Tungsten Fabric & Akraino

SDN and NVF for Edge

Sukhdev Kapur, Distinguished Engineer
Juniper Networks

Beijing June 20th 2019

Agenda

- Tungsten Fabric Community
- Tungsten Fabric Architectural Overview
- Remote Compute (Edge Compute)
- Akraino Community
- Tungsten Fabric + Akraino
- Q & A

Tungsten Fabric Mission

+

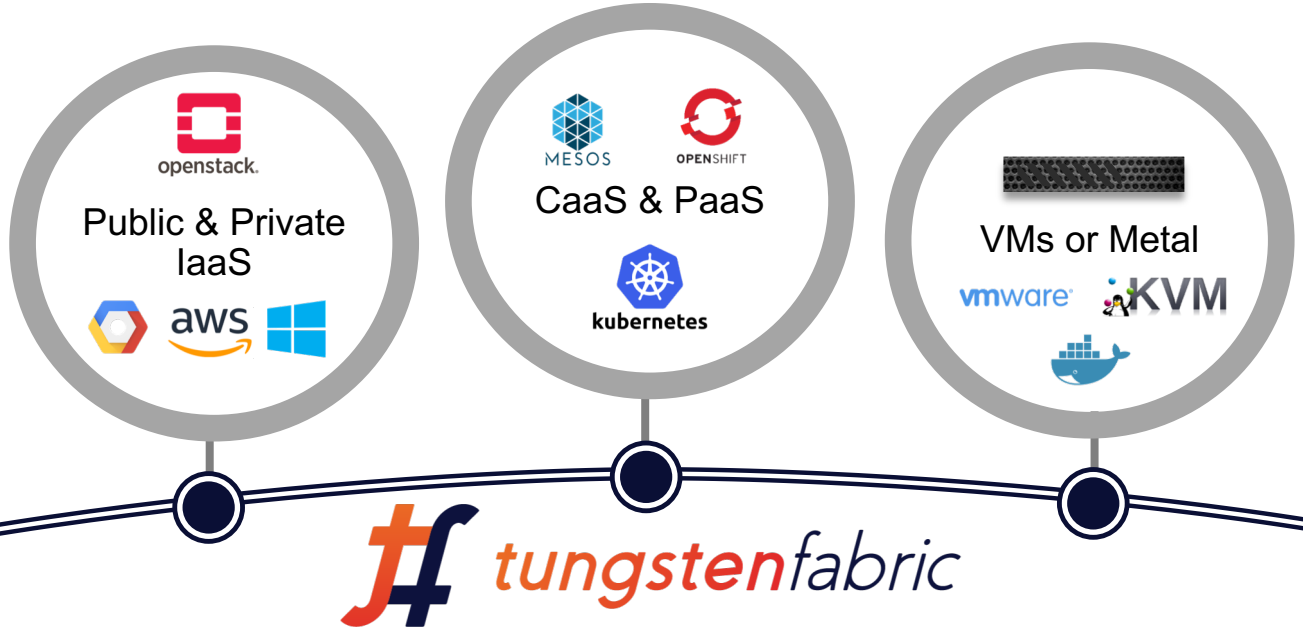
Build the world's most ubiquitous, easy-to-use, scalable, secure, and cloud-grade SDN stack, providing a network fabric connecting all environments, all clouds, all people.

