

New OCP Reference Architectures for AI Networking

OCP Educational Webinar Series

June 2, 2026



Community-driven hyperscale innovation for all

New OCP Reference Architectures for AI Networking

James Kelly

VP Market Intelligence
Open Compute Project



Alan Weckel

Founding Analyst
650 Group



Marc Austin

CEO
Hedgehog



Agenda Overview

AI Networking Market

Challenges with Proprietary Networks

Open Network Abstraction

CloudCo Use Case

Hedgehog Overview

Take Aways

Live Q&A

Agenda: AI Networking Market

AI Networking Market

Challenges with Proprietary Networks

Open Network Abstraction

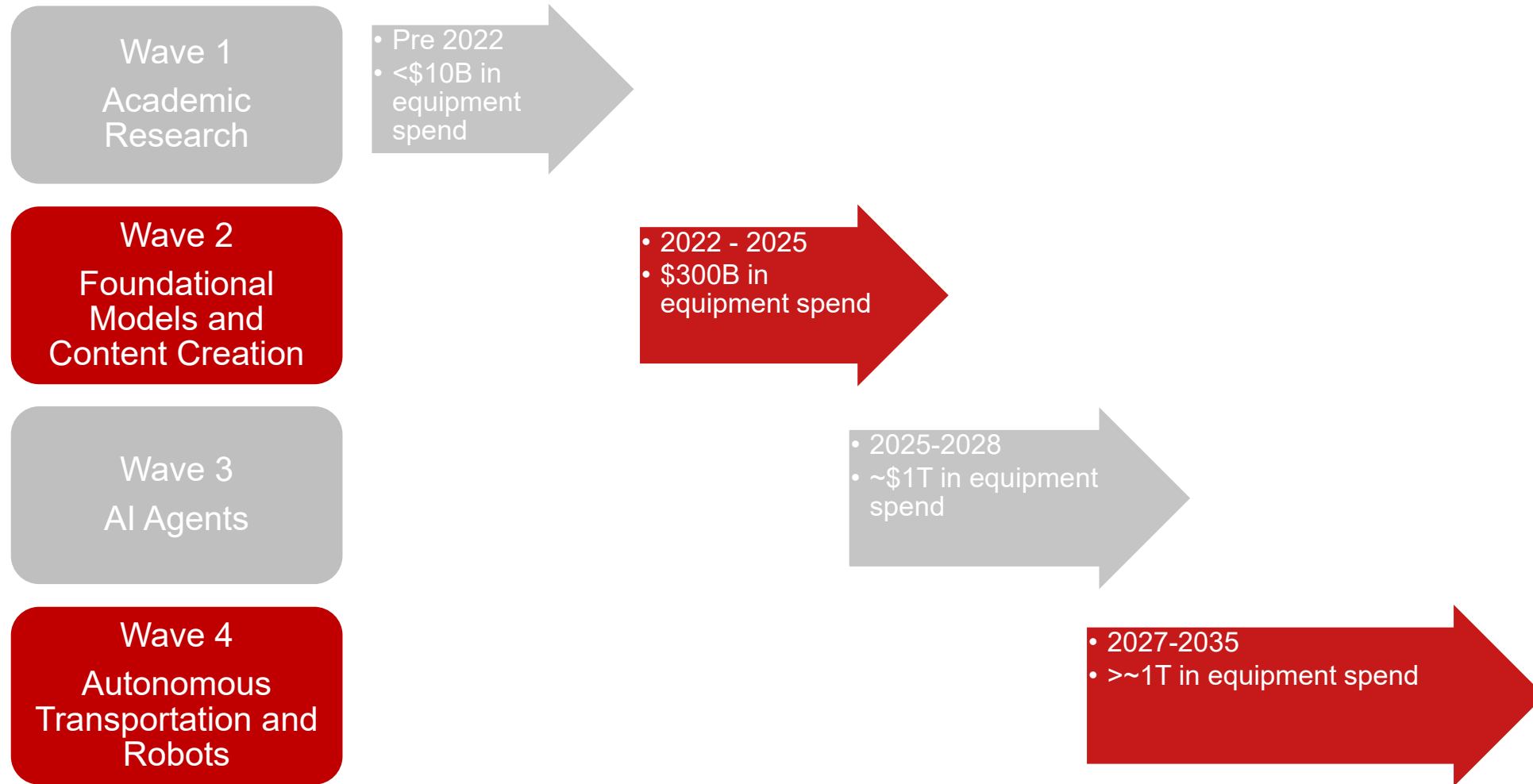
CloudCo Use Case

Hedgehog Overview

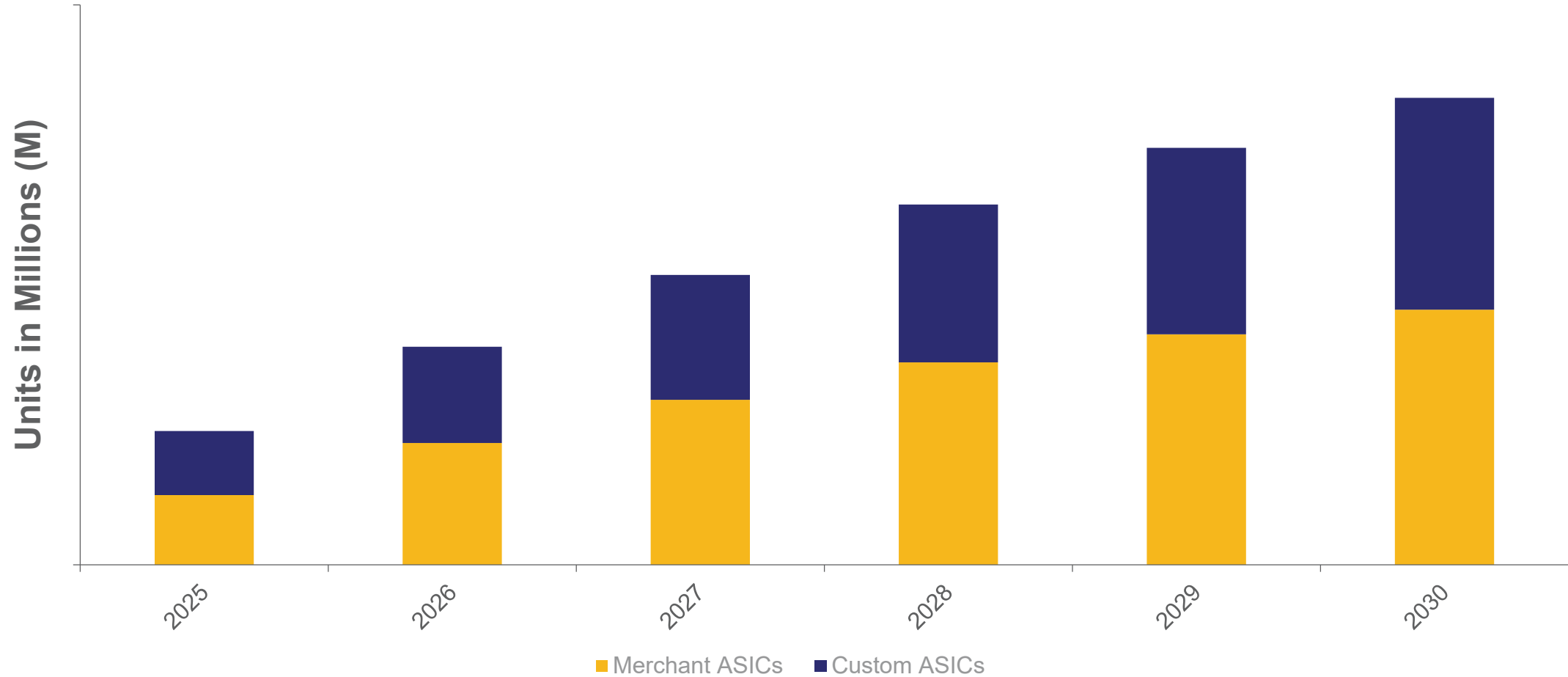
Take Aways

Live Q&A

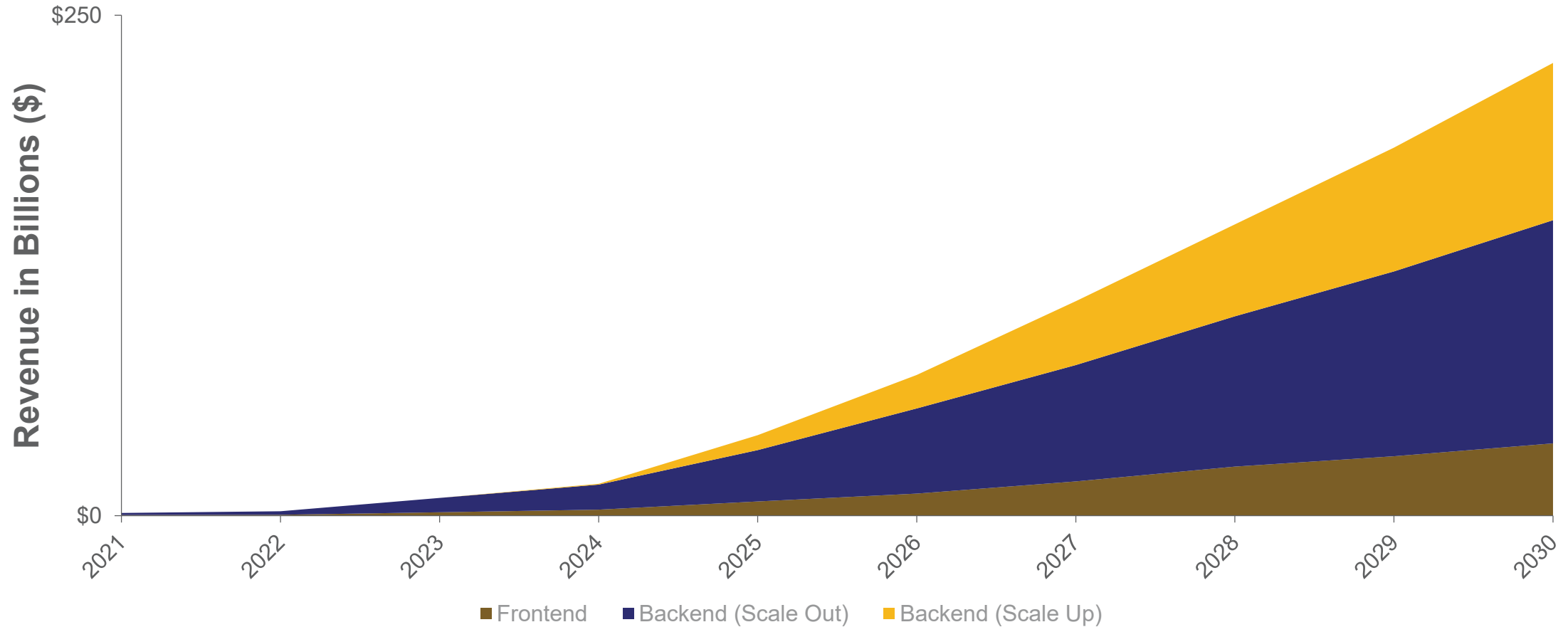
AI Waves



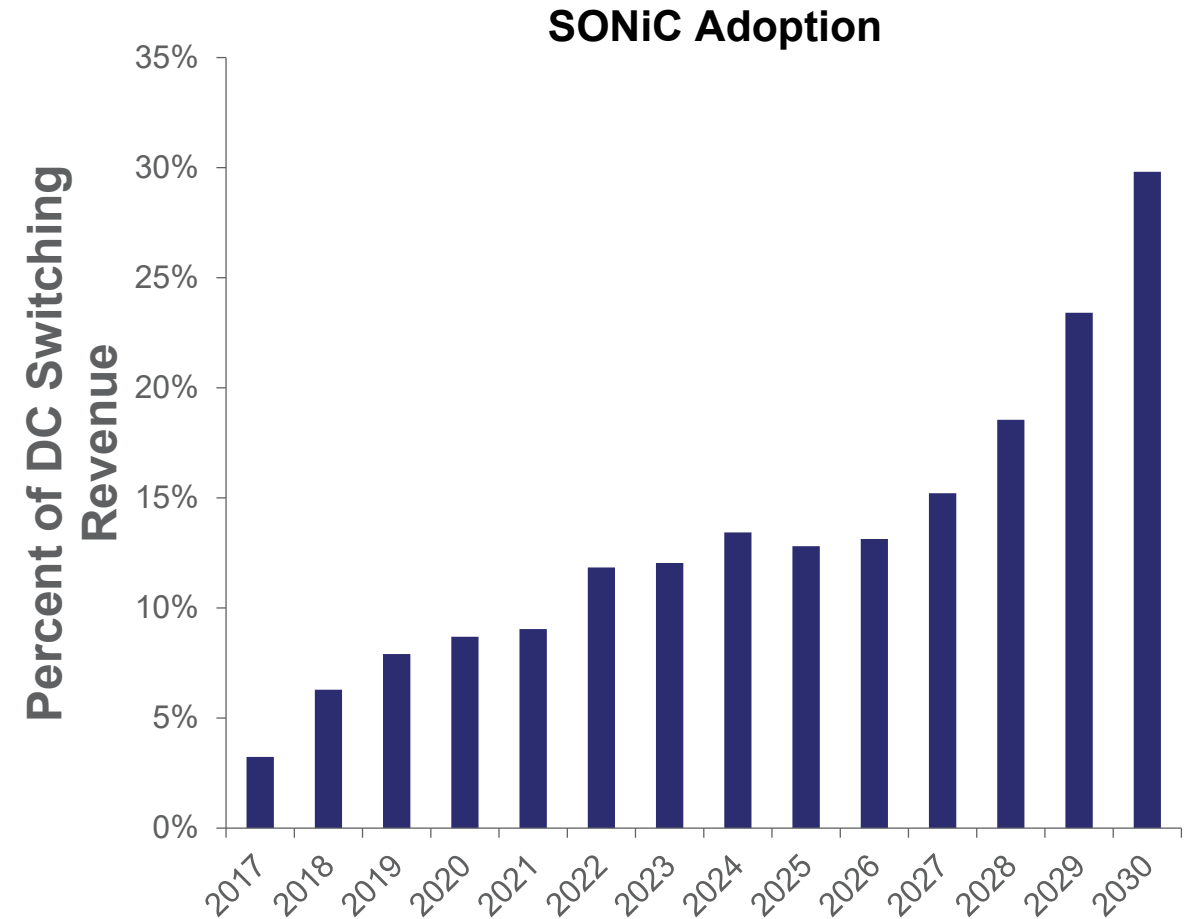
