Concrete to Code: Engineering AI Infrastructure at Gigawatt Scale

Santosh Janardhan

Head of Global Infrastructure, Meta





OCP GLOBAL SUMMIT

October 12-15, 2026
San Jose, California

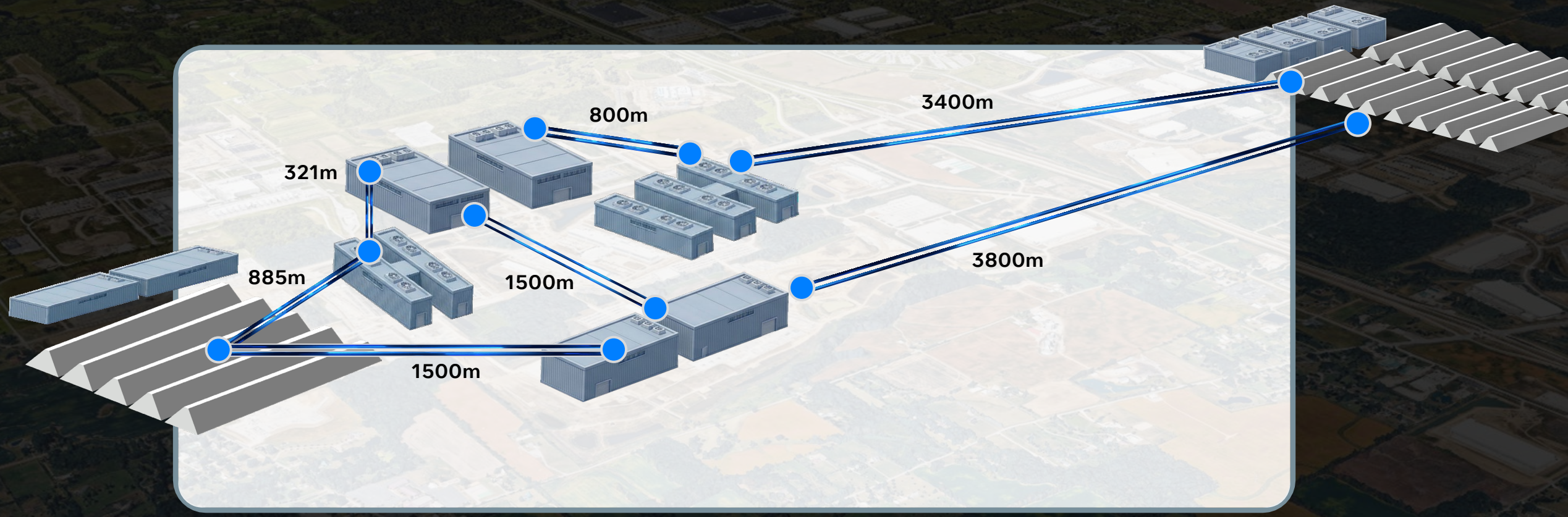
Assumption 1

A cluster lives under one roof.

