Betting on the Network: Meta's Case for Full-Stack, Open AI Networking

Omar Baldonado

Senior Director, Data Center & AI Networking

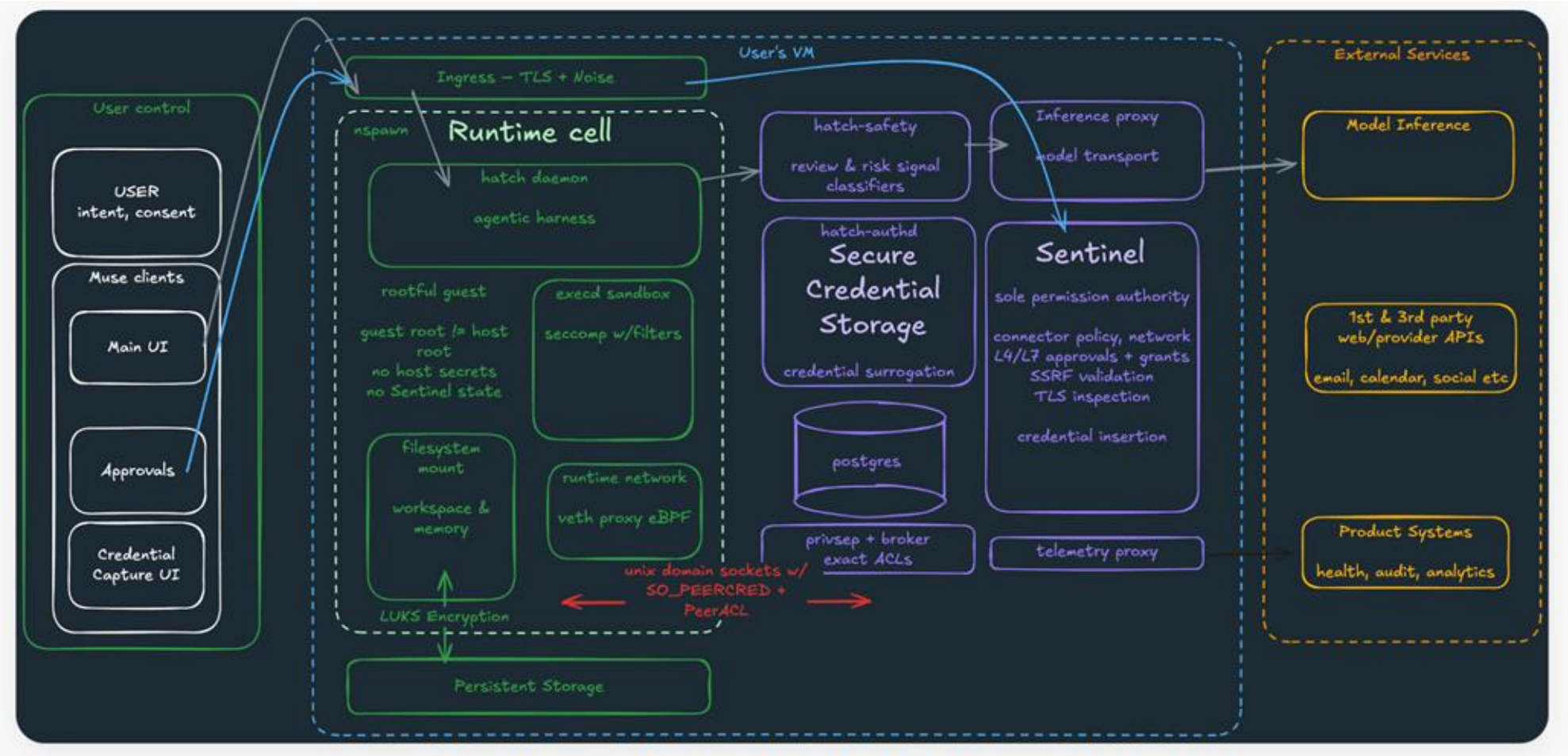
Meta

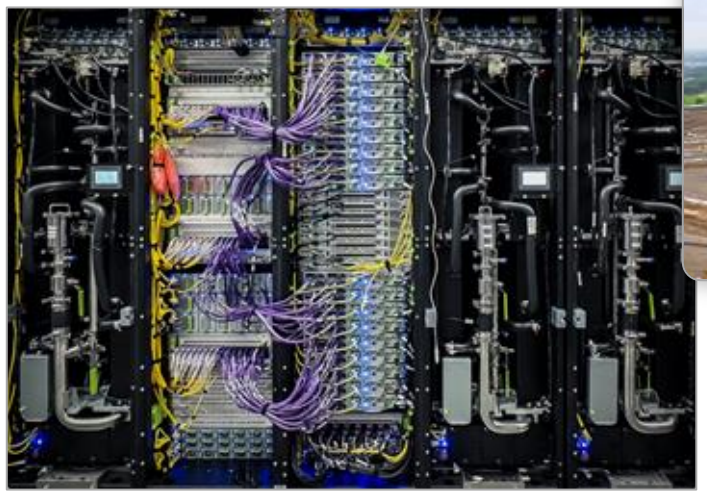
The Leap: Muse, an AI agent for everyone