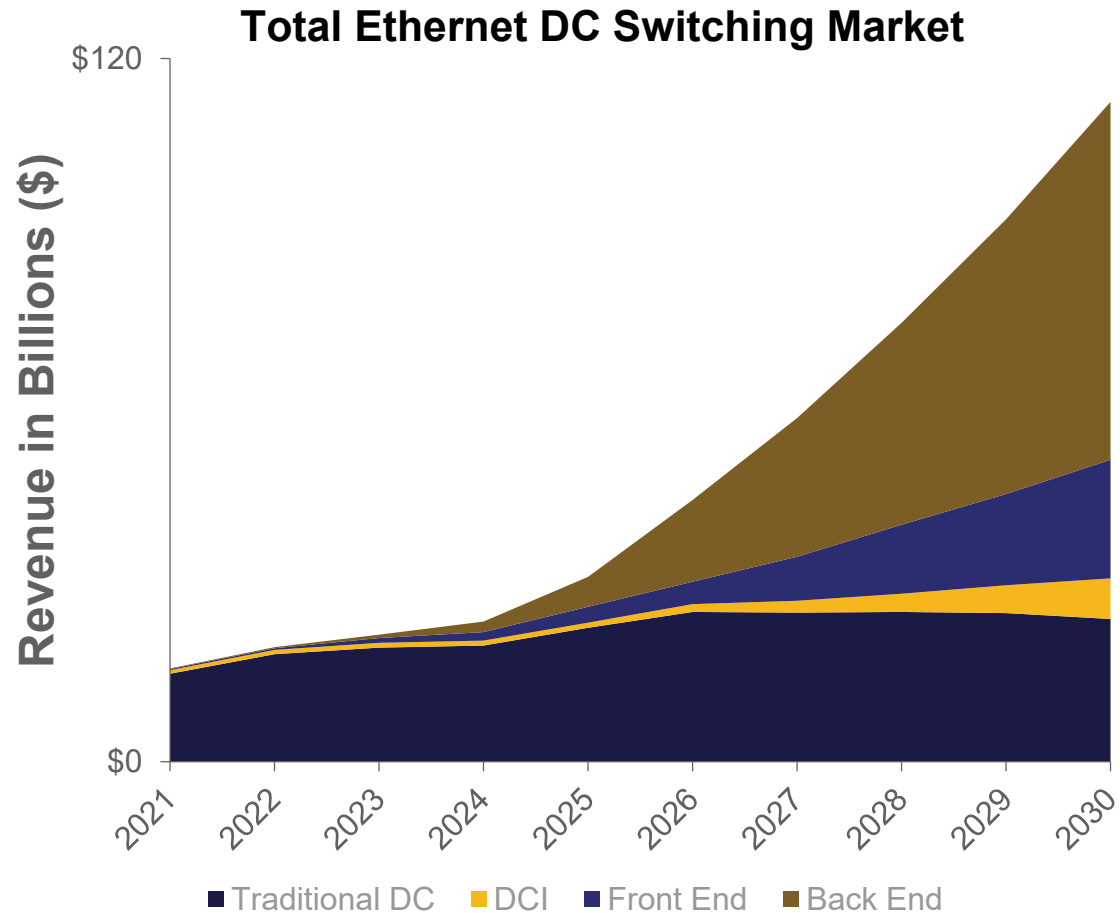
GPU/XPU Shipments



AI Networking Tied to GPU/XPU Growth



Ethernet Revenue and SONiC Growth



Audience Poll Question 1

- Are you planning to operate AI infrastructure for multiple tenants?
 - Single Tenant Only
 - Multi Tenant
 - Both
 - Don't Know

