSVP-TE, no BGP-LU, no MPLS OAM)
  - Delivery of all network services (IP, MPLS, Ethernet, Private Line, Wave) over IP
  - Automatic <50 ms protection (topology independent), automated traffic steering
- Mass scaling
  - Reduces core network state (no stateful RSVP-TE tunnels required for TE/FRR)
  - Enables route summarization between domains/areas (SRv6), On-Demand BGP Next-hops
- Advanced capabilities
  - Advanced TE (flow-based, ECMP-aware, multi-domain, disjointness, circuit-style, low latency)
  - Network slicing, service chaining, data plane monitoring, IP performance measurements

* Note, SRv6 provides maximum simplicity, scale and capabilities

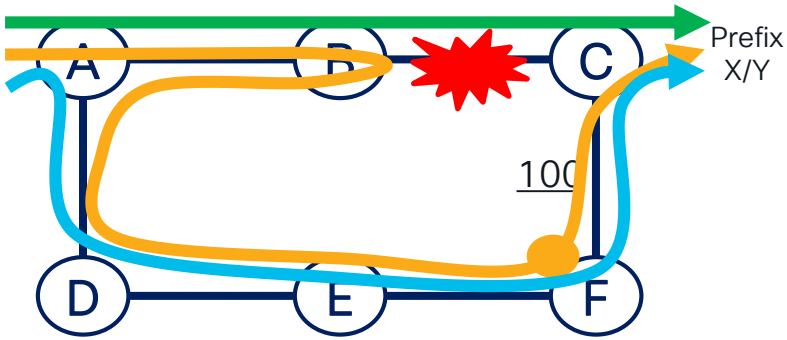
Fast Re-Route (FRR) Protection

- IGP convergence times vary, typically falling under 2-3 seconds
  - Sub 1 second convergence is achievable using modern hardware and the optimized (IOS XR default) timers
- To minimize traffic disruption further, requires FRR protection
- Variety of techniques available to attain 50 msec. FRR protection
  - RSVP-TE/FRR – often requires many stateful core tunnels
  - Per-Prefix LFA – cannot guarantee FRR coverage (e.g., box topology)
  - Remote LFA – requires targeted LDP sessions be established
  - TI-LFA – Topology independent, no stateful tunnels, no targeted LDP sessions (enabled by SR)

* LFA (Loop Free Alternate)

TI-LFA FRR Protection

```
router ospf 100 area 0 fast-reroute per-prefix
router ospf 100 area 0 fast-reroute per-prefix ti-lfa enable
```



Default metrics: 10

IGP path before link failure

TI-LFA protect path (SID list @ B = [F, F→C])

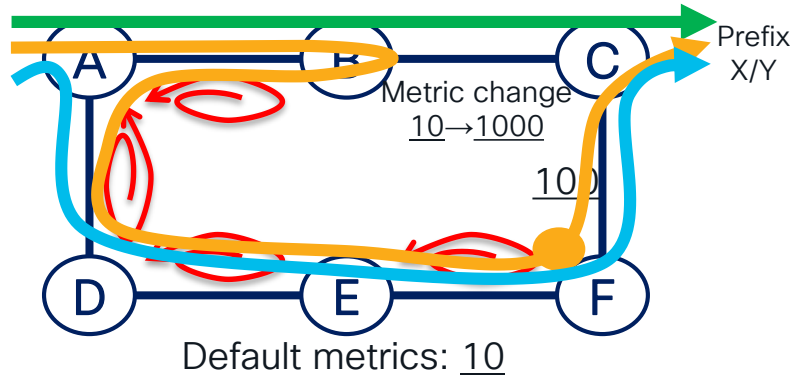
Post IGP convergence path

- <50 ms protection for link, node and SRLG failures
- Simple to operate and understand given it's automatically computed by the IGP:
  - IGP on Node B computes shortest path to Prefix X/Y via Node C (active path in FIB)
  - IGP on Node B also computes LFA to prefix X/Y via Node F (LFA path in FIB)
  - When Node B detects link failure to Node C, it FRR protects traffic using LFA path in FIB
  - LFA path is used until IGP reconverges, thereby, minimizing traffic loss
- No stateful core tunnels required
- 100% topology coverage / independent

