Kubernetes (K8s) Infrastructure Connectivity for the Modern Data Center

Camillo Rossi
Technical Marketing Engineer

Session ID: BRKDCN-2663

Cisco Webex App

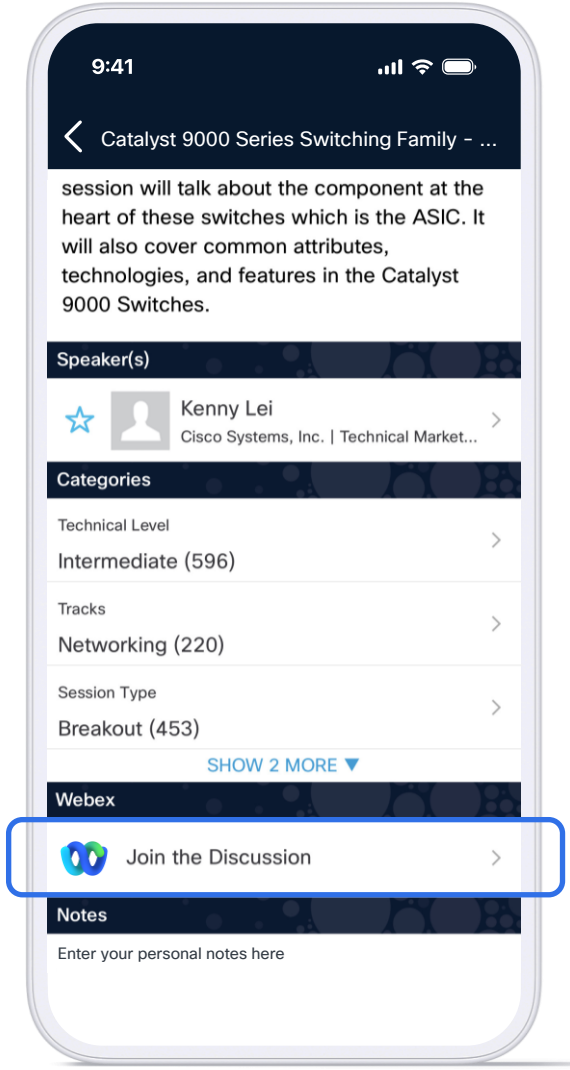
Questions?

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 14 November 2025.



https://ciscolive.ciscoevents.com/
ciscolivebot/**#BRKDCN-2663**

© 2025 Cisco and/or its affiliates. All rights reserved.

Agenda

01 Kubernetes Refresh

02 Network Challenges

03 BGP Based Architecture

04 Cilium

05 ACI Design

06 NX-OS Design

07 Demos

08 Q&A

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Kubernetes Refresh

What is Kubernetes?

Kubernetes (K8s) is an open-source container orchestration platform. It is designed to automate the deployment, scaling, and management of containerized applications. Originally developed by Google, Kubernetes is now maintained by the Cloud Native Computing Foundation (CNCF).

1. Container Orchestration:
   - Manages the lifecycle of containers across a cluster of machines.
   - Ensures containers are deployed, running, and healthy.
2. Self-Healing:
   - Automatically restarts containers that fail, replaces unhealthy nodes, and reschedules workloads as needed.
3. Scalability:
   - Supports horizontal scaling (adding/removing replicas of applications) based on traffic or resource demands.

4. Load Balancing:
   - Distributes traffic across multiple containers to ensure high availability and performance.
5. Declarative Configuration:
   - Users describe the desired state of the system (e.g., "I want 5 replicas of this app"), and Kubernetes ensures the actual state matches the desired state.
6. Portability:
   - Works across cloud providers, on-premises data centers, and hybrid environments.

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Key Kubernetes Concepts

- Pods:
  - Smallest deployable unit.
  - Encapsulates one or more containers with shared storage/network.
- Deployments:
  - Declarative updates for Pods.
  - Automatic Rollouts and Rollbacks
  - Scaling: Allows you to define how many instances of your app should run
- Services:
  - Expose your application to external/internal traffic.
  - Types: ClusterIP, NodePort, LoadBalancer, etc.
- ConfigMaps & Secrets:
  - Manage configuration data.
  - Secrets securely store sensitive information.
- Namespaces:
  - Logical isolation within a cluster.
  - Useful for separating environments (e.g., dev, p

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Network Challenges

Network Challenges – In General

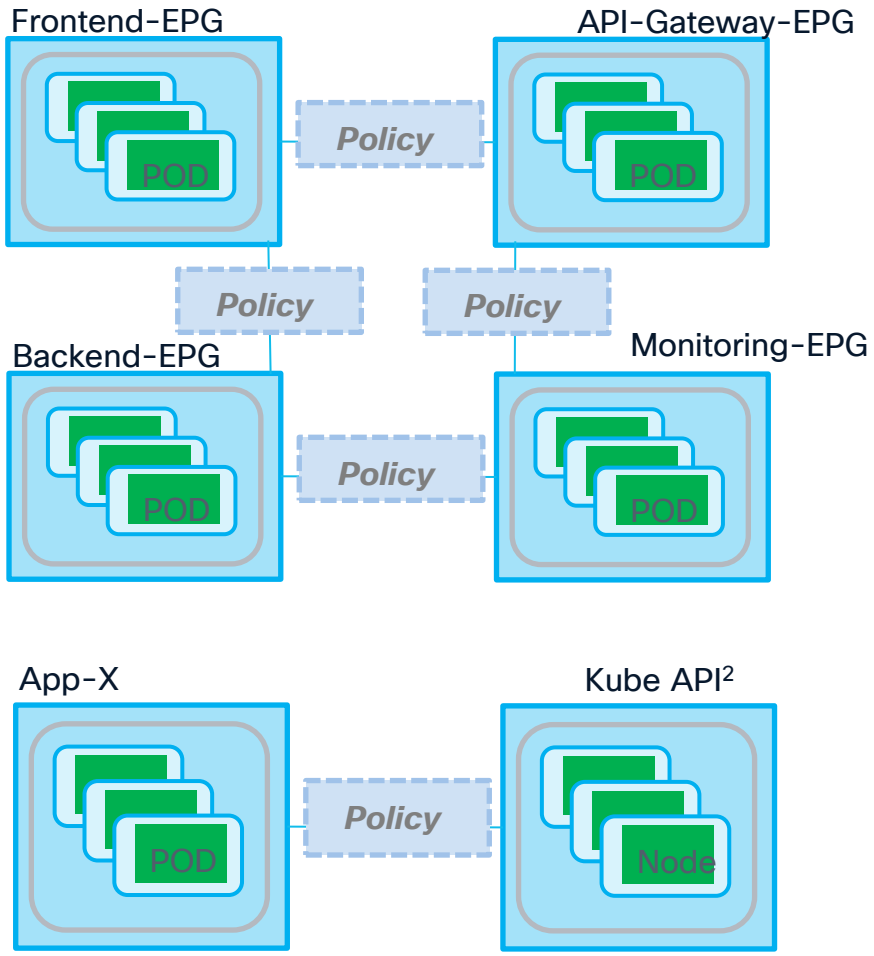
- Kubernetes networking involves different layers and components, such as pods, services, ingress, network policies, and more. Each component has its own networking requirements, making it complex to manage and understand.
- Kubernetes itself does not provide built-in networking functionality. Instead, it relies on the CNI specification to allow third-party plugins to handle the networking. This approach makes Kubernetes modular and flexible, enabling users to choose the best networking solution for their specific needs however has also led to the proliferation of several CNI Plugins all with different capabilities.
- Navigating the maze of Kubernetes CNI options can be challenging, as each plugin offers unique features and trade-offs that must align with your specific networking and security needs.