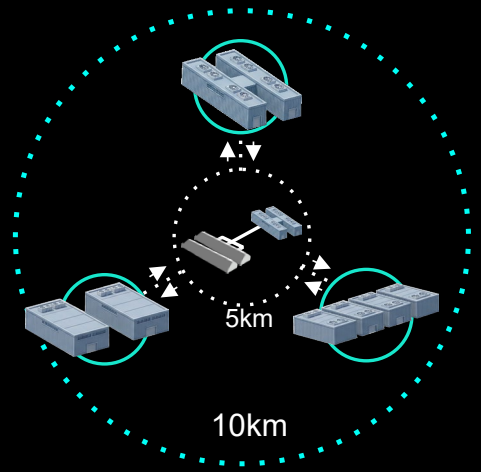
Meta

PROMETHEUS

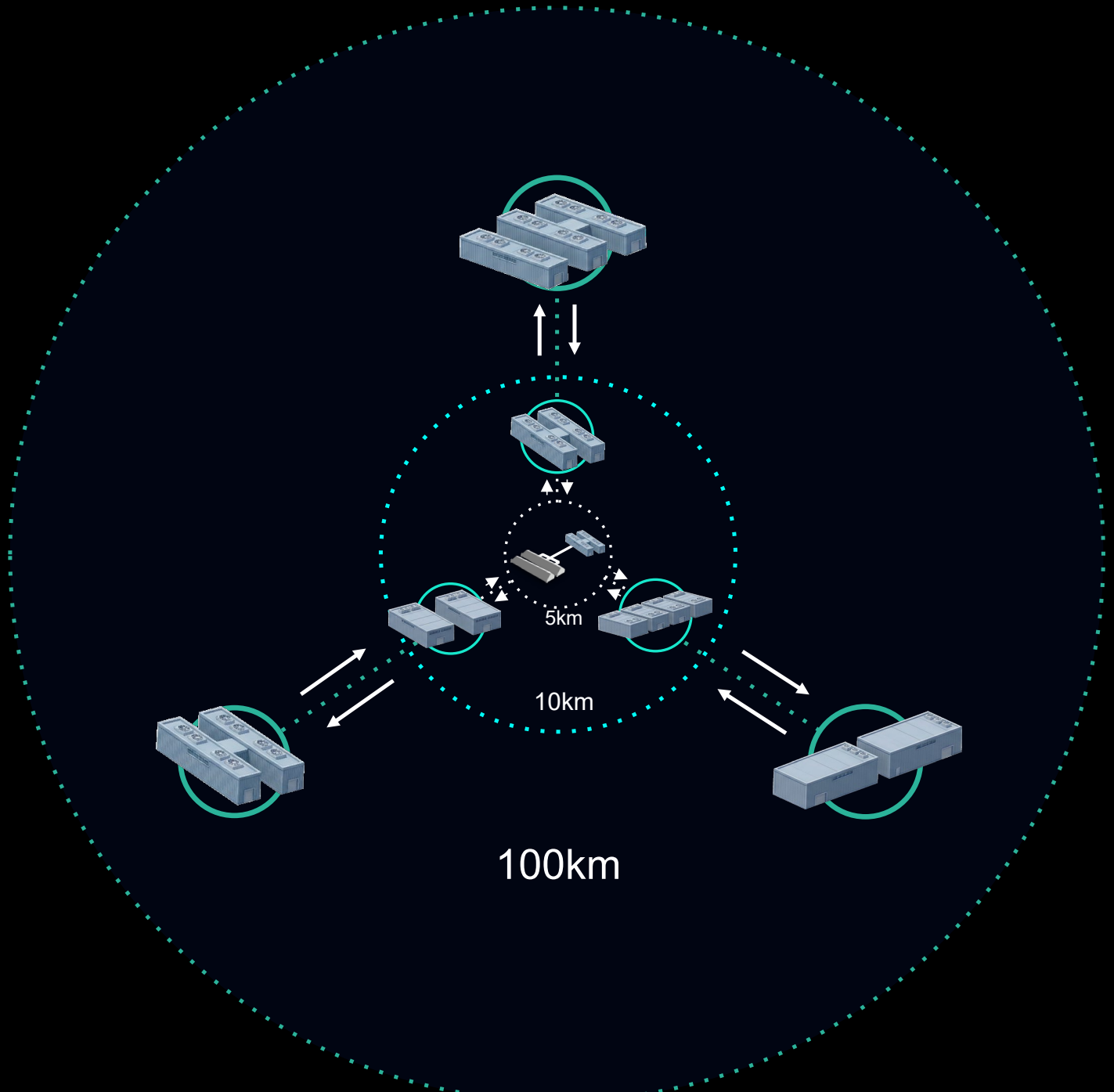
Our multi-region AI training cluster



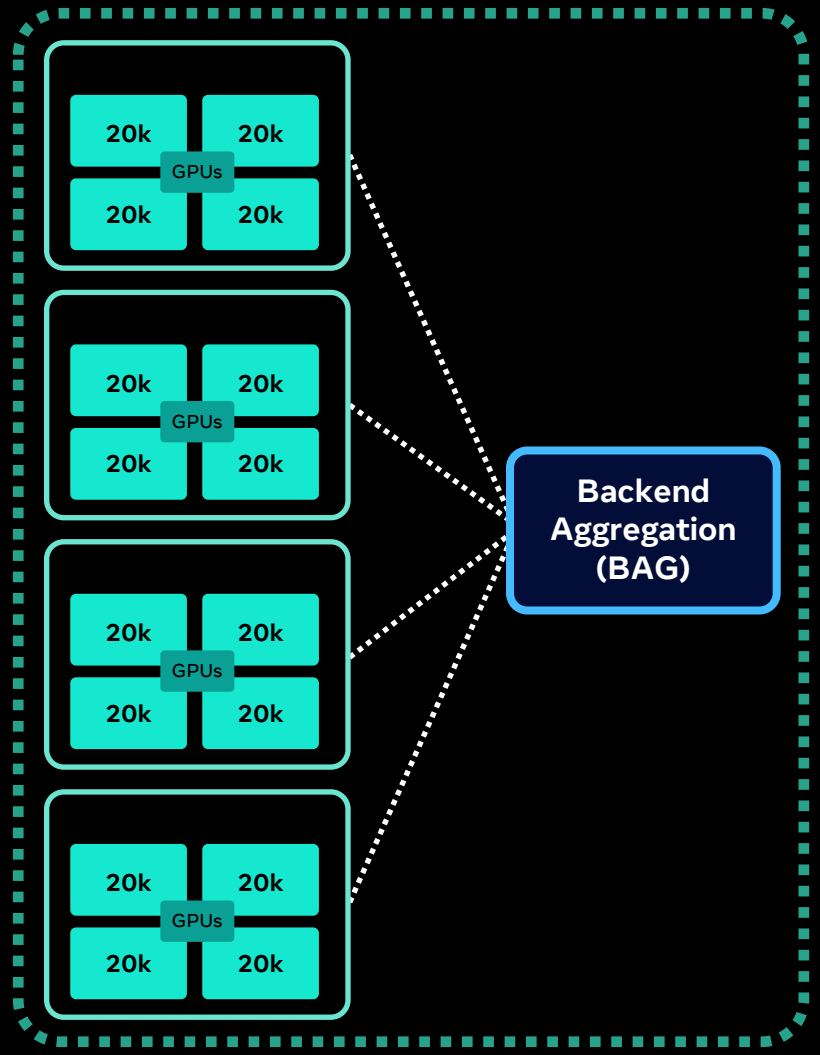
