Edge-corE NETWORKS

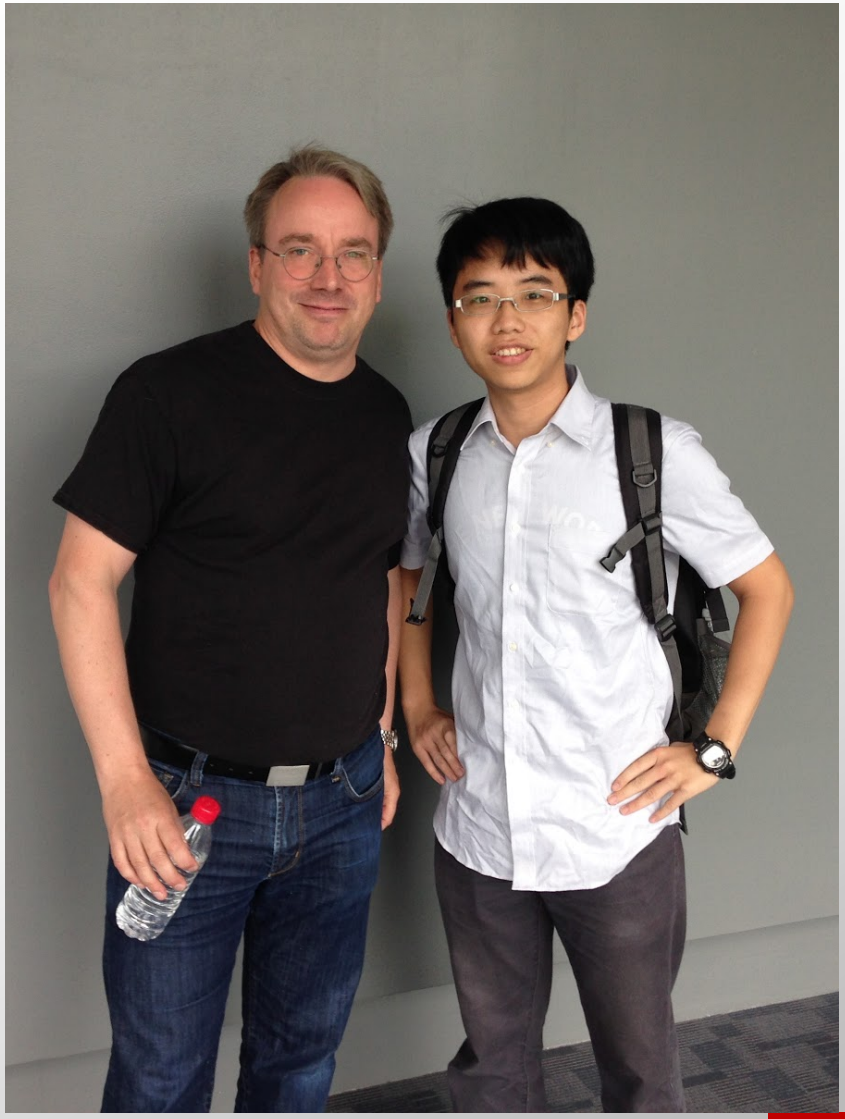
Next Generation IDC Network Solution Overview

An **Accton Group** Company, Q4 2019

Date Huang, Presales Engineer, Edgecore Networks

About Me: Date Huang

- **Speaker:**
  - 2019 OpenInfra Day Taiwan Speaker
    - Massive Bare-Metal Operating System Provisioning Improvement
  - 2019 OSC Tokyo Fall Speaker
  - 2019 COScon '19 Speaker
  - 2019 Hong Kong Open Source Conference Speaker
    - De-centralized Bare-Metal Operating System Provisioning
  - 2018 ISC High Performance Project Poster Demo
    - The Design and Implementation of Bare Metal Cluster Deployment Using BitTorrent
  - 2017 Open Source Summit North America co-Speaker
    - Building Cloud Infra using cost-effective ARM Boards
  - 2017 OpenStack Day Taiwan Speaker
    - Combine Continuous Integration (CI) with OpenStack
  - 2016 OpenStack Day Taiwan Invited Speaker
    - OpenStack on ARM64
- **Projects:**
  - Dozen Cloud
    - Mitaka Release Openstack Cloud Apps VPS by ARM64 CPU
  - EZIO
    - BitTorrnet-based OS Bare-metal Provisioning



