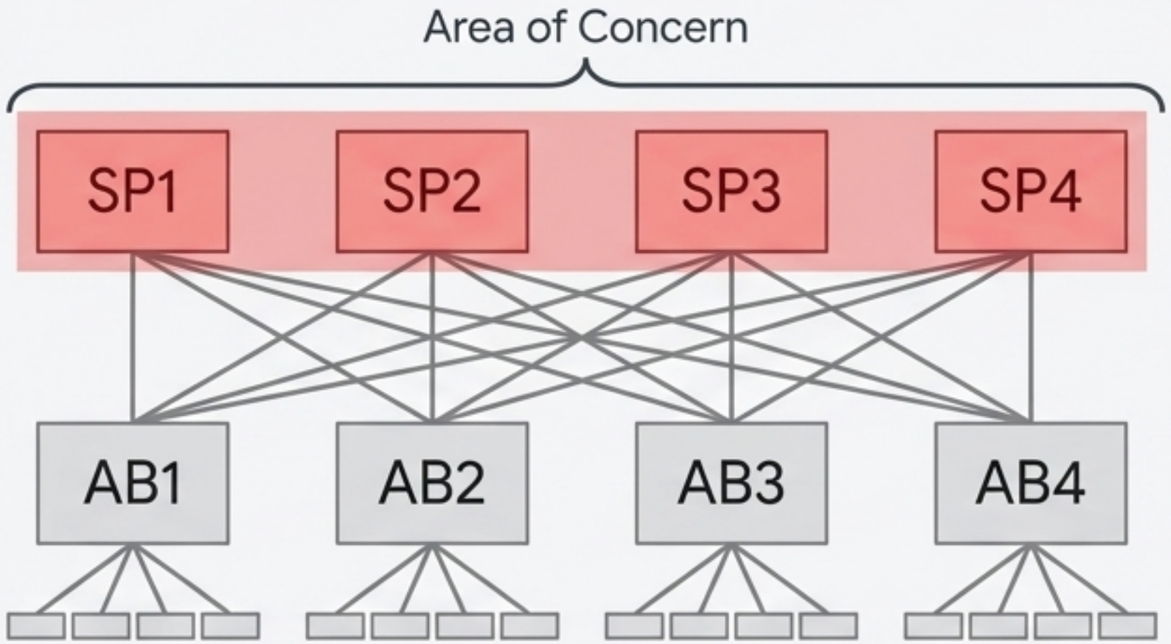
# Mission Apollo: Landing Optical Circuit Switching at Datacenter Scale

How Google reinvented its network backbone by trading electrons for photons, enabling a decade of unprecedented growth and performance

# Traditional Networks Were Hitting a Wall

As demand for services like Search, Cloud, and especially Machine Learning exploded, the conventional electrical packet-switched (EPS) network spine became a bottleneck for cost, power, and agility.




## Cost

Spine layer switches and their associated optics represent a significant and recurring capital expense.

## Power

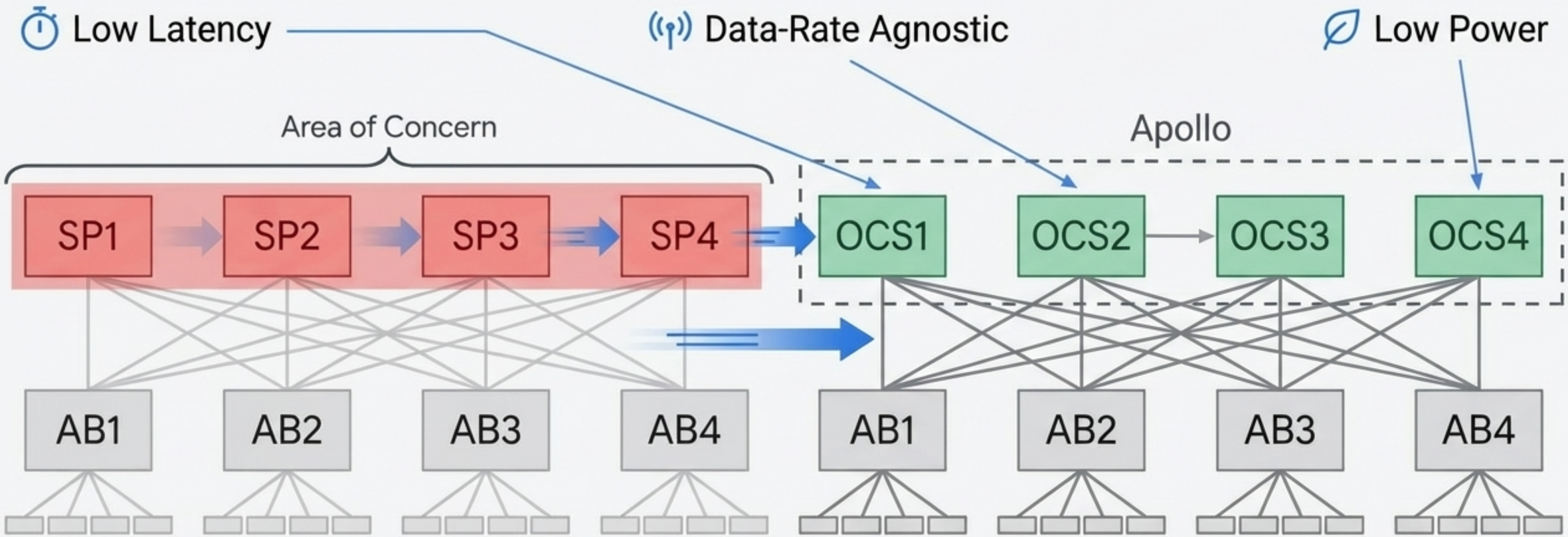
Electrical packet switches consume substantial power for packet processing, even for static, long-lived traffic flows.

