Distributed Hyperconvergence

Pushing Openstack and Ceph to the Edge

Sean Cohen, Giulio Fidente, Sébastien Han

Agenda

- Reality check
- Edge landscape and 5G
- Edge factors
- Distributed requirements and Edge form factors
- Where we are today
- Distributed HCI
- Stein release focus
    - Edge working group model
    - Deployment at the edge
- Roadmap and future use cases
- Key takeaways

We’re getting smarter

Smart Countries


smacc
Smart Country Convention
20.-22. November 2018 — CityCube Berlin
smacc
Smart Country Convention

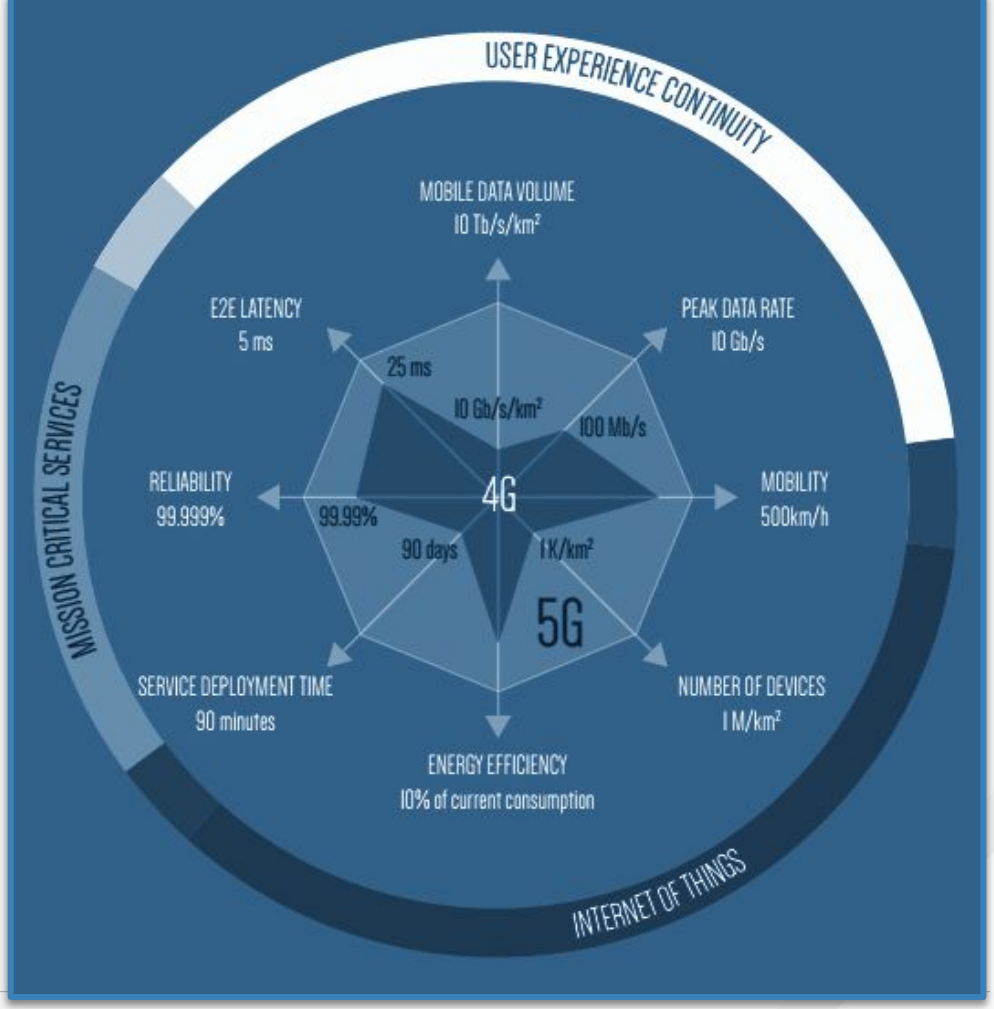
Smart Cities

Smart Cities

