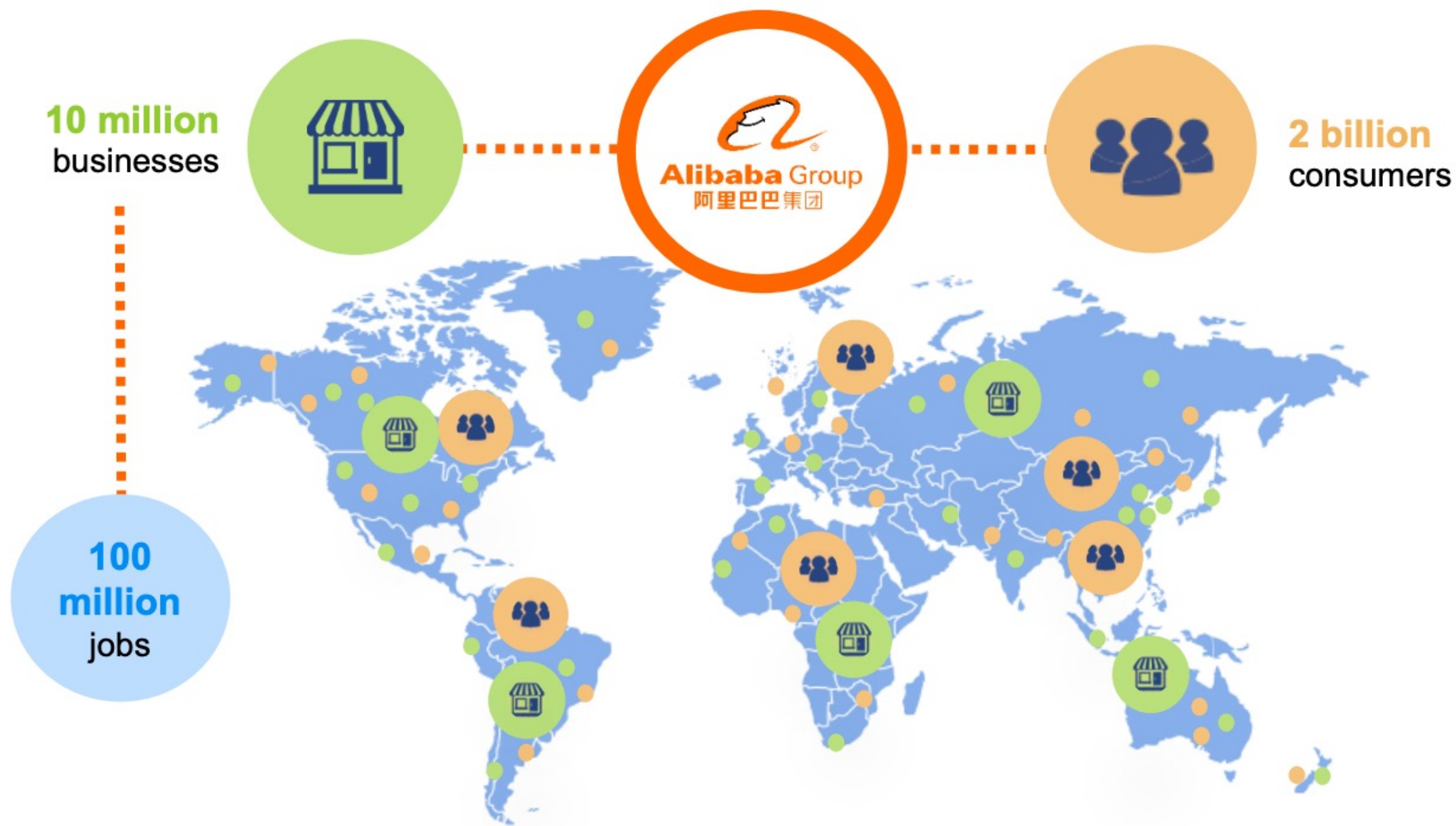


The Journey towards Predictable Network in Alibaba Cloud

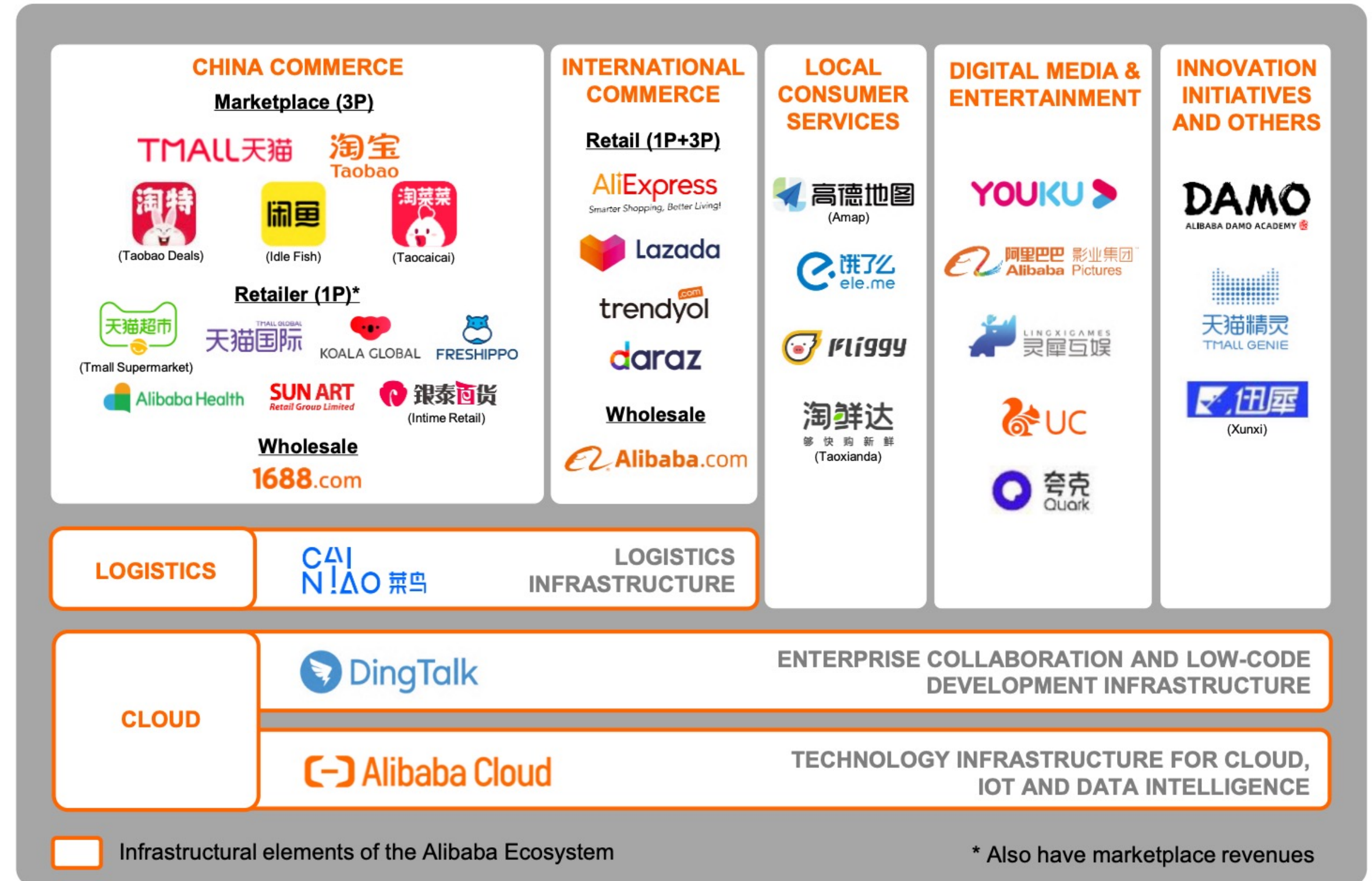
Dennis Cai
Head of Network Infrastructure