## Inflexibility

Rigid Clos topologies are difficult and disruptive to expand or upgrade, often requiring “rip and replace” overhauls that can take months.

# Introducing Apollo: A Dynamic Optical Circuit Switching Layer

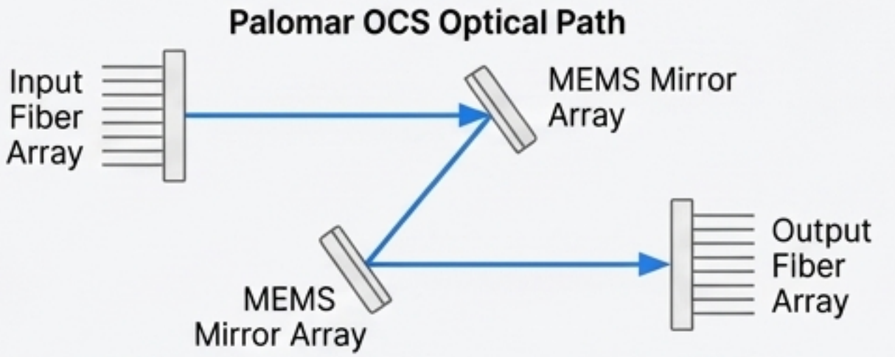
We replaced the power-hungry electrical spine with Apollo, a “cut-through” optical layer. Instead of processing packets, Apollo physically steers light, creating direct, ultra-low latency, data-rate-agnostic paths between network blocks.

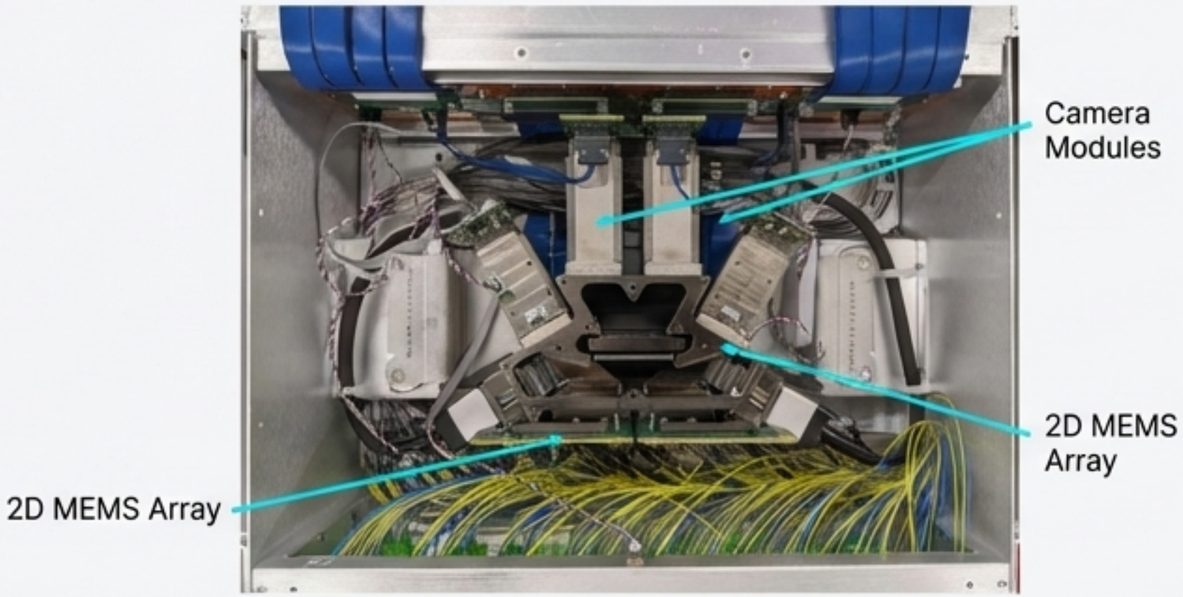


# The Apollo Toolkit (1/3): The Custom-Built Palomar OCS

No off-the-shelf OCS met our requirements for scale, reliability, and cost. We developed Palomar, a 3D MEMS-based switch, in-house to achieve our goals.

- **Scale & Performance:** 136x136 non-blocking ports.
- **Low Loss:** Typical insertion loss <2dB.
- **High Precision:** Stringent return loss of <-38dB to support bidirectional links.
- **Key Innovation:** Camera-based image processing for mirror control greatly simplifies manufacturing and ensures reliability at massive scale.