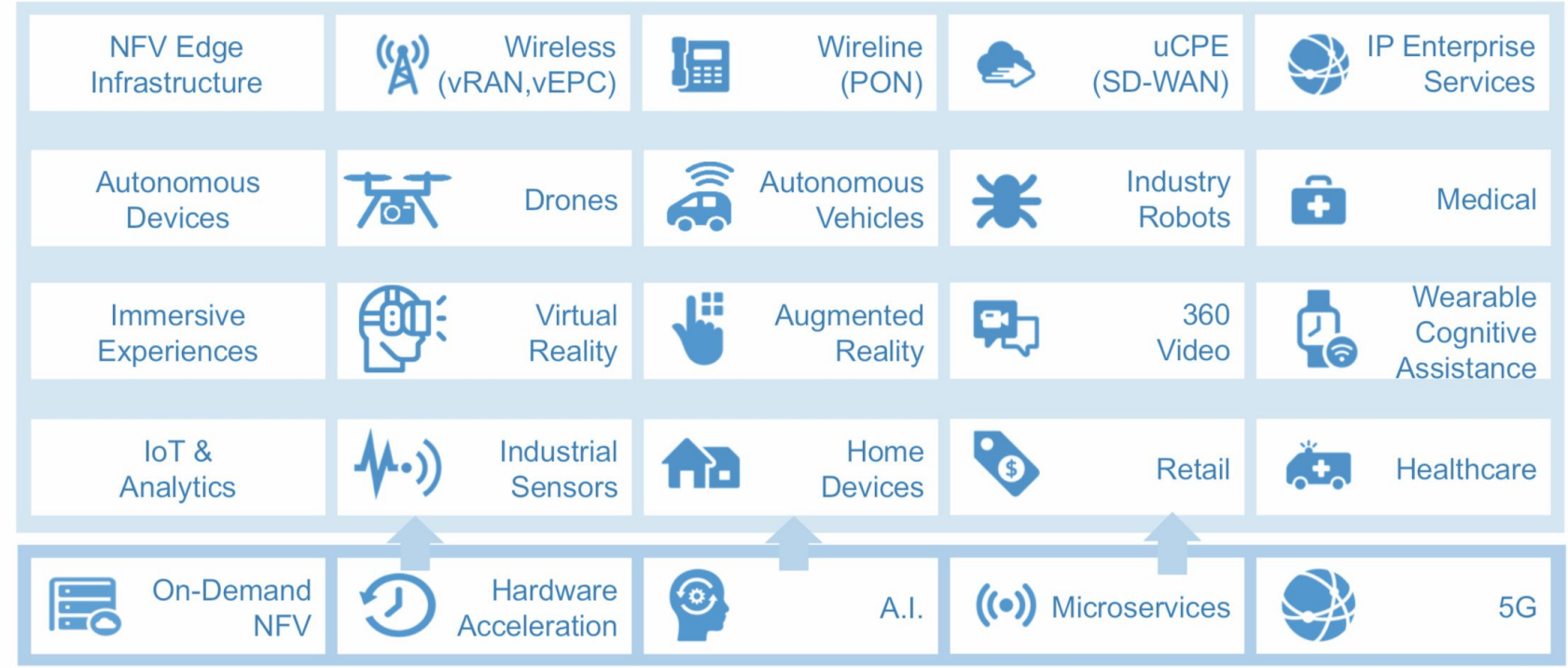
Smart Cars

Reality check

Distributed cloud is part of the standard



Edge use cases

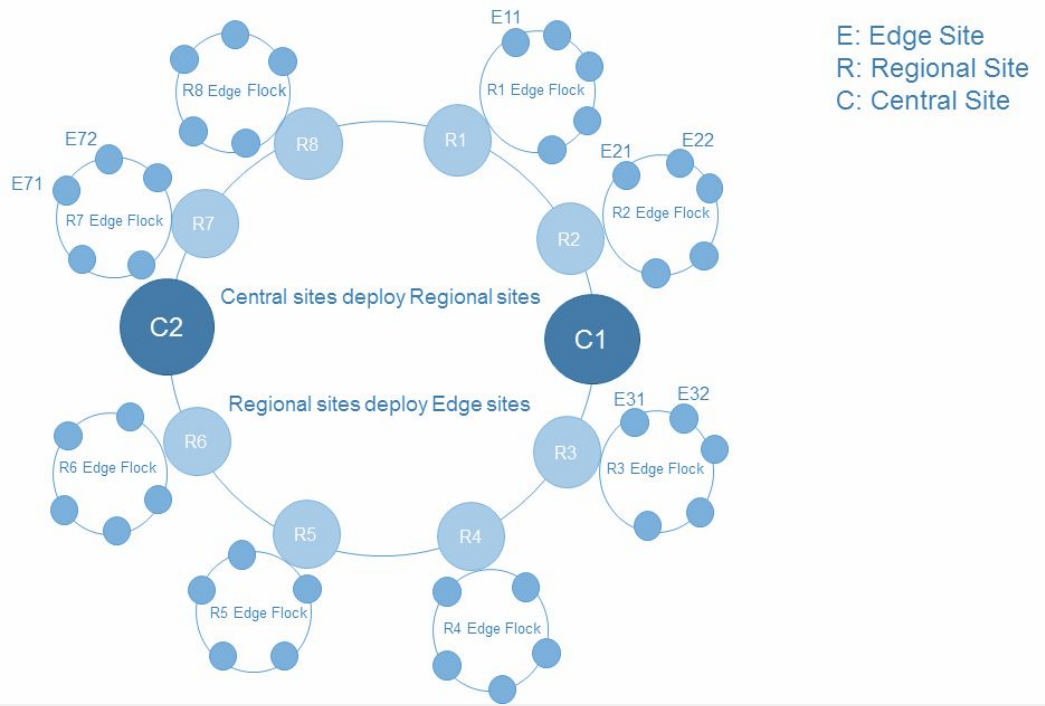


Edge - 5G Constraints

Latency	Bandwidth	Resilience
consumer ⇔ service RTT ≃1ms ➔ distance ≃100km	funnel of 1M's of 1Gb/s connections	make edge sites autonomous, minimize failure domain to 1 site

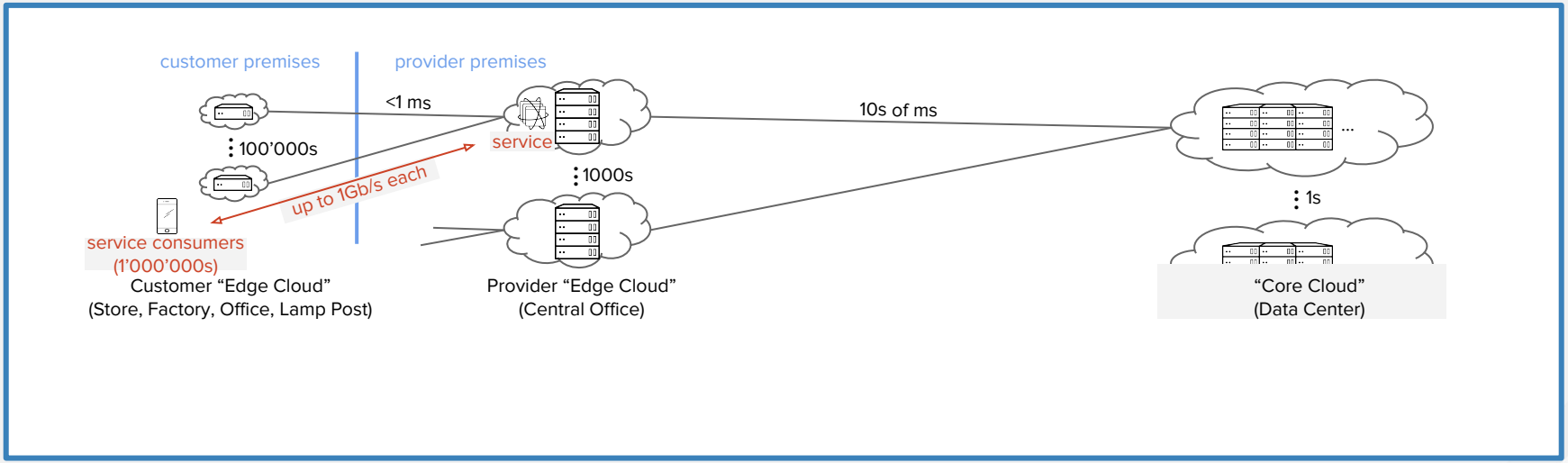
Regulations	Distribution	Scale
keep sensitive data on-site / within regulatory region	Need to deploy and perform a the lifecycle management of distributed systems	need to manage 10-100k's of sites with 1-10's of nodes each