© 2025 Cisco and/or its affiliates. All rights reserved.

Different Traffic Flows: Different Needs

East – West

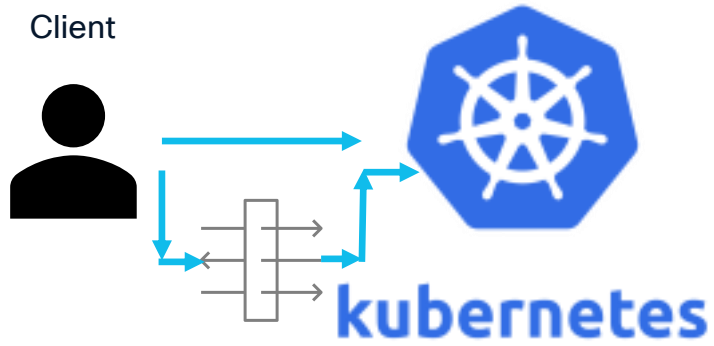
- These are communication flows inside the cluster:
  - Pod-to-Pod
  - Pod-to-Node
- This traffic is automatically handled by the CNI plugin of choice
- Kubernetes Network Policies can be used to secure the traffic in a Cloud native way[1]
- **Main Challenges:**
  - **Visibility:** The flows are invisible to the fabric due to overlays making troubleshooting hard especially if Kubernetes Network Policies are used
- **Solutions:**
  - Use a CNI that provides in depth visibility like Isovalent and Hubble
  - Remove the Overlay to allow the fabric to see the traffic flows



© 2025 Cisco and/or its affiliates. All rights reserved.

North -> South (ingress)

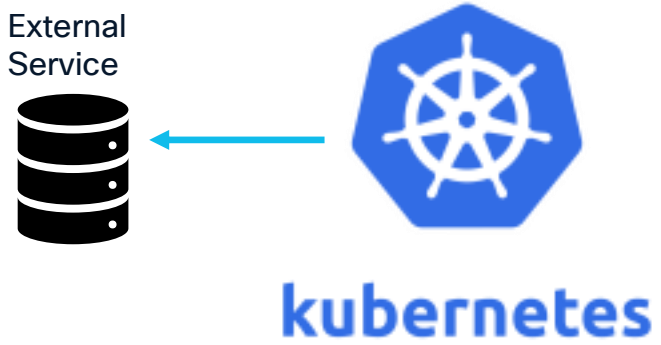
- These communication flows originate from entities outside the cluster that consume its services.
- Exposing services to external users or systems is complex
- **Main Challenges:**
  - **Service Advertisement:** How are the Services Advertised to the network?
  - **LoadBalancing :** How is the traffic Load Balanced effectively between the K8s Nodes hosting the PODs?
  - **Security:** How do we insert a Perimeter Firewall?[1]
- **Solutions:**
  - Running eBGP between the Kubernetes
  - Service insertion can be employed to seamlessly insert a perimeter firewall between the client and the Kubernetes cluster.



© 2025 Cisco and/or its affiliates. All rights reserved.

South -> North (egress)

- These communication flows are initiated by entities within the cluster (Pods) that consume external services. (Pod-Initiated Traffic)
- By default, POD initiated traffic is NATed to the Node IP
- In most cases using only K8s Network policies for security is not sufficient
- **Main Challenges:**
  - **Security:** Securing pod-initiated traffic to an external systems is difficult:
    - If the Pod IP is NATed to the Node IP, the best identifiable source becomes "the cluster" as individual Pod-level identity is lost in the process.
    - If the Pod IP is routable (advertised to the fabric), its ephemeral nature poses significant challenges for configuring external firewalls.
    - How do we insert a Perimeter Firewall?
- **Solutions:**
  - **Isovalent Egress Gateway**[1]:
    - Allows to NAT POD traffic leaving the cluster with predictable IPs associated with the gateway nodes.
    - Flexibility to select egress IP based on labels, or destination CIDRs
    - Ability to Advertise the Egress IP over eBGP



© 2025 Cisco and/or its affiliates. All rights reserved.