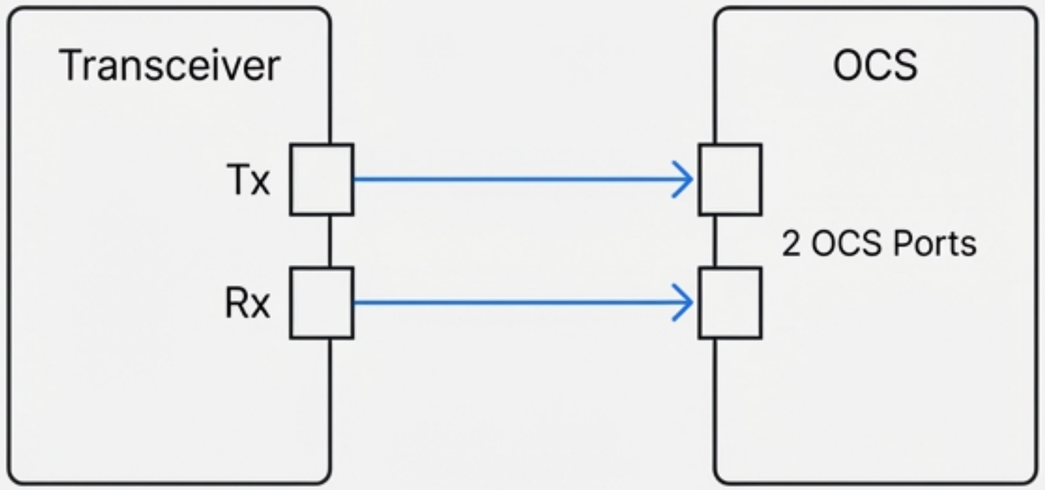


# The Apollo Toolkit (2/3): Doubling Our Fabric with Optical Circulators

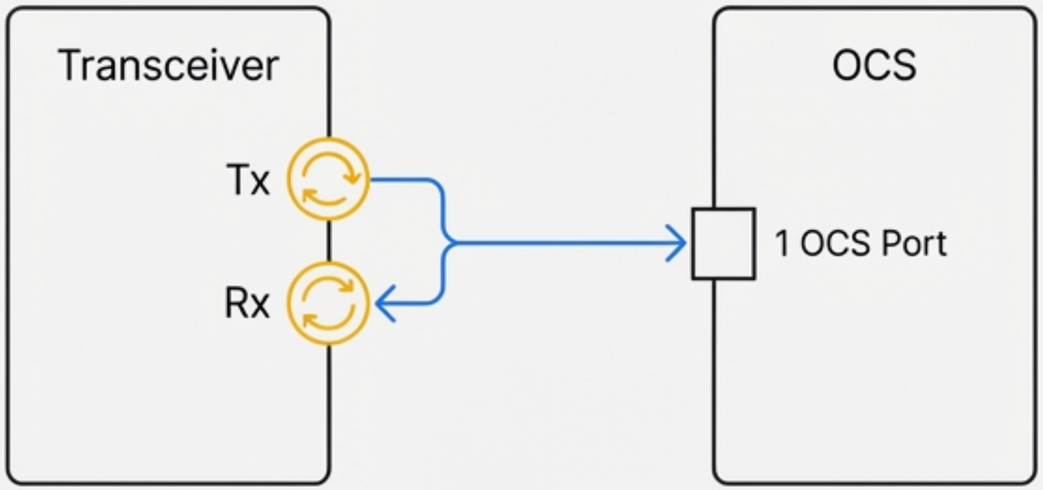
Circulators are the key to unlocking radical efficiency. They enable bidirectional traffic over a single fiber, effectively doubling the port capacity of each OCS and halving the required fiber plant.