The classical Edge topology

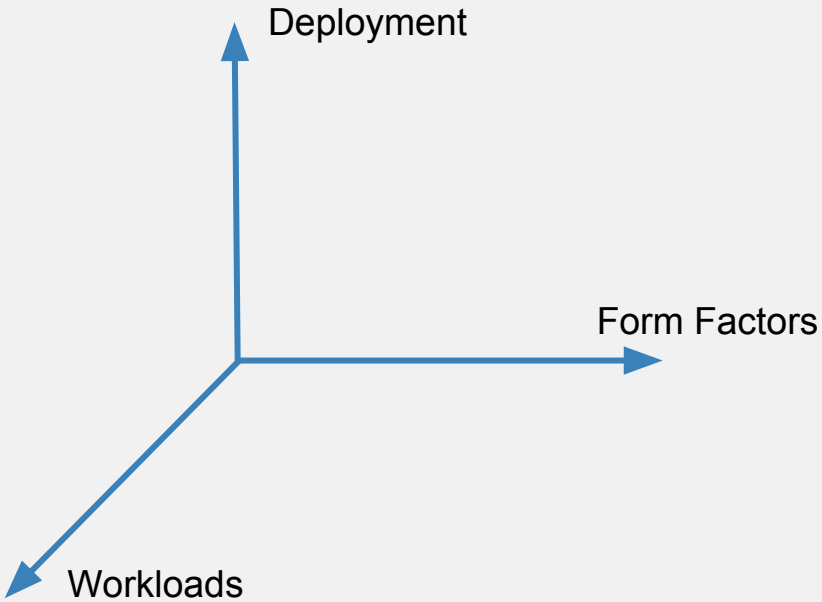


Source: https://wiki.akraino.org/display/AK/Akraino+Edge+Stack

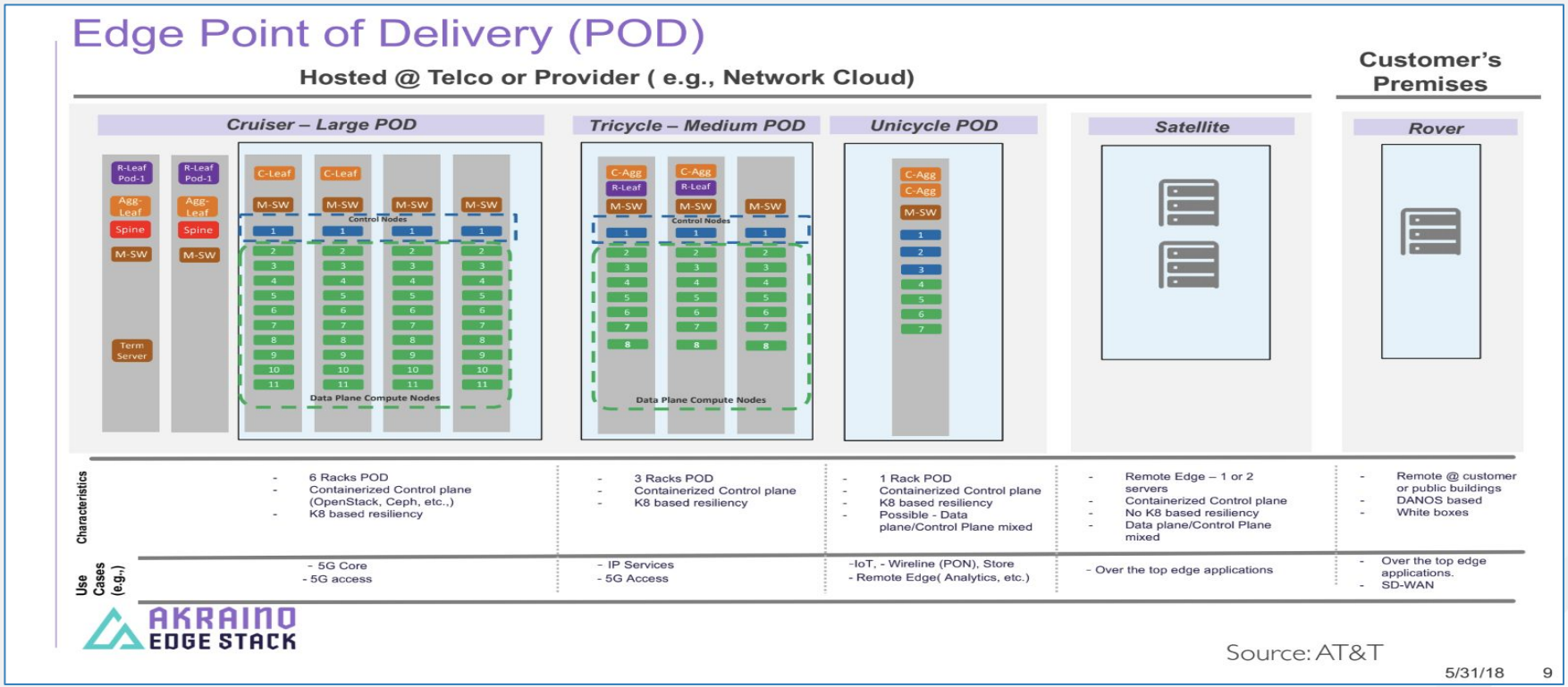
Edge - 5G Constraints



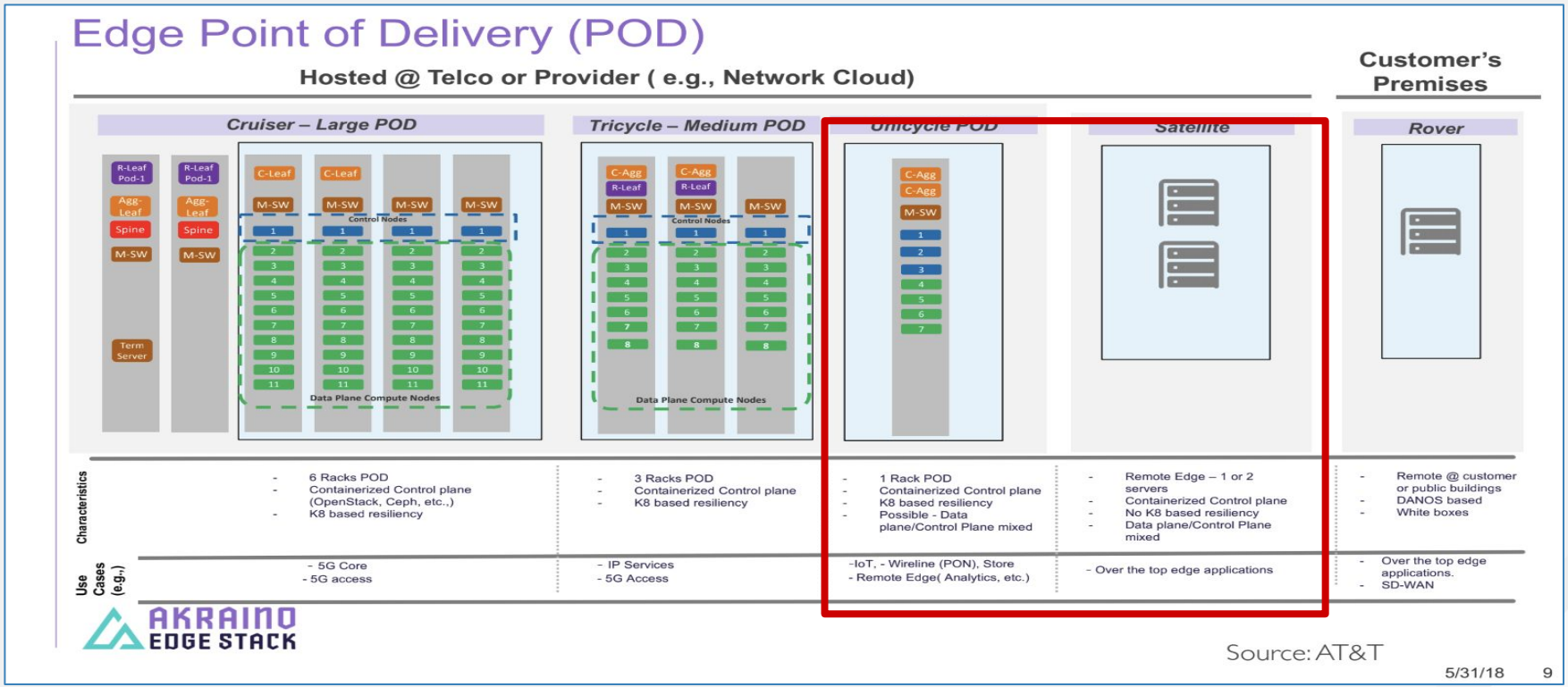
The Edge Factors



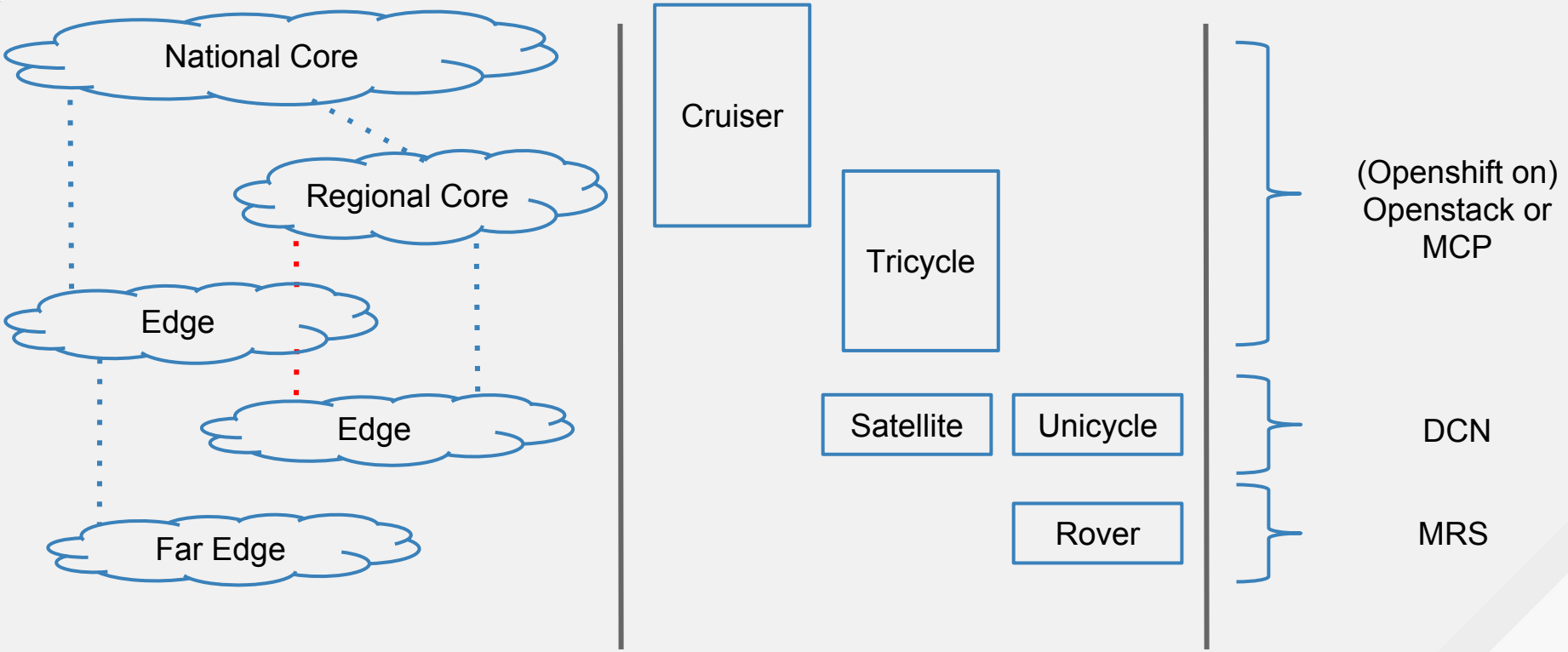
Edge Form Factors



Ceph HCI at the Edge



Edge Form Factors Deployments



Running Ceph at the Edge

Compute and Storage supporting key Apps at the edge

Why a hyper-converged infrastructure?