BGP Based Architecture

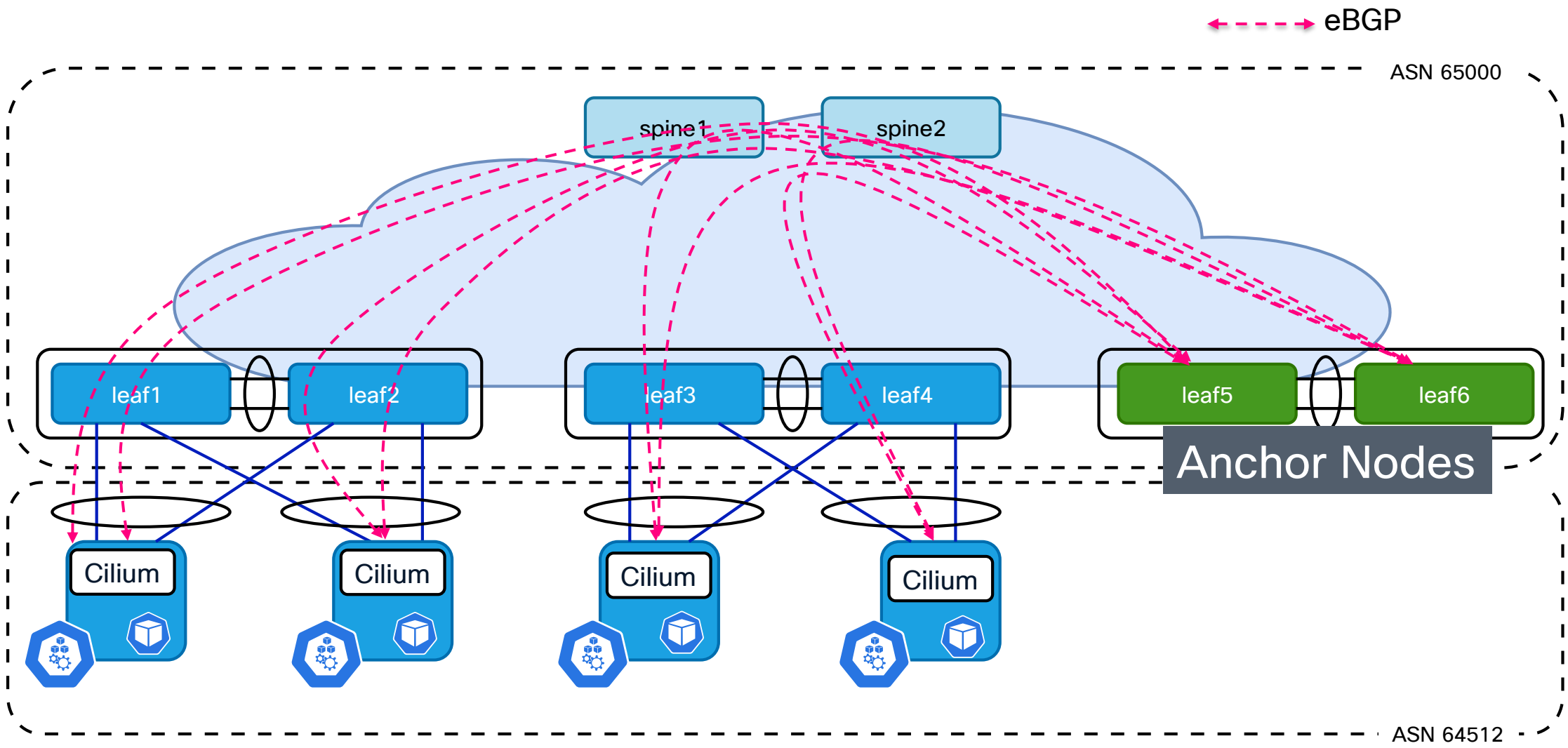
BGP Based Integration

- Identical Design for ACI and NX-OS VXLAN EVPN
- Leverages a Proven Protocol: Built on BGP for reliable and scalable routing.
- Unified L3 Networking:
  - Seamless Layer 3 connectivity for Kubernetes Nodes, Pods, and Service endpoints.
  - Enables effortless communication both within and outside the fabric.
- Enhanced Security (ACI/NX-OS):
  - Security group classification to secure traffic across Node, Pod, and Service subnets.
  - Service Insertion for external Firewall via Service Graphs or PBR Redirection
- Optimized Performance:
  - High-performance, low-latency connectivity without the overhead of overlays.
- Scalable Load Balancing:
  - Hardware-assisted ECMP (Equal-Cost Multi-Path) routing, ensuring efficient traffic distribution and scalability.
- Flexible Deployment Options:
  - Compatible with Any Hypervisor or Bare Metal, enabling a mix of deployment form factors for maximum flexibility.

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Centralized Routing Peering



© 2025 Cisco and/or its affiliates. All rights reserved.

BGP Tuning

- BGP Dynamic Neighbours
- BGP Graceful Restart
- BFD (Requires Isovalent Enterprise)
- Increase Max BGP ECMP paths for better load balancing

© 2025 Cisco and/or its affiliates. All rights reserved.

BGP Hardening

- Enabled BGP password authentication
- Set the maximum AS limit to one
- Configure BGP import route control to accept only the expected subnets from the Kubernetes cluster:
  - Pod subnet(s)
  - Node subnet(s)
  - Service subnet(s)
- Set a limit on the number of received prefixes from the nodes.

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Expected Routing Behavior

- Nodes service IPs (and optionally pods)Subnets will be advertised to ACI fabric
- Exposed Services will be advertised as host routes from every nodes that has a running POD associated to the service.
- Optimal Routing in place with ACI 6.1(3), NX-OS 10.5 or later

© 2025 Cisco and/or its affiliates. All rights reserved.