**BEFORE**: Duplex Connection




2 Fibers, 2 OCS Ports

**AFTER**: Bidirectional Connection



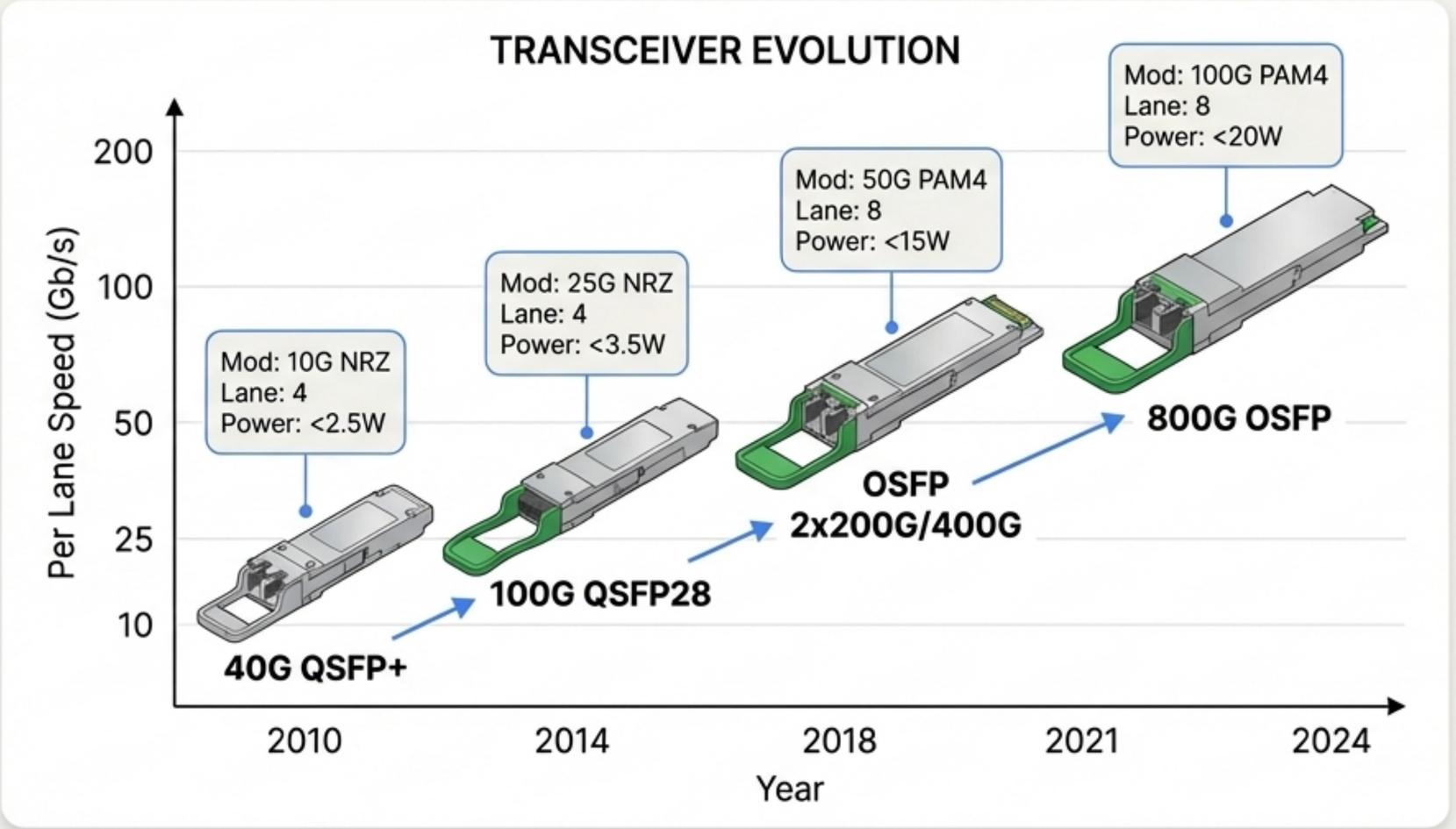

1 Fiber, 1 OCS Port

# The Apollo Toolkit (3/3): Future-Proofing the Network with Custom Optics

We co-developed four generations of WDM transceivers (40G to 800G+) specifically for the Apollo fabric. This ensured backward compatitility and handled the unique physical impairments of OCS-based links.

## KEY FEATURES

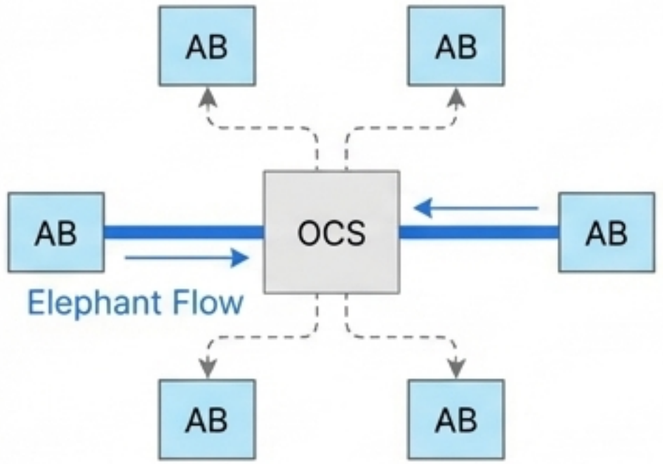
- **Handling Impairments:** Custom DSPs and FEC mitigate higher link loss and multipath interference (MPI) from reflections inherent in the bidirectional design.
- **Generational Interoperability:** All generations use the same CWDM4 wavelength grid, allowing new and old hardware to communicate seamlessly through the data-rate-agnostic OCS.



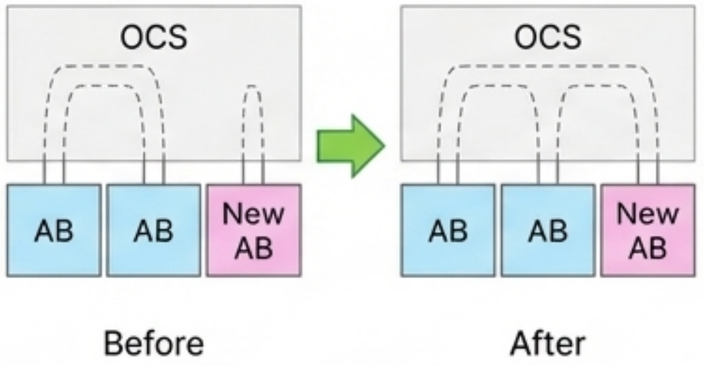
# Use Case 1: Building a Datacenter That Evolves on Demand

**Key Message:** Apollo transforms the network from static infrastructure into a dynamic, software-defined fabric, enabling unprecedented operational agility.

## Topology Engineering




## Pay-as-you-grow Expansion




## Rapid Tech Refresh

# The Agile Datacenter: A Decade of Proven Results

For nearly a decade, Apollo has served as the backbone for all of Google's datacenters, delivering significant, quantifiable improvements.

### 'Integer factor improvements' in cost, latency, and power.

Fundamentally rebalanced the economics and efficiency of the network fabric.