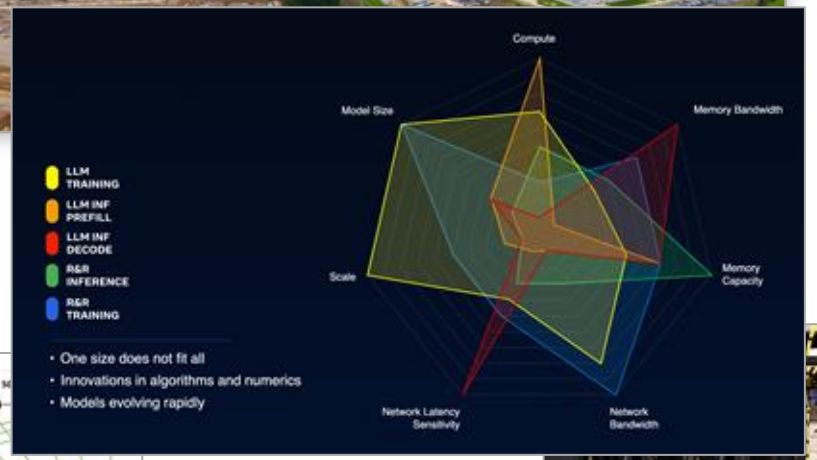


AI Infra has become the Infra

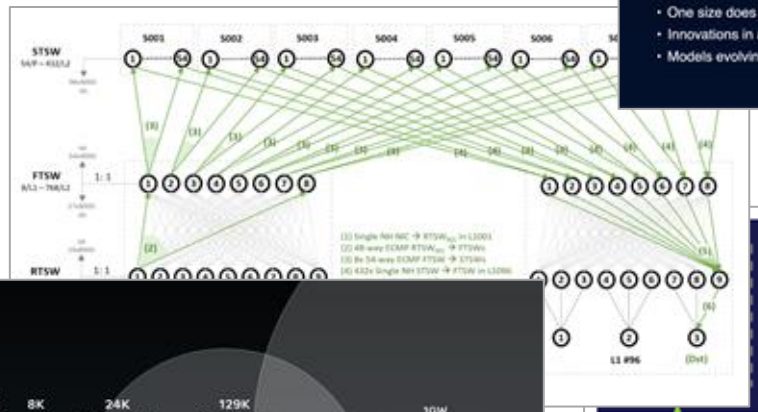






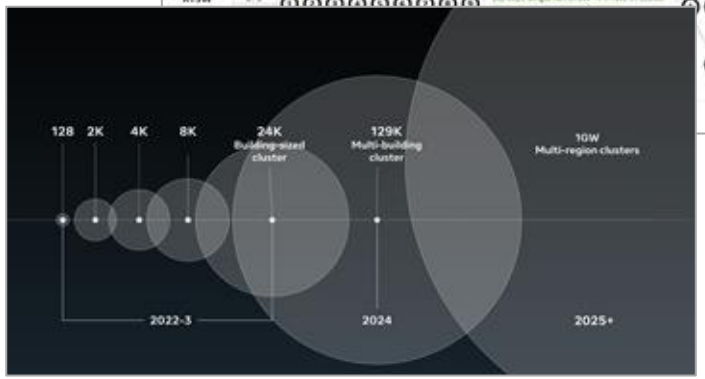

Compute
Model Size
Memory Bandwidth
Scale
Memory Capacity
Network Latency Sensitivity
Network Bandwidth
LLM TRAINING
LLM INF PREFILL
LLM INF DECODE
R&R INFERENCE
R&R TRAINING
• One size does not fit all
• Innovations in algorithms and numerics
• Models evolving rapidly

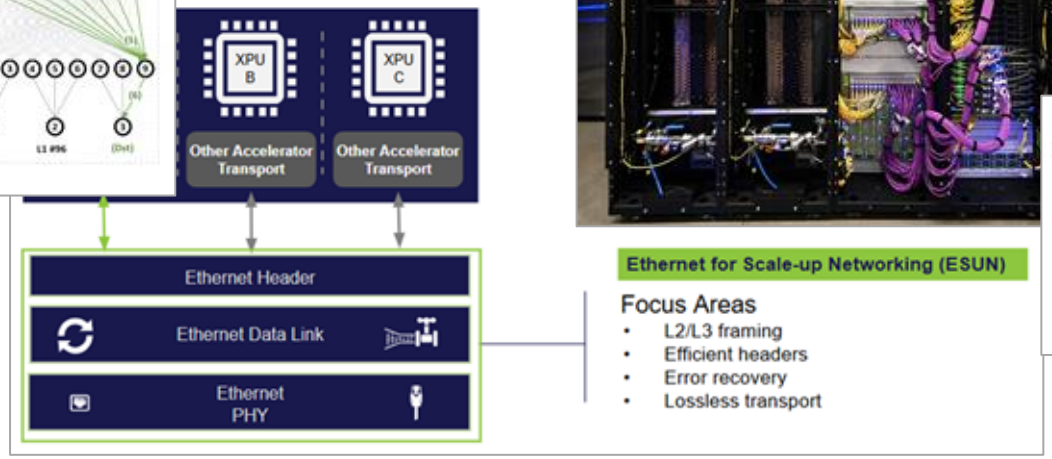

FBOSS
sai
Tomahawk 5
CISCO
G200
51.2T
Minipack3
Cisco 8501


STSW
FTSW
RTSW




NVIDIA


128
2K
4K
8K
24K
Building-sized cluster
129K
Multi-building cluster
1GW
Multi-region clusters
2022-3
2024
2025+