- Compute AND storage on the same node, better hardware utilization
  - Fits our use case since performance (big computation workloads) is not in the scope

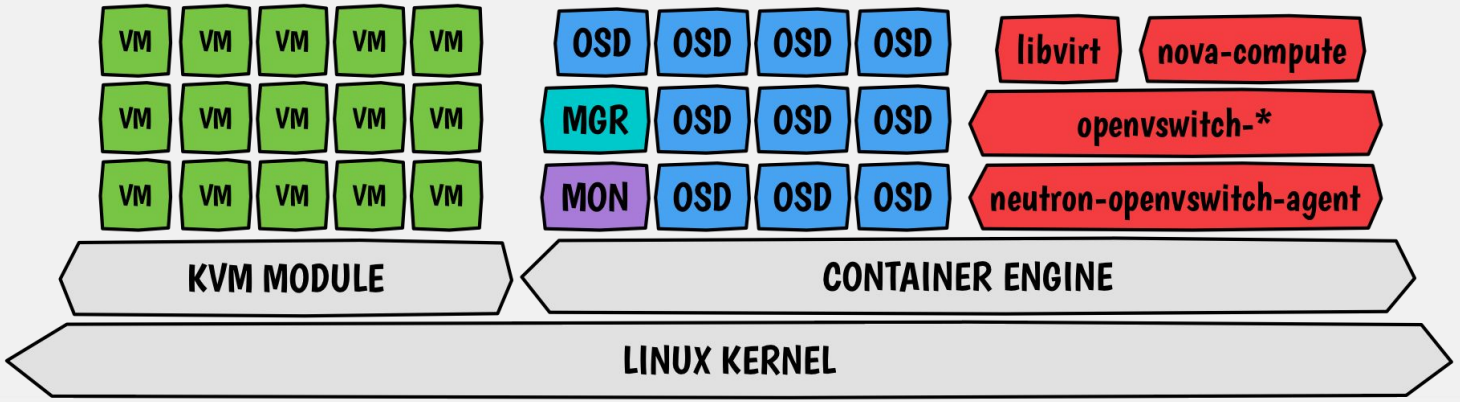
- For small/dense clusters to deliver NFV applications to the edge of your network:
  - Caching and low latency lookups
  - Logging with analytics
  - VNF hosting requiring low network latency

- They also allow for a smaller initial investment for a PoC, Pilots and Dev environments
  - Minimal environment
  - Handy to explore service APIs and storage interfaces

Distributed Compute Nodes with Ceph and containers

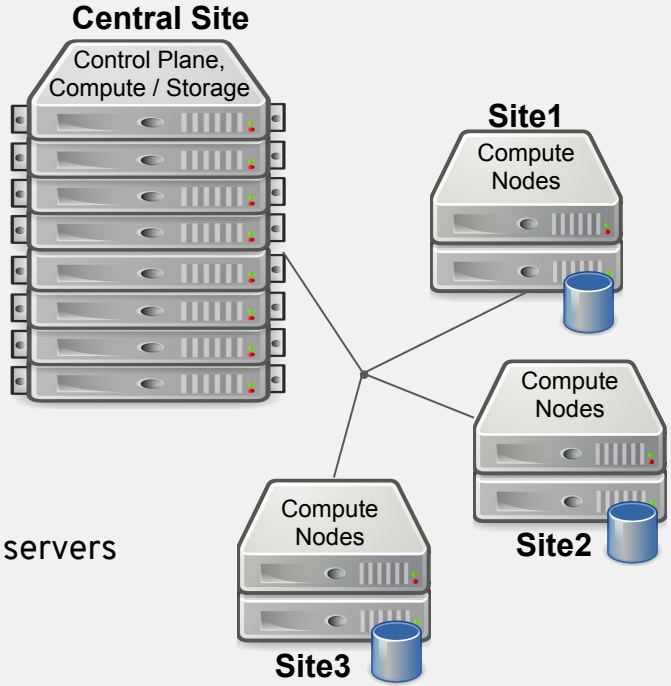
Container's goodness:

- Better control of resource utilization (memory and CPU cap)
- Isolation (with namespaces)
- In-place upgrade/rollback with images



What does Distributed HCI look like?

Distributed Compute Nodes



What:

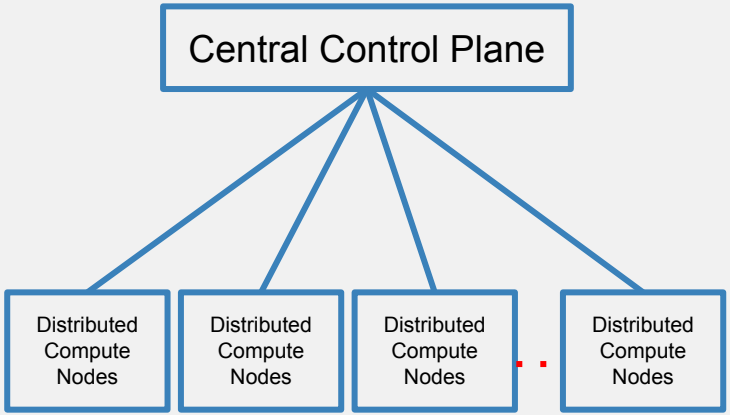
- Centralize control plane on a “Regional” data center
- Deploy smaller HCI remote sites (PoP) close to the edge

Why:

- Reduce Hardware footprint and cost at the edge
  - No OSP control plane needed for a small number of servers
- Utilize compute/storage resources at the edge
  - Put resources closer to where they are needed
- Reduce deployment complexity
  - Only one OSP deployment
- Reduce operational maintenance
  - Same control plane
  - Unified life cycle (scaling, updates, upgrades, etc)

DCN - Image distribution testings

Images will remain on the control place but must be distributed on the DCNs.



How to distribute Cloud images?

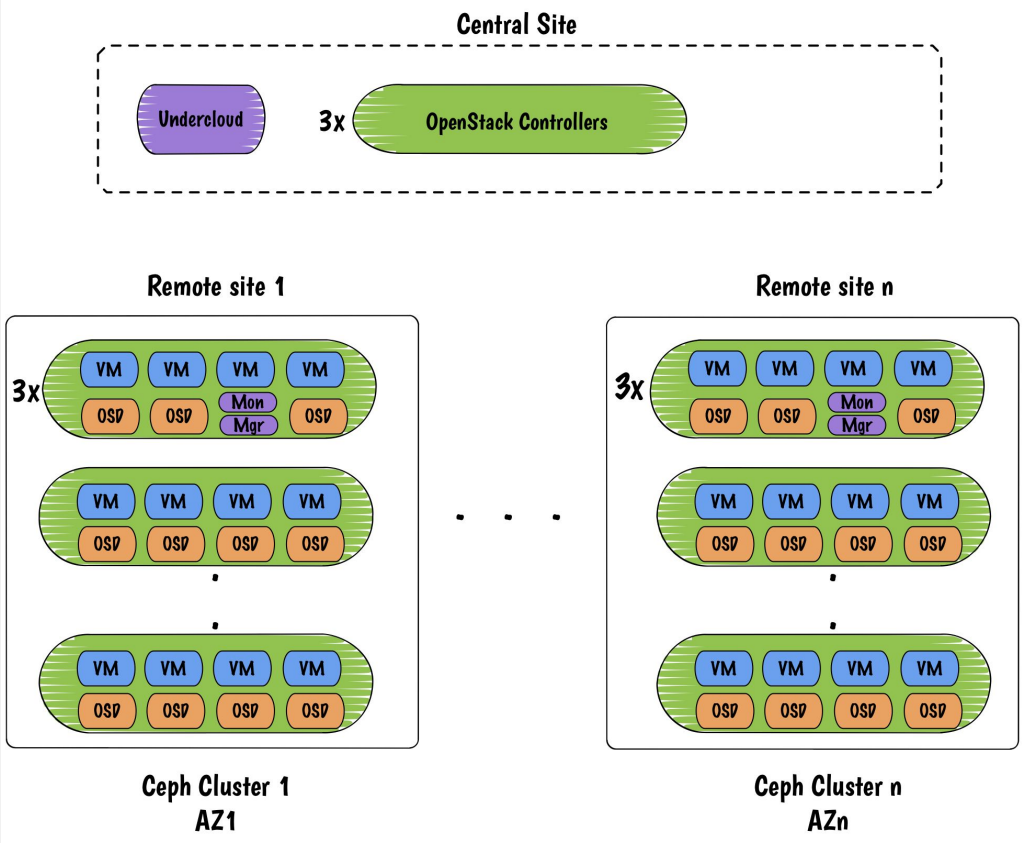
Objectives:

- **1 central control Plane and N remote sites with Compute or HCI nodes**

Testing finds:

- ➢ **Latency - errors over 50ms**
- ➢ **Image Size - errors over 2Gb**

Distributed Compute Nodes with Ceph



How do we get there?

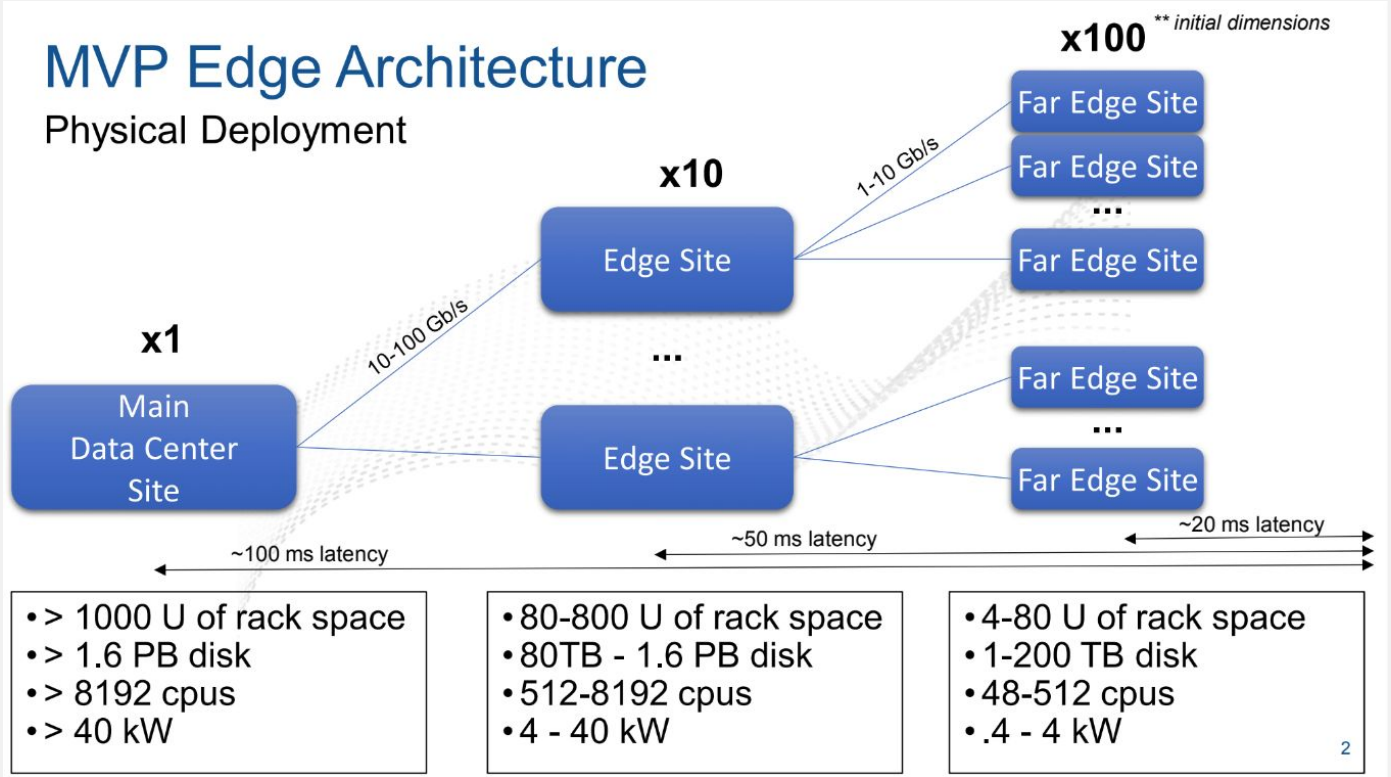
Edge Working Group

At the Denver 2018 PTG the Edge Working Group proposed diagrams for the Mobile service provider 5G usecase: https://wiki.openstack.org/wiki/Edge_Computing_Group

Let's review the design and consider an implementation of it

Note that these slides use the same terms that the Edge Working Group proposed in order to avoid confusion by creating competing terms which might have the same meaning

Edge Working Group: Physical Deployment

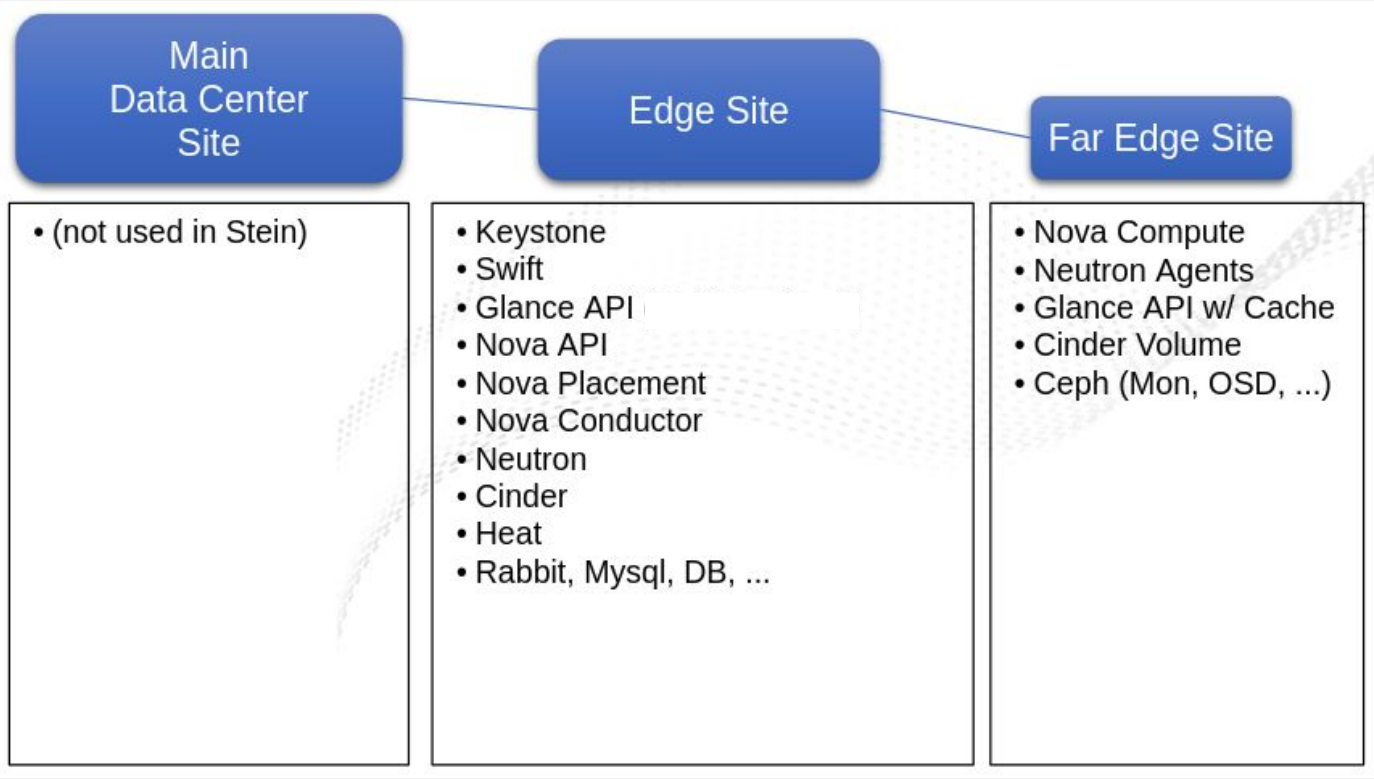


Edge Working Group: Services per Site

Observations

- Main Data Center Site
  - Identity Provider (IdP)
  - Federated Keystone
  - Glance API (for metadata centralization)
- Edge Site
  - The controlplane (basically a TripleO Controller node)
- Far Edge Site
  - Service delivery (basically a TripleO Compute node + persistent storage)