https://tungsten.io/

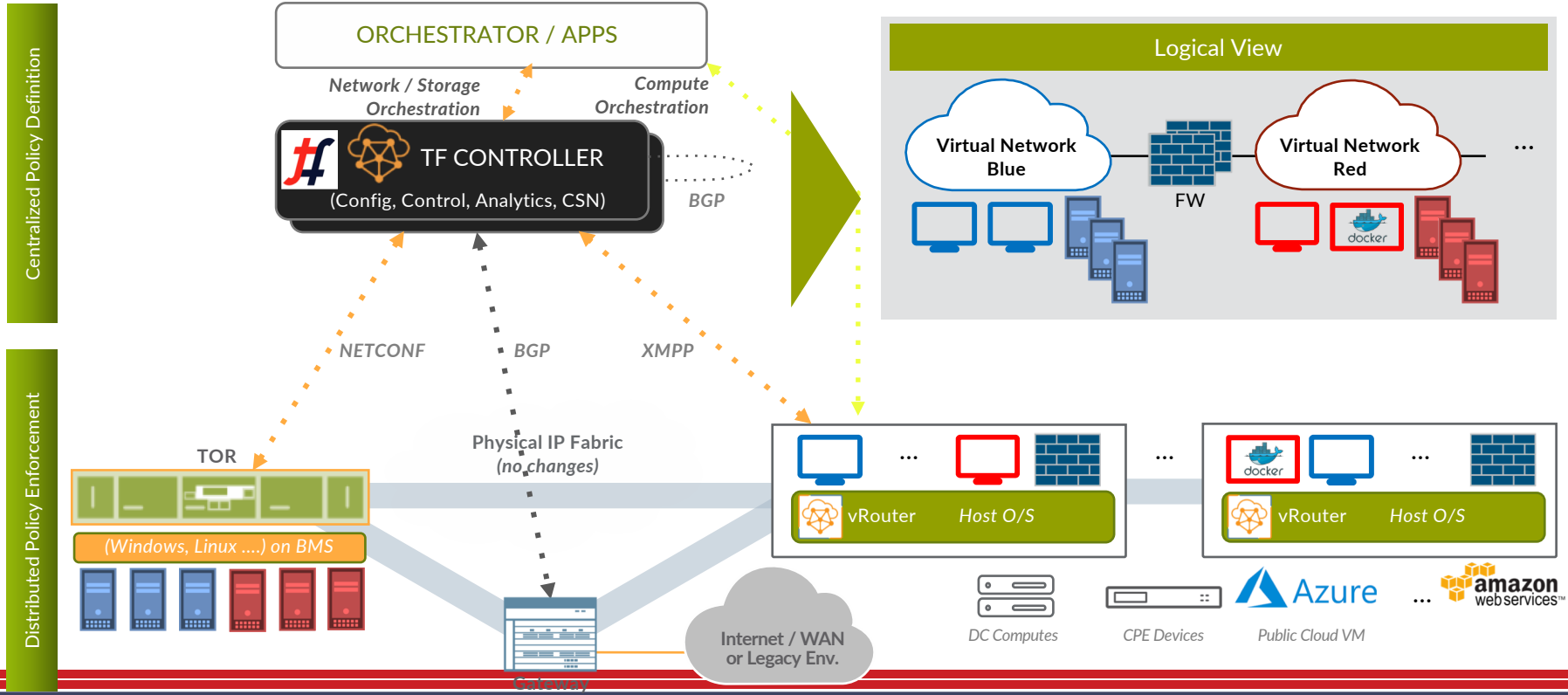
Tungsten Fabric as SDN Controller

RULE THEM ALL WITH ONE

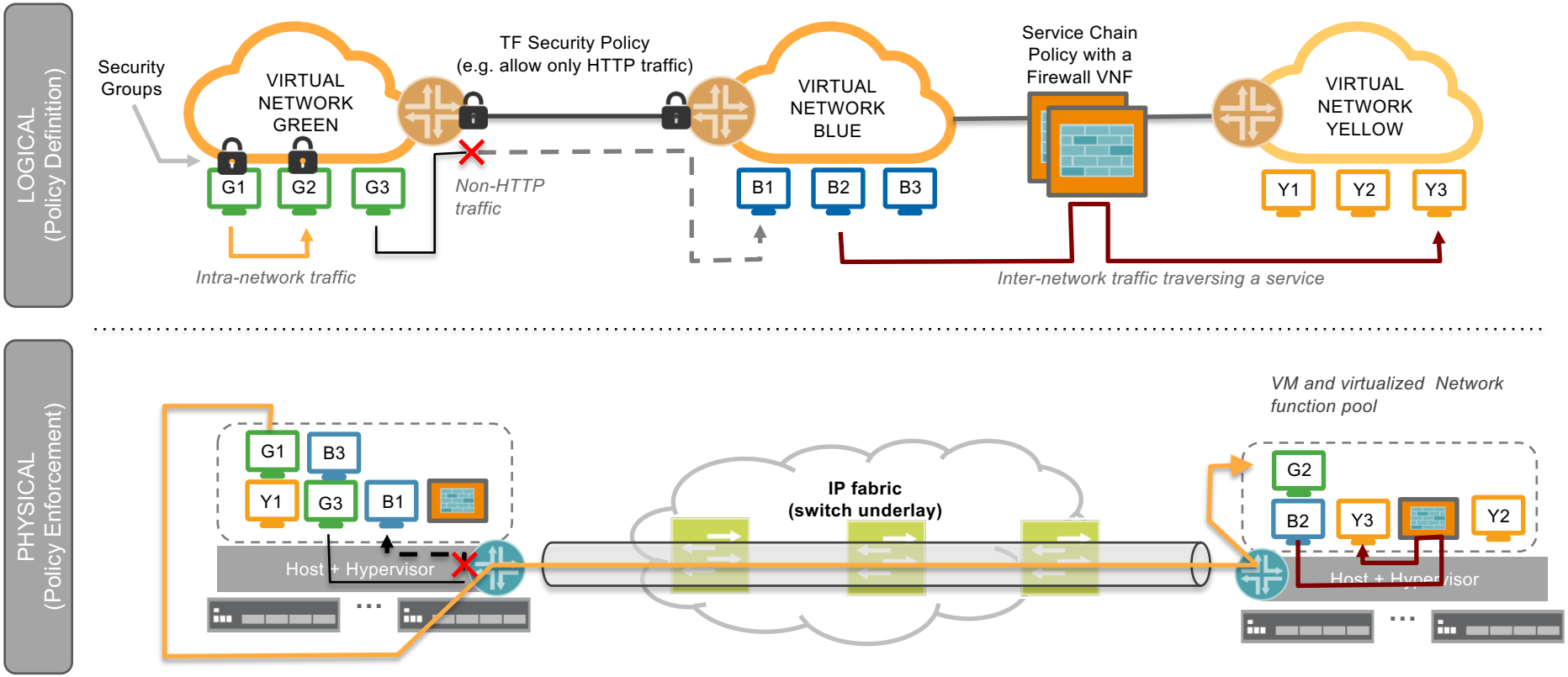
automated secure open SDN Controller



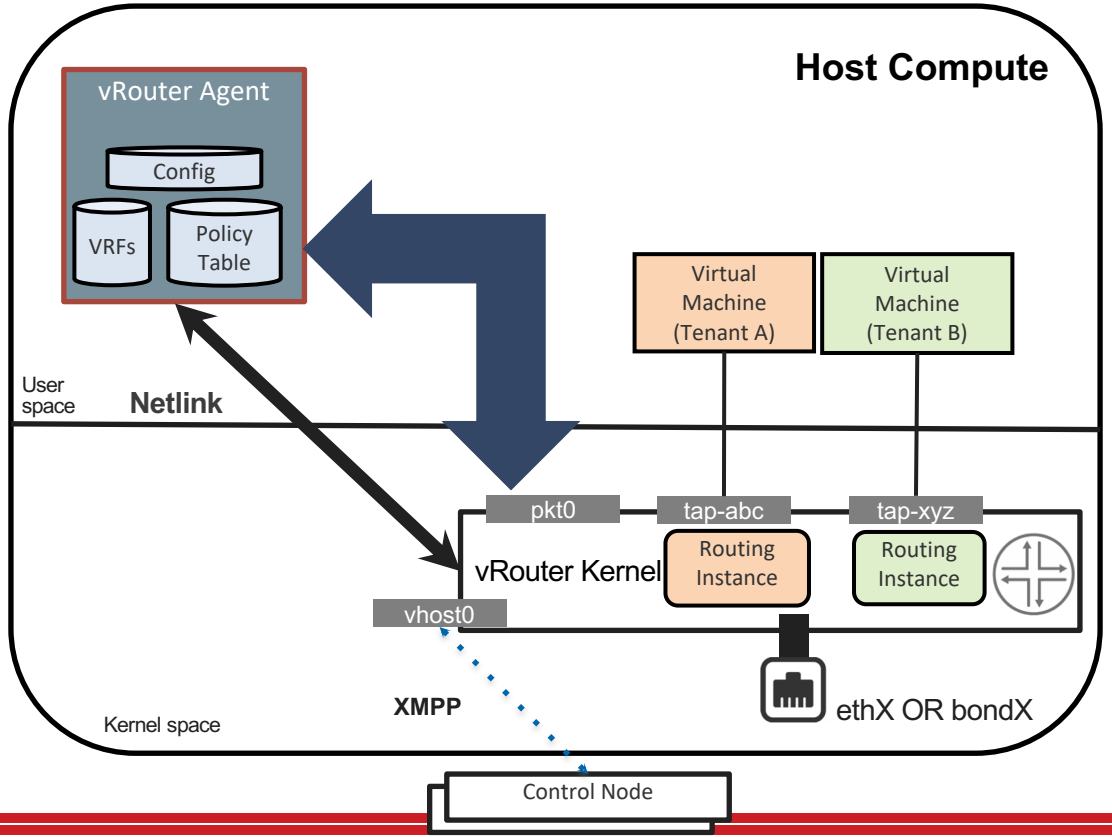
Architecture Overview



Visualizing Tungsten Fabric's Operational Effects



Tungsten Fabric vRouter Architecture & Overview



vRouter Agent

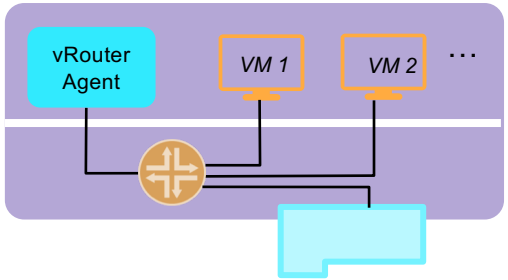
- Exchanging control state such as routes with the Control nodes using XMPP.
- Receiving low-level configuration state such as routing instances and forwarding policy from the Control nodes using XMPP
- Reporting analytics state such as logs, statistics, and events to the analytics nodes.
- Installing forwarding state into the forwarding plane
- Discovering the existence and attributes of VMs in cooperation with the Nova agent.
- Applying forwarding policy for the first packet of each new flow and installing a flow entry in the flow table of the forwarding plane.
- Proxying DHCP, ARP, DNS

vRouter Kernel/DPDK

- Encapsulating packets sent from the overlay network and de-capsulating packets received for the overlay network.
- Packets received from the overlay network are assigned to a routing instance based on the MPLS label or Virtual Network Identifier (VNI).
- Doing a lookup of the destination address of the in the Forwarding Information Base (FIB) and forwarding the packet to the correct destination. The routes may be layer-3 IP prefixes or layer-2 MAC addresses.
- Doing RPF check before sending Virtual machine traffic to destination. This is configurable.

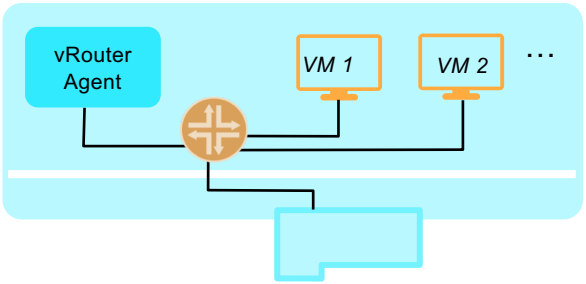
vRouter Deployment Models

KERNEL VROUTER



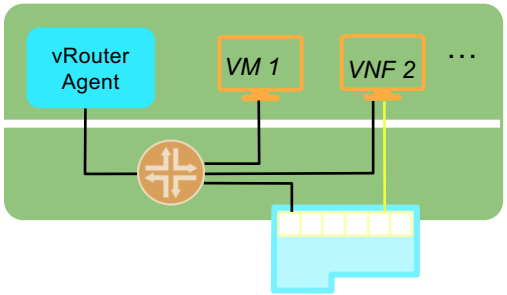
- *This the normal operation where fwding plane of vRouter runs in the kernel and are connected to VMs using TAP interface (or veth pair for containers)*
- *vRouter itself is enhanced using other performance related features:*
  - *TSO / LRO*
  - *Multi-Q Virtio*

DPDK VROUTER



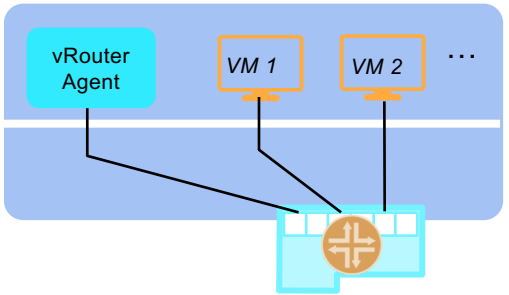
- *vRouter runs as a user space process and uses DPDK for fast path Packet I/O.*
- *Full set of SDN Capabilities Supported*
- *Requires the VMs to have DPDK enabled for performance benefits*

SRIOV/ VROUTER COEXISTENCE



- *Some workloads can directly SRIOV into the NIC, while others go through the vRouter*
- *Sometimes a VNF can have multiple interfaces some of which are SRIOV-ed to the NIC*
- *Interfaces that are SRIOV-ed into NIC don't get the benefits / features of vRouter*

SMARTNIC VROUTER



- *vRouter fwding plane runs within the NIC*
- *Workloads are SRIOV-connected to the NIC*

Tungsten Fabric Kubernetes Support

Tungsten Fabric Integration with k8s

** Contrail-Kube-manager listens to K8s API Server and conveys the API request to Contrail Controller*

kubectl
(user commands)

Namespace: kube-system

API Server | Controller/Replication Manager | etcd

Scheduler | Discovery | Dashboard | …

Contrail-kube-mgr

Contrail Controller

Contrail Analytics

POD 1

POD 2

POD 3

POD 4

Kubelet

CNI Plugin

vRouter
(replaces kube-proxy)

Compute Node-01

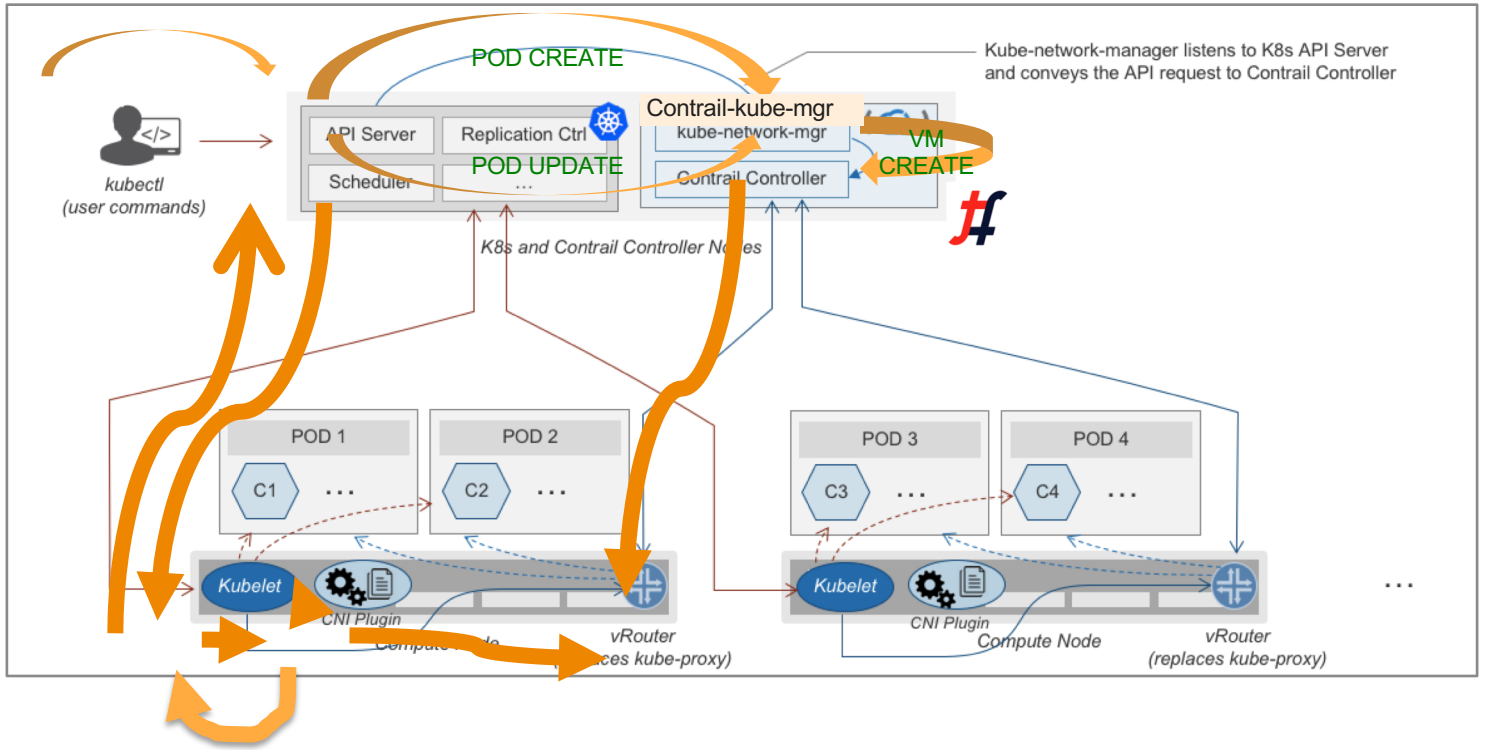
Kubelet

CNI Plugin

vRouter
(replaces kube-proxy)

Compute Node-02

Pod Creation



Tungsten Fabric Evolution to Microservices

- Contrail-Control (5 daemons)
- Contrail-Config (8 daemons)
- Contrail-Analytics (5 daemons)
- Contrail-WebUI (4 daemons)
- Contrail-DB (3 daemons)
- Contrail-vRouter (3 D) + Kernel/DPDK (FP)

Contrail Controller: 2n+1

OR

BMS

Contrail 1.X/2.X/3.X BMS or VMs base (SDN Controller)

Multiple Process running in one Container (FAT Containers)

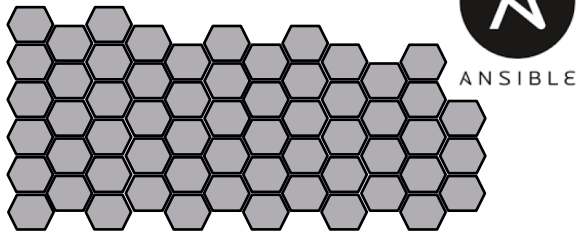
Analytics DB | Config + Control | Analytics

Kube MGR | HA Proxy | vRouter Agent

Contrail 4.X (Containers) BMS or VMs base (SDN Controller)

DaemonSet, Ingress Services with Host Networking
with choice of run single or multiple containers per PODs

27-30 Containers Images

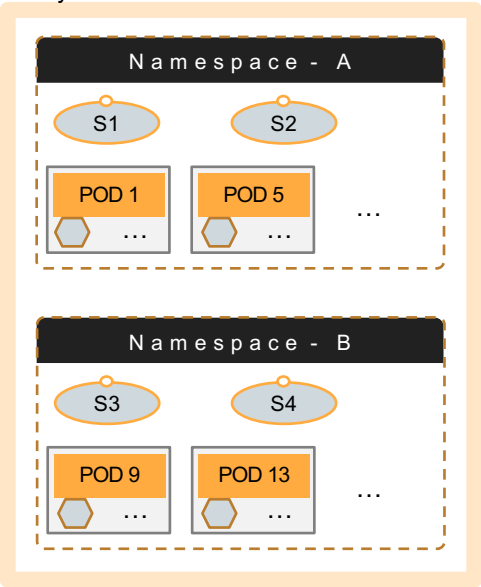


Contrail 5.X (Containers) Microservices (SDN Controller)

Levels of Isolation - Multitenancy

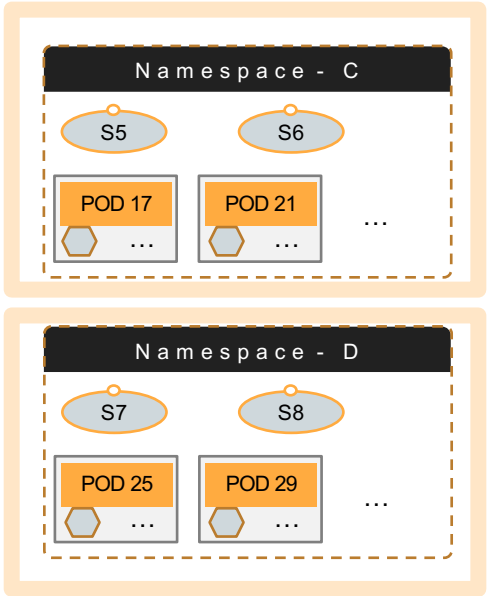
DEFAULT CLUSTER MODE

- This is how Kubernetes networking works today
- Flat subnet where -- Any workload can talk to any other workload