SR Microloop Avoidance

- Hop-by-hop IP routing may induce transient microloops during convergence events
- Microloops can lead to increased packet loss which is obviously undesirable
- SR microloop avoidance prevents microloops for isolated convergence events
- When a node learns of a topology change and then computes new paths for its destinations:
  - If the node sees that transient microloops are possible for a destination, then it constructs a SID-list to steer traffic microloop-free
  - SID list @ A, B, D, E for Prefix X/Y = [F, F→C]

```
ipv4 unnumbered mpls traffic-eng Loopback0
router ospf 100 microloop avoidance segment-routing
```



IGP path before link metric change

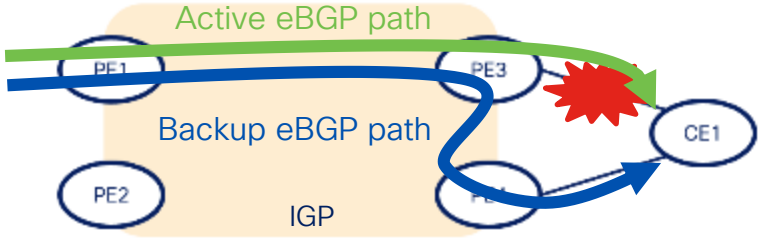
IGP microloop due to link metric change

SR microloop avoidance

Post IGP convergence path

BGP PIC Edge

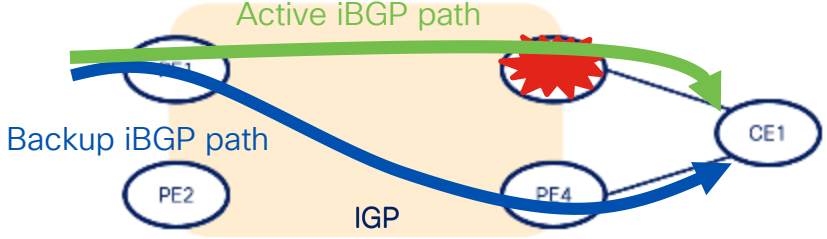
BGP PIC Edge Link Protection



- Egress PE-CE link failure triggers BGP PIC Edge Link Protection on egress PE (PE3)
- PE3 has backup path via PE4 pre-programmed in its FIB in advance of the failure
- PE3's BGP convergence onto its PE4 backup path is prefix independent and pre-programmed, hence, fast (~sub-second)