Introduction

Edgecore Open Networking Attributes

- Open Networking
- Open Software
- Open Hardware

} **Dis-aggregation!**

Edgecore Open Networking Benefits

Freedom **Control** **Innovation**

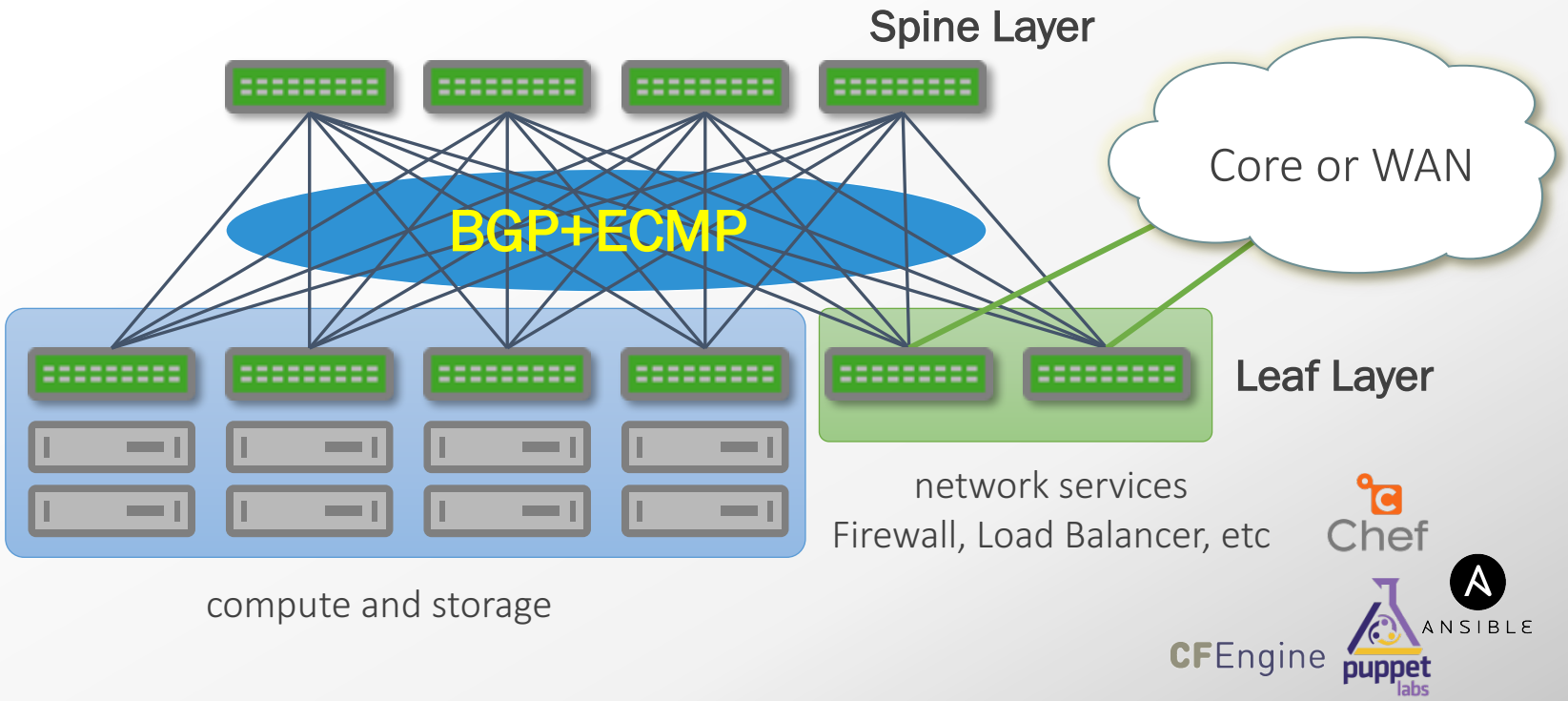
Benefits:

- Disaggregation provides **FREEDOM of Choice** and removes vendor lock-in
- Greater **CONTROL** over Network Infrastructure through open software platforms
- Rapid **INNOVATION** through a community & DevOps approach
- Reduced **CAPEX** and **OPEX**