XPU B
XPU C
Other Accelerator Transport
Other Accelerator Transport
Ethernet Header
Ethernet Data Link
Ethernet PHY
Ethernet for Scale-up Networking (ESUN)
Focus Areas
• L2/L3 framing
• Efficient headers
• Error recovery
• Lossless transport



Open
Full-Stack
Networking
Innovation

The Scope of Network



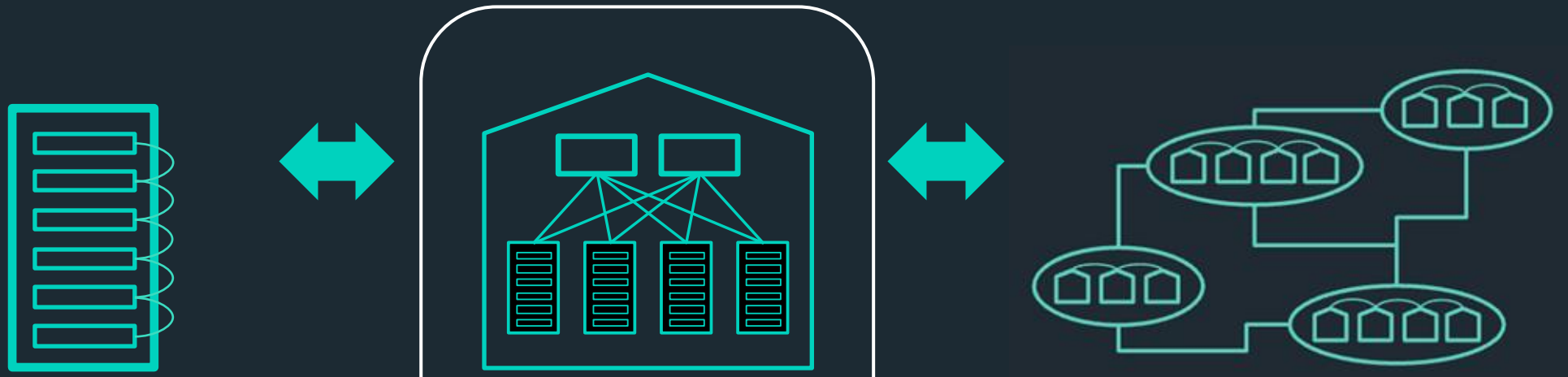
Scale-up /
Rack-level

Scale-out /
Building-level

Scale-across /
Beyond building

Shared hardware & software

Scale-out



Scale-up /
Rack-level

Scale-out /
Building-level

Scale-across /
Beyond building

Shared hardware & software

102.4T Scale-Out: IcePack & IceCube / N-IceCube

- **IcePack**: our first liquid-cooled network rack for high power density & thermal efficiency (based on ORv3-HPRv1)

- **IceCube / N-IceCube**: our 102.4T switch, with **Broadcom** TH6 / **NVIDIA** Spectrum-6, running FBOSS + SAI software (systems from **Celestica** and **Accton)**