Agenda: Challenges with Proprietary Networks

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

CloudCo Use Case

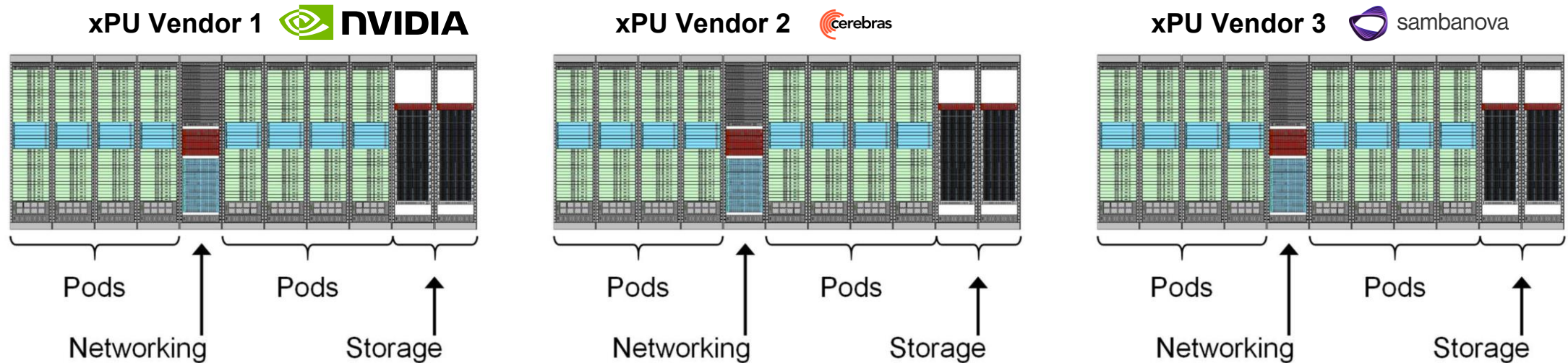
Hedgehog Overview

Take Aways

Live Q&A

Proprietary Networks Create Cluster Complexity

- xPU vendors want to sell their AI accelerators as POD-scale solutions
- PODs with proprietary networks create vertically integrated silos
- POD silos fail to operate as a single CLUSTER



Audience Poll Question 2

- How many accelerator types are you planning to deploy over time?
 - ☐ 1
 - ☐ 2
 - ☐ 3
 - ☐ 4
 - ☐ 5 or more

Agenda: Open Network Abstraction

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

CloudCo Use Case

Hedgehog Overview

Take Aways

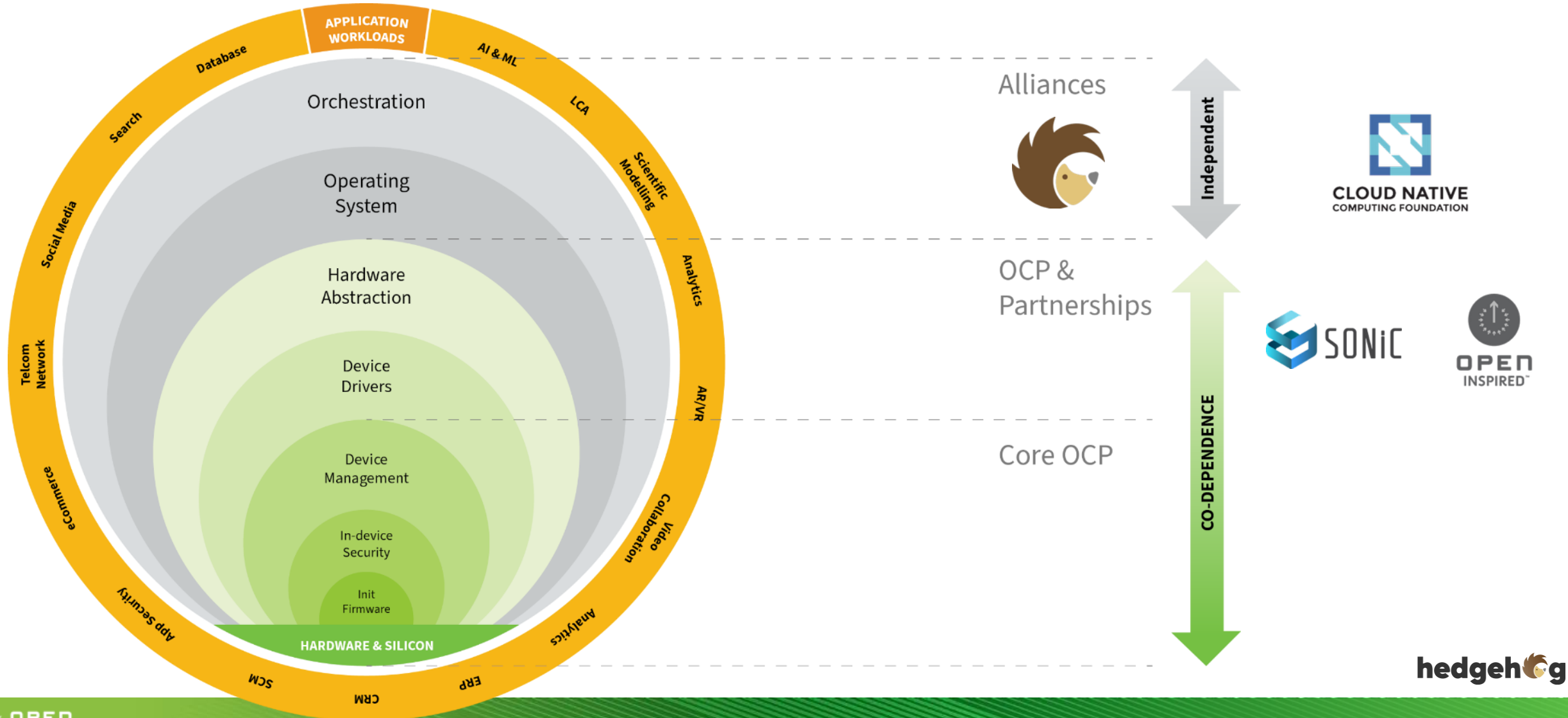
Live Q&A

Solution: Open Network Abstraction

- Hyperscalers have solved multi-vendor complexity with open networking and VPC abstractions
- Tenants use VPCs with multiple xPU instance types