See also: IETF draft-ietf-rtgwg-bgp-pic-20

BGP PIC Edge Node Protection

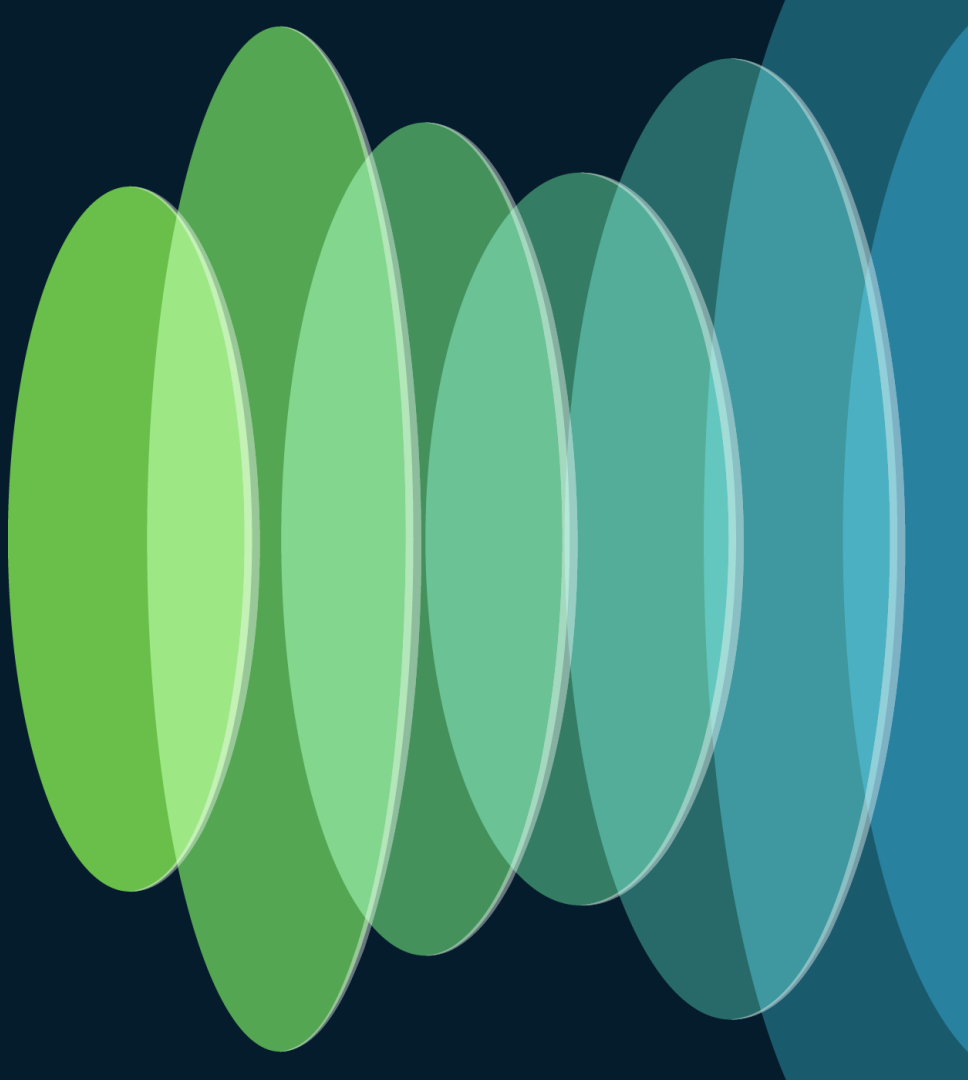


- Egress PE node failure (PE3) triggers BGP PIC Edge Node Protection on ingress PE (PE1)
- Triggered by removal of PE3 from PE1's IGP data base
- PE1 has backup path via PE4 pre-programmed in its FIB in advance of PE3 node failure
- PE1's BGP convergence onto PE4 path is prefix independent and pre-programmed, hence, fast but dependent on IGP convergence (removal) of PE3

IP Data Plane Best Practices

- Interface ACLs / packet filtering at edge – automation is key in maintaining accuracy
- If providing Internet services, Unicast RPF or ACL to mitigate IP source address spoofing (IETF BCP 38 and BCP 84) and facilitate traceback of security attacks
- ICMP best practices (e.g., no ip unreachables, no ip redirects, no IP→MPLS TTL propagation to prevent TTL expiry attacks)
- Network-wide MTU sizing to avoid IP fragmentation and/or ICMP 'Fragmentation Needed and Don't Fragment was Set'

Operational Considerations



Operational Process Rigor

- Entails the ability to execute network changes without adversely affecting network stability and performance
- All network change procedures should be explicitly:
  - Documented: Step by step including pre-checks, change procedures, expected impacts, post-checks and roll-back procedures – no ambiguity!
  - Reviewed and lab tested: Written by one person/group, Tested by another person/group. Testing includes all operational tools and scripts involved. Test lab / sandbox should mimic the production network – differences may introduce risk
  - Recommend vendor(s) services team also review to verify and look for any potential collateral issues

Operational Process Rigor

- After lab verification, first time deployment of network change should be deployed manually (step by step)
- All network changes that may affect traffic forwarding should be performed during a scheduled maintenance window
- After change procedure successfully deployed by hand, relevant tools and scripts may be incrementally utilized for larger scale deployment (i.e., crawl, walk, jog, run, fly)
- Every network event necessitates a Root Cause Analysis (RCA) and suggestions for future mitigation measures (i.e., continuous improvement)

Other Operational Considerations

- Retire end-of-life (EOL) and end-of-support (EOS) equipment with known software defects, security vulnerabilities and no available software updates
  - Know EOL/EOS of your software ahead of time
- To maintain stability, IP routing nodes should not exceed their scale limits

Planned Maintenance Events

- Optimal techniques available to gracefully migrate traffic via alternate nodes/paths (if available) during maintenance activities:
  - IS-IS set-overload-bit
  - OSPF max-metric
  - BGP graceful maintenance (or comparable BGP policy)
  - EVPN cost-out
- Verify that both in-band and out-of-band remote terminal access are operational beforehand to ensure uninterrupted management access during a maintenance activity

Best Practices for Network Management & Tooling

- Network Documentation
  - Maintain up-to-date documentation of your network topology, device inventory, IP addresses etc. Use automated documentation tools wherever possible to keep records current
- Network Monitoring
  - Continuous monitoring for network performance, availability, and security events
  - Use network monitoring tools to identify and alert on abnormal conditions
- Configuration Management
  - Use configuration management tools to maintain consistent configurations across network devices and to automate the deployment of changes
  - Keep a version-controlled backup of all configurations to facilitate quick recovery