Meta



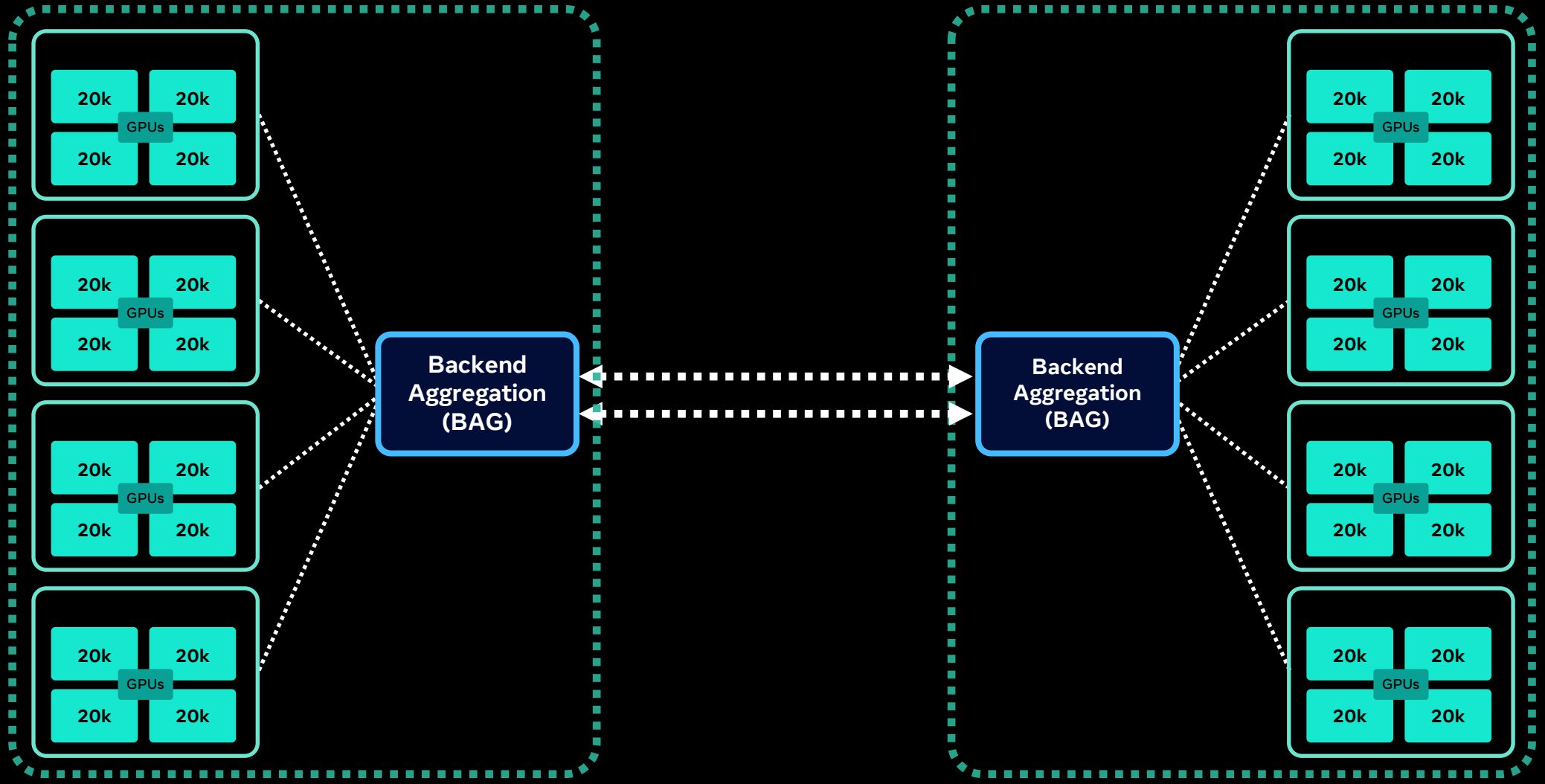
Meta



Meta


20k
20k
GPUs
20k
20k
20k
20k
GPUs
20k
20k
20k
20k
GPUs
20k
20k
20k
20k
GPUs
20k
20k
Backend Aggregation (BAG)