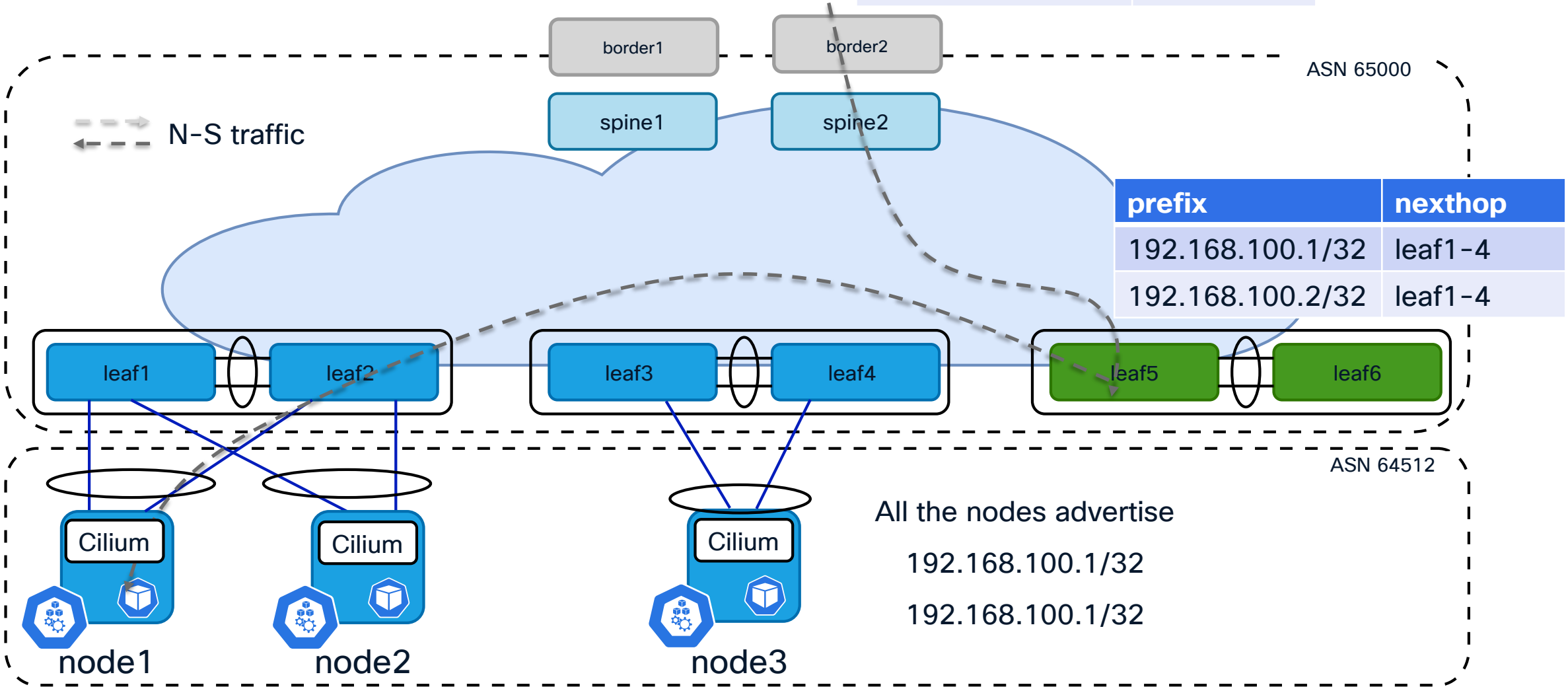
cisco

Traffic Flow Optimizations

Centralized Routing Peering

Problem: Asymmetric and Unbalanced traffic

prefix	nexthop
192.168.100.1/32	leaf5,leaf6
192.168.100.2/32	leaf5,leaf6



prefix	nexthop
192.168.100.1/32	leaf1-4
192.168.100.2/32	leaf1-4

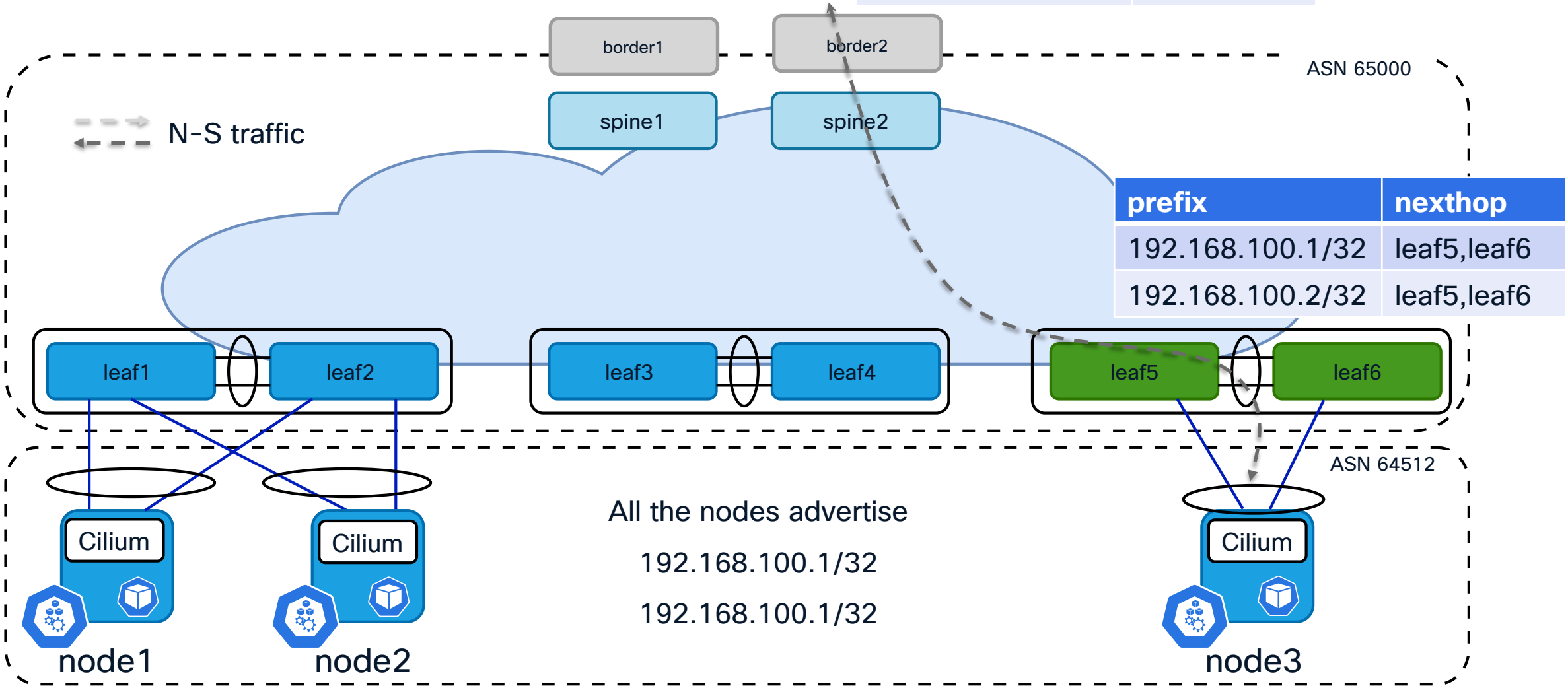
© 2025 Cisco and/or its affiliates. All rights reserved.