The world is accelerating into digitalization

Alibaba's Vision for Fiscal Year 2036



The Alibaba Ecosystem



Alibaba Group's MISSION IS TO MAKE IT EASY TO DO BUSINESS ANYWHERE

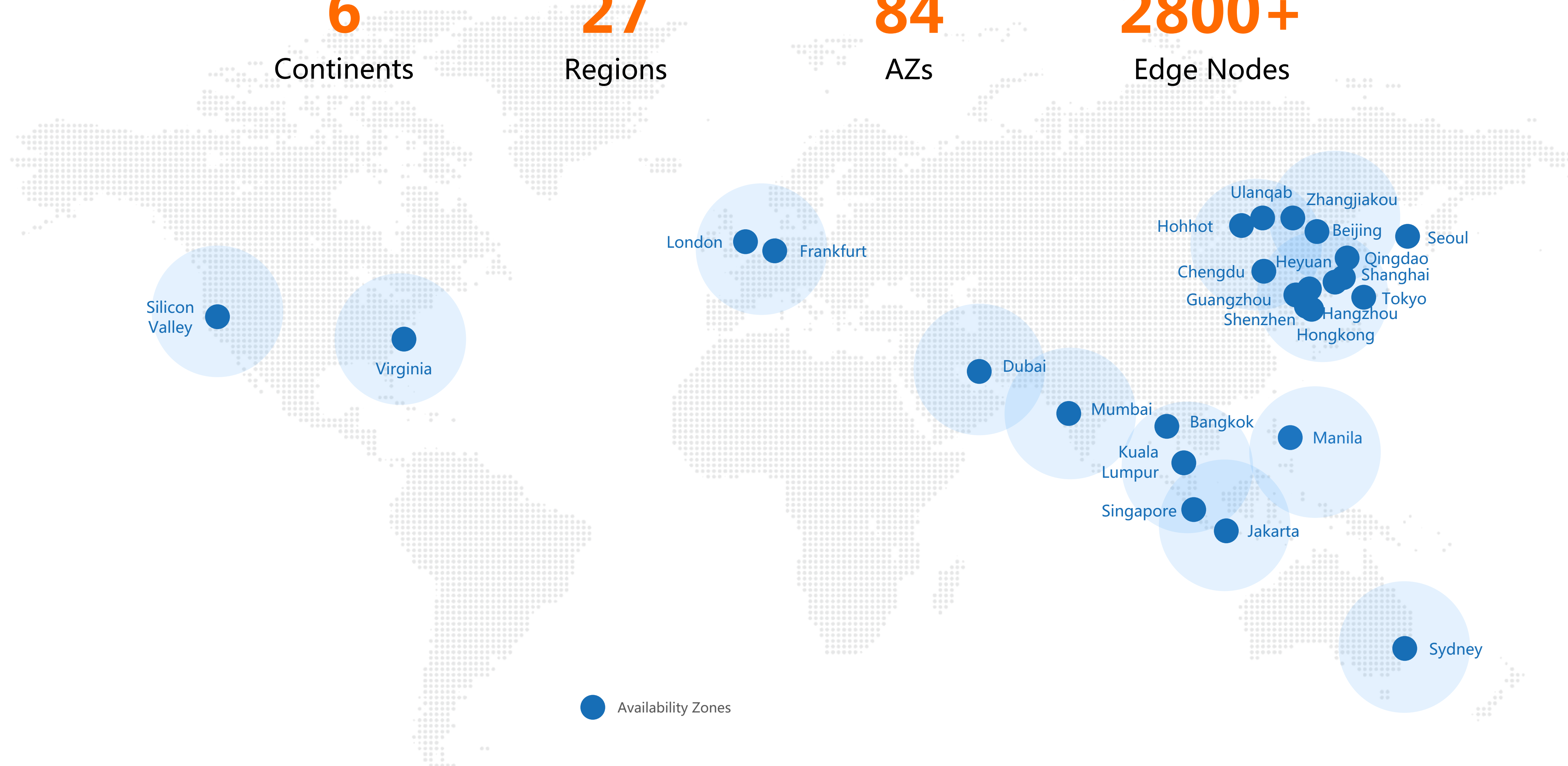
Alibaba Cloud global infrastructure

6
Continents

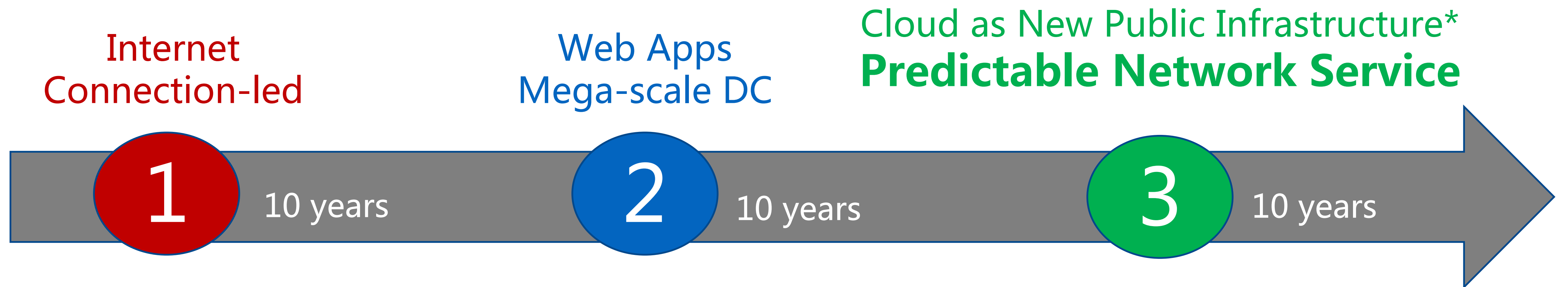
27
Regions

84
AZs

2800+
Edge Nodes



Our Vision : Towards Predictable Network



- Connection-Led
- Scale up
- Closed systems
- Human managed

- Mega-scale DC
- Scale out
- OPEN, disaggregation
- Automated
- SDN, NFV

- Mega-scale to next level: from central to the edge
- Big data and AI applications and workloads
- SDN 3.0: Programmable network with advanced hardware chips

* Super reliable and secure , Operation efficiency , Predictable performance

- Our journey toward predictable DC network
- Our journey towards predictable edge cloud network