### Elimination of disruptive 'rip and replace' upgrades.

Amortizing fabric cost over the multi-decade life of the building, not just a few years.

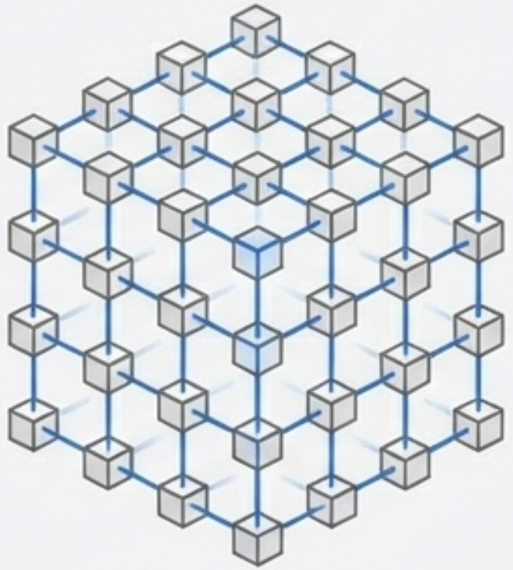
### Faster adoption of latest-generation networking.

Allowing incremental, non-disruptive rollouts of new technology as it becomes available.

# Use Case 2: Supercharging Machine Learning

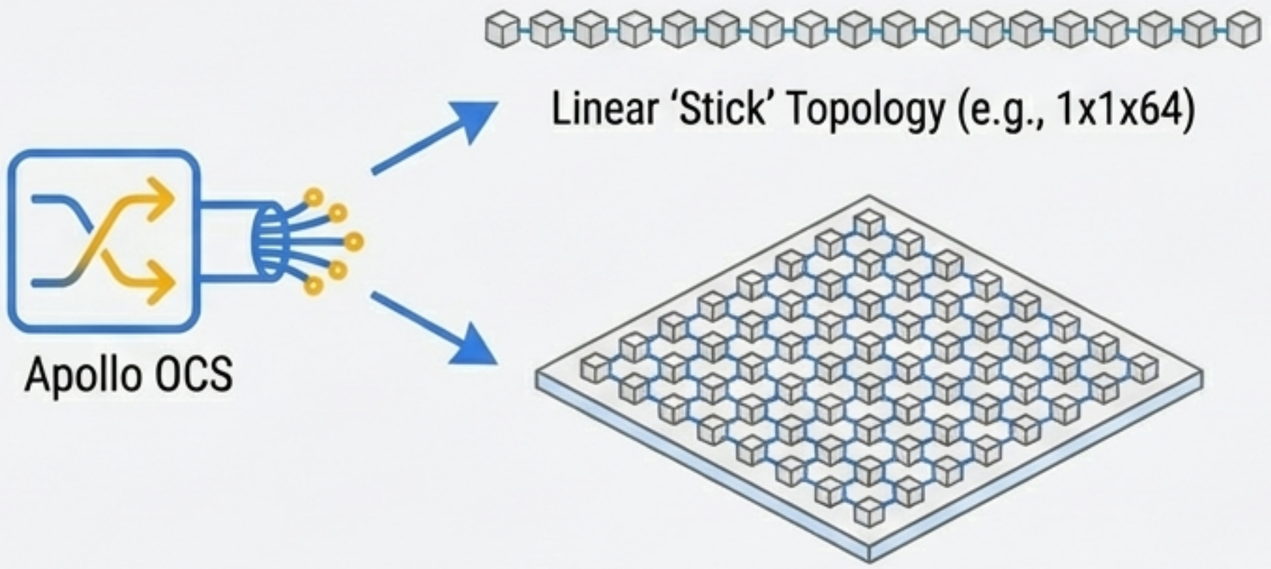
**Key Message:** Large ML models have predictable, intense communication patterns. Apollo allows us to reconfigure the network topology into the ideal shape for any given model, dramatically accelerating training.

## Standard Topology



Standard Topology

## OCS-Enabled Topologies



# Up to 3.3x Faster Training with Flexible Topologies

The ability to create optimal network shapes provides a massive performance boost, all while being more cost-effective and power-efficient than a traditional electrical fabric.