Centralized Routing Peering

Problem: Only the nodes connected to Anchor Nodes are used

prefix	nexthop
192.168.100.1/32	leaf5,leaf6
192.168.100.2/32	leaf5,leaf6

ACI ONLY



prefix	nexthop
192.168.100.1/32	leaf5,leaf6
192.168.100.2/32	leaf5,leaf6

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Centralized Routing Peering

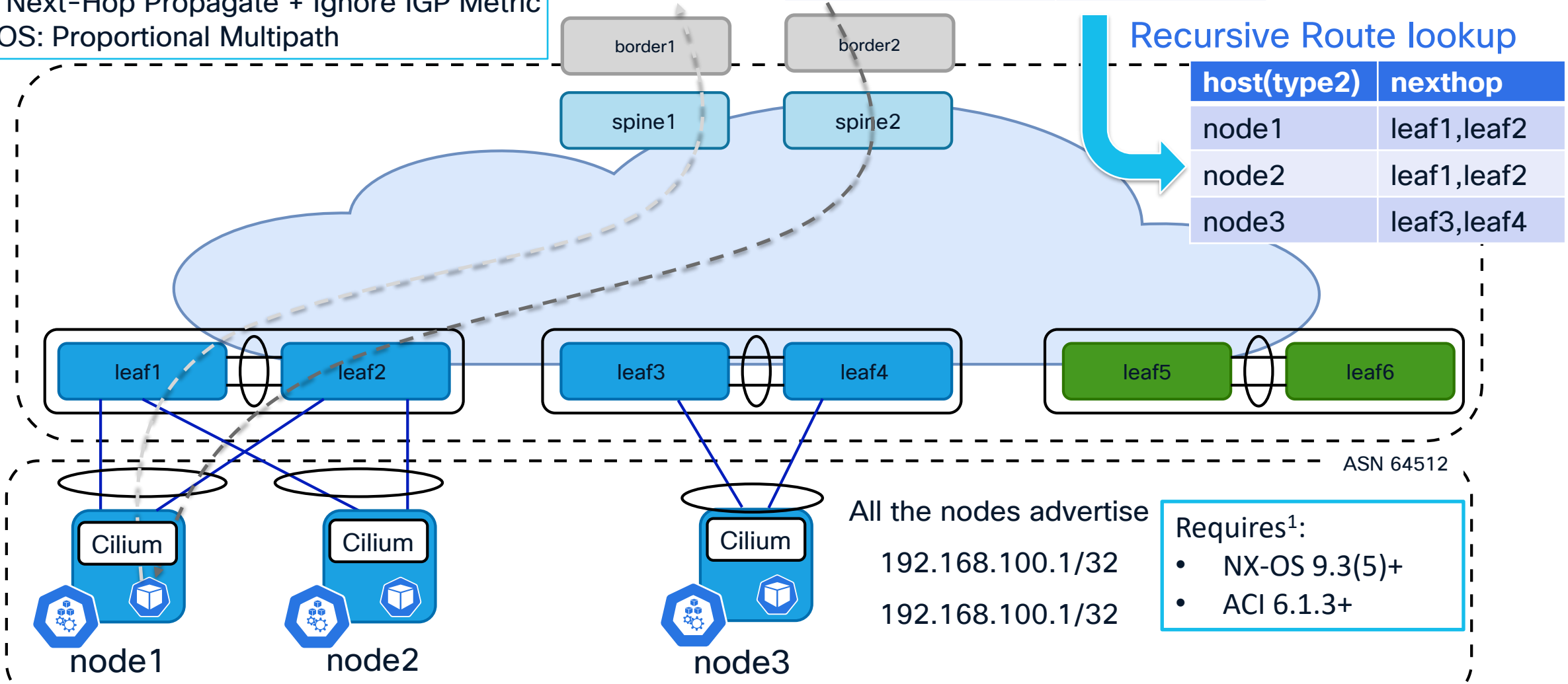
Solution

ACI: Next-Hop Propagate + Ignore IGP Metric
NX-OS: Proportional Multipath

prefix	nexthop
192.168.100.1/32	node1-3
192.168.100.2/32	node1-3

Recursive Route lookup

host(type2)	nexthop
node1	leaf1,leaf2
node2	leaf1,leaf2
node3	leaf3,leaf4



All the nodes advertise

192.168.100.1/32

192.168.100.1/32

Requires[1]:

- NX-OS 9.3(5)+
- ACI 6.1.3+

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Isovalent Networking for Kubernetes (Cilium)

CISCO Live !

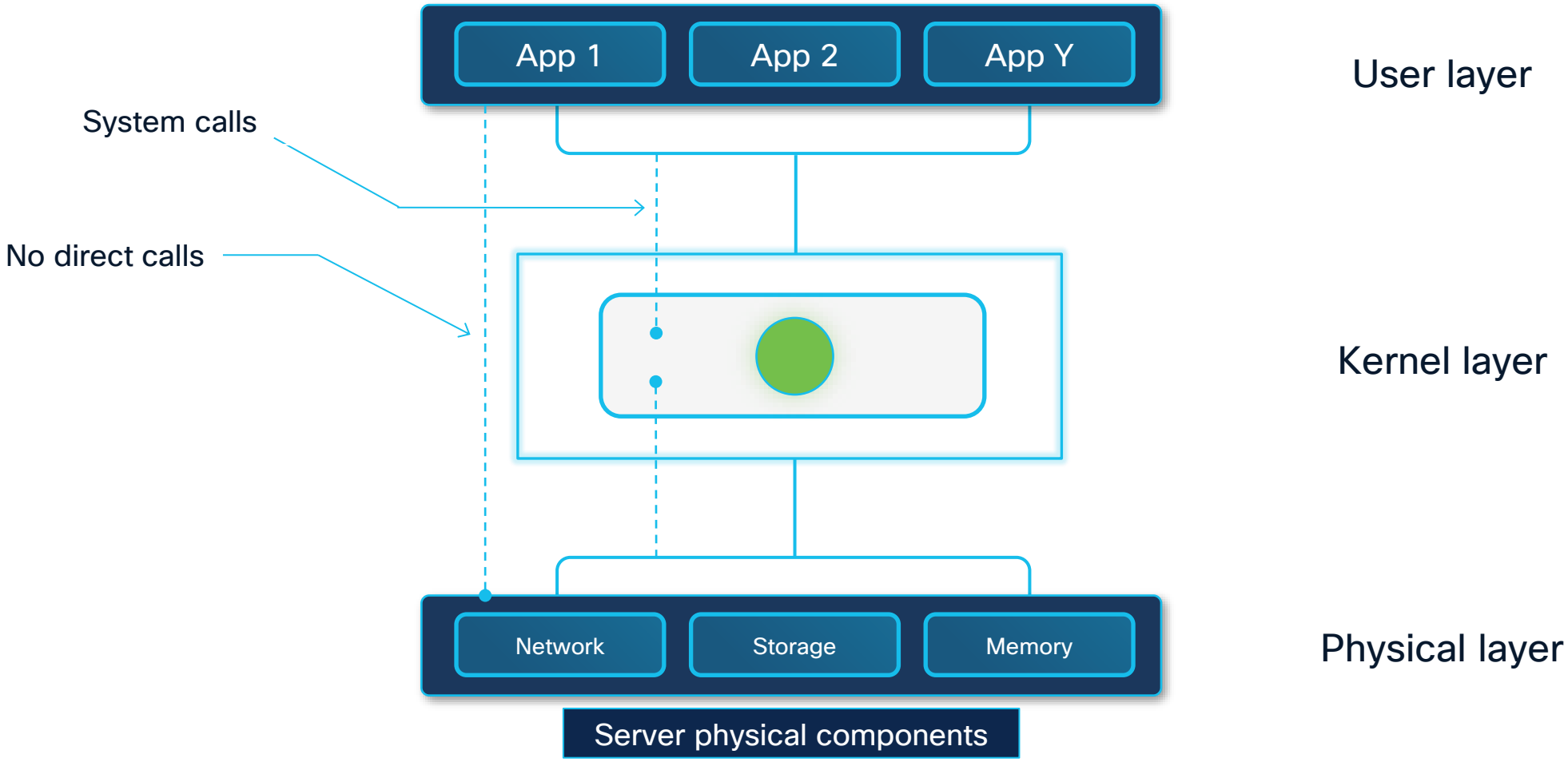
What is it?

- Cilium is an open source, cloud native solution for providing, securing, and observing network connectivity between workloads.
- Cilium was created by Isovalent (joined Cisco via acquisition in April 2024.) and is currently the only CNCF graduated networking plugin project.
- The enterprise version is called **Isovalent Networking for Kubernetes**
- **Advanced Network Security**: fine-grained security policies that can be defined based on application-layer identities rather than just IP addresses. This allows for more precise control over communication between microservices.
- **Scalability and Performance**: By using eBPF it achieves high performance and scalability, as it can run network policies directly in the Linux kernel without the need for complex user-space processing.
- **Service Mesh Integration**: integrated with service mesh architectures to provide transparent, high-performance networking and security features that complement service mesh capabilities.
- **Observability**: Robust observability tools, allowing users to monitor and troubleshoot network traffic and security policies with detailed metrics and visibility into application communication patterns.
- **Kubernetes Native**: Designed with Kubernetes in mind seamlessly integrates with Kubernetes to provide native support for network policies and service discovery

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

How do operating systems work?



© 2025 Cisco and/or its affiliates. All rights reserved.