The foundation towards Predictable DC Network

Mega-scale, super reliable, operation efficiency, low cost

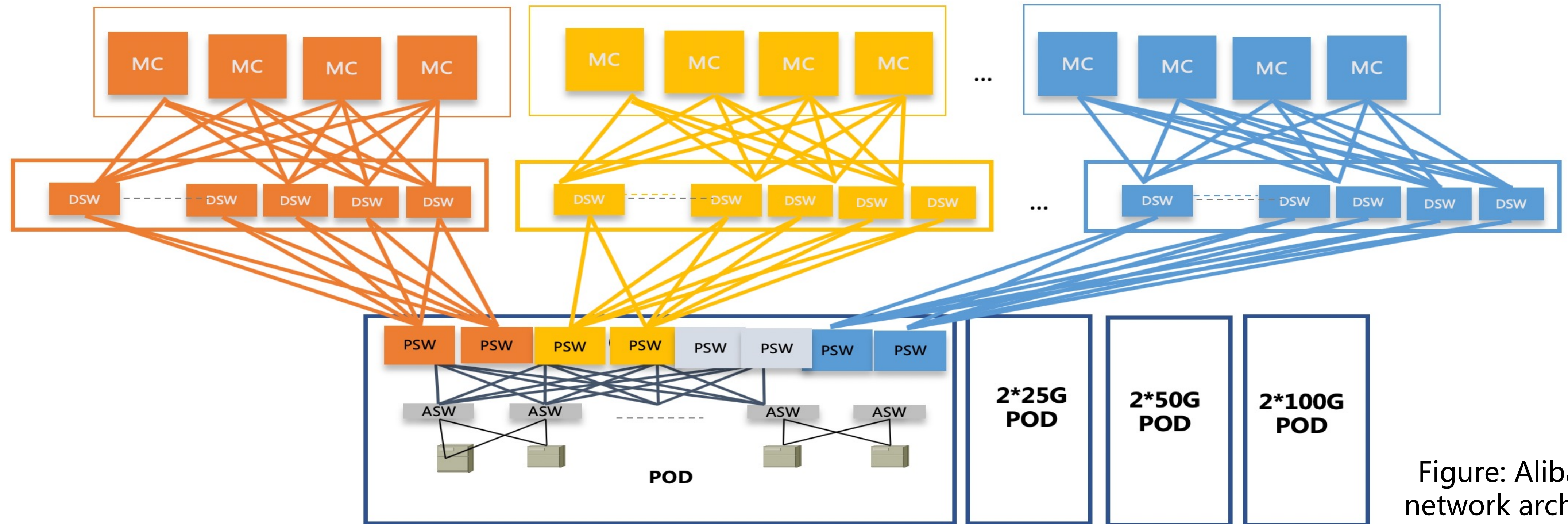
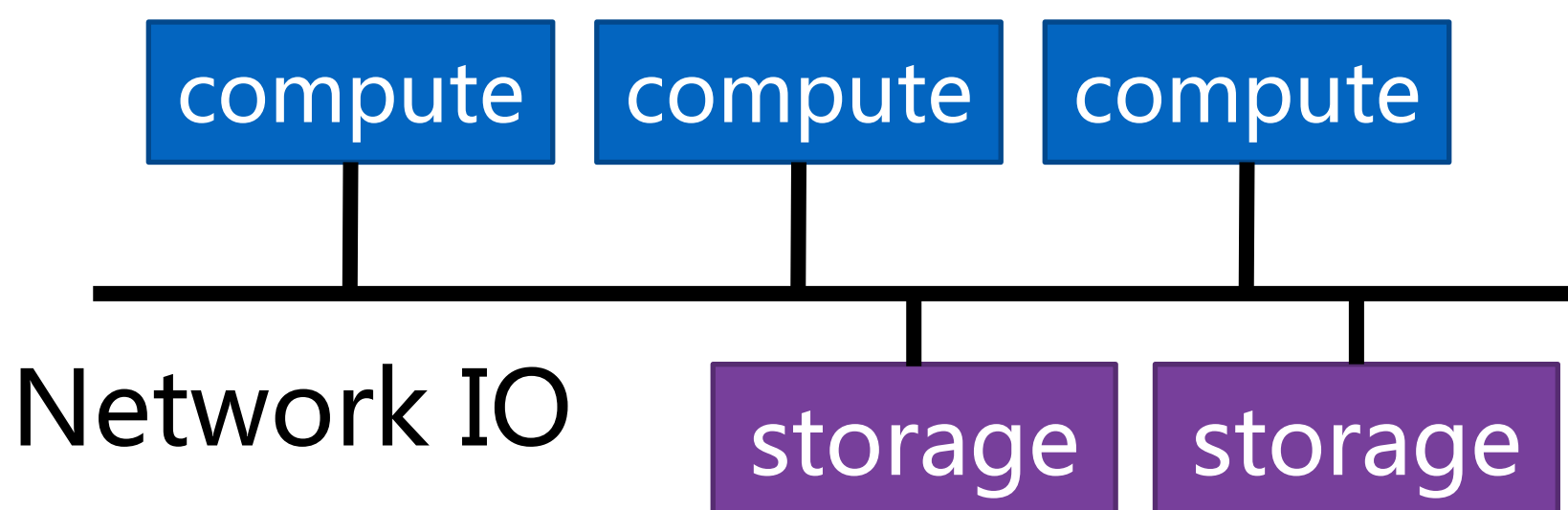


Figure: Alibaba DC network architecture

- Scale-out Architecture
Sing-chip switches
- In-house Built
Switches, Optics, NICs
- HA
Server dual-homing
Advanced monitoring
- Operation Efficiency
Network telemetry

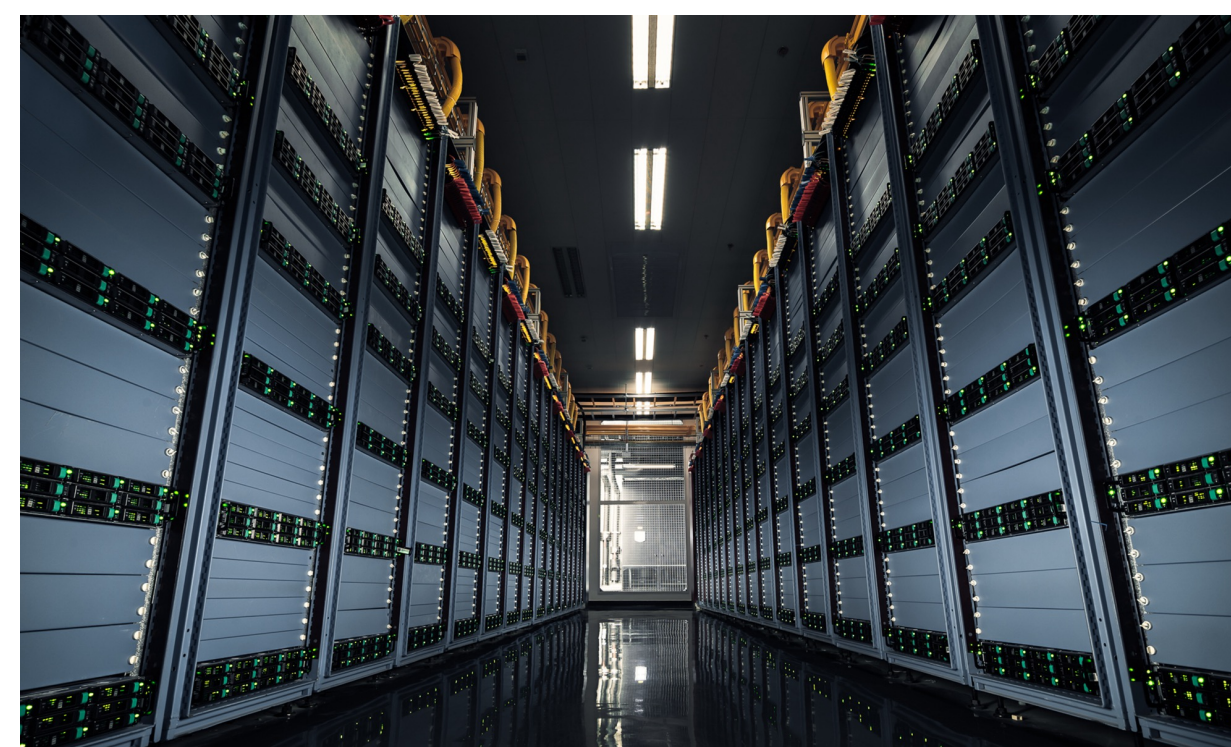
Applications demand Predictable DC Network