TripleO, Relative to the Edge Working Group



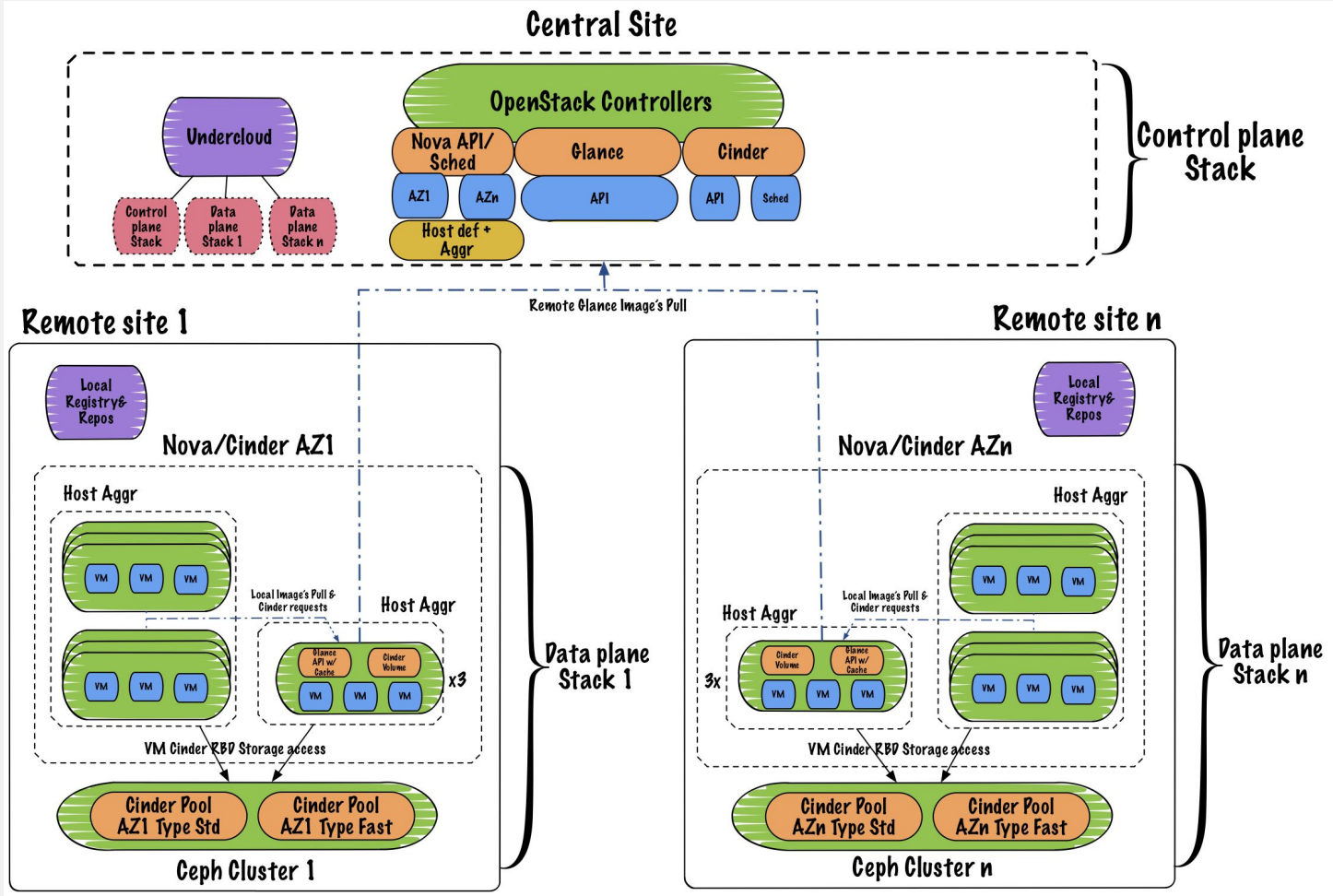
Block Storage in the Far Edge Site

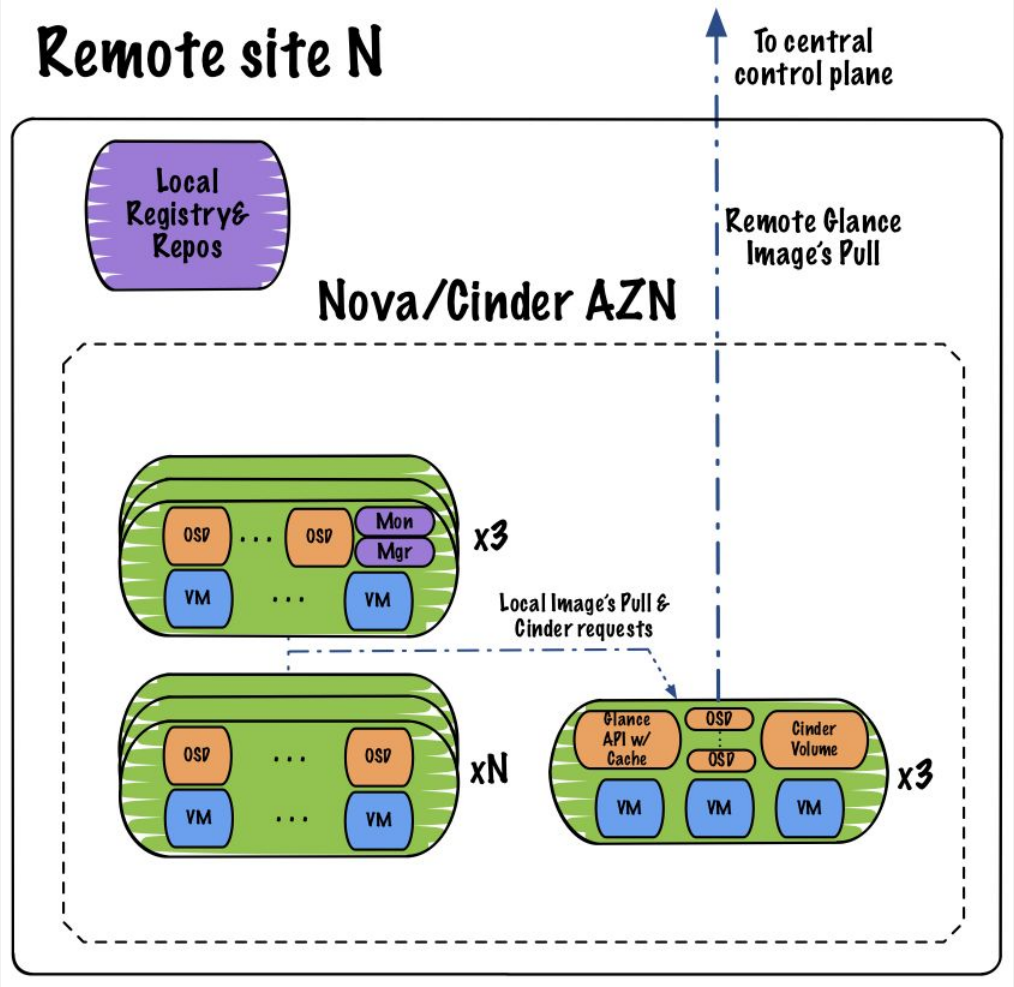
Hyperconverged servers allow us to add storage using existing compute nodes