Meta

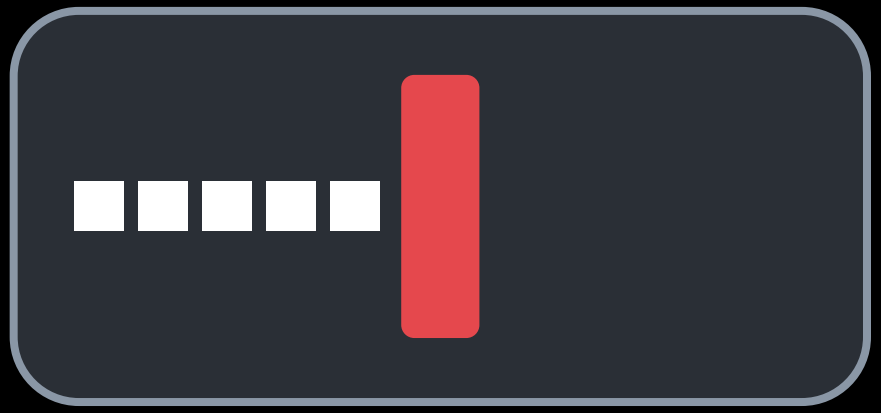

20k
20k
GPUs
20k
20k
Backend Aggregation (BAG)
Backend Aggregation (BAG)

Meta

MetaRoCE

A transport that can see every path.