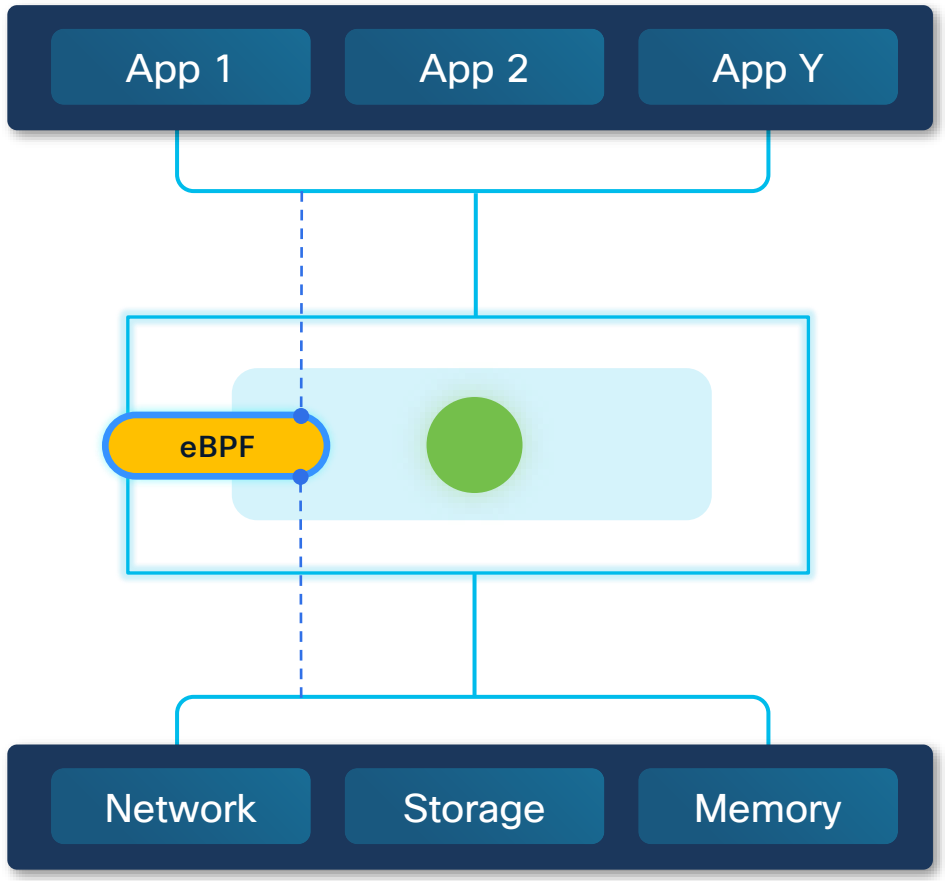
eBPF – Foundation of Cilium

Network filtering

Observability

Security policy

What Javascript is
to the browser,
eBPF is to the kernel.



User layer

Kernel layer

Physical layer

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

eBPF – Foundation of Cilium

- Kubernetes networking
- Load balancing
- Kubernetes services
- Identity-based security
- L7 policies

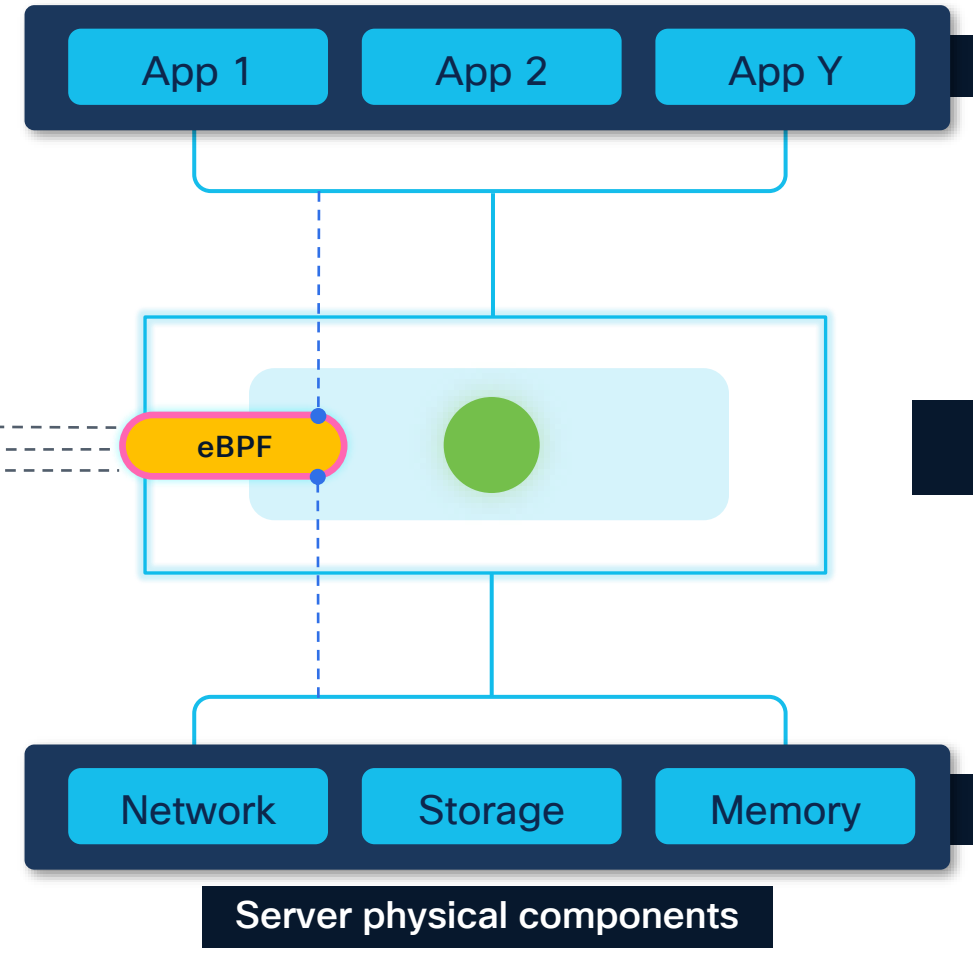
Network filtering

- Dependencies map (service and flows)
- Monitoring and alerting
- App monitoring

Observability

Security policy

- Monitor process execution
- Runtime security policies
- Real time enforcement



© 2025 Cisco and/or its affiliates. All rights reserved.

How to BGP with Cilium?

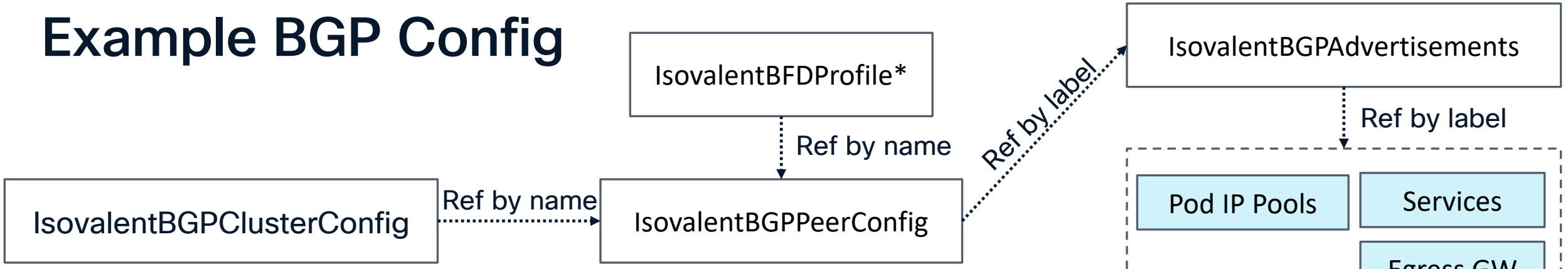
We only need 4 configs

- **IsovalentBGPClusterConfig**: Defines BGP instances and peer configurations that are applied to multiple nodes.
- **IsovalentBGPPeerConfig**: A common set of BGP peering setting. It can be used across multiple peers.
- **IsovalentBGPAdvertisements**: Defines prefixes that are injected into the BGP routing table.
- **IsovalentBFDProfile**: Defines the BFD parameters