**3.3x**

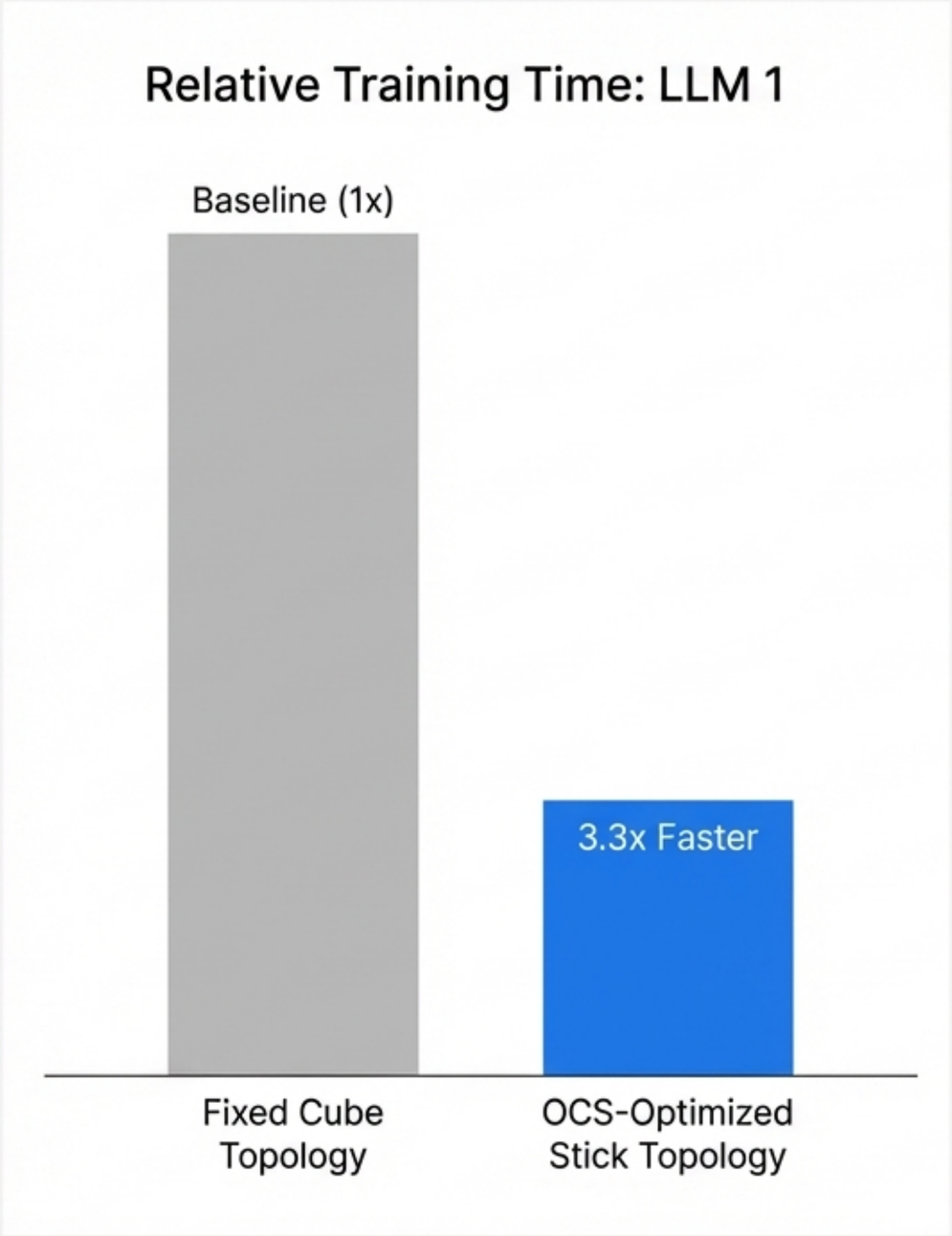
Speedup for a specific LLM on an optimal “stick” topology versus a standard cube.

**1.5x**

Speedup for another model on a “deca-card” configuration.