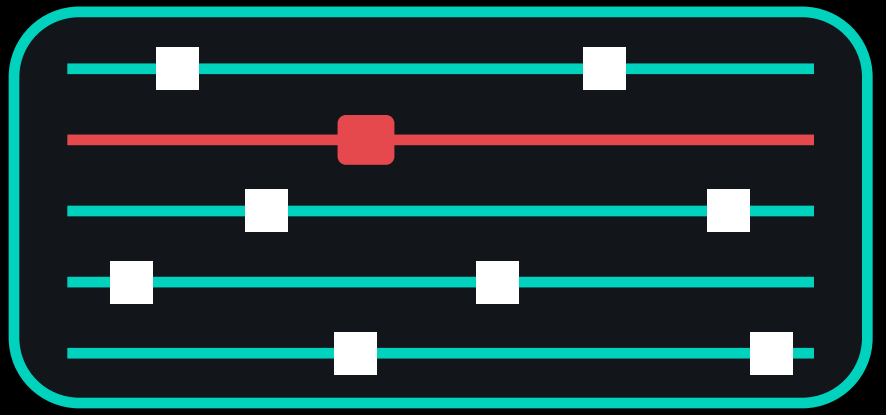
TODAY



Treats the Network as one single path.
Congestion stalls everything.

Coarse signals · Resend everything

MetaRoCE



Treats the Network as several paths.
Routes around congestion.

Per-path signals · Resend only what's lost

Meta




Turner



Assumption 2

Data centers are built from concrete and steel.