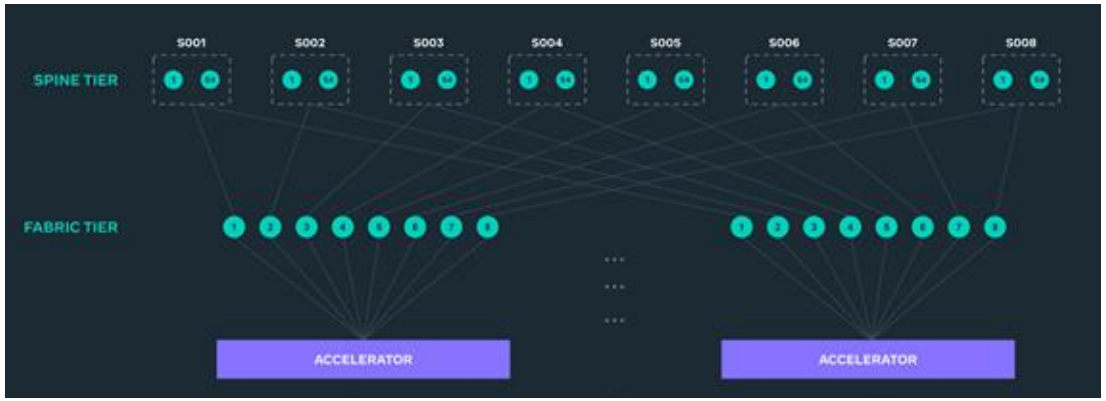
IcePack



IceCube / N-IceCube

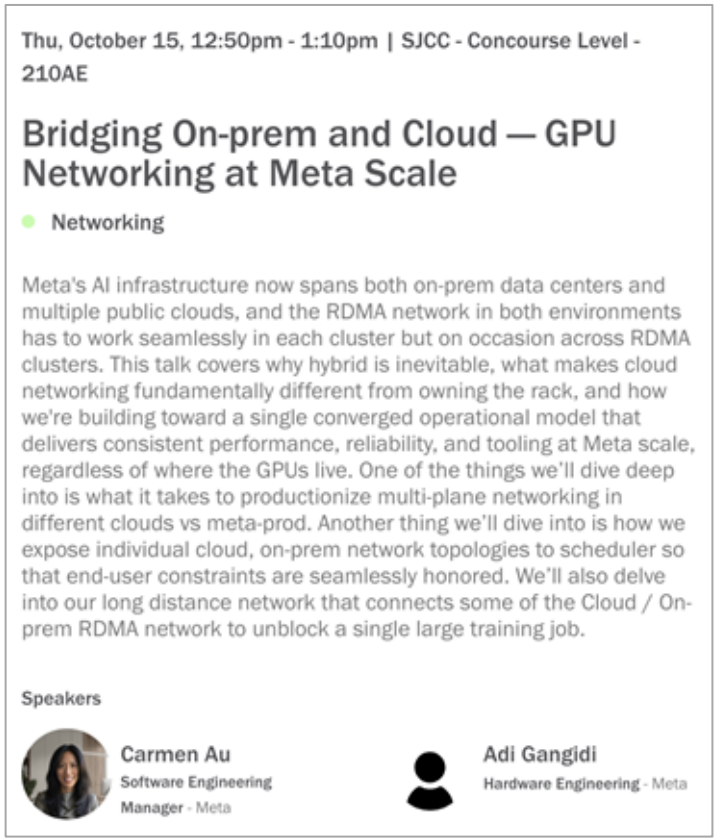
Flattened Planar Fabric: Our next scale-out fabric

- **“Multi-plane fabric”** that connects accelerators to multiple fabric planes (removing top-of-rack)
- 2-tier => fewer switches/optics & lower latency
- NIC-level load balancing
- Combined switch & host control plane
- Close partnership with **NVIDIA**



Working with NVIDIA & clouds on multi-plane

- Leveraging Meta's unique position as both **hyperscaler with 1P infra** and **cloud customer**
- Saw heterogeneous multi-plane implementations in clouds, e.g., load balancing, topologies, failure domains
- Sharing design & operational learnings from Meta's and **NVIDIA's** FPF development
  ⇒ Benefit joint cloud partners with higher reliability and performance



Thu, October 15, 12:50pm - 1:10pm | SJCC - Concourse Level - 210AE

Bridging On-prem and Cloud — GPU Networking at Meta Scale

Networking

Meta's AI infrastructure now spans both on-prem data centers and multiple public clouds, and the RDMA network in both environments has to work seamlessly in each cluster but on occasion across RDMA clusters. This talk covers why hybrid is inevitable, what makes cloud networking fundamentally different from owning the rack, and how we're building toward a single converged operational model that delivers consistent performance, reliability, and tooling at Meta scale, regardless of where the GPUs live. One of the things we'll dive deep into is what it takes to productionize multi-plane networking in different clouds vs meta-prod. Another thing we'll dive into is how we expose individual cloud, on-prem network topologies to scheduler so that end-user constraints are seamlessly honored. We'll also delve into our long distance network that connects some of the Cloud / On-prem RDMA network to unblock a single large training job.

Speakers

Carmen Au
Software Engineering Manager - Meta

Adi Gangidi
Hardware Engineering - Meta

Scale-across



Scale-up /
Rack-level

Scale-out /
Building-level

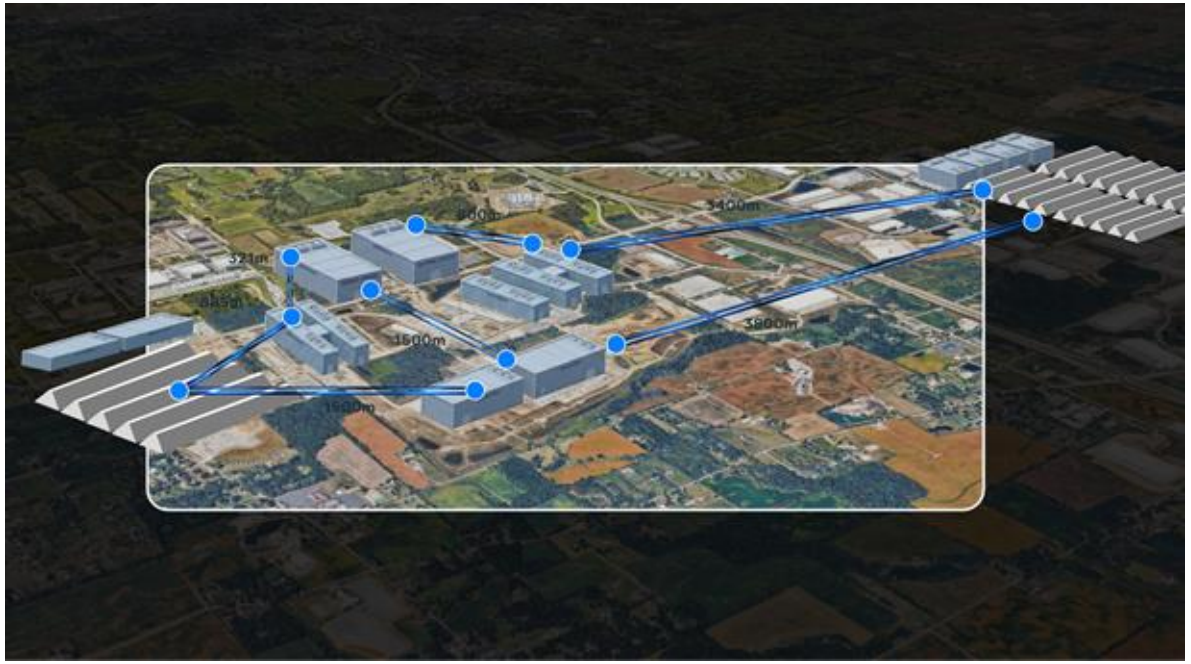
Scale-across /
Beyond building

Shared hardware & software

Prometheus: 1+ GW site (partial zoom-in)