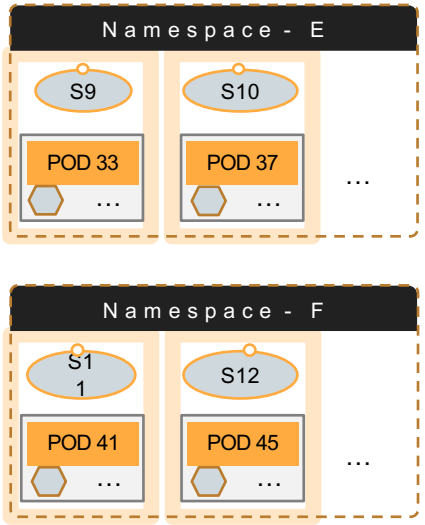
NAMESPACE ISOLATION

- In addition to default cluster, operator can add isolation to different namespaces transparent to the developer



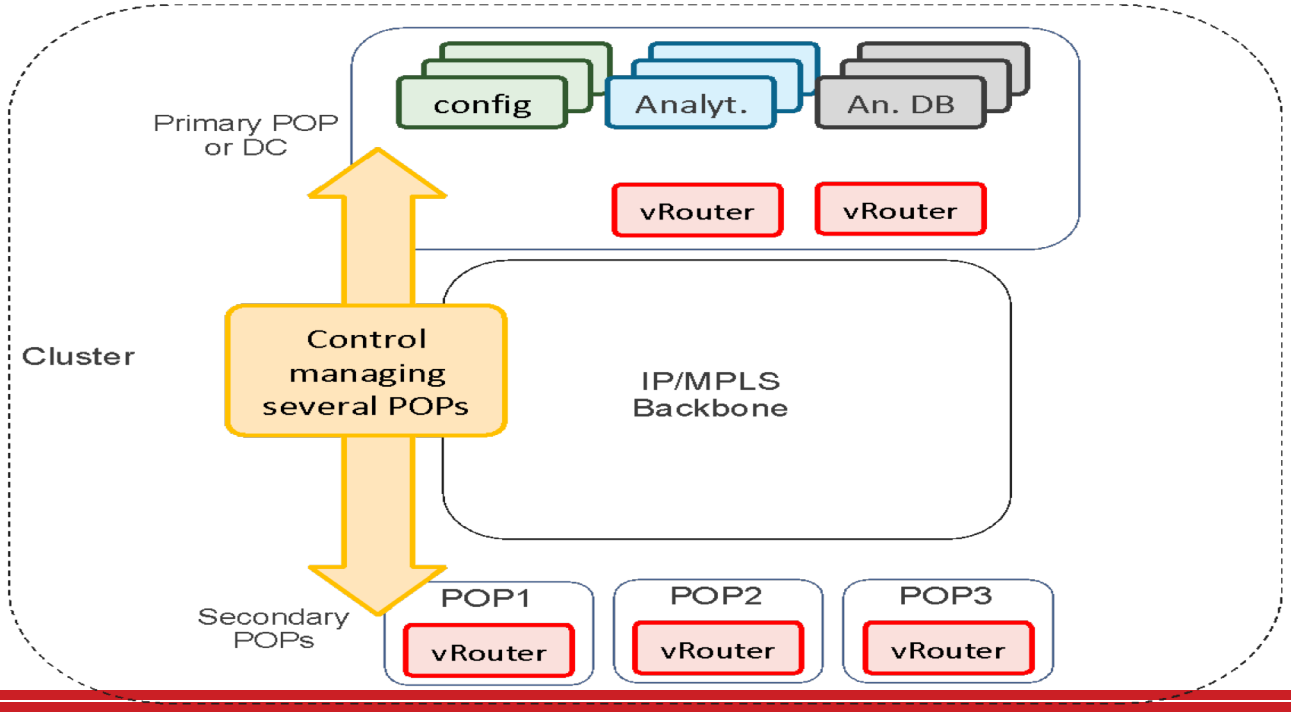
POD / SERVICE ISOLATION

- In this mode, each POD is isolated from one another
- Note that all three modes can co-exist

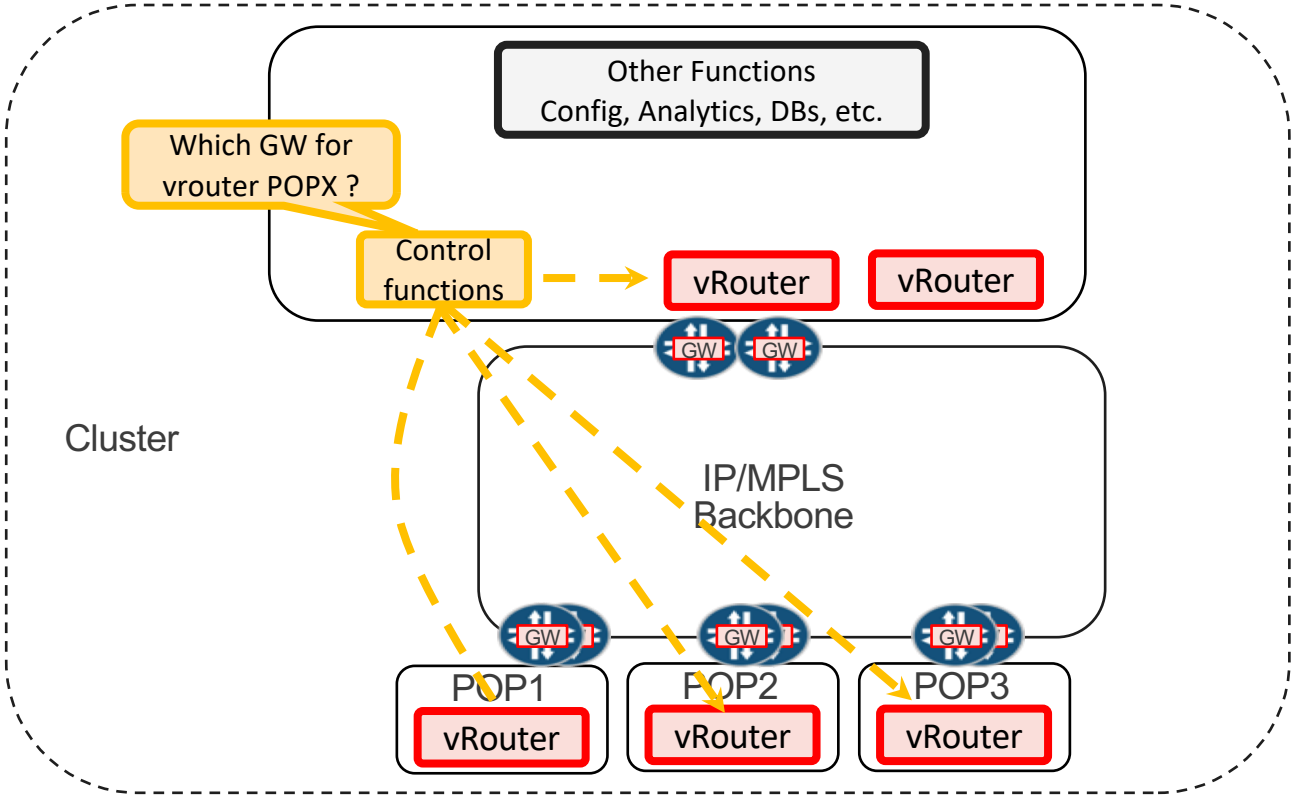


Remote Compute (Edge Computing)

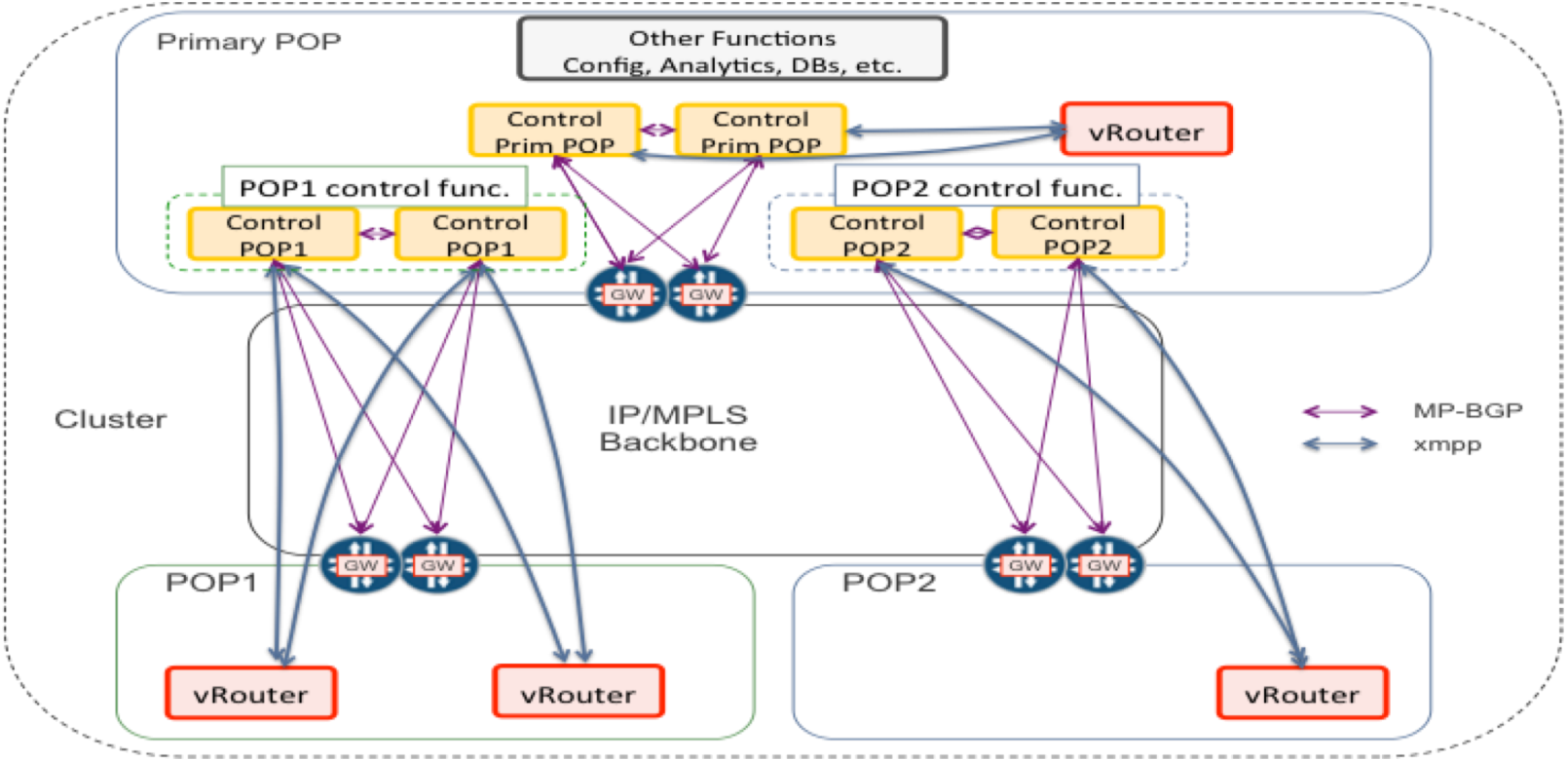
Objective: spread a Single Cluster over several POPs