- A new type of TripleO role for running the following services
  - Nova Compute
  - Ceph OSD, Mon, Mgr
  - Cinder Volume
    - We run this service in active/active mode, (vs active/passive when into controlplane) to not impose hardware requirements necessary to pacemaker
- Each Far Edge Deployment will
  - Stand up its own independent Ceph cluster
  - Be its own Availability Zone for its Edge Controller
  - Scale roles independently from the Edge Site

Image Storage in the Far Edge Site (Stein)

- Deploy Glance on the Edge Controller Node
  - Set the Glance backend to Central Site
- Deploy Glance-API with Cache on the Far Edge HCI Nodes
- Images may be pulled to the far edge via HTTP and then cached
- Plan is for TripleO to be able to deploy this Glance scenario for Stein
  - But it’s a building block to a more advanced image management after Stein


Central Site
Undercloud
Control plane Stack
Data plane Stack 1
Data plane Stack n
OpenStack Controllers
Nova API/ Sched
Glance
Cinder
AZ1
AZn
API
API
Sched
Host def + Aggr
Control plane Stack
Remote Glance Image's Pull
Remote site 1
Local Registry& Repos
Nova/Cinder AZ1
Host Aggr
VM
Local Image's Pull & Cinder requests
Host Aggr
Glance API w/ Cache
Cinder Volume
x3
VM Cinder RBD Storage access
Cinder Pool AZ1 Type Std
Cinder Pool AZ1 Type Fast
Ceph Cluster 1
Data plane Stack 1
Remote site n
Local Registry& Repos
Nova/Cinder AZn
Host Aggr
Local Image's Pull & Cinder requests
Host Aggr
3x
Cinder Volume
Glance API w/ Cache
VM Cinder RBD Storage access
Cinder Pool AZn Type Std
Cinder Pool AZn Type Fast
Ceph Cluster n
Data plane Stack n


Remote site N
To central control plane
Local Registry& Repos
Remote Glance Image's Pull
Nova/Cinder AZN
OSD
OSD
Mon
Mgr
x3
VM
VM
Local Image's Pull & Cinder requests
OSD
OSD
xN
VM
VM
Glance API w/ Cache
OSD
OSD
Cinder Volume
x3
VM
VM
VM

Roadmap and Future Use Cases

Long-term considerations and roadmap for Stein

- Workload survives temporary Edge/Far Edge disconnects
- No storage requirements in the “Far Edge” sites
- HCI with Ceph Monitors using containers resource allocations
- Ability to deploy multiple Ceph clusters with TripleO
- Deploy Glance Cache with TripleO
- Glance Image synchronization and replication mechanisms

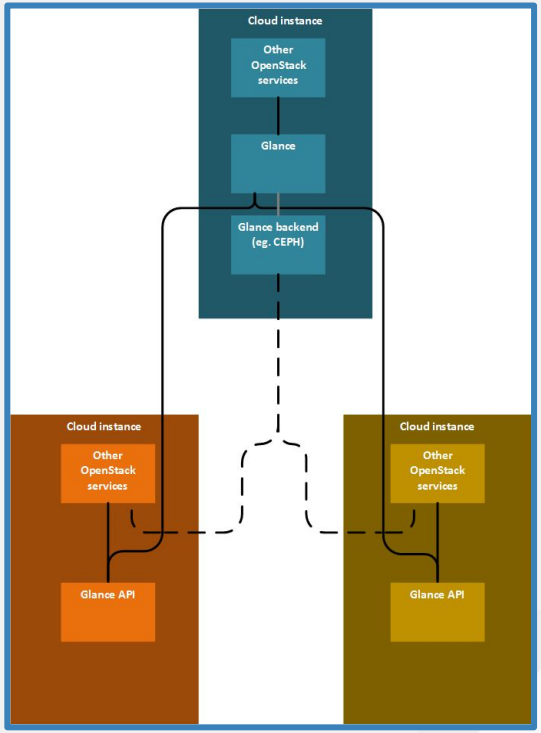
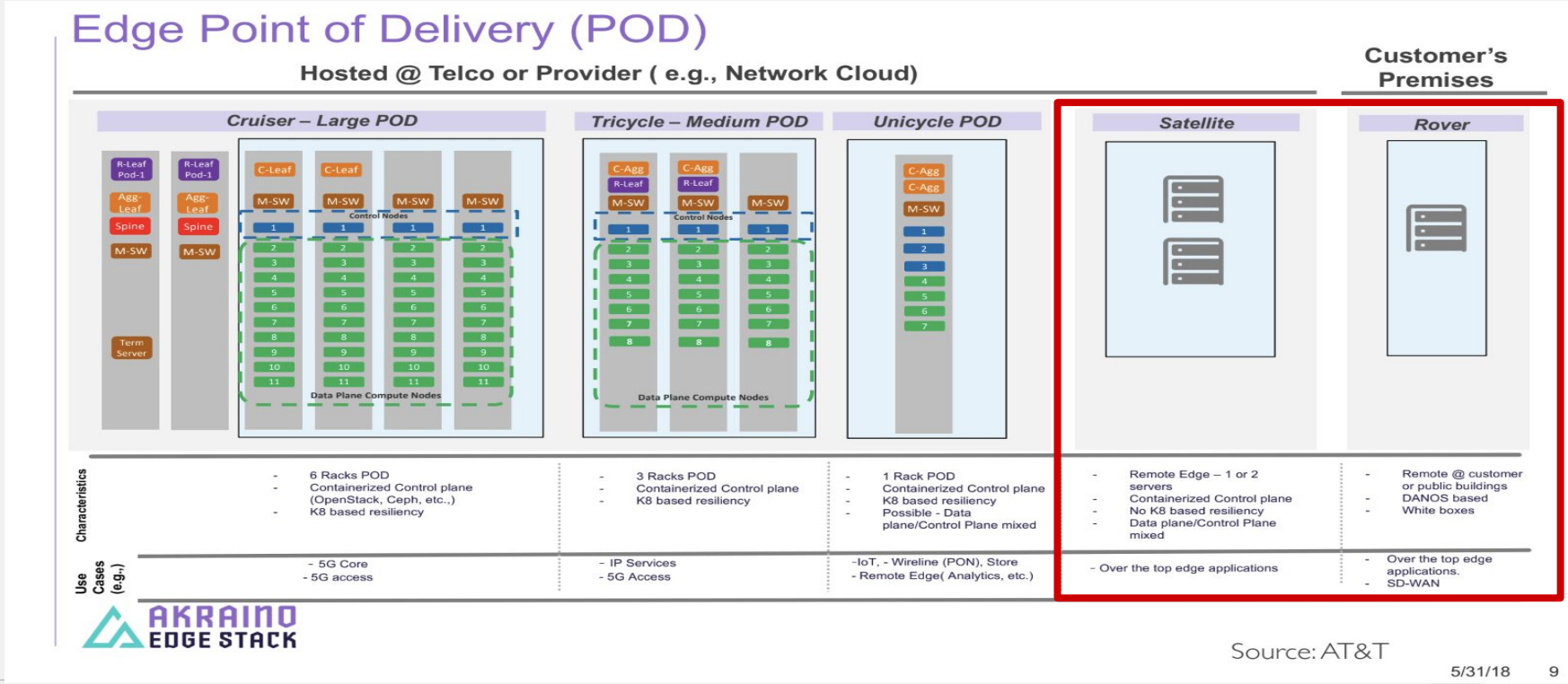
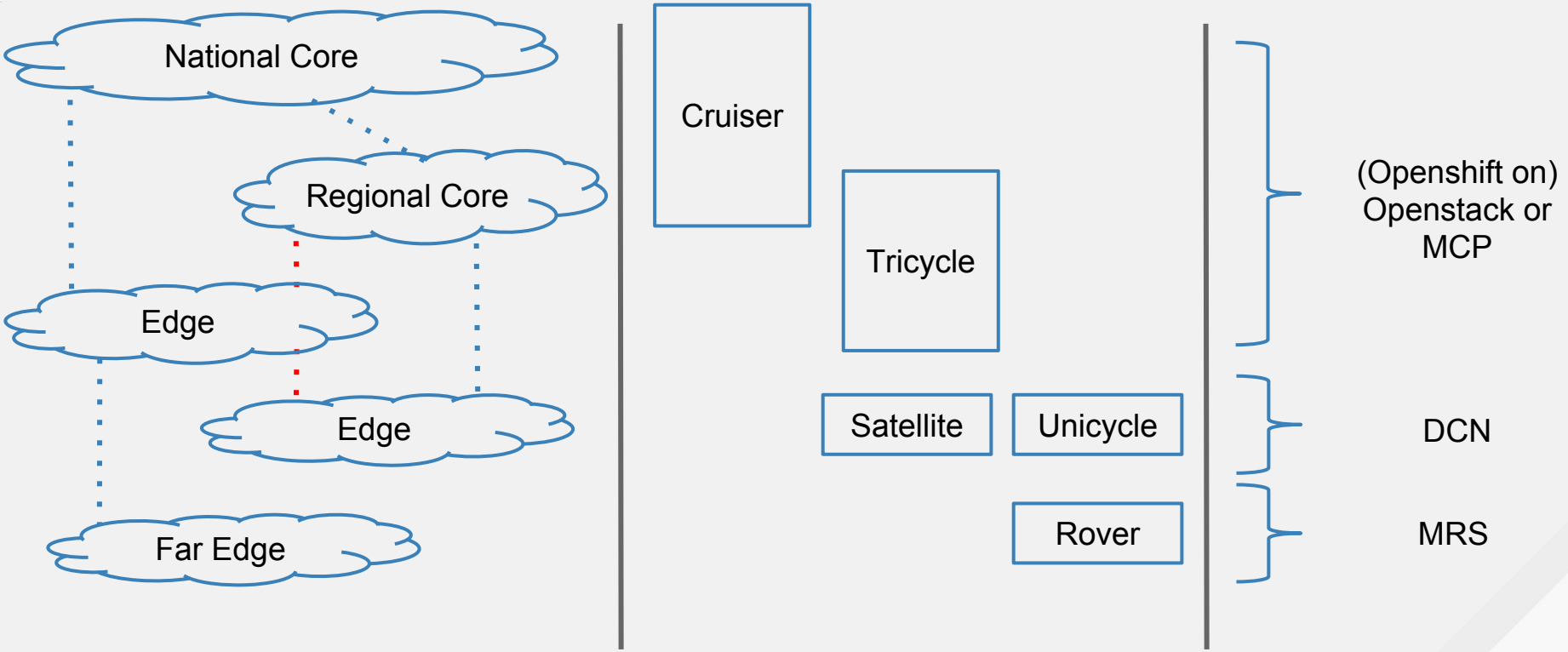