Meta

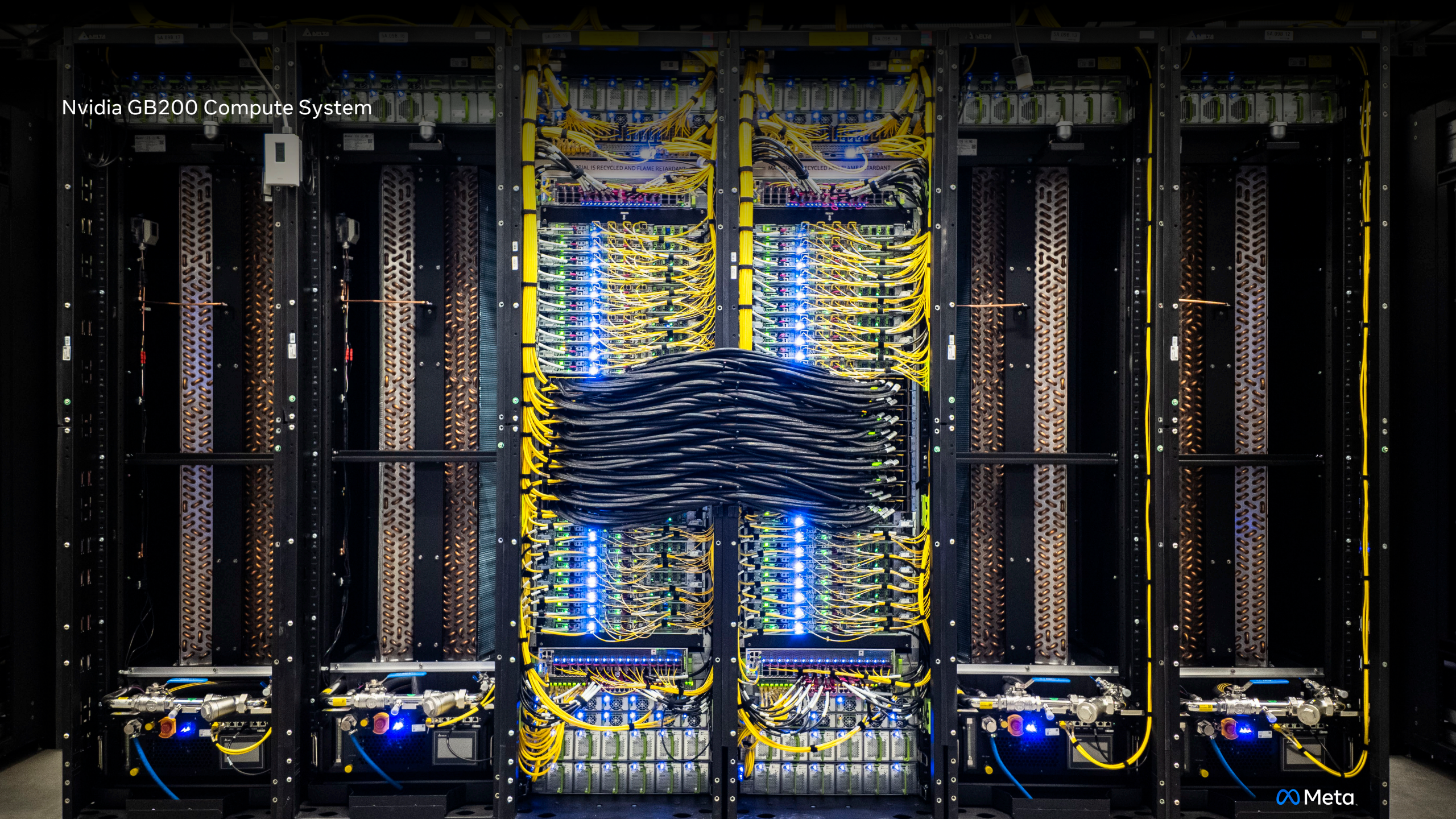

Meta



Assumption 3

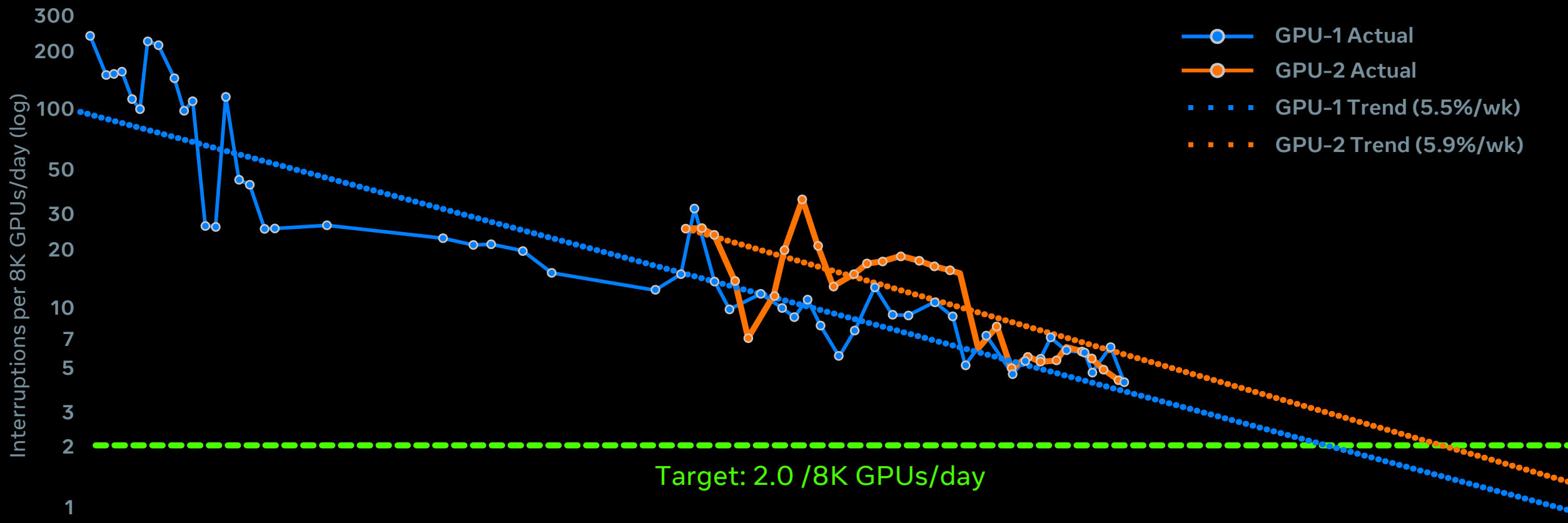
Once it’s built, it just works.