High-performance storage



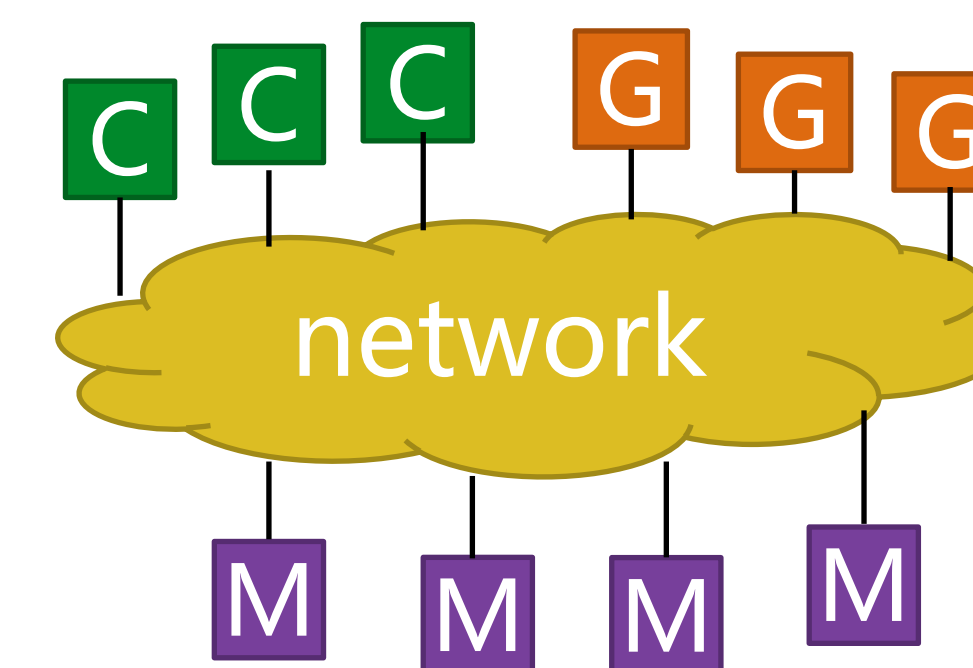
- HDD → SSD → PMEM
- Higher-throughput, lower latency
- 1M IOPS / 10~100us

HPC/ML



- CPU → GPU, TPU, DPU, FPGA
- Faster compute, lower latency
- e.g. 10K-100K Accelerators, latency <10us