Best Practices for Network Management & Tooling

- Performance Management
  - Regularly analyze network performance metrics to identify bottlenecks or degradation of service
  - Implement Quality of Service (QoS) to prioritize critical traffic and ensure bandwidth allocation aligns with business needs
- Change Management
  - Use a formalized change management process to control and record any modifications to the network environment
  - Conduct pre-and post-change analysis to understand the impact and verify success

Best Practices for Network Management & Tooling

- Disaster Recovery and Business Continuity
  - Develop and regularly test disaster recovery plans to ensure network services can be quickly restored after an outage
- Capacity Planning
  - Proactively plan for future network growth by regularly evaluating bandwidth usage trends and predicting future needs
- Automation and Orchestration
  - Implement automation to streamline repetitive tasks, reduce errors, and increase efficiency
  - Expedites fault isolation and Mean Time To Repair (MTTR) by correlating events, logs, and KPIs

Other Important Considerations

- When selecting network management tools, consider factors like scalability, integration capabilities, user-friendliness, and support
- Remote Access
  - Have separate networks for In-Band, Ethernet OOB and console OOB ports
  - Ensure console access is not dependent on either the Management Ethernet network or the In-Band network
  - Use dedicated circuits if necessary

Stats Collection via Telemetry

- Stats collection via Dial-In and Dial-Out
- Periodic or event based
- Customized to recipient - subscribe to any YANG subtree
- Benefits over SNMP
  - Increased operational data and visibility (far beyond SNMP MIBs)
  - High scale due to push model (versus SNMP pull model) - significantly less CPU intensive
  - Network monitoring cadence reduced from 5mins to 10secs
  - Structured data is software friendly – easier to read & automate

Provisioning via Network Automation

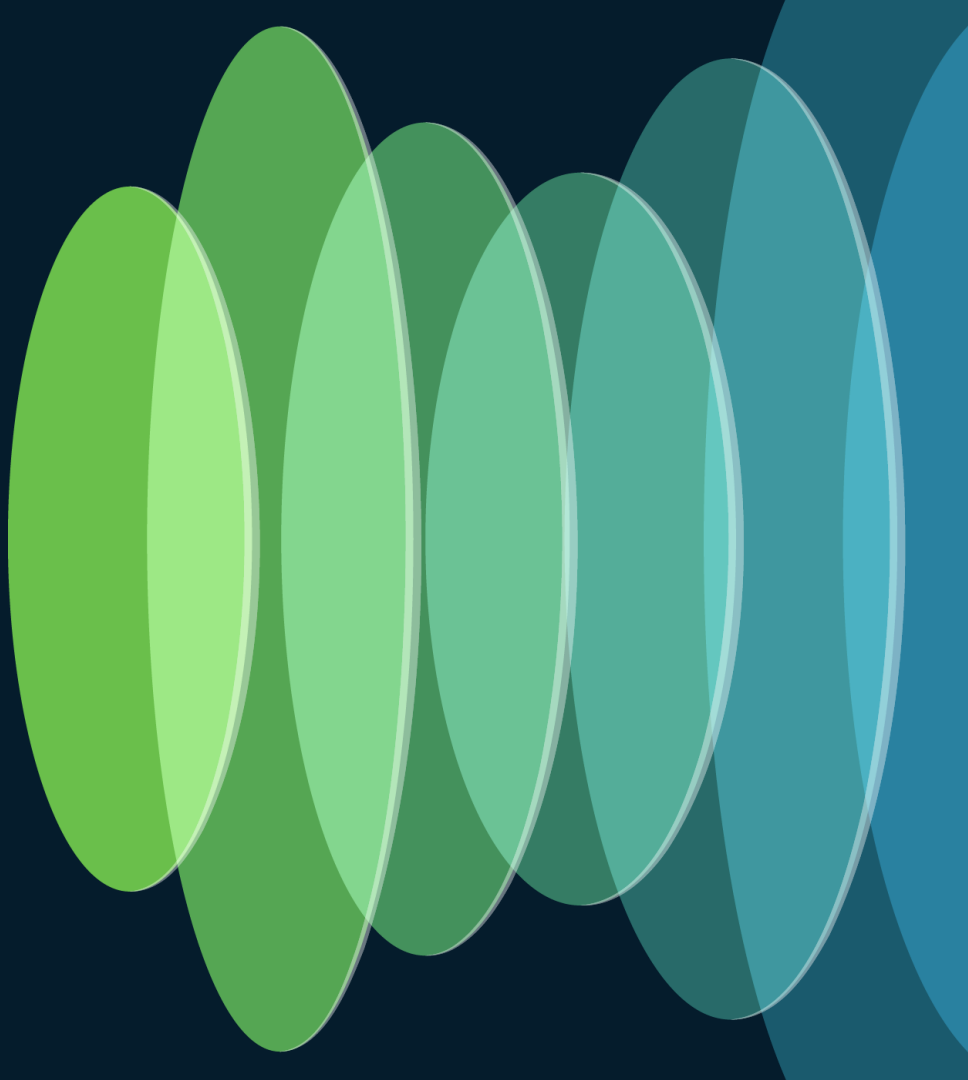
- Leverage Zero Touch Provisioning for Day0 configuration
  - ZTP script can be enhanced to take care of Day0 code upgrades.
- Golden ISO image
  - Build GISO image based on the required package/SMUs
  - GISO simplifies code upgrade process, reduces the complexity of managing individual packages and maintains standardization across the network infrastructure
- Automate Day2 changes
  - Yang models provide standard way to define the network configuration and data
  - Tools like NSO/Ansible/Puppet etc can be leveraged to automate Day2 changes