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Example BGP Config




```yaml
apiVersion: isovalent.com/v1alpha1
kind: IsovalentBGPClusterConfig
metadata:
  name: cilium-bgp
spec:
  bgpInstances:
  - name: "instance-65113"
  localASN: 65113
  peers:
  - name: "leaf201"
    peerASN: 65002
    peerAddress: 192.168.2.43
    peerConfigRef:
      name: "cilium-peer"
```

```yaml
apiVersion: isovalent.com/v1alpha1
kind: IsovalentBGPPeerConfig
metadata:
  name: cilium-peer
  spec:
    ebgpMultihop: 1
    gracefulRestart:
      enabled: true
   families:
     - afi: ipv4
       safi: unicast
       advertisements:
       matchLabels:
         advertise: "bgp"
  bfdProfileRef: tor-bfd-profile
```

```yaml
apiVersion: isovalent.com/v1alpha1
kind: IsovalentBGPAdvertisement
metadata:
  name: bgp-advertisements
  labels:
    advertise: bgp
  spec:
    advertisements:
    - advertisementType: "Service"
      service:
        addresses:
          - LoadBalancerIP
        selector:
          matchLabels:
            advertise: bgp
```

© 2025 Cisco and/or its affiliates. All rights reserved.

Addressing the challenges

East – West Visibility

Hubble

- Hubble is a fully distributed networking and security observability platform. It is built on top of Cilium and eBPF to enable deep visibility into the communication and behaviour of services as well as the networking infrastructure in a completely transparent manner.
- Service dependencies & communication map
- Network monitoring & alerting
- Application monitoring
- Security observability

© 2025 Cisco and/or its affiliates. All rights reserved.

Demo

North -> South

Running BGP!

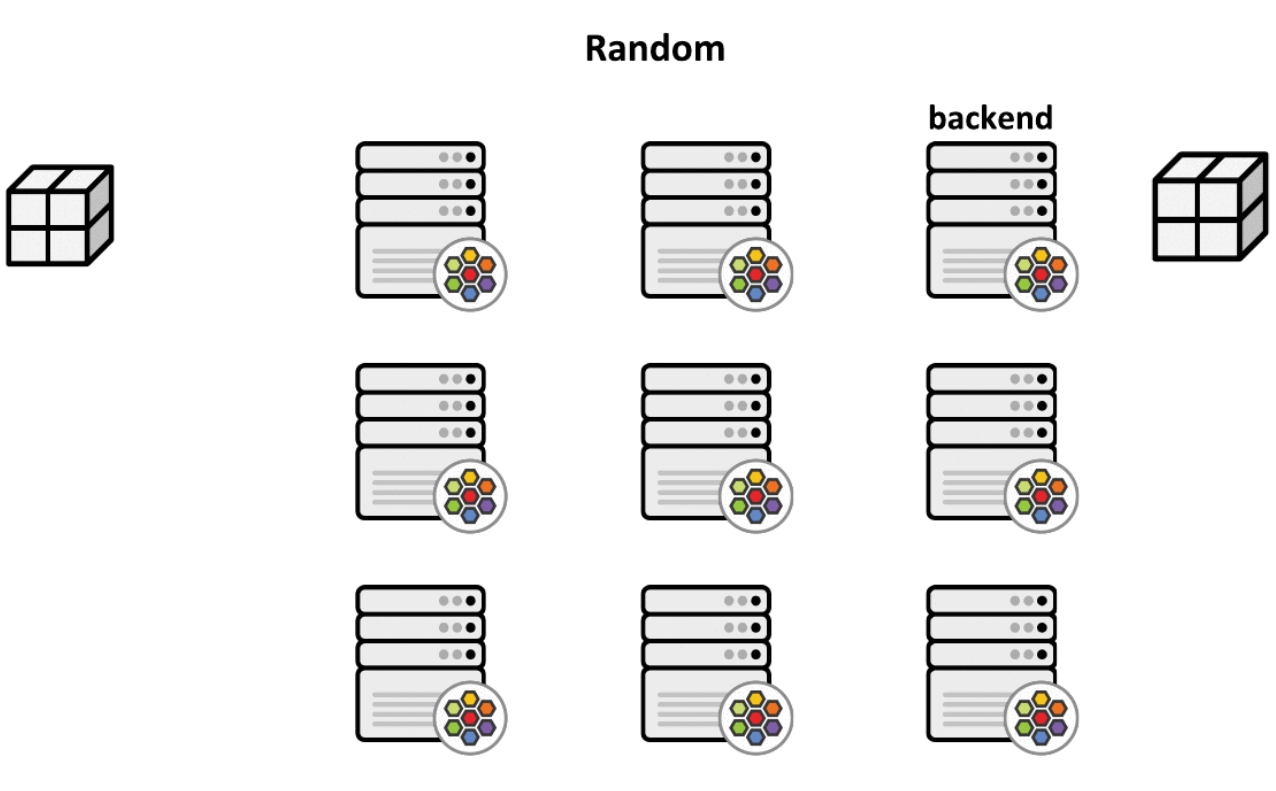
- By running BGP we are addressing all the challenges we had:
  - The services are advertised, load balanced with ECMP and inserting a firewall become a trivial task.
- Bonus: Use ACI ExtEPG or NXOS GPO for Fabric Security
- **However, there is one issue we didn’t discuss yet…**

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

ECMP Resilient Hashing or the lack of it...

- ECMP resilient hashing is not available on all platforms, and this can result in TCP flows reset.
- Maglev consistent hashing minimizes disruptions by ensuring all load balancing nodes share a consistent backend lookup table. Using the packet's 5-tuple hash, traffic is reliably forwarded to the same backend without state synchronization. This enhances resiliency during failures and ensures effective load balancing when new nodes are added.