Open Source Software



Community-driven hyperscale innovation for all

Open Specification Hardware Choice

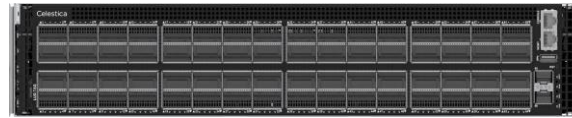
2U 64-port 800GbE Data Center Switch

AI / ML & Big Data Analytics

Hyperscale Data Centers &
Cloud Computing

High-Performance
Computing (HPC)

Network Backbone (800GbE
Data Center Leaf / Spine)



Celestica DS5000



Dell Z9684F-ON



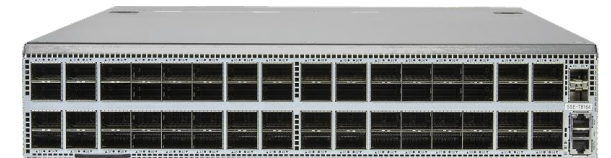
NextHop AI 4010



Cisco Nexus 9364-E-SG2



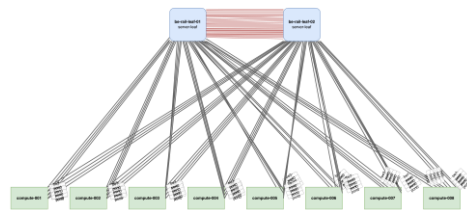
Edgecore AIS800-64D



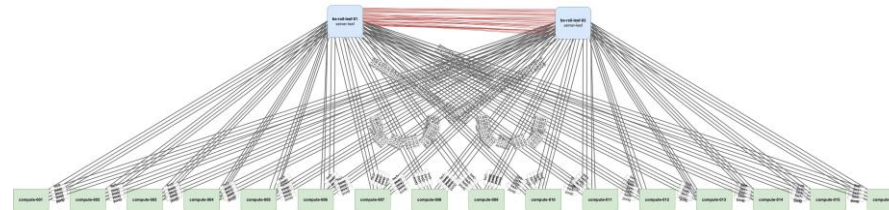
Supermicro SSE-T8164

Everyone Can Network Like a Hyperscaler

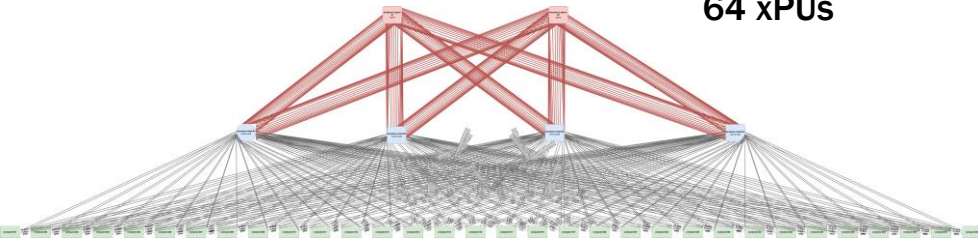
- OCP RAs provide a prescriptive scale-out Ethernet fabrics with Open Pod Group scalable units ranging from 64-1024 xPUs
- Includes RoCEv2 QoS, rail-optimized topology, dual-plane resilience, VPC abstraction, full lifecycle automation, and roadmap to Ultra Ethernet



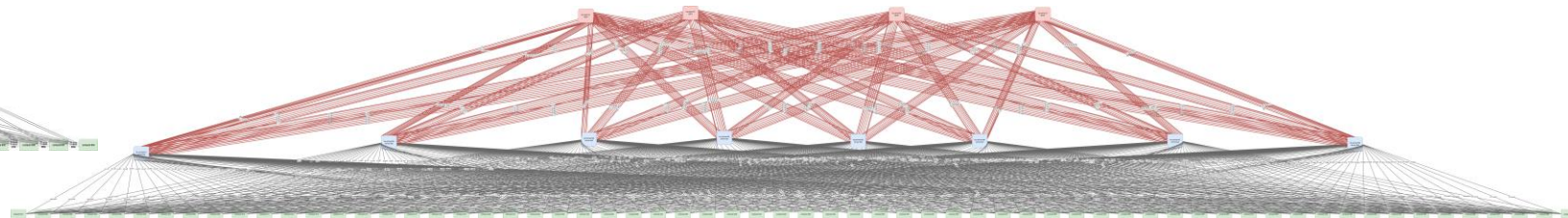
64 xPUs



128 xPUs



256 xPUs



512 xPUs

Complete BOM Available on OCP Contributions & Marketplace

- Reference architecture white papers for training and inference
- Topologies for common scalable units
- OCP hardware
- Open-source software appliance
- Infrastructure as code