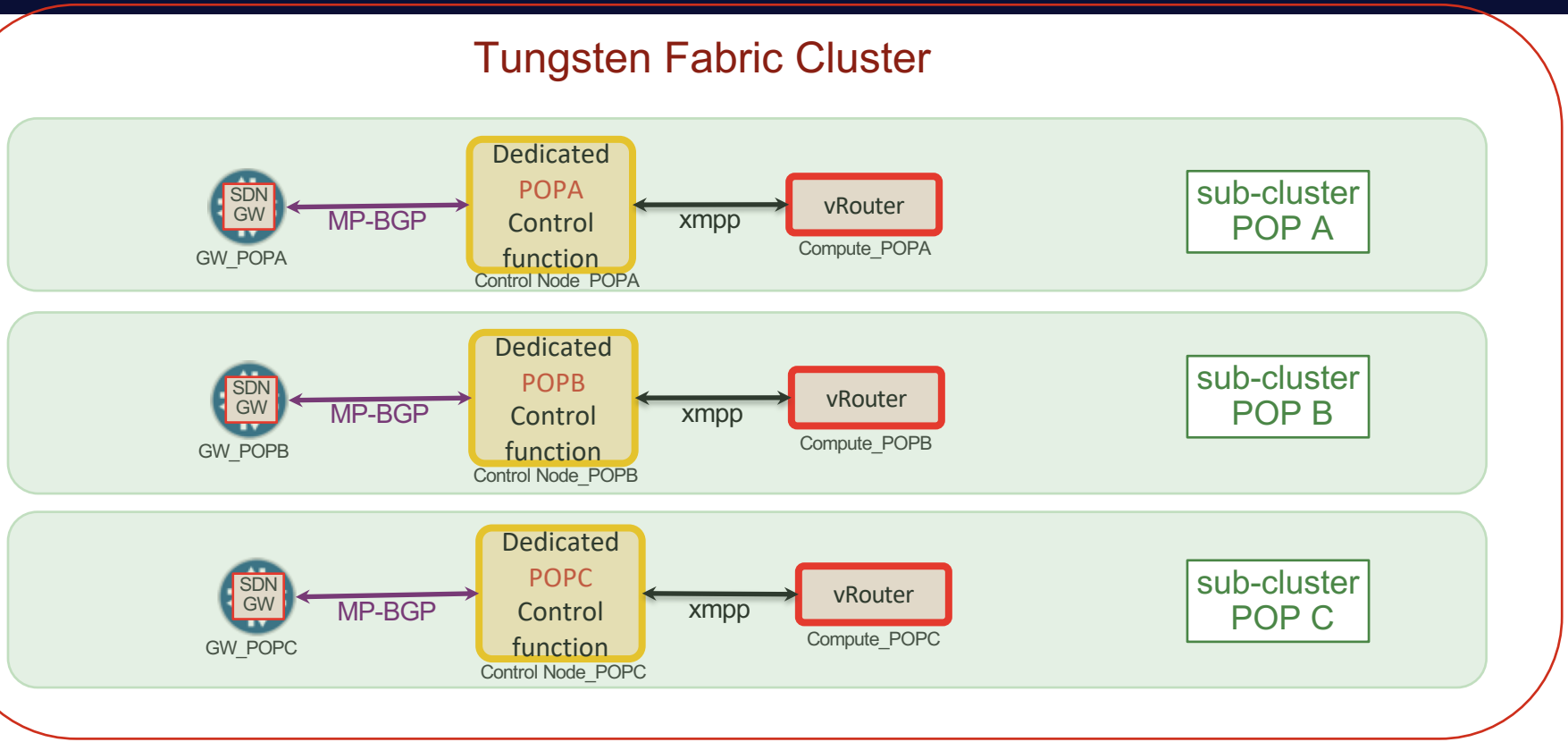
Remote Compute (Edge Computing)



Logically Distributed Control Plane



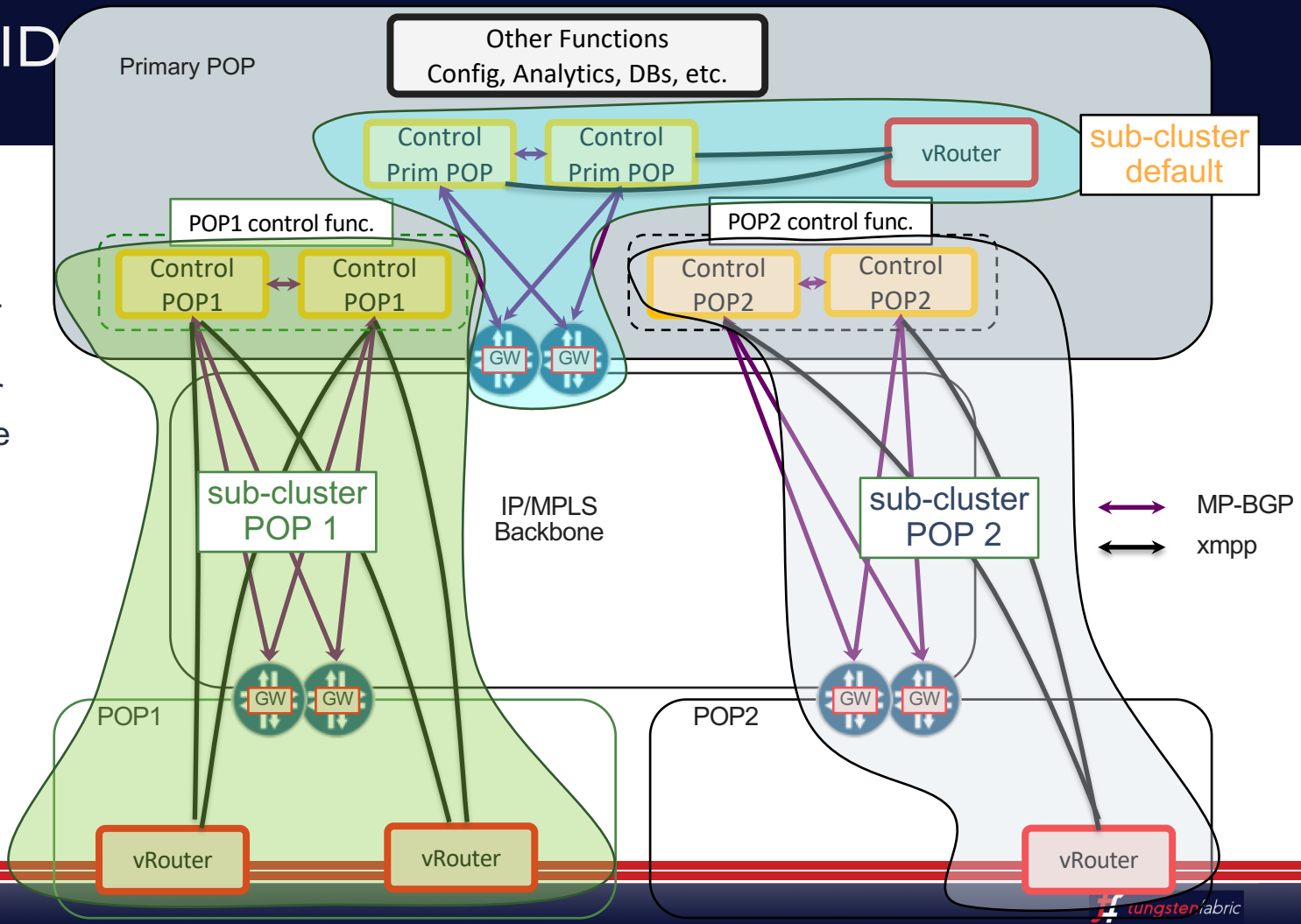
Remote Compute – “Sub-Cluster ID” Concept



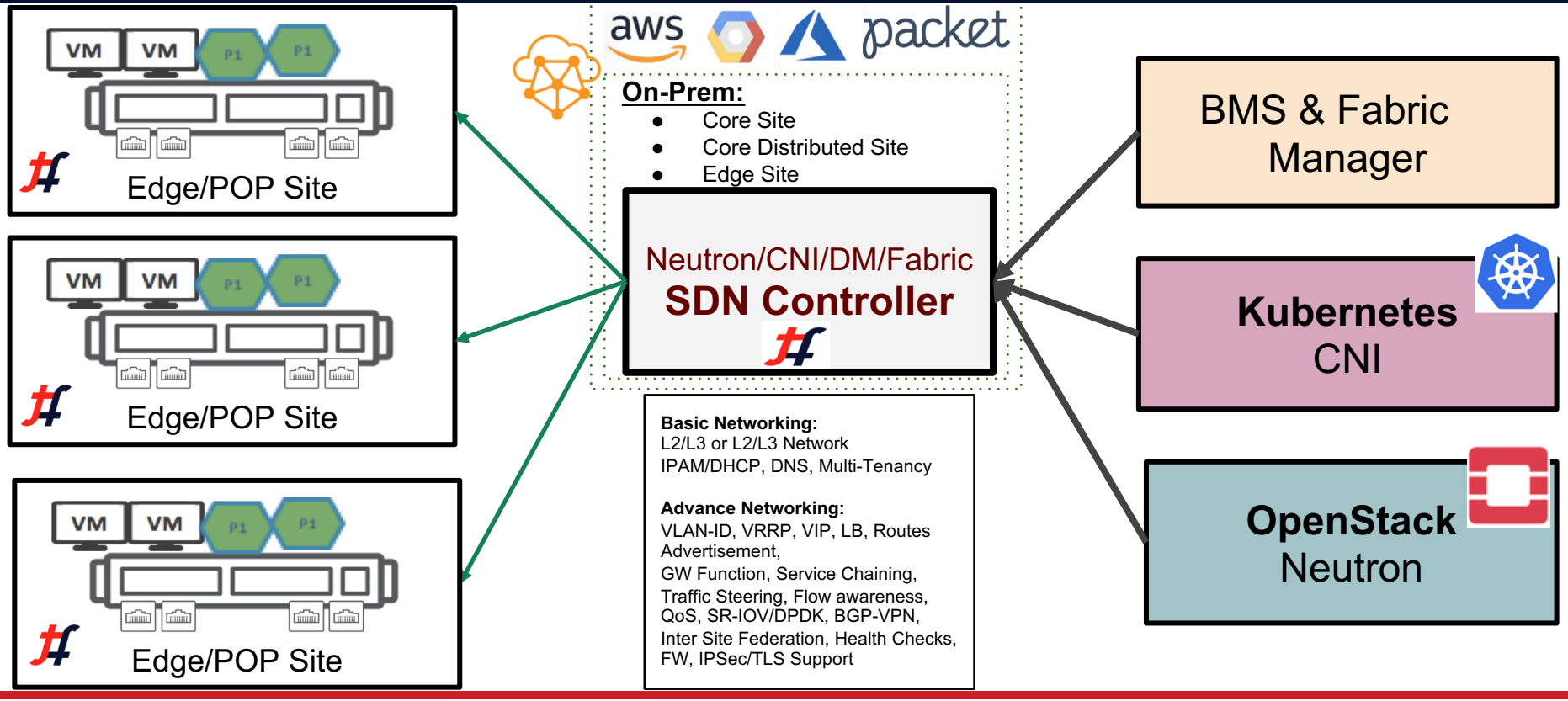
Sub-Cluster ID

3 Sites: 1 Primary POP and 2 small POPs (POP1 and POP2)

- CN, GW and Vrouter for a POP belong to the same sub-cluster
- All Control Nodes are deployed on a same site (CN are “logically” in a POP from a routing point of view)