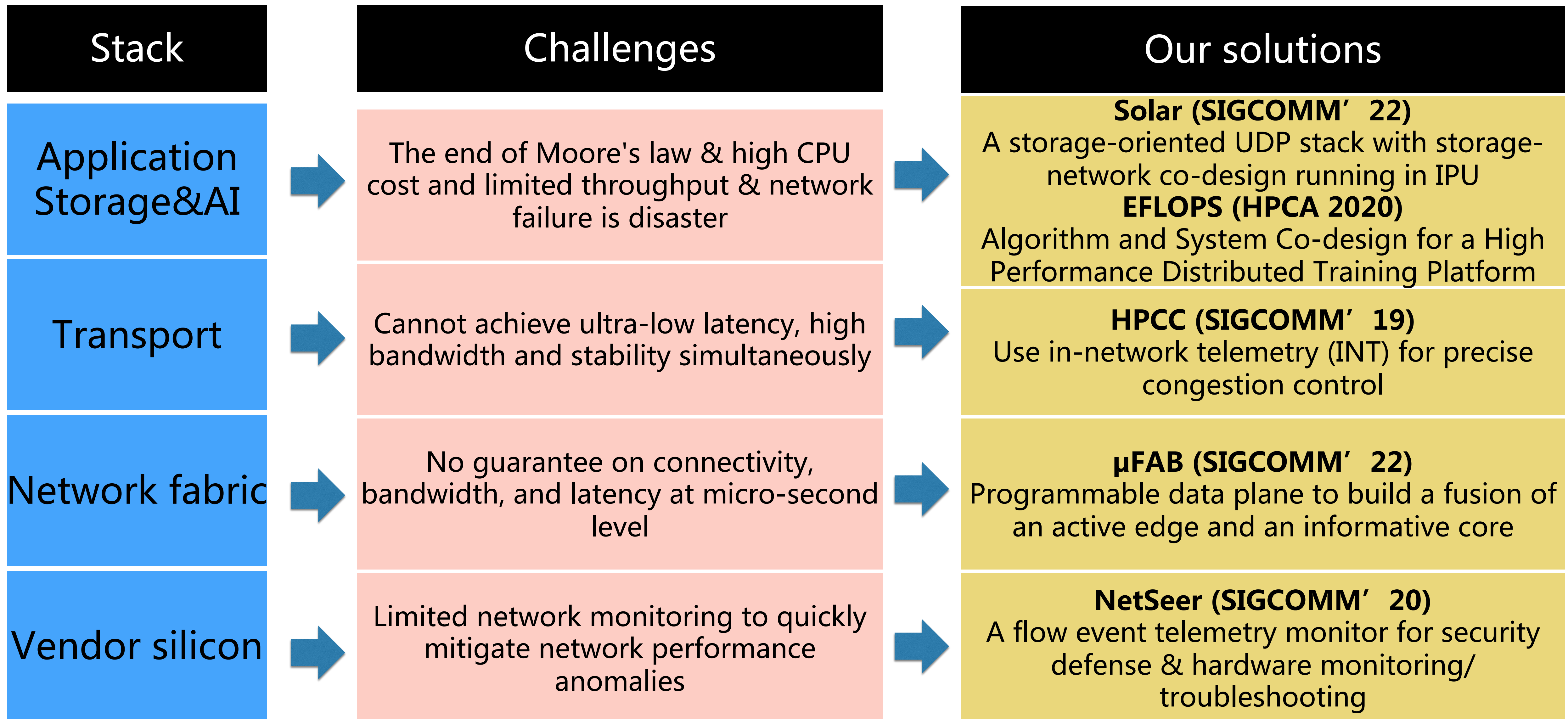
Composable DC Resource disaggregation



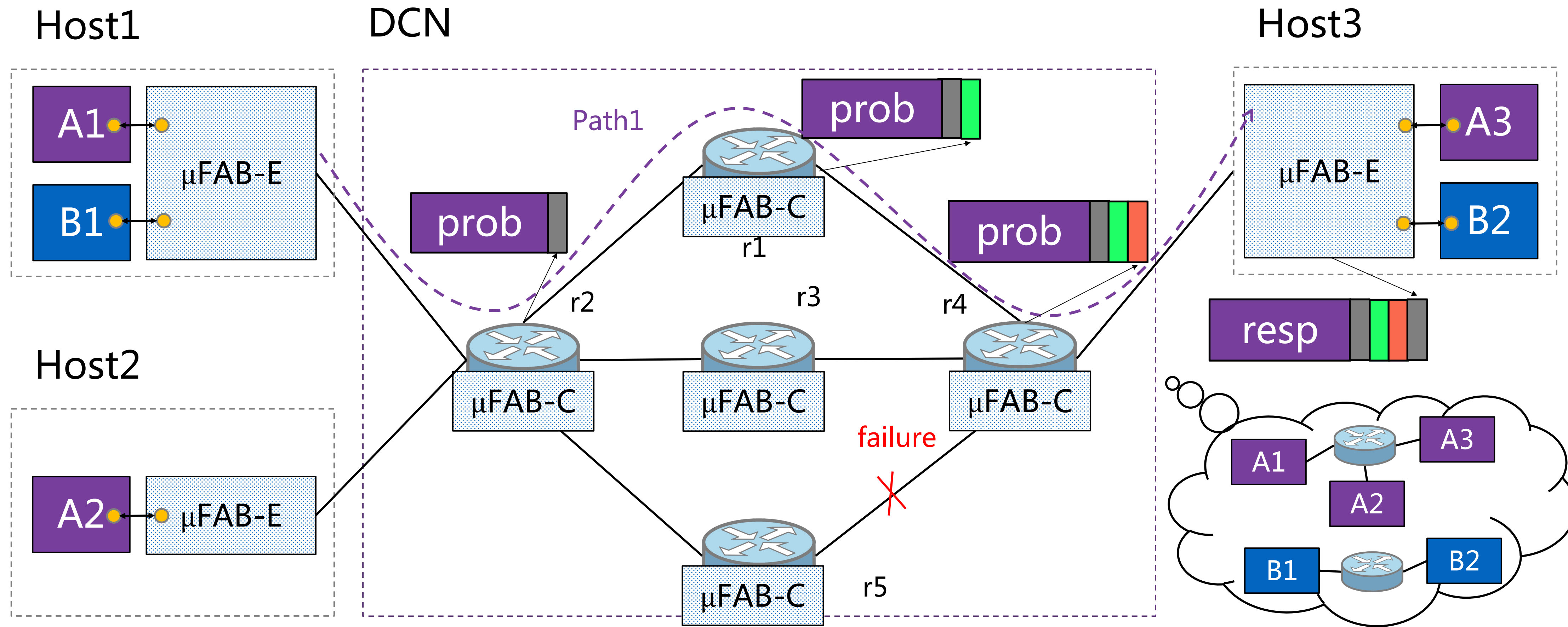
- Remote Memory Access (RMA)
- More network load
- Need ultra-lower latency: 3-5us

Applications become increasingly sensitive to performance anomaly at **micro-second granularity**

Our journey toward Predictable DC Network



μFAB: Predictable μFabric on Informative Data Plane



uFab-E

- Send probes with fetch core information back
- Schedule packets to paths
- Control sending rate of each path with back-pressure

uFab-C

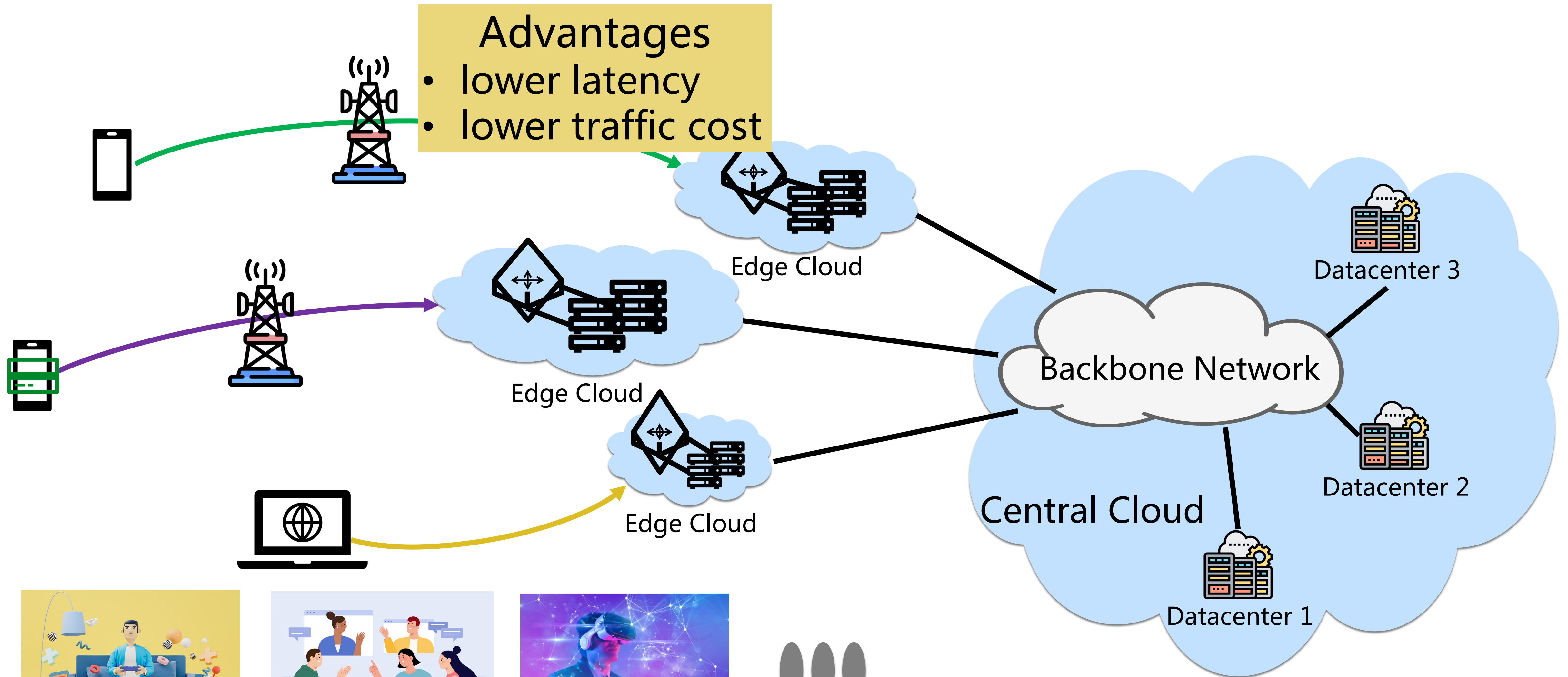
- Check failures locally and in neighbors
- Summarize total bandwidth subscription on each link
- Piggyback the preceding information with INT

- Our journey toward predictable DC network
- Our journey towards predictable edge cloud network