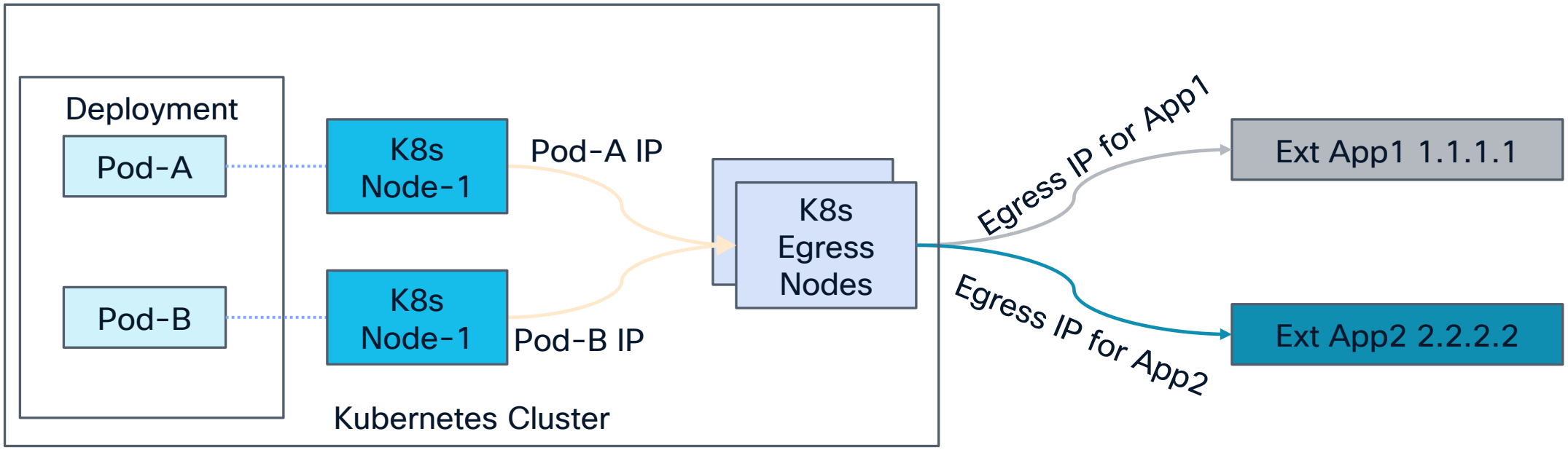
© 2025 Cisco and/or its affiliates. All rights reserved.

Demo

South -> North

Isovalent Egress Gateway

- The Egress Gateway features allows for redirecting traffic originating from pods and destined to specific CIDRs outside the cluster to be routed through particular nodes.
- When the egress gateway policies are in place, packets leaving the cluster are masqueraded with selected, predictable IPs (egressIP) associated with the gateway nodes. This allows for network administrators to apply network controls to pods and or namesapces establishing outbound connectivity.



© 2025 Cisco and/or its affiliates. All rights reserved.

Demo

Conclusion

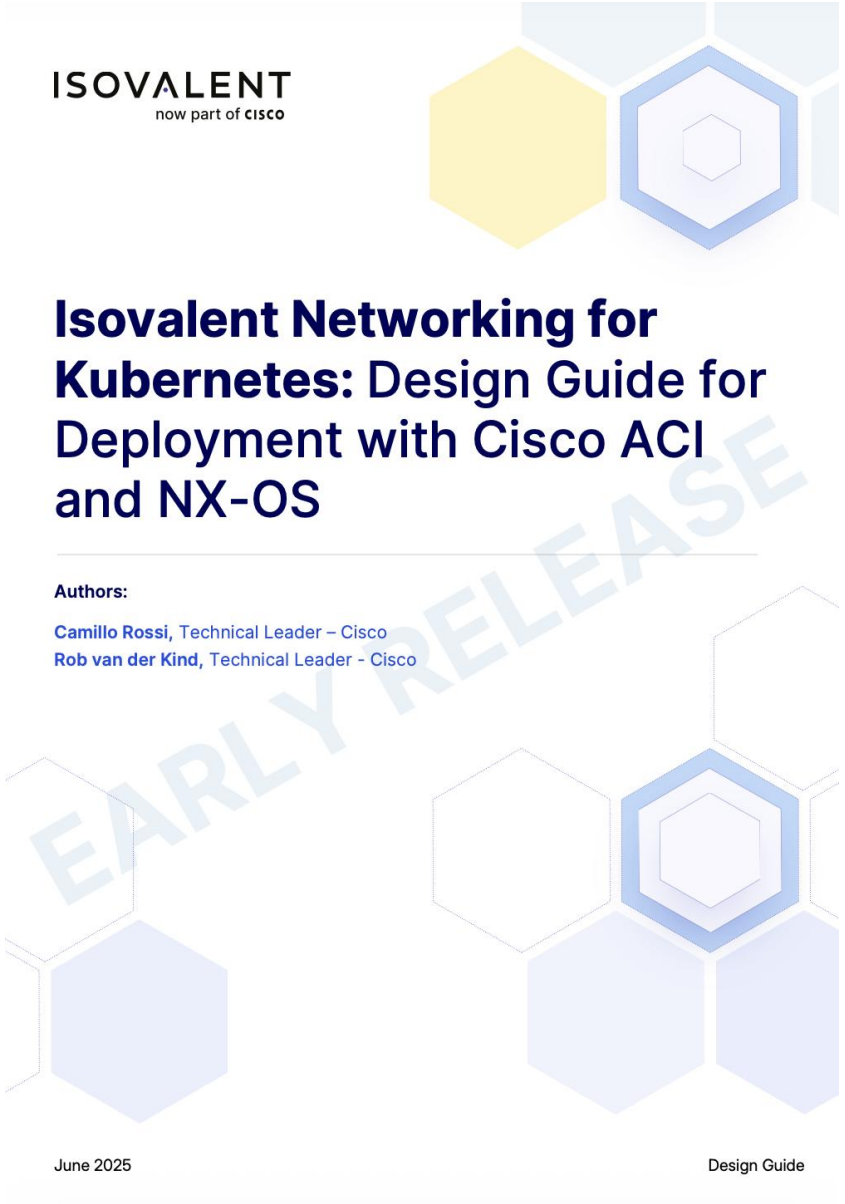
Isovalent and Cisco DC Fabric – Better Together

- Easy and Scalable:
  - BGP is a well establish protocol easy to configure and troubleshoot
  - High Scale with thousand nodes per cluster
- Visibility:
  - Every communication flow can be monitored with Hubble
  - Traffic Traversing the network can be monitored with conventional tools
- Security:
  - Could Native Security enforced by Isovalent using eBPF
  - Network level security (Node to Node, N/S or S/N) leveraging fabric contracts or service insertion
- NetOps Ready:
  - Kubernetes ACI and NX-OS can be driven by API, ansible or terraform ensuring automation is a breeze

© 2025 Cisco and/or its affiliates. All rights reserved.

Get the design guide!

- Free Download at: https://isovalent.com/briefs/designing-isovalent-enterprise-for-cisco-aci-and-nexus/



© 2025 Cisco and/or its affiliates. All rights reserved.

Complete your session evaluations

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

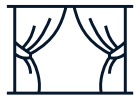
Earn up to 800 points by completing all surveys and climb the Cisco Live Challenge leaderboard.

Level up and earn exclusive prizes!

Complete your surveys in the Cisco Live Events app.

© 2025 Cisco and/or its affiliates. All rights reserved.

Continue your education



Visit the Cisco Stand for related demos

Book your one-on-one Meet the Expert meeting

Attend the interactive education with Capture the Flag, and Walk-in Labs

Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

© 2025 Cisco and/or its affiliates. All rights reserved.

Thank you

CISCO