Prometheus Inter-Site Connectivity: BAG (Backend Aggregation) Network





Making network perform for large Muse models

1. **Keep availability high**: interruptions slow down jobs, and frequent interruptions at this scale (millions of components across all scale-up/out/across networks)

1. **Implement systems to handle faults:** Parallelism-Aware Fault Tolerance (PAFT) system isolates restarts to smaller clusters; optimized checkpointing

1. **Iterate quickly on co-design of models & algorithms:** reduce exposed communications through tight measurement ⇒ optimization loop

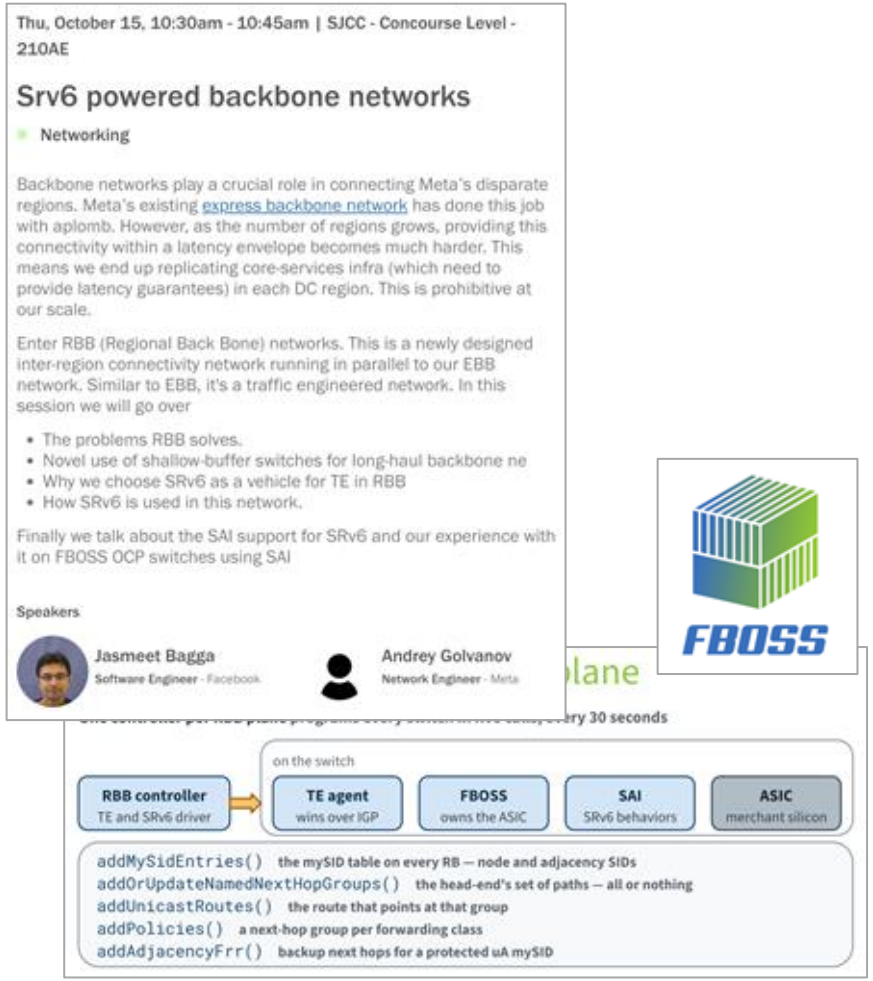


Building regional backbones (RBB)



Using SRv6 on FBOSS + SAI for RBB

- Implemented Segment Routing over IPv6 (SRv6) in FBOSS using SAI
- Motivated by speed, control, and flexibility
- Leveraged shallow buffer switches, starting with **Cisco** Silicon One for initial production
- Design to production in 5 months, followed by 3 months of long-haul and traffic engineering/SDN controller production