The commercialization of edge cloud is accelerating

Advantages

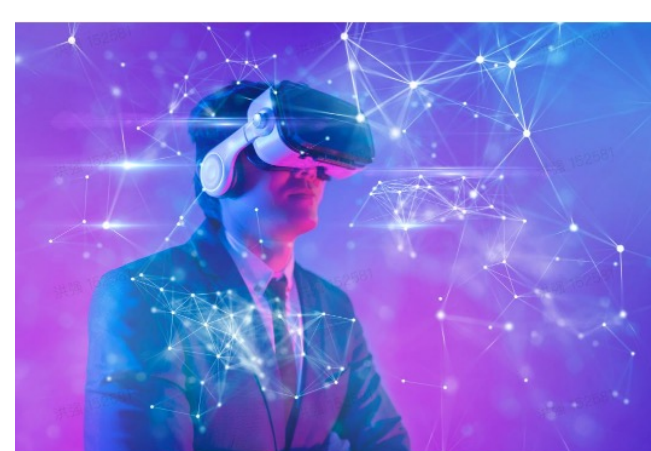
- lower latency
- lower traffic cost



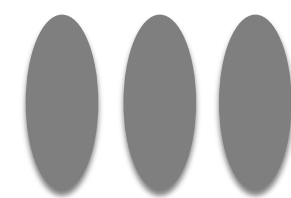
Cloud Gaming



Video Conferencing



XR & Metaverse



The challenges to build edge clouds



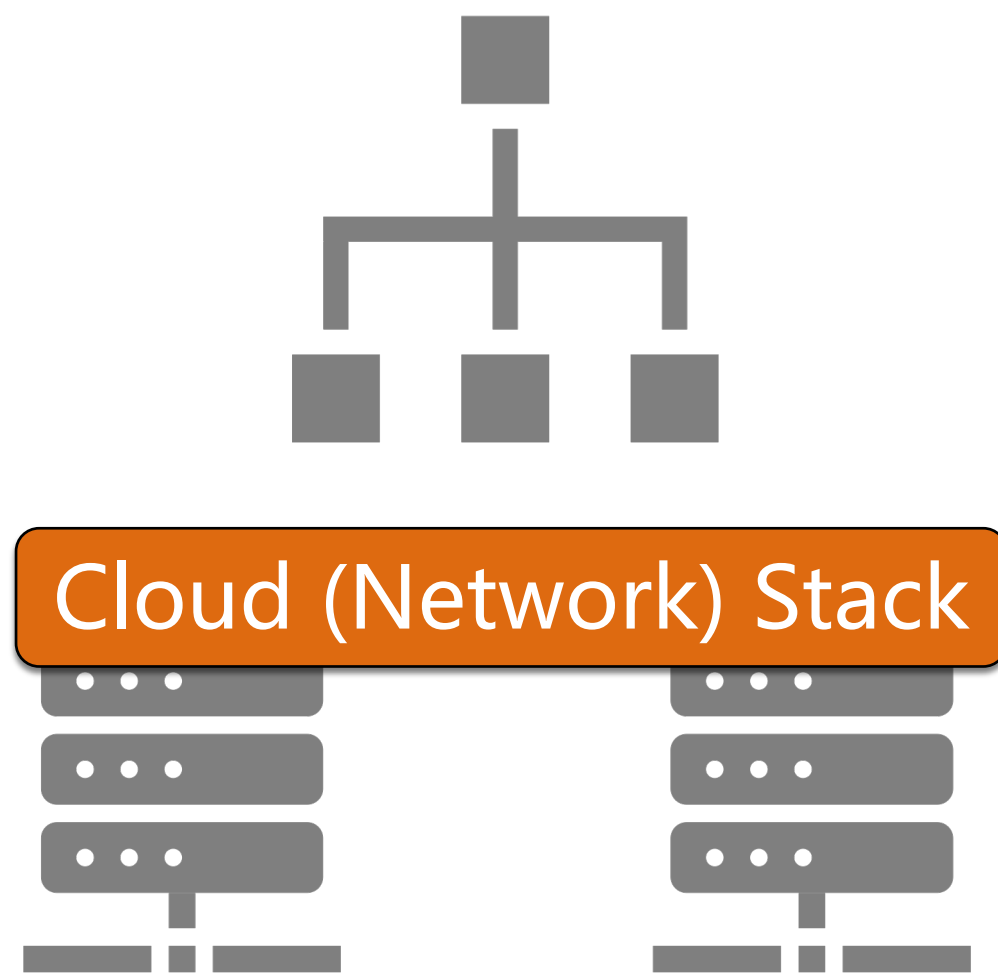
High Performance
due to application demands

Low Host Intrusion
due to heterogenous hardware

Low "Cloud Tax"
due to limited
resources

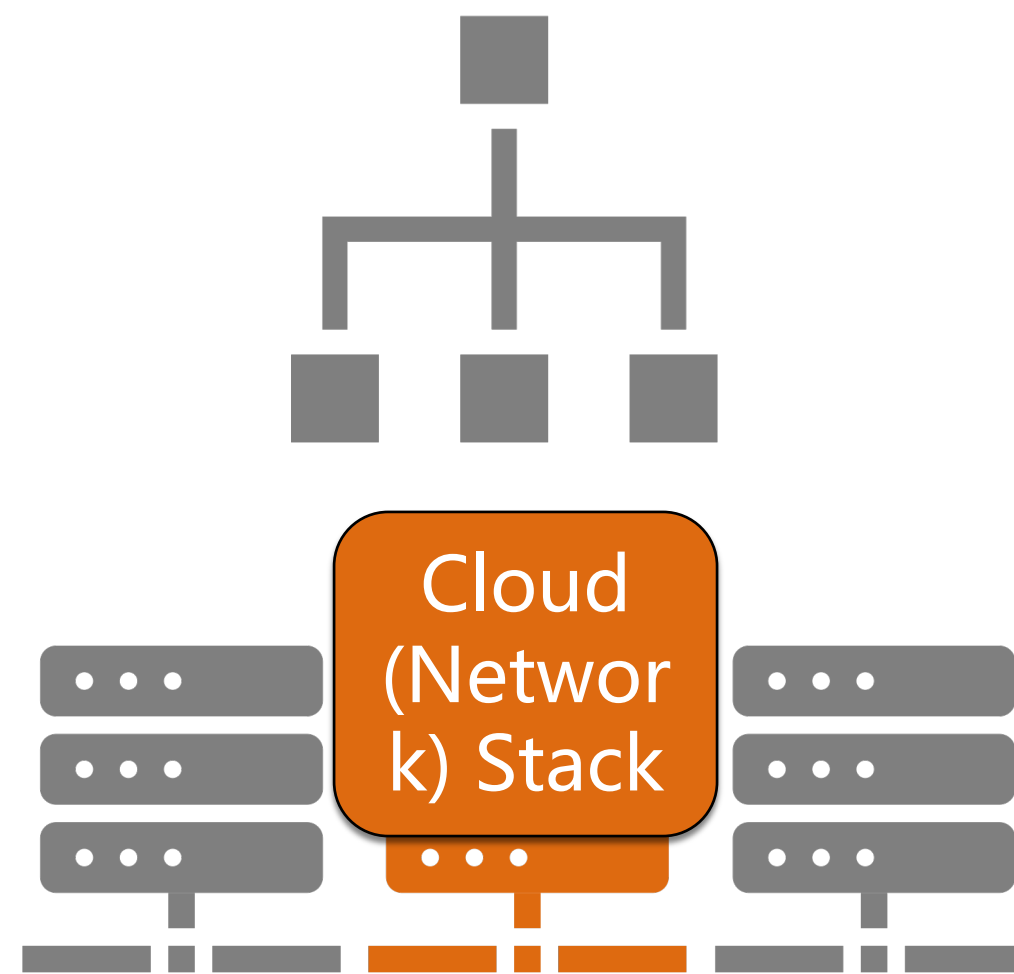
Our solution: edge cloud stack in the network