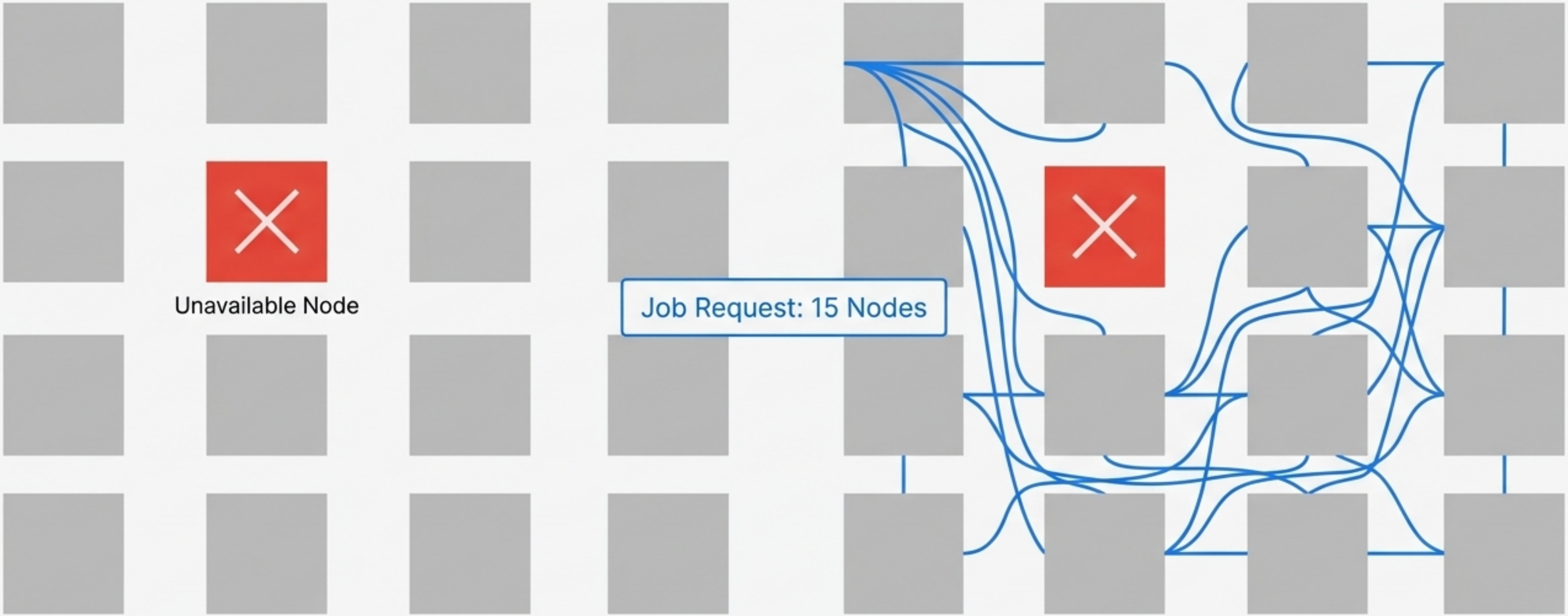
**24% & 10%**

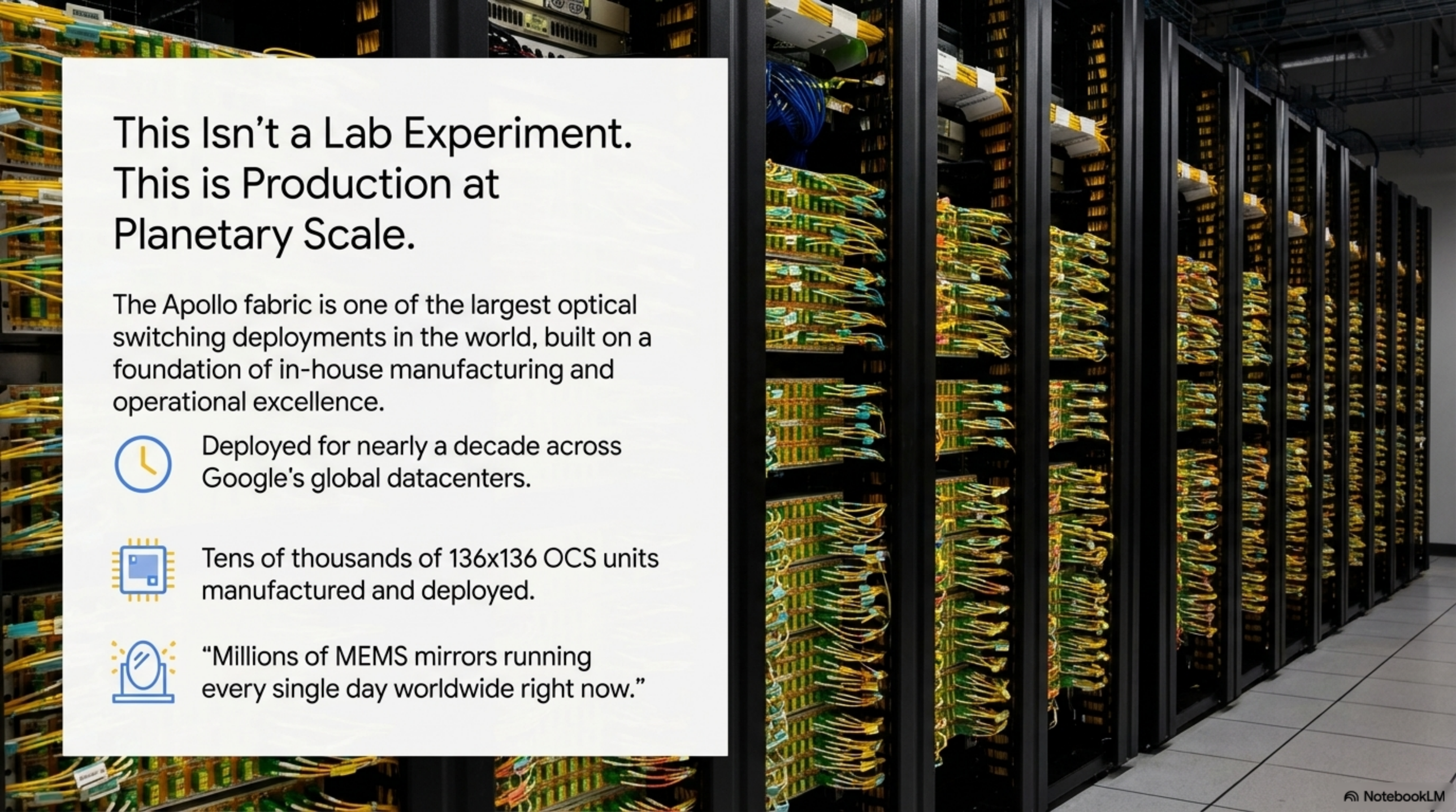
Cost Avoidance & Power Savings compared to a hypothetical EPS-based ML fabric.



# Eliminating Stranded Capacity and Maximizing Utilization

Because we can reconfigure the fabric logically, OCS allows us to dynamically route around failed or busy compute nodes, ensuring that large-scale jobs can always be formed from the pool of available resources.





# This Isn’t a Lab Experiment. This is Production at Planetary Scale.

The Apollo fabric is one of the largest optical switching deployments in the world, built on a foundation of in-house manufacturing and operational excellence.

Deployed for nearly a decade across Google's global datacenters.

Tens of thousands of 136x136 OCS units manufactured and deployed.