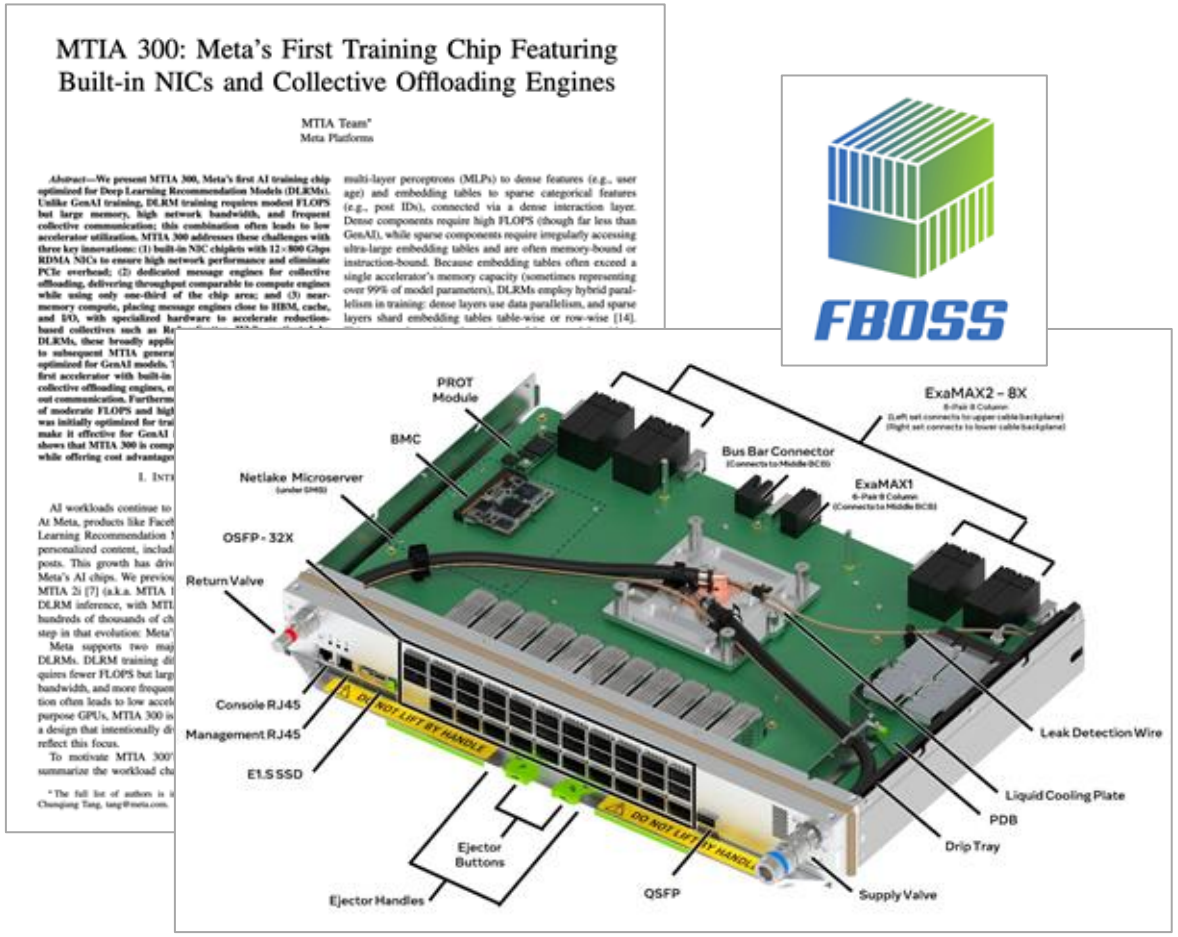


Scale-up


Scale-up /
Rack-level
Scale-out /
Building-level
Scale-across /
Beyond building
Shared hardware & software