Meta


Nvidia GB200 Compute System
Meta

GPU-1 vs GPU-2 Interruption Trends

Log Scale | Interruptions per 8K GPUs/day | Dotted = projection



Meta

Option A - 1 slide

NO ONE BUILDS THIS ALONE



DATA CENTERS NETWORK CHIPS & RACKS

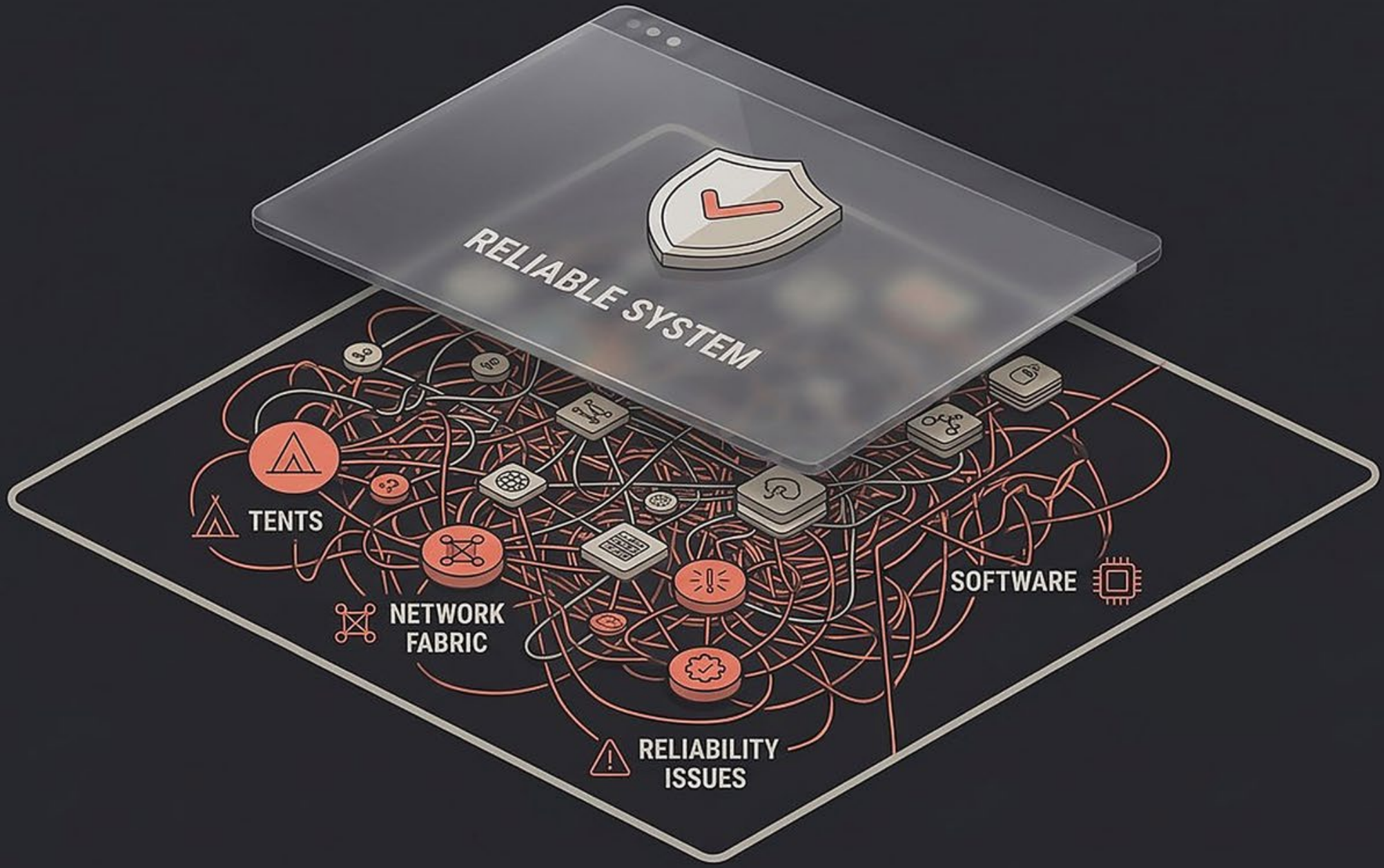
Co-designed at every layer, at every stage.