Contributions

Welcome to the OCP Contributions Database. All contributions listed here have undergone a comprehensive validation process and are available for you to utilize in your own infrastructure. See what you can learn from these contributions, build on their expertise, share your experiences and give back to the Community. [Join in the movement and contribute today!](#)

× Q

Go to the Marketplace >

Type —
☐ Reference Architecture (2)

Project —
☐ Strategic Initiatives > Open Cluster Designs for AI (2)

Contributor —
☐ Celestica (2)
☐ Hedgehog (2)

Sort by: Most Recent ▾

SI, Open Cluster Designs Aligned AI Inference Fabric Reference Architecture
Open Cluster Designs Aligned AI Inference Fabric Reference Architecture

Contributor: Hedgehog, Celestica
Family: Information
Project: Strategic Initiatives > Open Cluster Designs for AI
Release Year: 2026




Audience Poll Question 3

- How many xPU are you deploying?
 - 64 (or less)
 - 128
 - 256
 - 512
 - 1024
 - 2048 (or higher)

Agenda: CloudCo Use Case

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

CloudCo Use Case

Hedgehog Overview

Take Aways

Live Q&A

CloudCo Use Case

- Focus on low-cost, low-latency inference for multiple enterprise tenants
- Offers AI accelerators from xPU vendor 1, xPU vendor 2, xPU vendor 3
- Considered a proven proprietary network or new OCP network

Reference Architecture Business Case

\$2.2M capex savings, **\$42.2M** EBITDA impact, **10 mos** faster time to market*

Approach	Design	Procure	Deploy	Operate	Performance	Reliability	Security	Price
Do It Yourself	\$535K capex •\$321/hr •9.5 person months •3 mos recruiting •5 mos executing •Lead Architect •RDMA/Storage Specialist •Security Architect •Automation/SONiC Engineer •Validation/Perf Engineer	\$1.65M capex •Perimeter security required •200gbs gateway router required	\$720K capex \$10.5M opex •2 mos time to GPU value •\$3M GPU carry cost (I+D) •\$7.5M opportunity cost	\$2.2M opex •8 ops staff •2 SREs •2 netops engineers •2 automation engineers •2 remote hands	\$8M lost rev •70% GPU utilization = 85% peak utilization - 15% MFU penalty	\$1.6M lost rev •88hr ave MTTR •120 hrs congestion •24 hrs switch failure •72 hrs fabric perf	\$7.7M opex •\$6.2M downtime lost rev •\$1.5M IR + ransom •10 tenants •40 annual incidents •No network isolation •100% blast radius	\$45M revenue •ClusterMax Bronze •\$6/hr
Use OCP Reference Architecture	\$35K capex •.25 person months •Lead Network Architect	\$65K capex •200gpbs secure gateway included	\$0 capex \$1.5M opex •1 week time to GPU value •Free Hedgehog consulting services •\$300K carry cost •\$1.2M oppty cost	\$900K opex •4 ops staff •2 SREs •2 remote hands	\$0 opex •85% GPU utilization •See Semi Analysis FarmGPU NCCL test data presented at OCP Global Summit 2025	\$60K lost rev •2hr ave MTTR •2 hrs congestion •1 hr switch failure •6 hrs fab perf	\$300K opex •\$65K downtime lost rev •\$235K IR + ransom •10 tenants •40 annual incidents •VPC isolation •10% blast radius	\$60M revenue •33% goodput premium •ClusterMax Silver •\$8/hr
CapEx Impact	•\$500K	•\$1M	•\$720K					
EBITDA Impact			•\$9M	•\$1.3M	•\$8M	•\$1.5M	•\$7.4M	•\$15M
Time to Market Impact	•8 mos		•2 mos					

*figures assume 1024 B200 cluster with DIY Ansible automation and reliance on containers for tenant isolation



Plan Your Own AI Network

visit

<https://plan.hedgehog.cloud>

configure your own assumptions for DIY compared to NVIDIA or OCP reference architectures