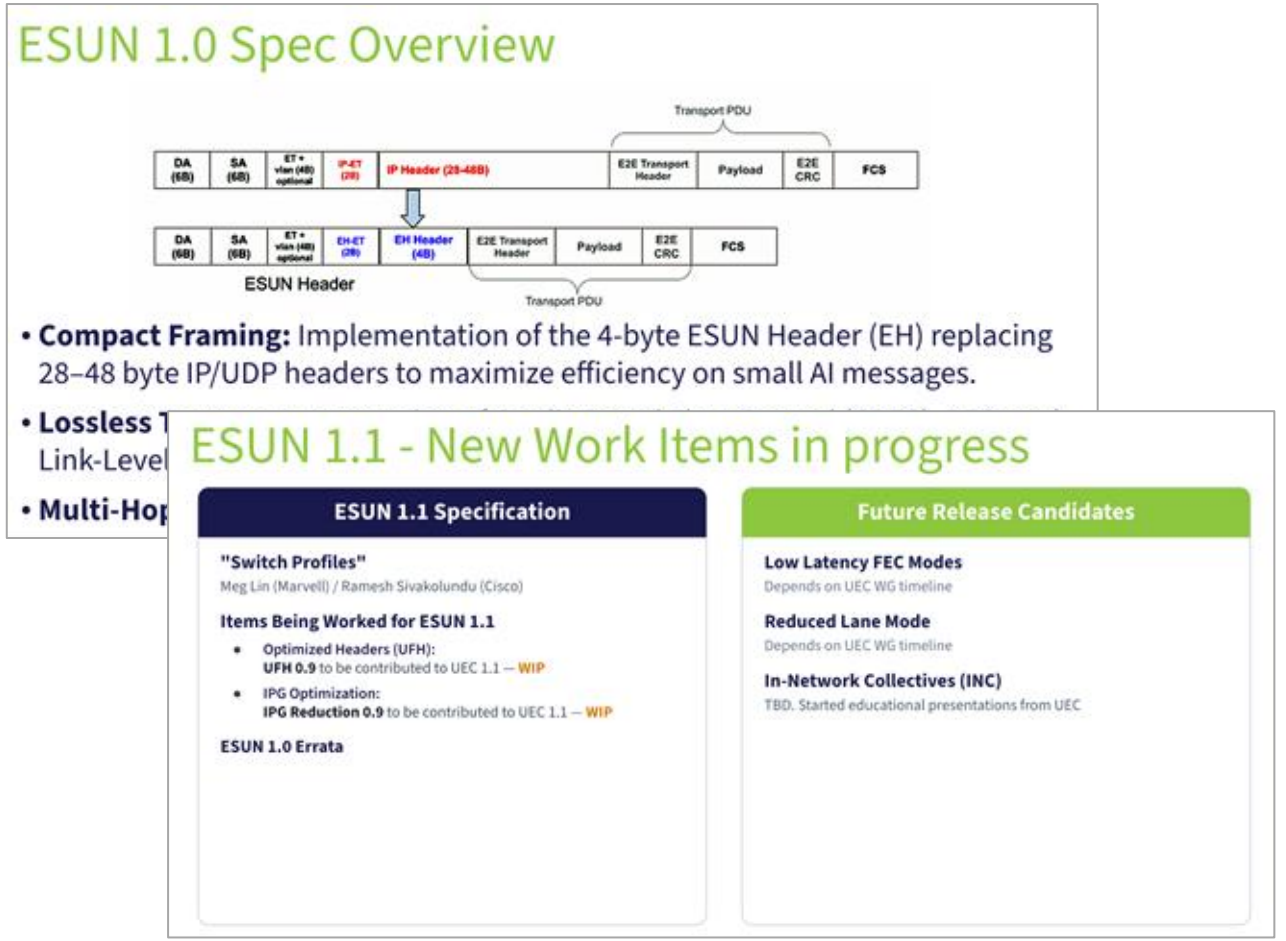
Ethernet for Scale-Up in MTIA: Tahan

- Tahan provided Ethernet-based scale-up for MTIA 300 (Meta Training & Inference Accelerators)
- Used **Broadcom** TH5 51.2T ASIC
- Leveraged FBOSS as NOS
- Looking ahead to **Broadcom** TH6 design
- Subset of features to optimize for scale-up, prepping for ESUN-compliant ASICs



OCP Ethernet for Scale-Up Networking (ESUN) progress

- Multi-vendor collaboration in OCP - thank you!
- 1.0 Spec published in March
- 1.1 Spec working on continued latency reductions



Shared HW & SW



Scale-up /
Rack-level

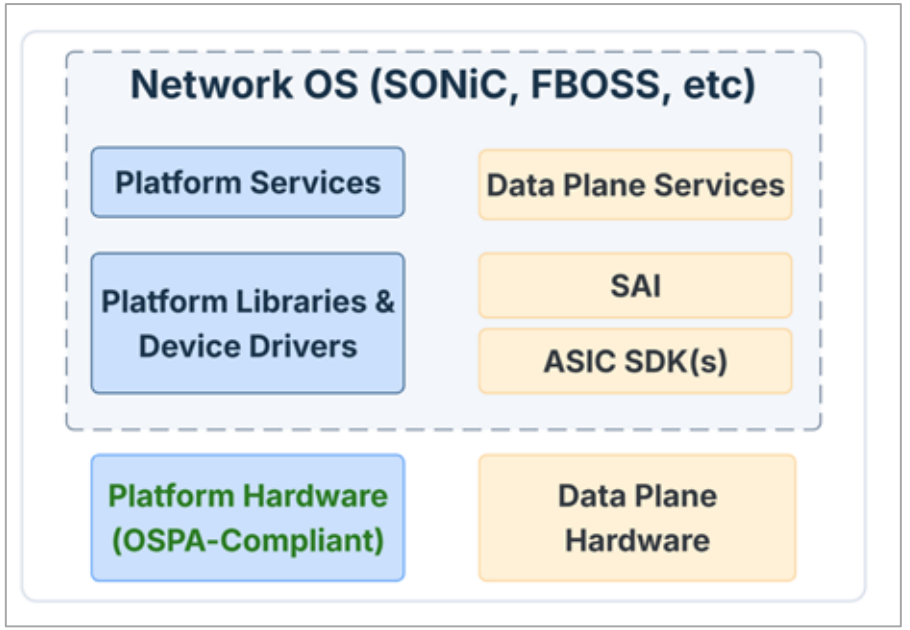
Scale-out /
Building-level

Scale-across /
Beyond building

Shared hardware & software

Open Switch Platform Architecture

- **Problem**: slow onboarding of new platforms, duplicate engineering, lifecycle cost
- **Solution**: vendor-neutral and NOS-neutral platform-layer specification across power, sensors, thermals, fan, inventory, out-of-band interfaces, firmware, troubleshooting
- **Community-driven**: started by Meta, Microsoft, Alibaba, now a dozen OCP members contributing!
- **Status**: Specification 1.0 in review



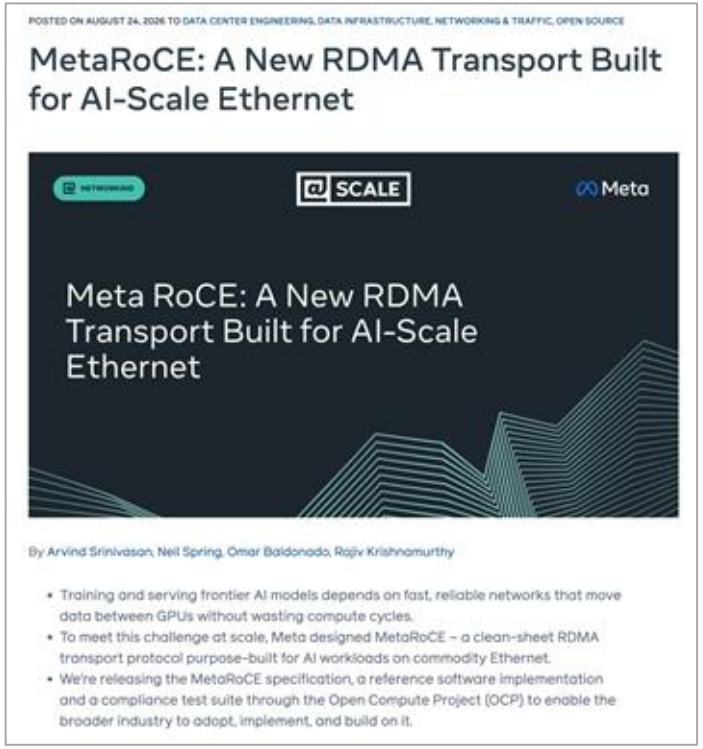
FBOSS: more open, config-driven, & AI-friendly