Incident Response

- Create Priority 1 (P1) “network down” incident response playbooks and conduct simulations / fire drills to prepare for potential failure scenarios
- Create applicable vendor remote access “accounts” in advance / proactively and “enable” on-demand only when required
  - Having to set all this up during a P1 outage wastes valuable time and diverts resources
- During P1 events:
  - Ticketing system should automate notifications and escalations to get resources to start troubleshooting as quickly as possible
  - Open vendor ticket(s) and/or engage all applicable vendors without delay
  - Establish one conference bridge for troubleshooting, RCA, mitigation procedures, etc.
  - Establish a second conference bridge for business leaders – this group should ideally not interfere with / distract the troubleshooting group

* See references [4] and [12] also

Conclusion



Conclusion

- Network failures, errors, security attacks and maintenance activities that impact network availability and performance may also affect an operator's ability to retain existing customers and attract new customers
- It may not be possible to predict or avoid all these events, however, operators can certainly reduce the risks and impact with:
  - Resilient architecture
  - Operational process rigor
  - Network management tooling
  - Ruthless automation
  - Readiness for responding to incidents
- Operators must balance risks, operational complexity and costs
- High availability also requires continuous improvements in all areas

Acknowledgements

- Phil Bedard, Chandu Bhupathiraju, Les Ginsberg, Jakob Heitz, Rajesh Patki, Serge Krier, Ketan Talaulikar

References (1)

1. Cisco IOS XR Deployment Best Practices for OSPF/IS-IS and BGP Routing
   - https://www.cisco.com/c/en/us/support/docs/ios-nx-os-software/ios-xr-software/IOS-XR-Best-Practices/IOSXR-Deployment-BestPractices.html
2. Cisco Security Advisories:
   - https://tools.cisco.com/security/center/publicationListing.x
3. Documentation blogs and tutorials on all things IOS XR: https://xrdocs.io/
4. Getting Started with Site Reliability Engineering for Autonomous Operations (Cisco Live 2020, DGTL-BRKPRG-1058)
5. Highly Available Wide Area Network Design (Cisco Live 2019, BRKRST-2042)
6. High Availability Network Fundamentals (Cisco Press 2001)

References (2)

7. Internet Peering Concepts and Trends (Cisco Live 2020, DGTL-BRKSPG-2003)
8. No Packet Left Behind: Minimising Packet Loss Through Automated Network Operations (NANOG 88, 13 June 2023)
9. Revisiting Recommended BGP Route Flap Damping Configurations
   - http://rfd.rg.net/
   - https://labs.ripe.net/author/clemens_mosig/route-flap-damping-in-the-wild/?pk_vid=794952f64168acc11662146987c053b8
10. Router Security Strategies: Securing IP Network Traffic Planes (Cisco Press 2008)
11. Segment Routing: www.segment-routing.net
12. Transforming network operations through reliability engineering (Cisco Live 2020, DGTL-BRKNWT-2153)

Complete Your Session Evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to **win 1 of 5 full conference passes** to Cisco Live 2025.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