Infra in every host



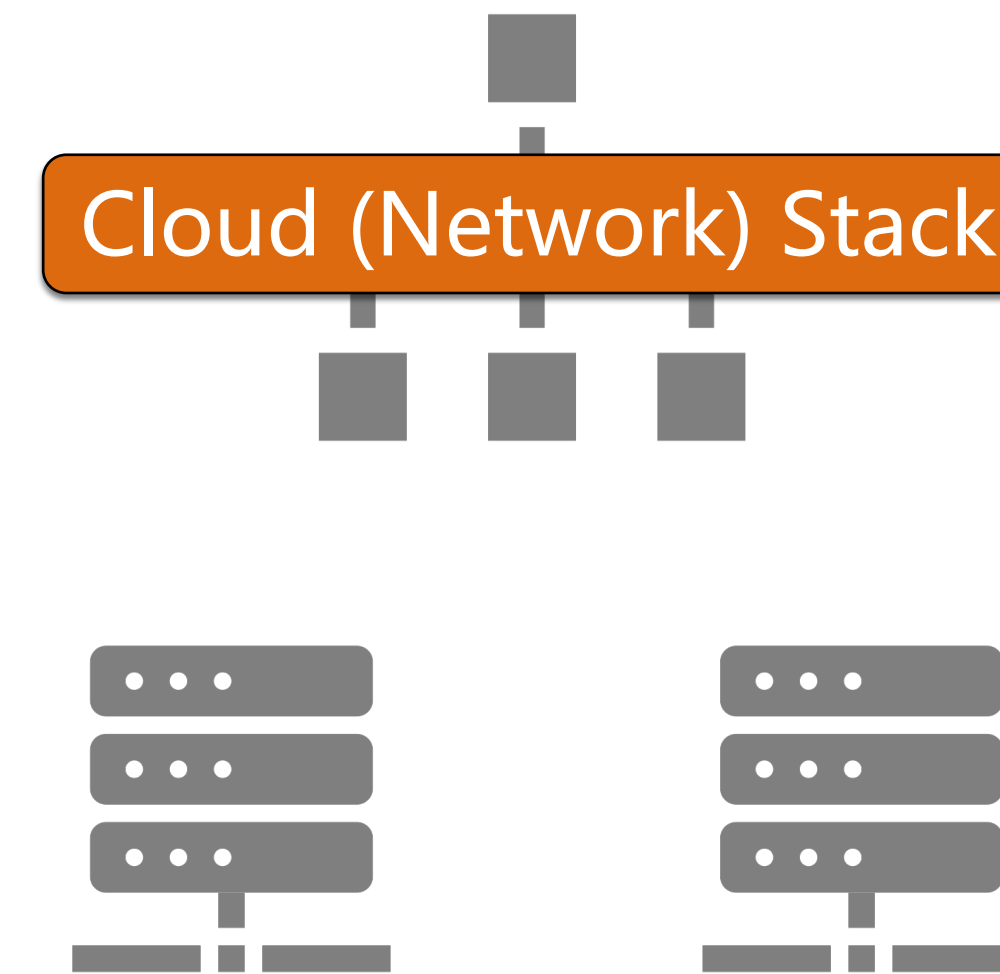
Host Intrusion

Infra in dedicated host



Cloud Tax & Performance

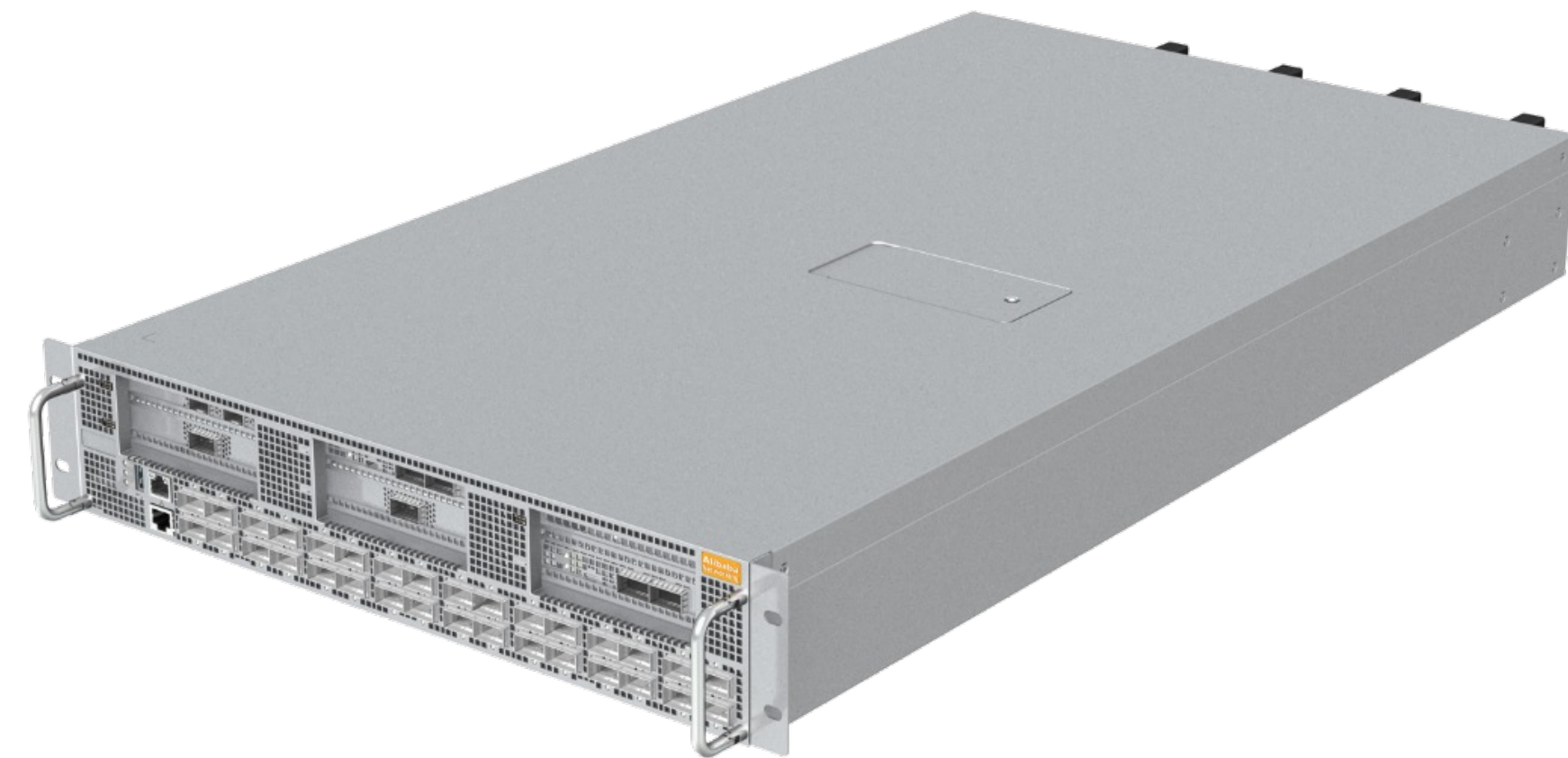
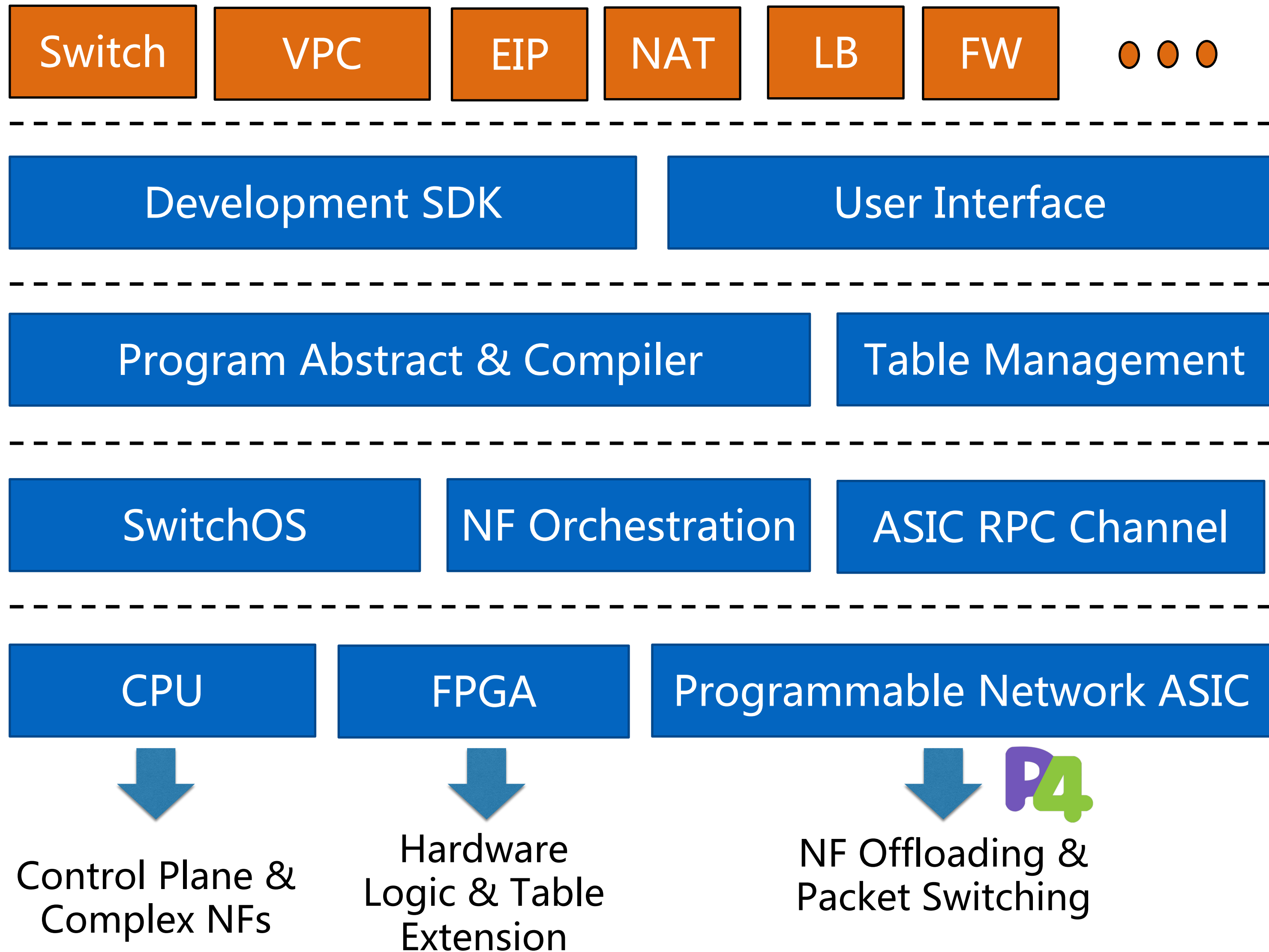
Infra in network