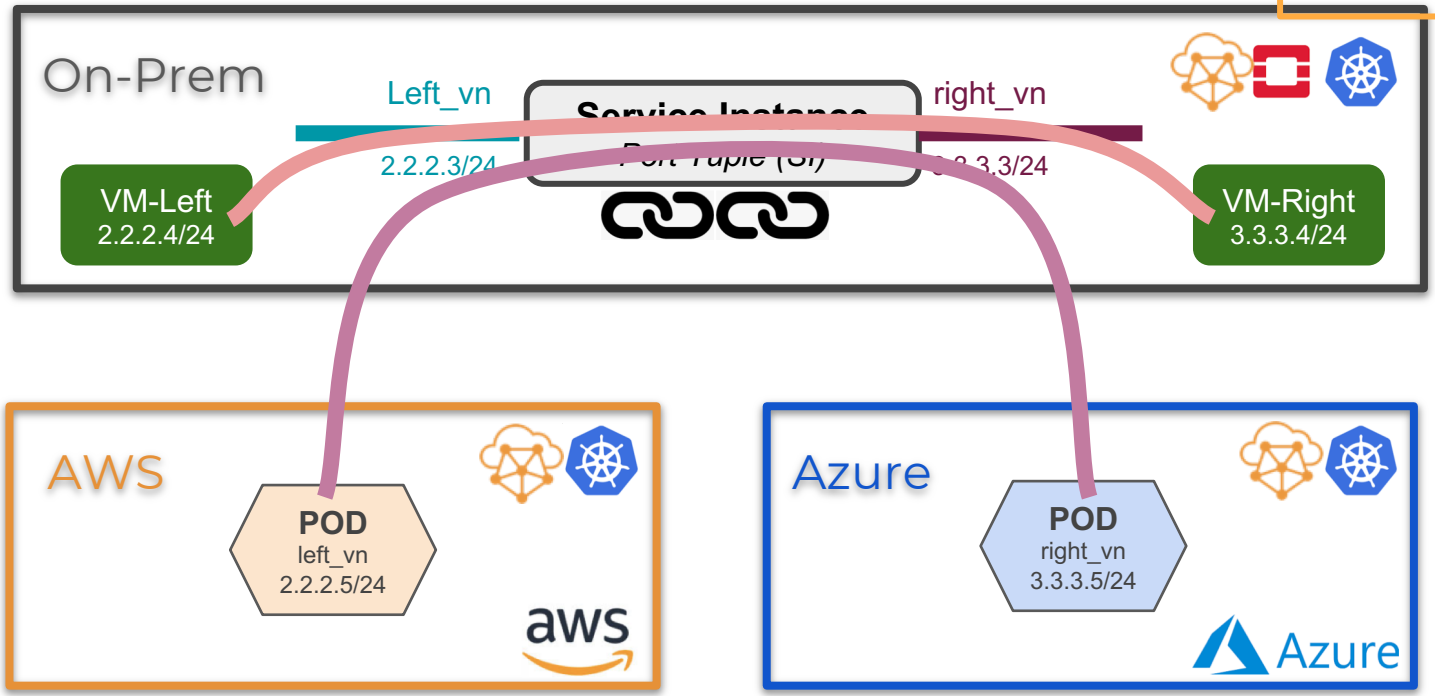


Tungsten Fabric – SDN for Distributed Networking for VMs, PODs, & BMS



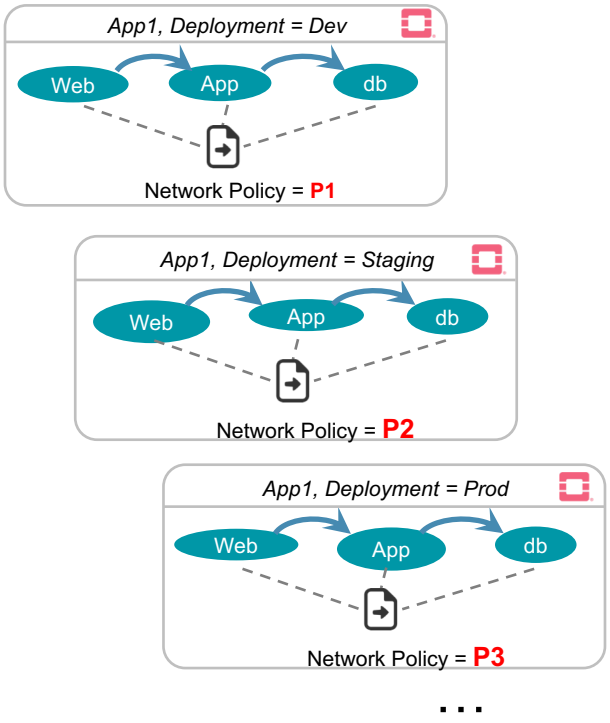
Multi Cloud Service Chaining

left_vn: 2.2.2.0/24
right_vn: 3.3.3.0/24



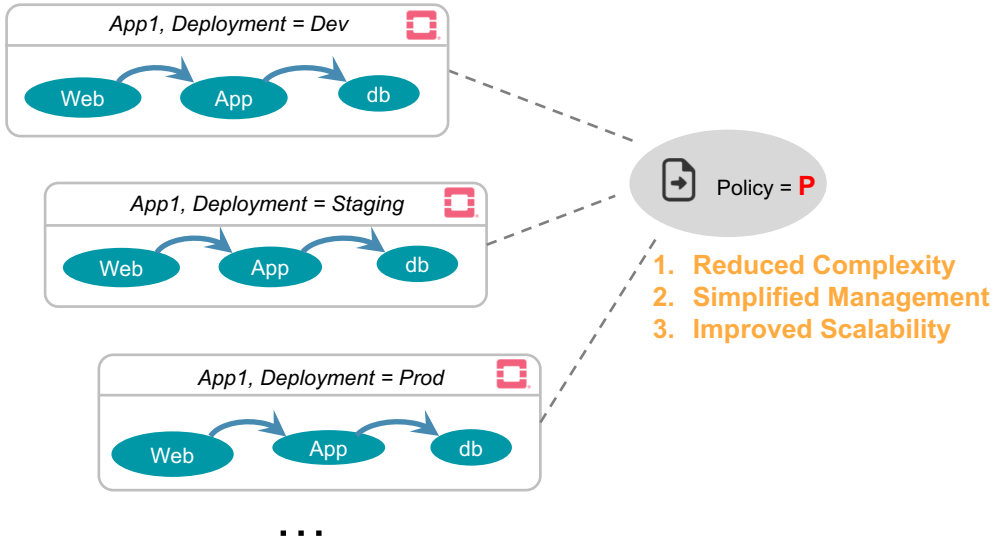
Policy Framework

Old Behavior



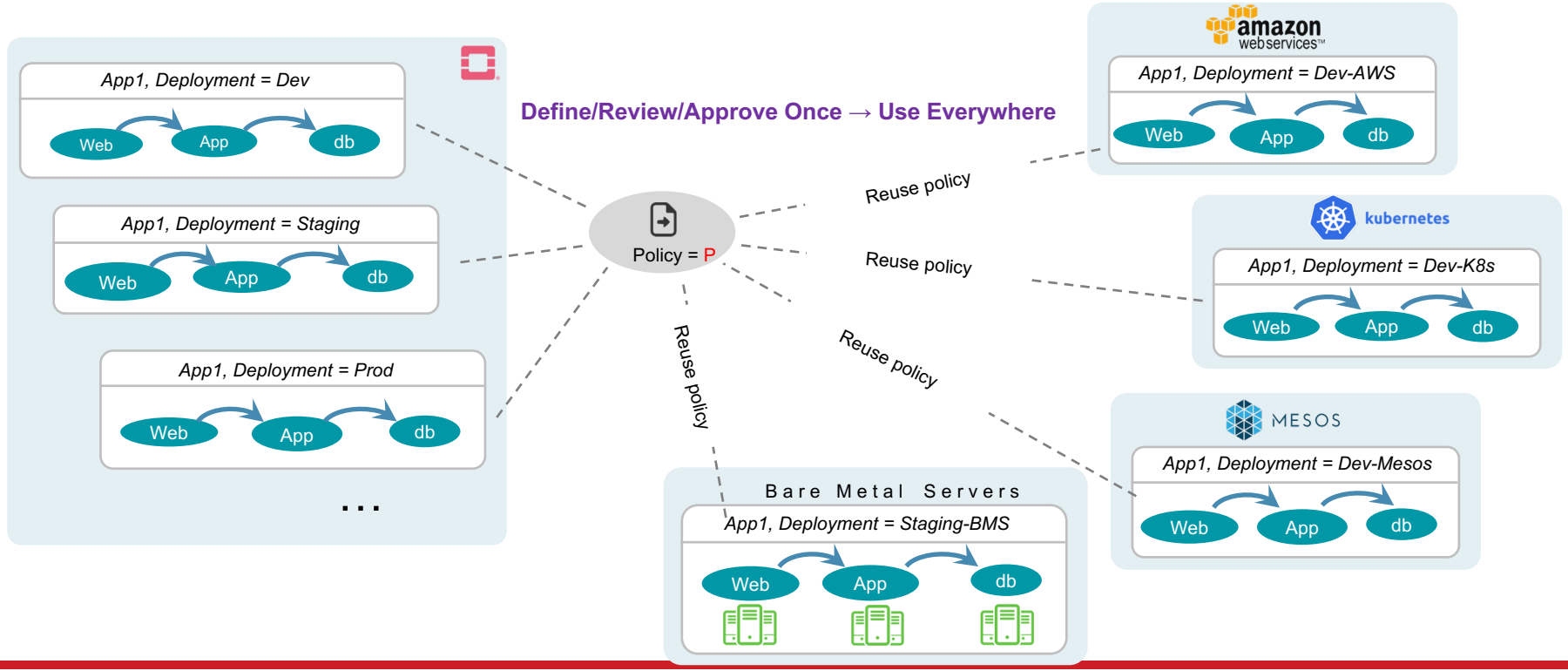
New Behavior

Can we use one policy to be applied in all the different deployments ?



Policy Framework

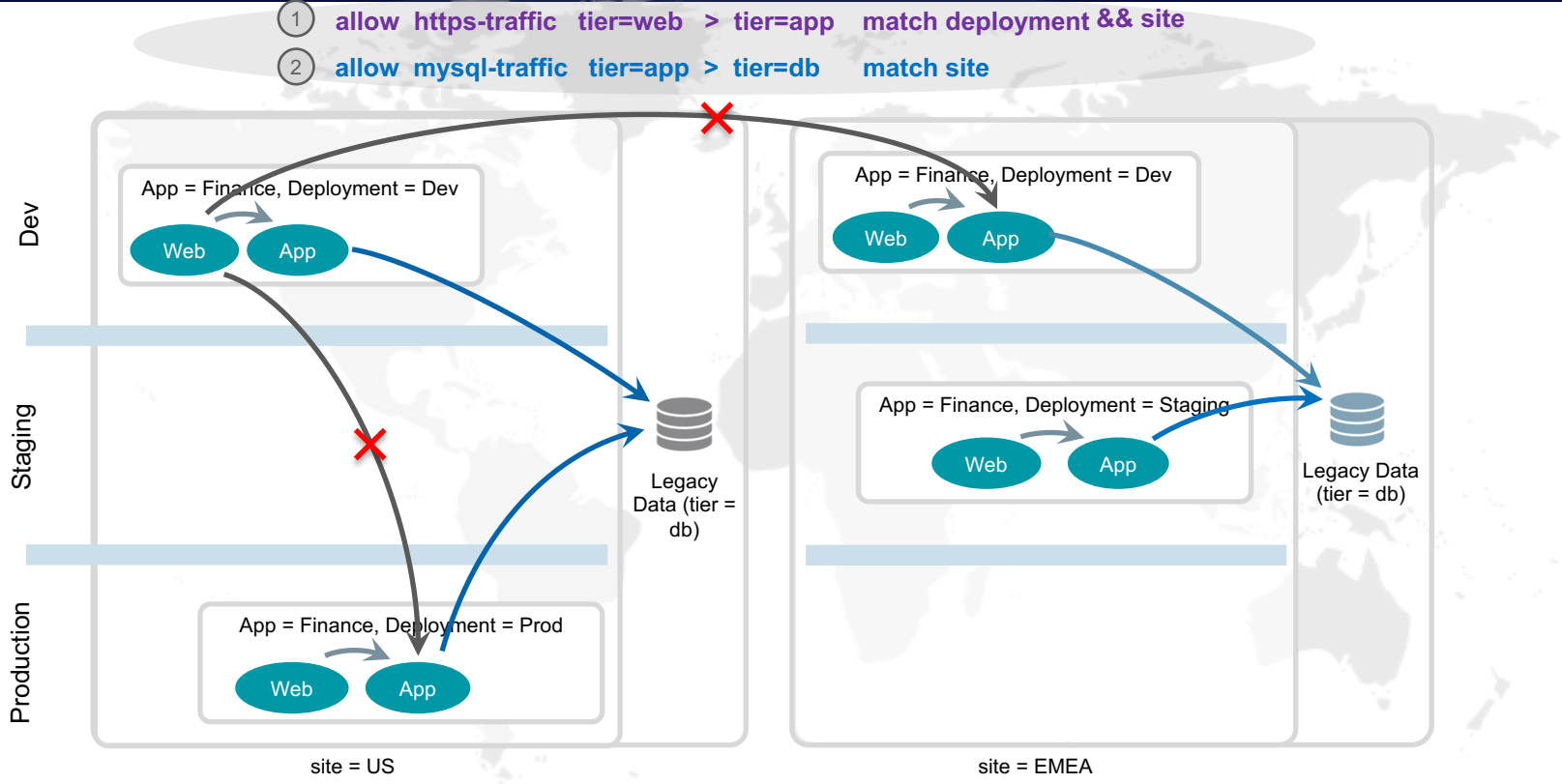
Reuse of policies across multiple clouds and with multiple orchestrators



Policy Framework – Use Case Example

Defn

Enforcement



Why Edge Computing?