Level up and earn **exclusive prizes!**

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

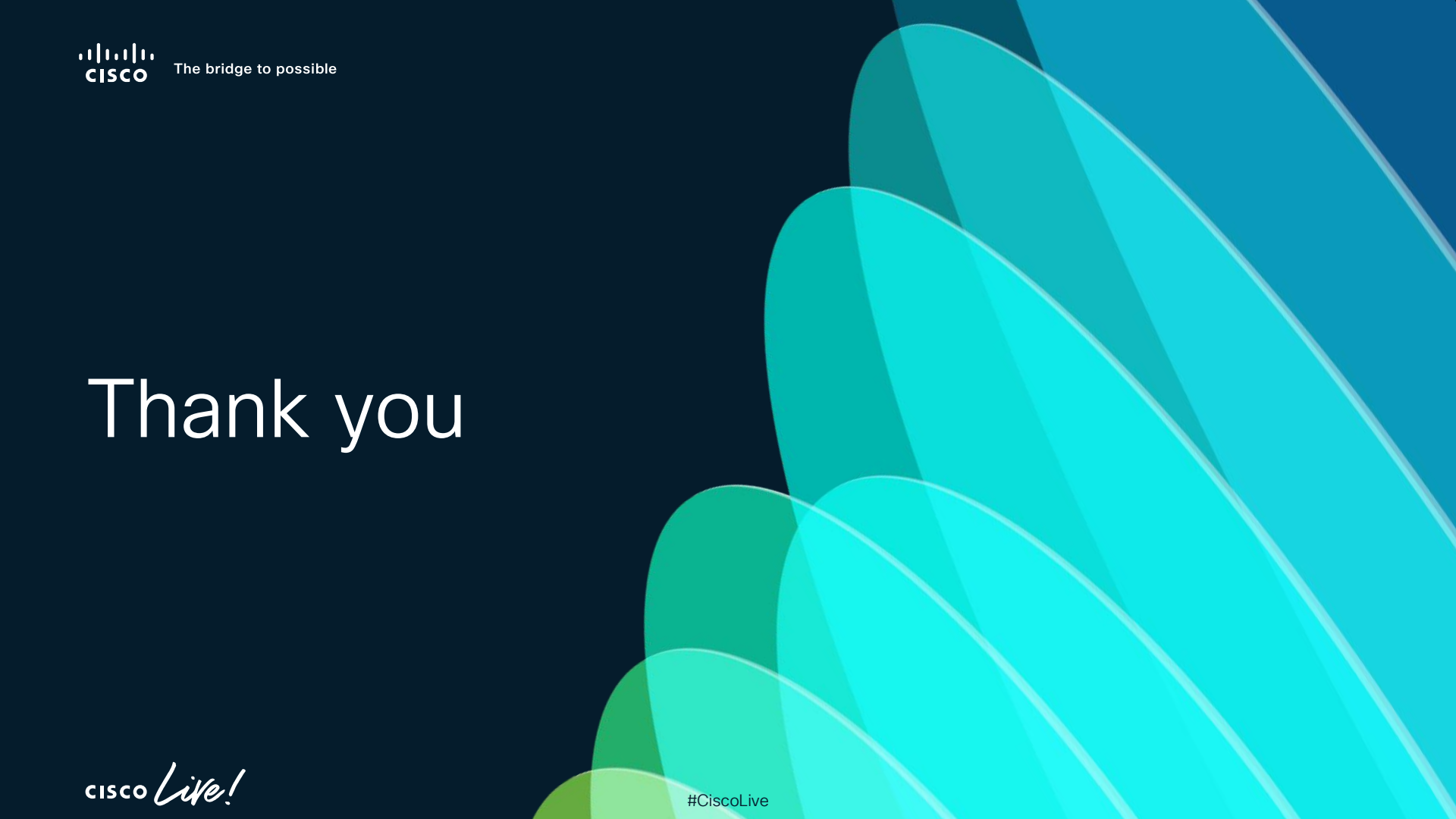
Recommended Sessions

Code	Title
BRKSPG-2029	Designing Routed Optical Networks: IP/MPLS Considerations
LTRMPL-1201	SR-MPLS 101: A Primer for Bigger Things to Come
LTRSPG-2003	IOS XR High Availability IP/MPLS Routing Lab
BRKOPT-1007	Optical Networking Fundamentals
BRKMPL-1203	SR-TE 101: Driving a New Traffic School of Thought
BRKMPL-1205	Efficiency in Every Packet: A Study in Segment Routing Automatic Self-Correction Capabilities
BRKMPL-2103	Mastering BGP: Deep Dive into Basics & Design Best Practices for BGP and L3VPN
BRKMPL-2129	SR IGP Flex-Algo
BRKMPL-2135	Preparing for a Successful Segment Routing Deployment
BRKMPL-2203	Introduction to SRv6 uSID Technology

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

Contact us at: djsmith@cisco.com and lokhanna@cisco.com


The bridge to possible
Thank you
cisco Live!
#CiscoLive