Image handling in edge environment

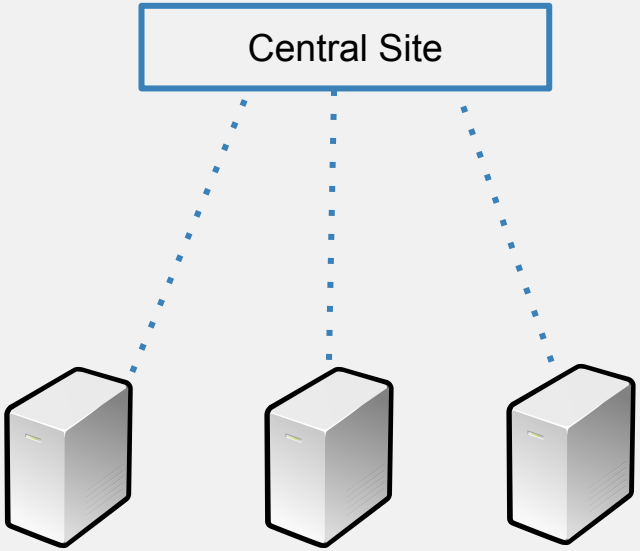
Ceph HCI at the Edge



Edge Form Factors Deployments



MRS - Multiple Remote Standalone (Rover)

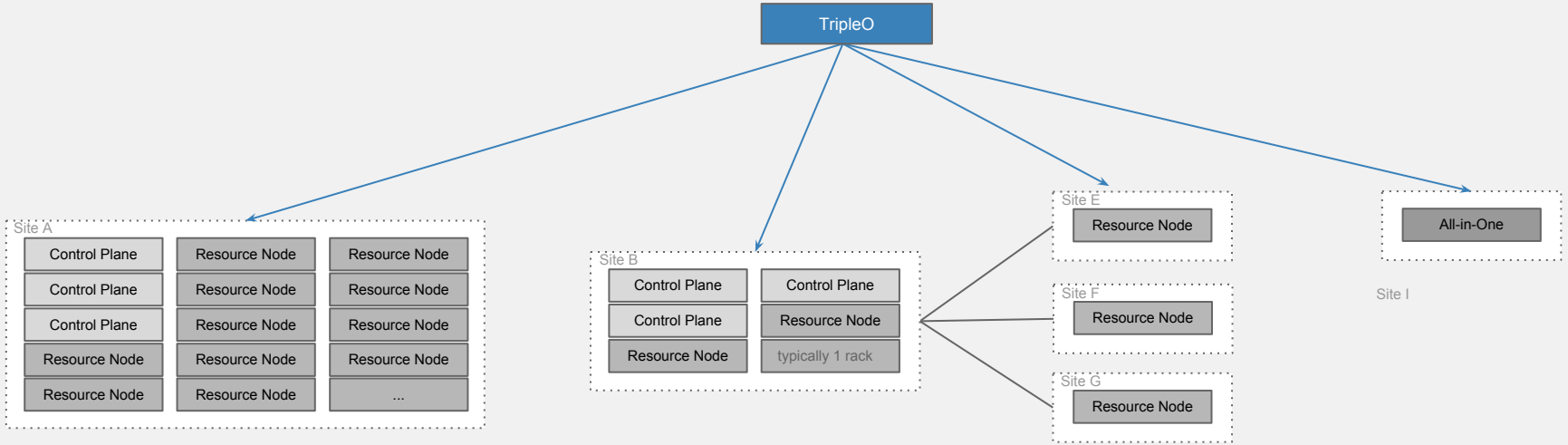


Objectives:

- **Multiple standalone server deployed from 1 single location**
- **Connect to Central Site on-demand to re-synchronise meta-datas**
- **Standalone Servers (control + compute + storage)**

Initial Limitations:
Impact on Lifecycle
No HA

OpenStack at the Edge - Deployment



Large Clusters

5G Core & Access, Data Mining

Close Edge

IP Services, 5G Access, POP, Branch

Distributed Nodes

IoT, Analytics, Wireline, Store

Stand-Alone

In field restore cloud

CLOSE EDGE — FAR EDGE

Takeaways

Wanna learn more?

Resources:

- Edge Woking Group
  - IRC Channel on Freenode - #edge-computing-group
  - Mailing list - http://lists.openstack.org/cgi-bin/mailman/listinfo/edge-computing
  - Wiki - https://wiki.openstack.org/wiki/Edge_Computing_Group
- Stein Specs
  - TripleO Split Control Plane from Compute/Storage Support
  - Image handling in edge environment
  - Distributed Ceph HCI
- Stein PTG Edge pads
  - https://etherpad.openstack.org/p/EdgeComputingGroupPTG4
  - https://etherpad.openstack.org/p/tripleo-ptg-stein-edge

@leseb_
@SeanCohen_RH
giuliofidente.com

redhat.

THANK YOU

plus.google.com/+RedHat

facebook.com/redhatinc

linkedin.com/company/red-hat

twitter.com/RedHatNews

youtube.com/user/RedHatVideos