Agenda: Hedgehog Overview

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

CloudCo Use Case

Hedgehog Overview

Take Aways

Live Q&A

Hedgehog

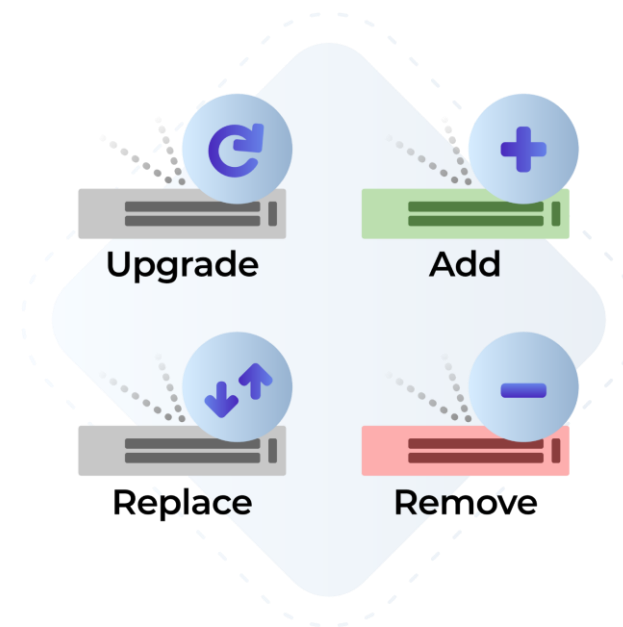
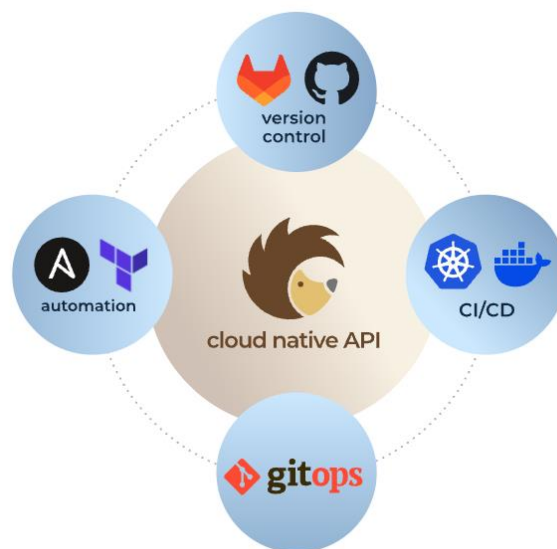
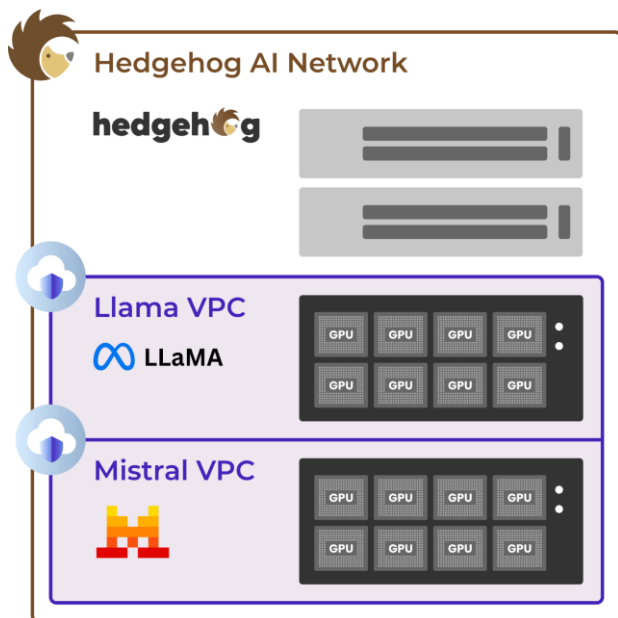
- Hedgehog offers a choice of validated AI network reference architectures



OPEN
Compute Project®

Hedgehog

- with a common VPC abstraction for multi-vendor xPU networks
- a cloud-native API for multi-tenant cloud user experience
- and full lifecycle automation



Agenda: Take Aways

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

Neocloud Use Case

Hedgehog Overview

Take Aways

Live Q&A

Take Aways

- SONiC and OCP enable an Ethernet networking fabric that is GPU/XPU agnostic
 - Easier for customers to deploy XPU from multiple vendors
 - Reduced network complexity and rework with each new accelerator
- Faster time to onboarding new XPU
 - Less human capital needed when the fabric stays the same
 - Consistent tools, security, and observability
- OCP Reference Architectures allow for consistent deployments
 - All customers can deploy like Hyperscalers
 - Reduced time before first token

Agenda: Live Q&A

AI Network Market

Challenges with Proprietary Networks

Open Network Abstraction

Neocloud Use Case

Hedgehog Overview

Take Aways

Live Q&A

Live Q&A

James Kelly

VP Market Intelligence
Open Compute Project



Alan Weckel

Founding Analyst
650 Group



Marc Austin

CEO
Hedgehog



Thank you for attending

New OCP Reference Architectures for AI Networking

OCP Educational Webinar Series

June 2, 2026



Community-driven hyperscale innovation for all