- **Similar problem**: reduce time to onboard a new platform to FBOSS itself
- **Solution**: make FBOSS more config-driven, more in open-source (e.g., routing protocols), and AI-friendly for coding and operations - partnership with **Nexthop AI**
- **Wins**:
  - Multiple platforms now coming up faster without Meta help ⇒ faster intro into Meta production
  - AI skills in open-source usable with any coding agent



MetaRoCE: Meta’s RDMA transport for AI

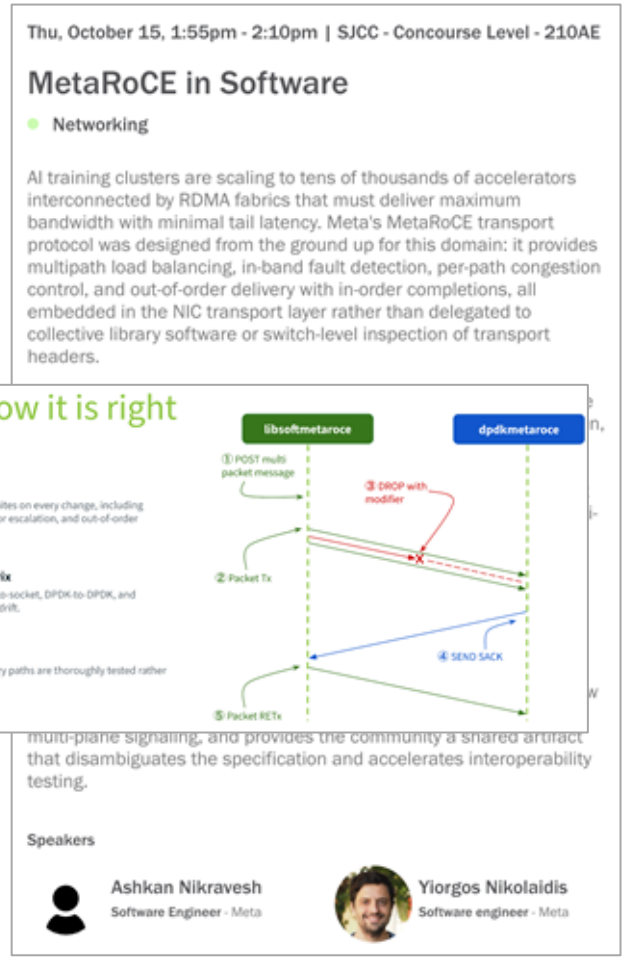
- Protocol announced in August
- Addresses RoCEv2 limitations with:
  - RDMA transport on best-effort Ethernet
  - Multi-path handling inside a queue-pair
  - Out-of-order tracking & stream multiplexing
- Already supported by **AMD** in Pensando NICs
  - Working with **NVIDIA**



MetaRoCE: now in open-source!

Today at OCP: Meta is contributing *two* implementations of MetaRoCE for hardware / protocol developers:

- **libsoftmetaroce**: user-space implementation
- **dpdkmetaroce**: performance with kernel bypass



Thu, October 15, 1:55pm - 2:10pm | SJCC - Concourse Level - 210AE

MetaRoCE in Software

Networking

AI training clusters are scaling to tens of thousands of accelerators interconnected by RDMA fabrics that must deliver maximum bandwidth with minimal tail latency. Meta's MetaRoCE transport protocol was designed from the ground up for this domain: it provides multipath load balancing, in-band fault detection, per-path congestion control, and out-of-order delivery with in-order completions, all embedded in the NIC transport layer rather than delegated to collective library software or switch-level inspection of transport headers.

How we know it is right

Continuous Integration

Both implementations run their suites on every change, including awkward corners like atomics, error escalation, and out-of-order delivery.

Cross-Implementation Matrix

Every pairing is exercised (socket-to-socket, DPDK-to-DPDK, and cross-pairs) to actively catch spec drift.

Fault Injection

The modifier layer ensures recovery paths are thoroughly tested rather than just assumed.

multi-plane signaling, and provides the community a shared artifact that disambiguates the specification and accelerates interoperability testing.

Speakers

Ashkan Nikravesh
Software Engineer - Meta

Yiorgos Nikolaidis
Software engineer - Meta

Congrats to teams, thanks to partners



Scale-up / Rack-level

Scale-out / Building-level

Scale-across / Beyond building

Shared hardware & software

The future is for everyone - let’s build it together!




Thank
You!
OCP
GLOBAL
SUMMIT
October 12-15, 2026
San Jose, California