Emerging technologies are demanding lower latency and accelerated processing at the edge

NFV Edge Infrastructure

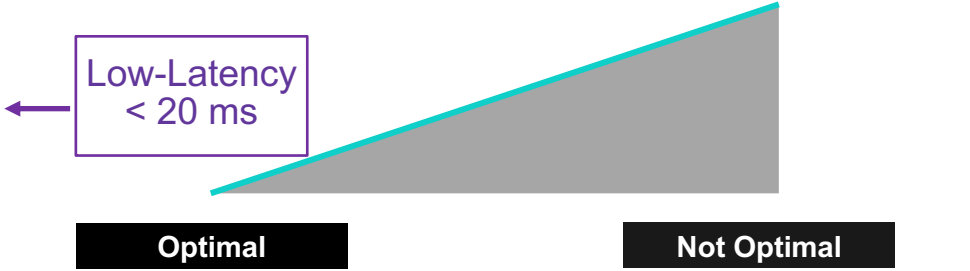
Autonomous Devices

Immersive Experiences

Industrial IOT

Edge Cloud

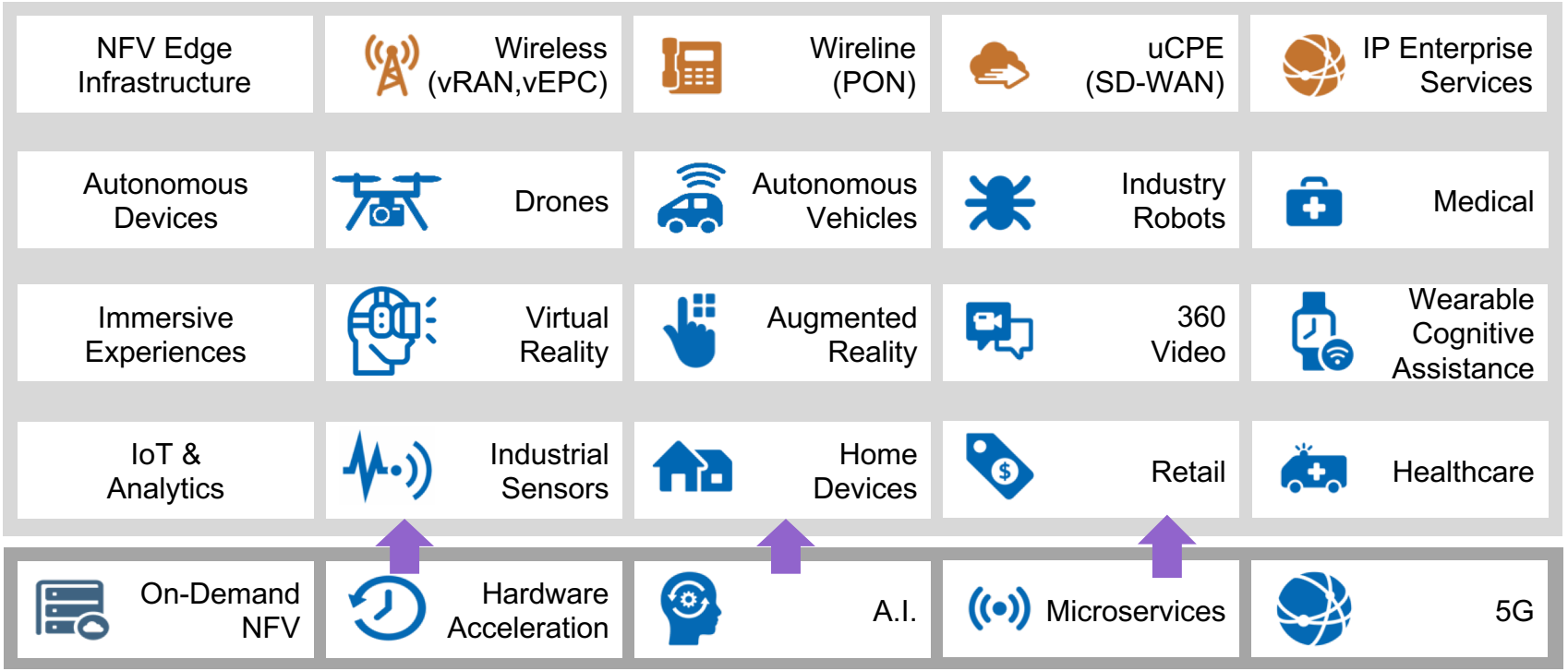
Perform data processing at the edge of the network, near data sources



Central Cloud

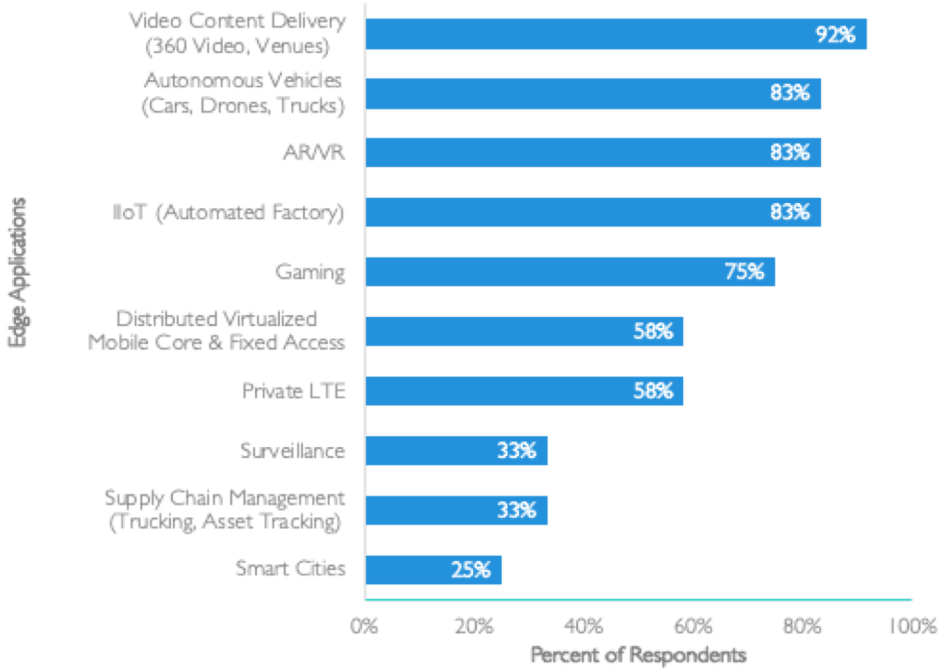
Highly centralized computing resources of cloud service providers

Emerging Edge Applications & Convergence of Technologies are demanding lower latency + accelerated processing



Edge Killer Apps: Non-traditional video + Connected things that move

Q: What are the top 5 (or more) edge services?

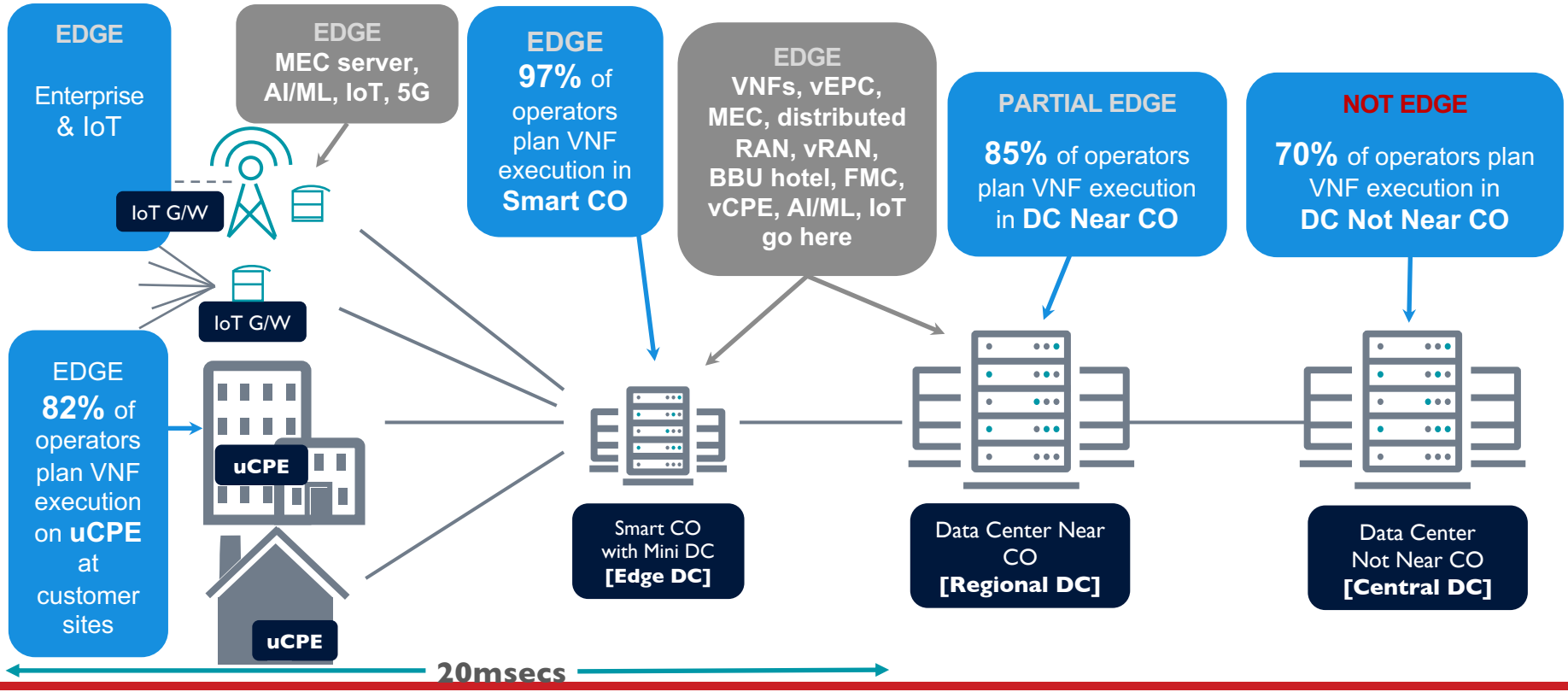


- Many metro IX locations within 20ms of parts of populations…
- Telcos have advantage of COs, cell sites, cell backhaul aggregation, fixed backhaul, street cabinets, etc. much closer to users
- Edge enhanced apps include many elements: natural language, facial recognition, immersive experience, swarming
- Big (too much) bandwidth top driver
- Our categories are a grouping of several applications; can be user delivered

Source: IHS Markit. *NFV Strategies: Global Service Provider Survey*, June 2017; *Respondents control 61% of global telecom capex*

Where are the edges?

Distributed cloud, edge compute, AI/ML, IoT, 5G, VNFs/NFV, FMC



LF Edge - Founding projects

Bringing several Edge verticals and domains under one umbrella

AKRAINO EDGE STACK

EDGE X FOUNDRY™

EDGE VIRTUALIZATION ENGINE

HOME EDGE

OPEN GLOSSARY OF EDGE COMPUTING

Platinum Members:

arm

AT&T

DELL EMC

ERICSSON

Hewlett Packard Enterprise

HUAWEI

IBM

intel

inwinSTACK

JUNIPER NETWORKS

MobiledgeX

NETSIA

NOKIA

NTT

Qualcomm

Radisys

redhat

SAMSUNG

SEAGATE

Tencent 腾讯

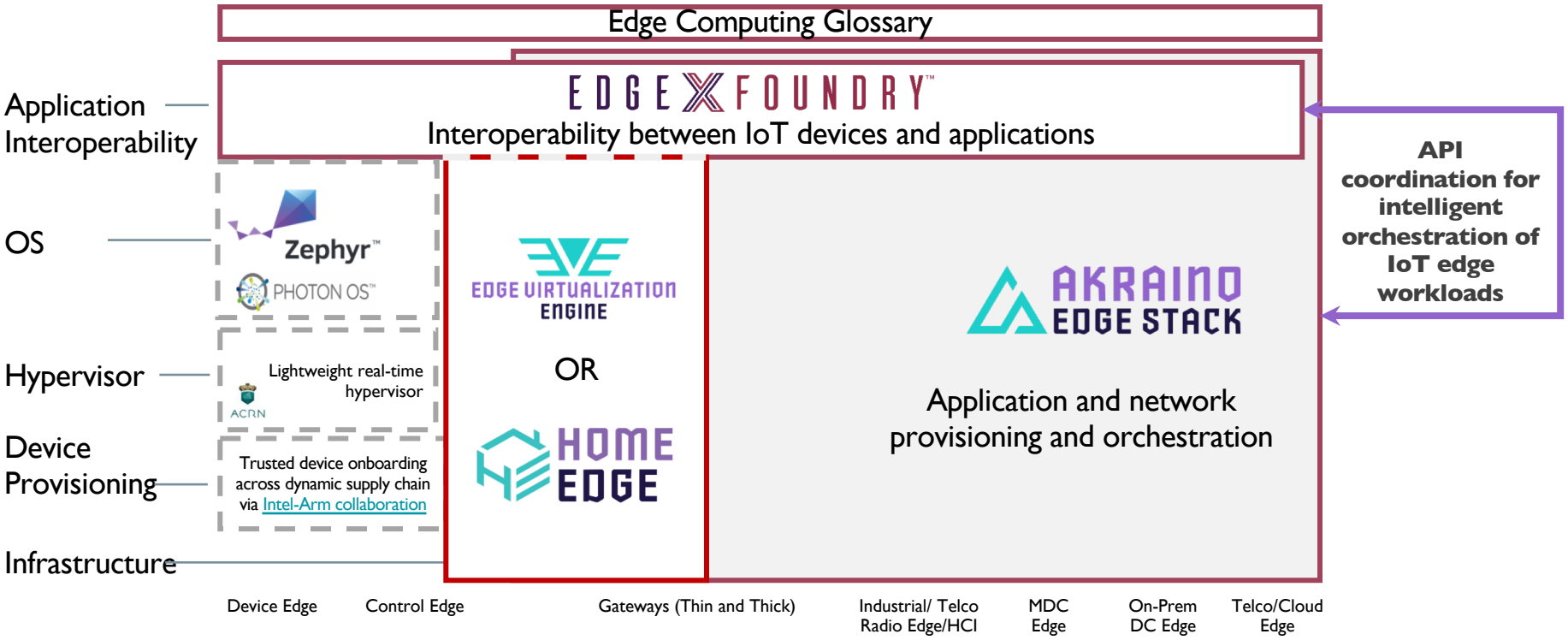
WIND

wipro

ZEDEDA

60 + Members already

Scope of LF Edge



Akraino Edge Stack Executive Summary

Zero Touch
Edge Cloud
Automation

Akraino is an Edge project targeted to

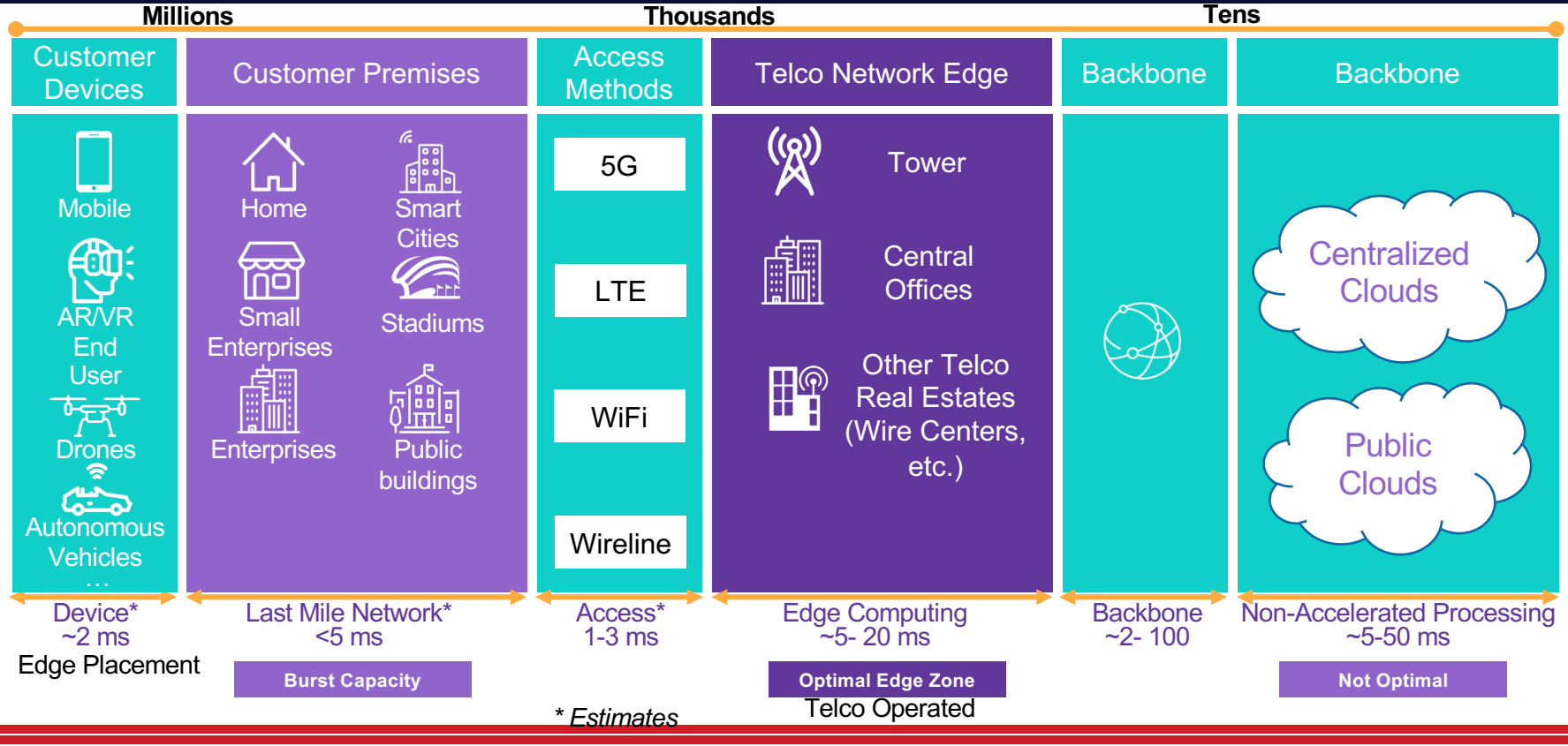
- Address Telco, Enterprise and Industrial IoT use cases

Mission:

1. Create end to end configuration for a particular Edge Use case which is complete, tested and production deployable meeting the use case characteristics {Integration Projects - Blueprints}
2. Develop projects to support such end to end configuration. Leverage upstream community work as much as possible to avoid duplication. {Feature Projects}
3. Work with broader edge communities to standardize edge APIs {Upstream Open Source Community Coordination - For example, Socialization, so community tools and Blueprints can interoperate. This work can be a combination of an upstream collaboration and development within the Akraino community [i.e. a feature project]}
4. Encourage Vendors and other communities to validate Edge applications and VNFs on top of Akraino blueprints {Validation Project - ensures the working of a Blueprint}

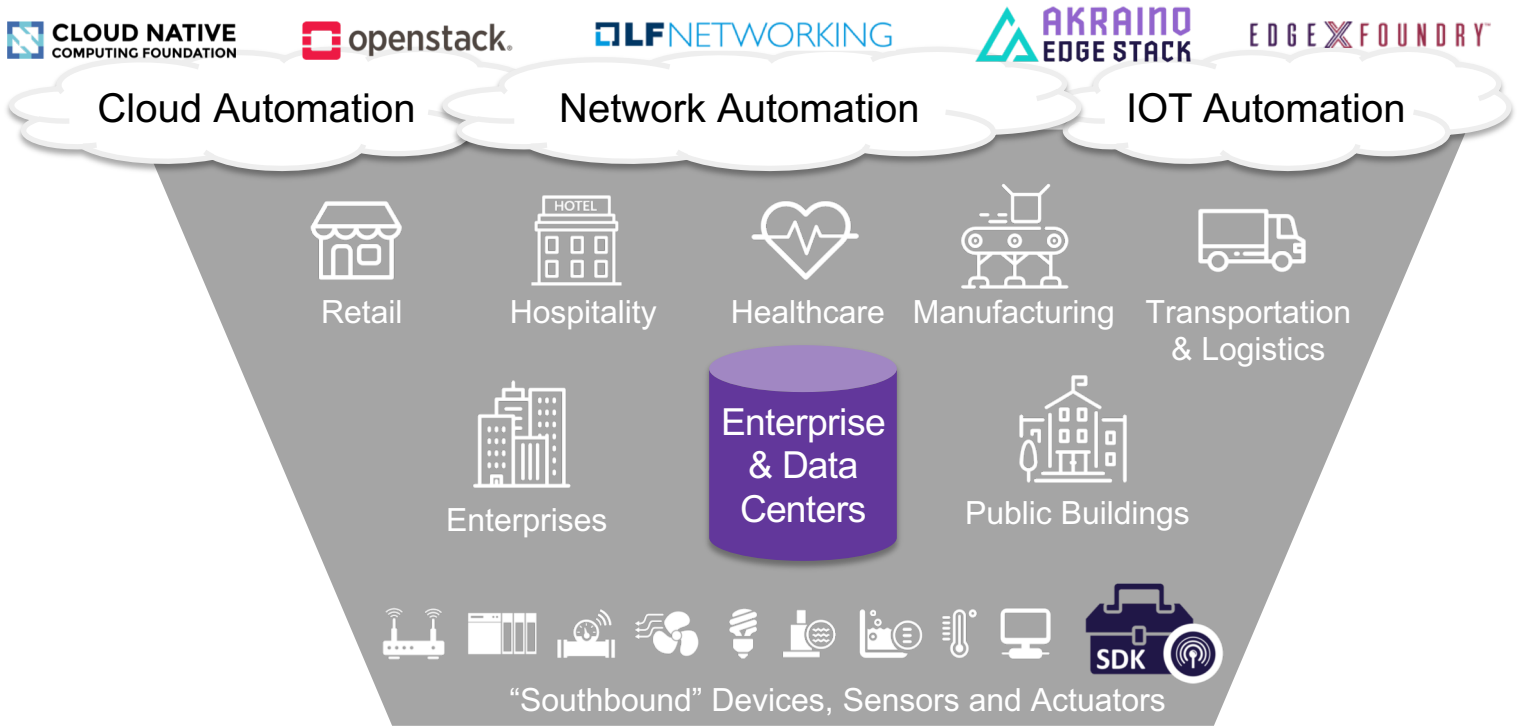
Use Case 1: Operator's Owned Network Edge

Optimal Zone For Edge Placement



Use Case 2: IOT Driving the New Edge for Enterprise

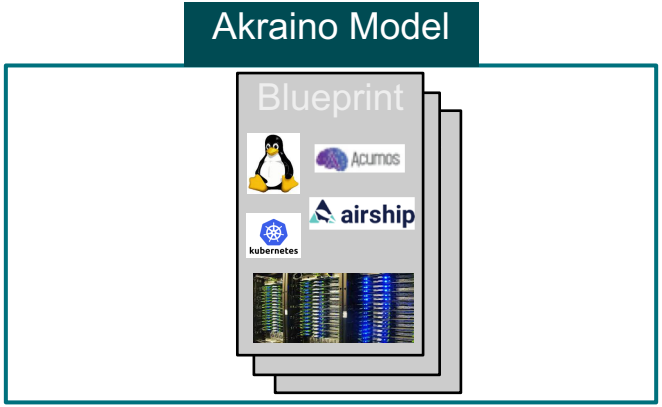
Retail, Transportation, Healthcare…



Akraino Edge Stack Blueprint

Before Akraino

Akraino Model



- User integrates multiple opensource
- Multiple gaps
- No integrated solution for Edge use cases
- Complex CI
- No guaranteed working solution

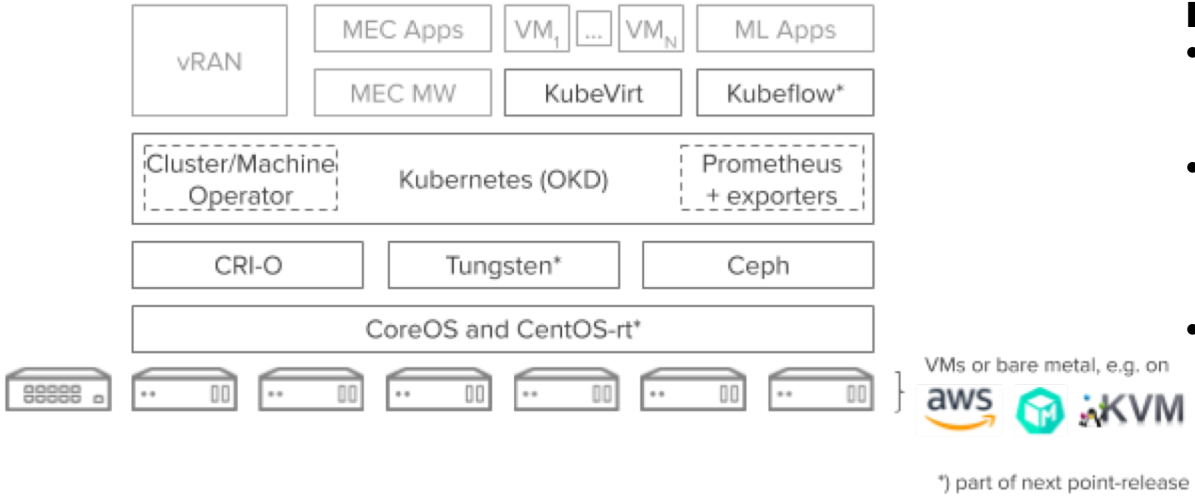
- Akraino Community Integrates multiple opensource for edge use cases.
- Bridge gaps (development of code in upstream and at Akraino)
- Fully integrated solution
- Simple CI
- Validated with multiple testing