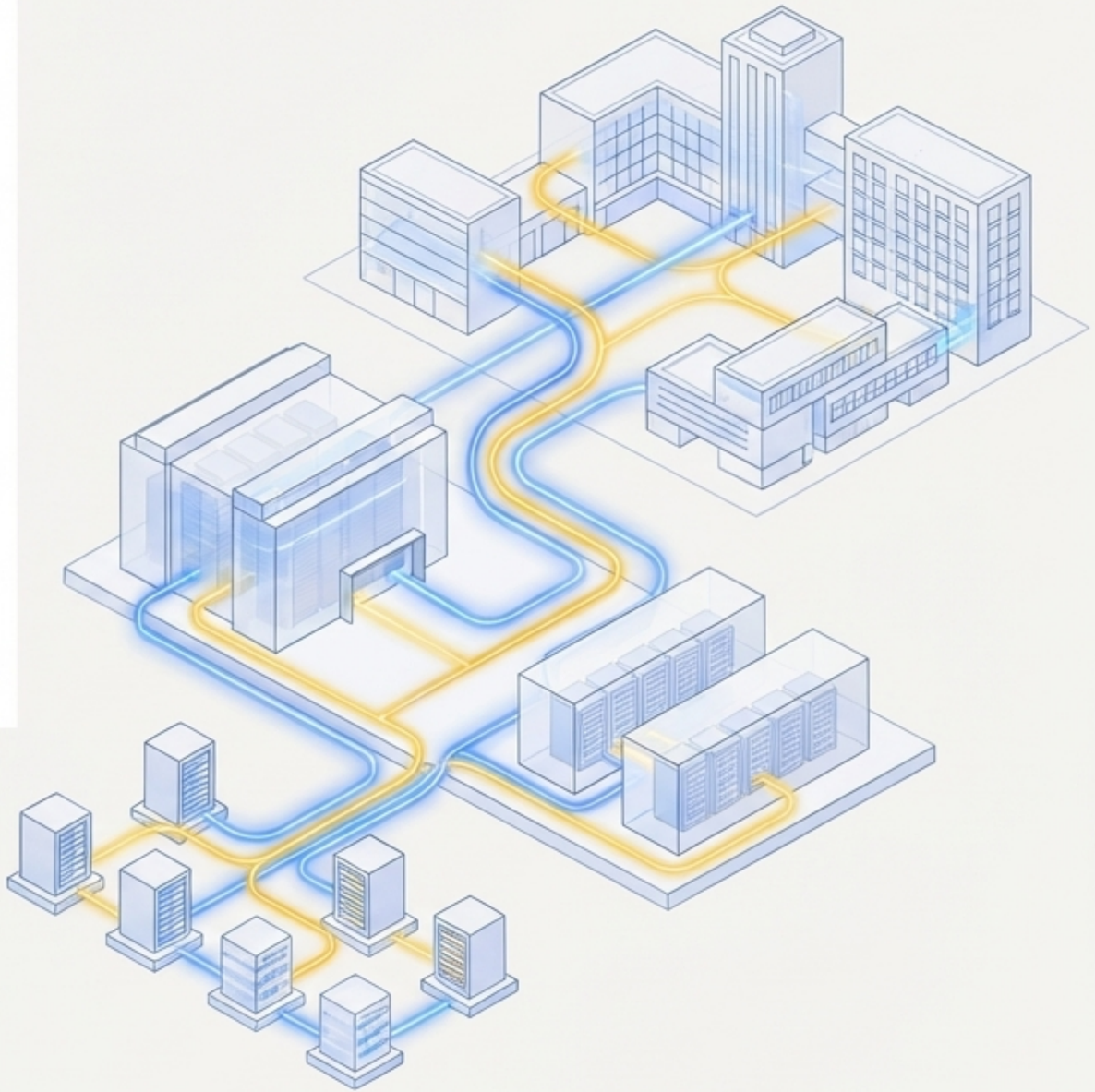
“Millions of MEMS mirrors running every single day worldwide right now.”

# The Future is Photonic: What's Next for Optical Switching?

**Apollo is the first major step. The long-term vision is to push optical technologies deeper into the network stack and across new domains.**

Future Directions:

- **Larger & Faster Switches**
  Scaling OCS port counts and exploring microsecond-scale switching for more dynamic, bursty workloads.
- **Beyond the Datacenter**
  Extending optical switching concepts to intra-campus and wide-area networks.
- **Deeper Integration**
  Bringing optical switching closer to the compute, potentially between Top-of-Rack switches and aggregation blocks.



# A Paradigm Shift Proven at Scale

By productionizing optical circuit switching, we built a more cost-effective, power-efficient, and agile network that accelerates both our core services and the frontier of AI.

## The Challenge

Solved the cost, power, and inflexibility crisis of traditional hyperscale networks.

## The Apollo Solution

Built a robust, vertically-integrated optical fabric with a custom OCS, efficiency-doubling circulators, and future-proof transceivers.

## The Impact

Achieved a dynamic, pay-as-you-grow datacenter fabric and unprecedented performance gains (up to 3.3x) for large-scale machine learning.

# Learn More

Original Research Paper
"Mission Apollo: Landing Optical Circuit Switching at Datacenter Scale"
ArXiv: 2208.10041v1

Practitioner's Talk
"Lightwave Fabrics: At-Scale Optical Circuit Switching for Datacenter and Machine Learning Systems"
OCP Future Technologies Symposium on YouTube

For specific questions, please contact the authors at the email address provided in the OCP talk.

Google | OCP