Enabler:
Programmable Network

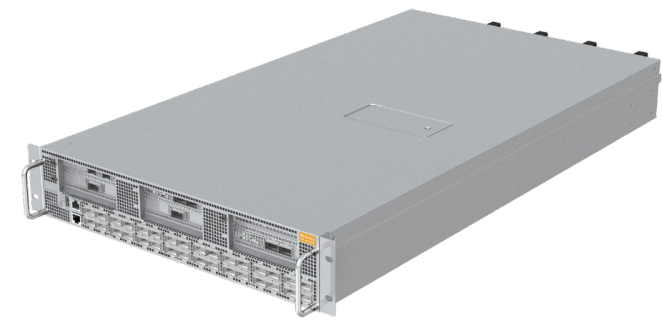


SNA*: a hyper-converged programmable gateway

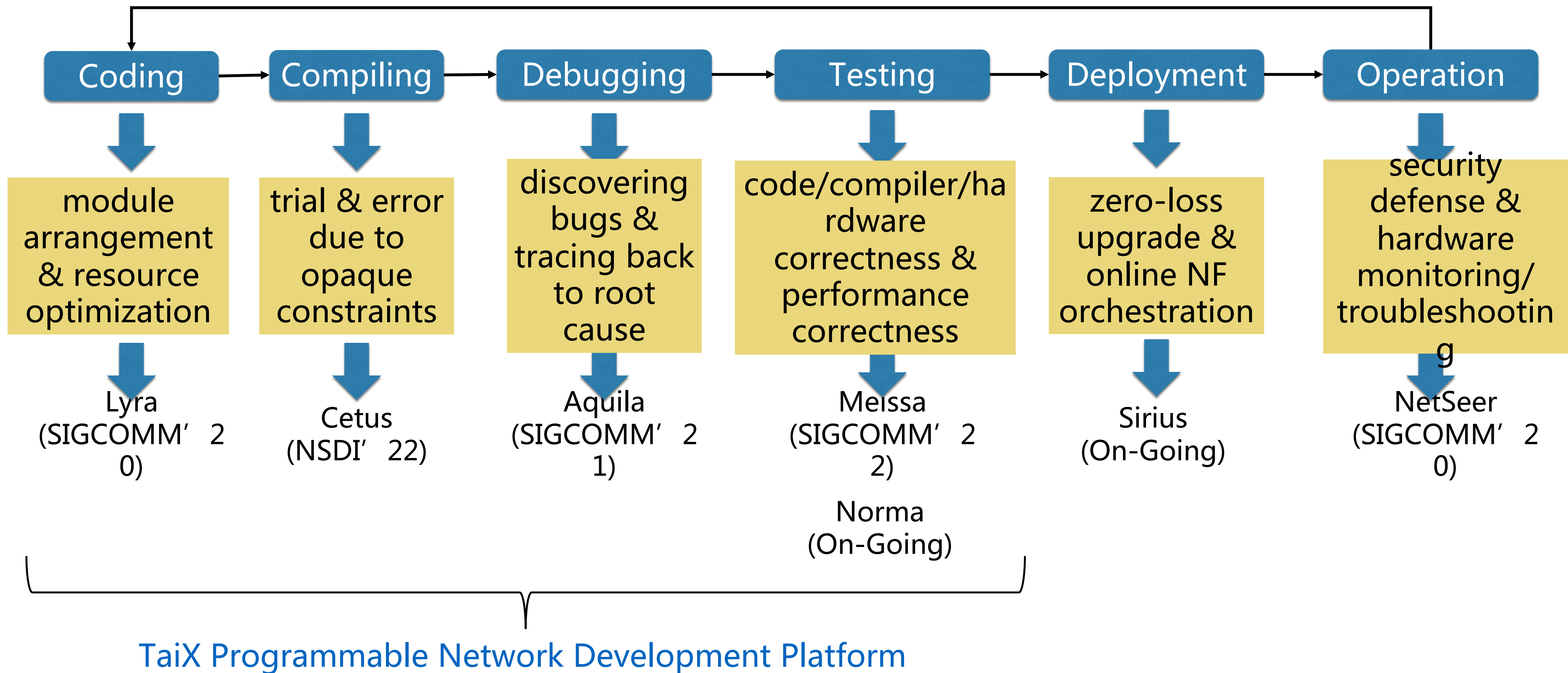


* Smart Network Appliance

Challenges to build and deploy SNA



Efficiency & Reliability
hard to achieve due to the pre-mature development tools in
the network-centric heterogenous computing environment



Conclusion

- The world is accelerating into digitalization. And the cloud computing is becoming a new public infrastructure to support the digital world
- Cloud network is well beyond what traditional network need to provide, from many angles, such as secure resource sharing, HA, performance, operation efficiency and etc
- **Our vision: network is evolving towards Predictable Network.** New emerging technologies will make this possible, such as end-to-end network programmability, programmable hardware switch chips and NICs, and the P4 eco-systems
- In Alibaba Cloud, we already started the journey 5 years ago, and made some significant progress, both in central DC and the edge cloud
- **But this journey is just the beginning**



Thank You