Meta

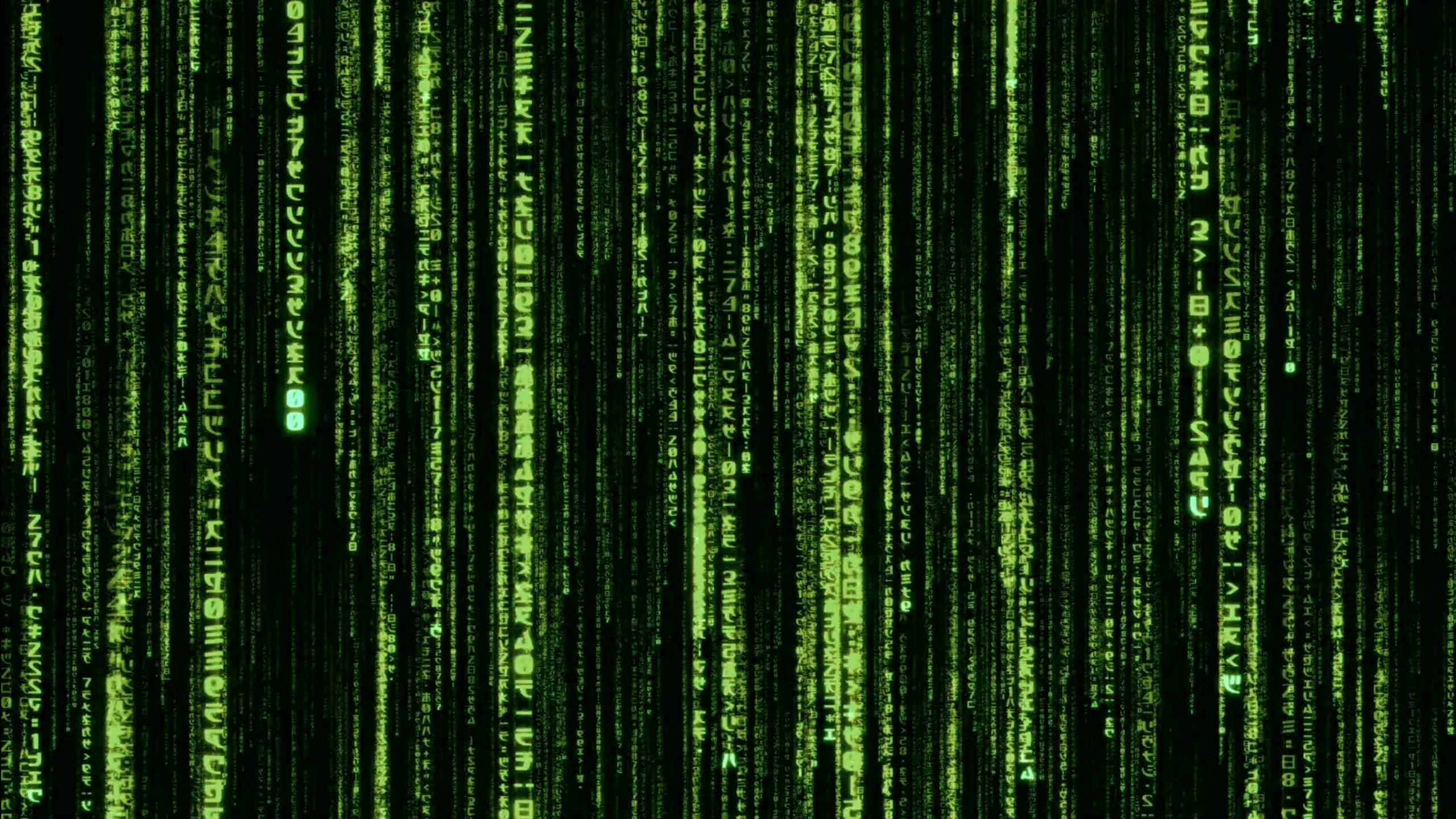
Assumption 4

Software is the easy part.

Meta


RELIABLE SYSTEM
TENTS
NETWORK FABRIC
SOFTWARE
RELIABILITY ISSUES

Meta



Our Titan AI Training Clusters



PROMETHEUS

1 GW+ cluster



HYPERION

5 GW over the next few years

Meta



INFRASTRUCTURE