Akraino R1 Blueprints

Blueprint Family	Blueprint	Primary Use Case	Industry Target	Blueprint Summary
Network Cloud (NC)	Unicycle with SR-IOV	Telco Edge use cases (Multi Server). Multiple applications	Telco, Enterprise	NC blueprint family enables hardware configuration and automated deployment of telco grade multiple edge sites from a remote regional controller.
	Rover	Telco Edge use cases (Single Server). Multiple applications	Telco, Enterprise	
	Unicycle with OVS-DPDK	Telco Edge use cases (Multi Server). Multiple applications	Telco, Enterprise	
Telco Appliance	**Radio Edge Cloud (REC)**	Appliance for Radio Access Network (RAN), RAN Intelligent Controller and Near realtime Edge MEC Appliance	Telco 5G, Enterprise	Appliance tuned to support the O-RAN Alliance and O-RAN Software Community's Radio Access Network Intelligent Controller (RIC)
Integrated Edge Cloud (IEC)	Type 1 (small Edge)	Telco or enterprise application deployment on Arm servers	Telco, IOT and Enterprise	IEC enables the new functionalities and deployment model on the network edge. It supports ARM processors and architecture.
	Type 2 (Medium Edge)	Telco or enterprise application deployment on Arm servers	Telco, IOT and Enterprise	
StarlingX	**Far Edge Distributed Cloud**	Enterprise edge and Far edge. Multiple applications	Enterprise & IOT	Addresses edge and Far edge use cases at high density locations such as malls, airports and sports stadiums to support value added services at these events and locations.
Kubernetes- Native Infrastructure for Edge	**Provider Access Edge**	vRAN and MEC (AR/VR, Machine learning, etc.,)	Enterprise & Telco	Blueprints in the Kubernetes-Native Infrastructure for Edge family leverage the best-practices and tools from the Kubernetes community to declaratively and consistently manage edge computing stacks from the infrastructure up to the workloads.
Edge Lightweight and IOT blueprint (ELIOT)	IOT Gateway	IOT	IOT & Enterprise	ELIOT targets on making the edge node a lightweight software stack which can be deployed on limited hardware capacity.
	uCPE	uCPE	Enterprise & Telco	

Akraino R1 Blueprint: Provider Access Edge

Kubernetes Native Infrastructure for Industrial Automation

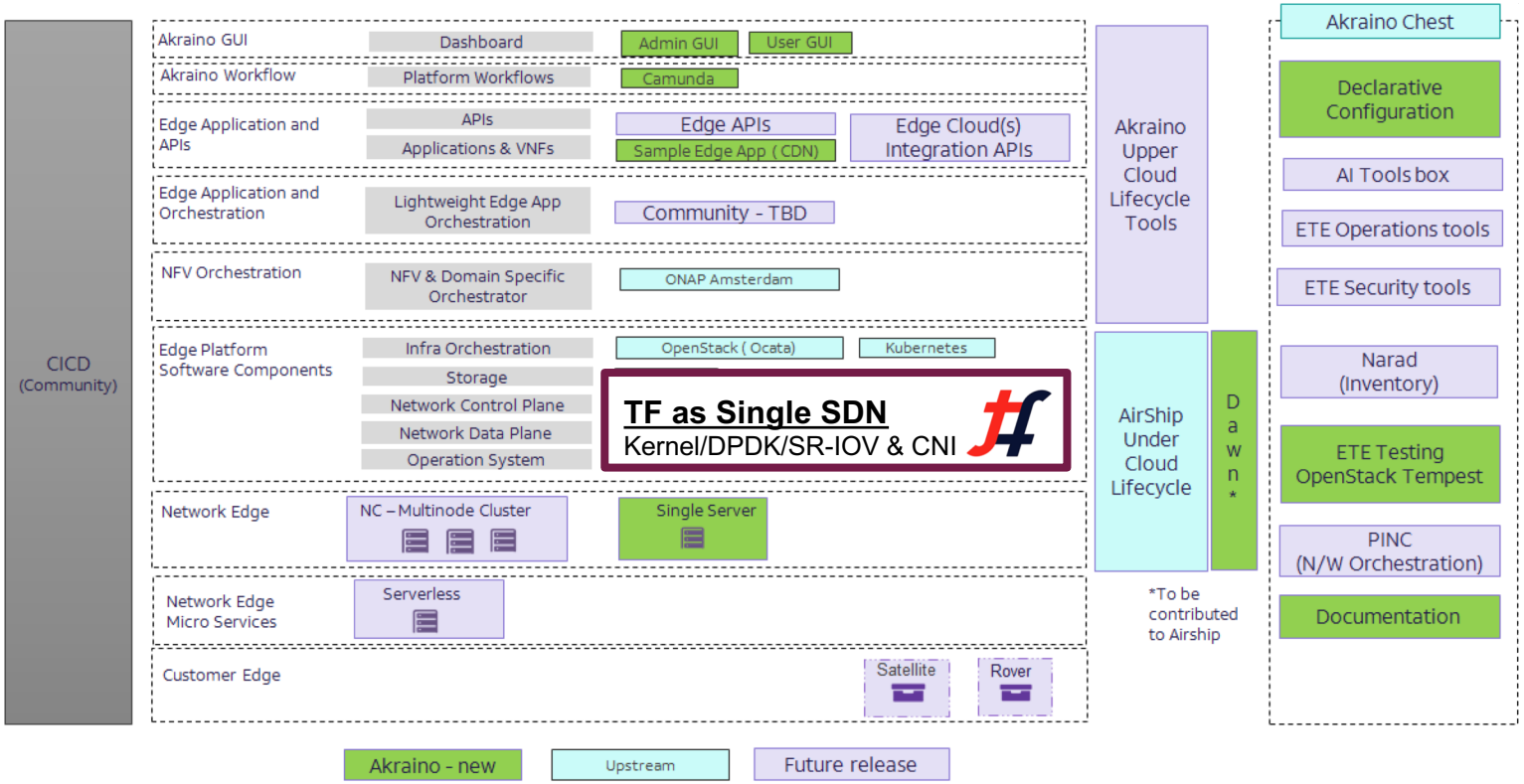


Purpose/Features

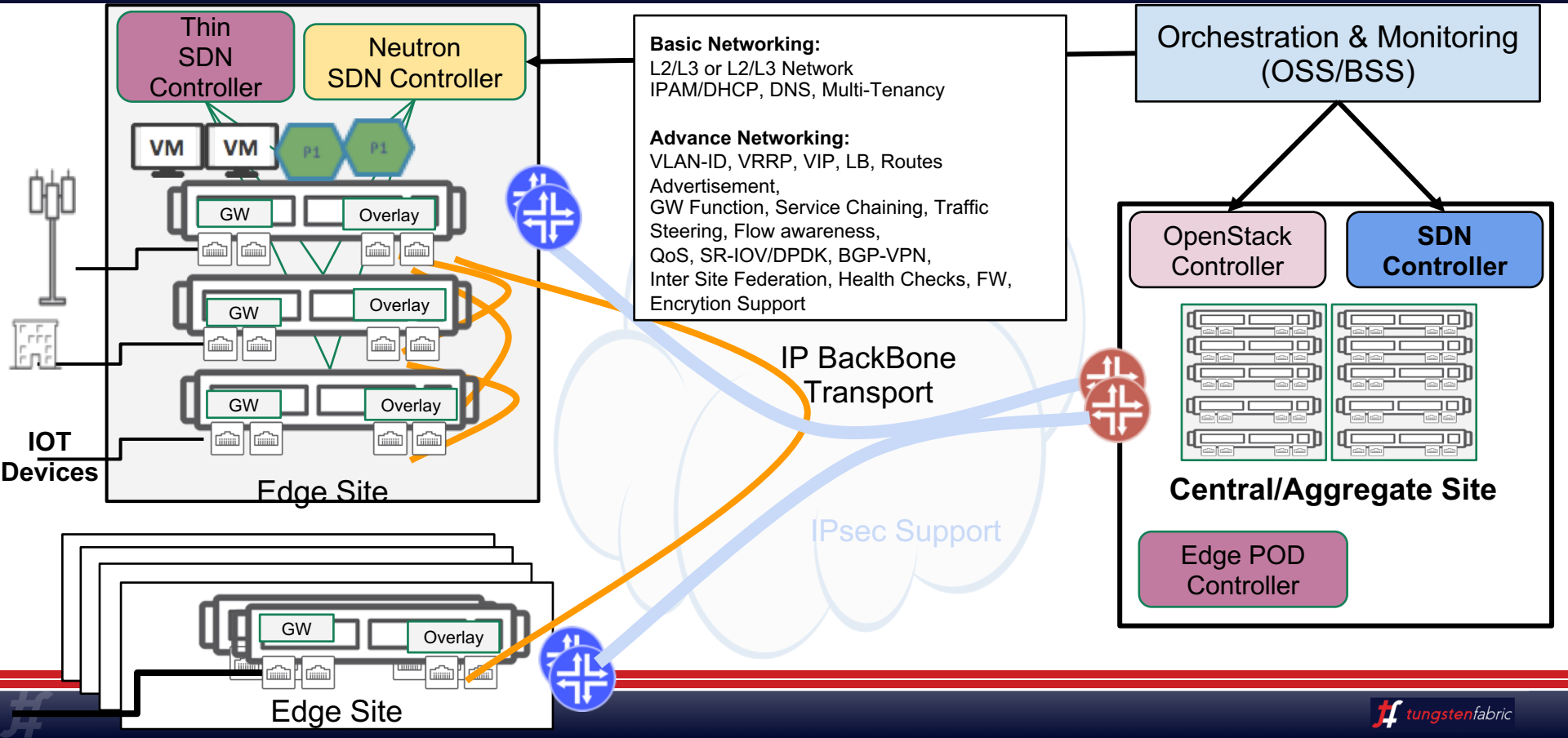
- Addresses generic Edge Use cases (small footprints deployments)
- Focused on Native Container workloads able to host NFV and MEC with no OpenStack
- Manage edge stacks at scale and with a consistent, uniform user experience from infrastructure up to workloads, on bare metal or public cloud

Target Industry: Telco, Enterprise

Akraino Network Cloud & TF Integration *(Blueprint)*



Remote Compute and Telco Cloud



AI/ML and AR/VR applications at Edge

Case Attributes	Description	Informational
Type	New Blueprint for enabling AI/ML and low latency AR/VR capabilities at the Edge	
Blueprint Family - Proposed Name	Integrated Edge Cloud	
Use Case	Programmability on switches and I/O Accelerations on programmable NICs & embedded FPGAs to deliver AI/ML workload placement and low latency demands of AR/VR applications onboarding edge stack	
Blueprint proposed Name	AI/ML and AR/VR applications at Edge	
Initial POD Cost (capex)	Leverage white boxes, standard NICs: The cost of POD will depend upon the hardware profiles and peripherals desired	
Scale & Type	Detailed in Resource Requirements slide	
Applications	AI/ML streaming workloads and AR/VR applications	
Power Restrictions	Less than 10Kw	
Infrastructure orchestration	Openstack Queens or above Docker 1.13.1 or above Container Orchestration –K8s 1.10.2 or above OS - Ubuntu 16.x, CentOS	
SDN	Tungsten Fabric, kernel/DPDK/SmartNIC offload vRouter	
Workload Type	Containers over VM or baremetal	
Additional Details	• Runs on Commodity HW • x86, ARM, SoC, ... • Multiple options for partial or full NIC offloads (Intel, Netronome, Mellanox) • Future Supports: • eBPF/XDP offload	See next slide for additional details

Try Tungsten Fabric

https://tungstenfabric.github.io/website/Tungsten-Fabric-15-minute-deployment-with-k8s-on-AWS.html

Thank You