SDDC Design Principles

Latest SDDC Design Principles

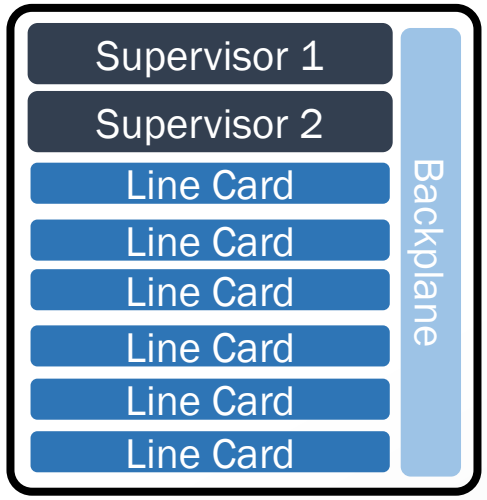
- CLOS for Scale-Out
- Vender agnostic
- Full Life Cycle Automation
- Real-time Monitoring
- Proactive Analytics



CLOS Networks

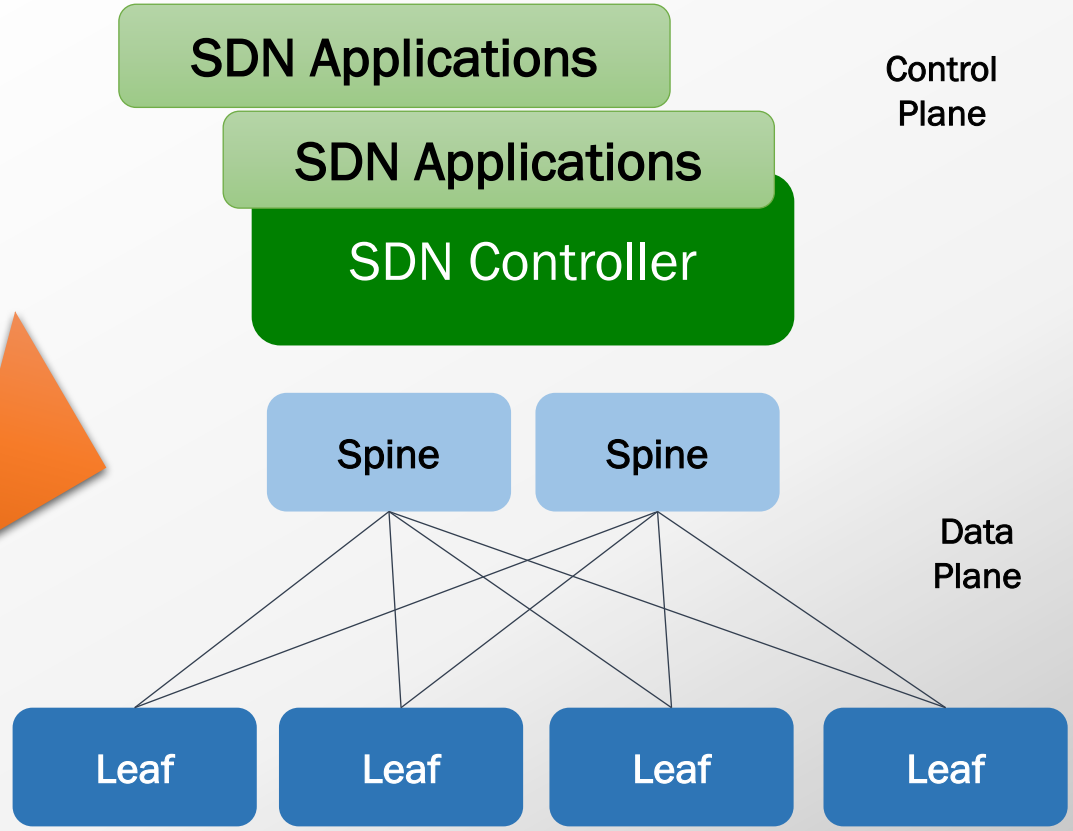
Proprietary Chassis vs. Open Leaf & Spine Fabric

Proprietary Chassis



- Closed & Proprietary Stack (ASIC, HW, SW)
- Expensive
- Complex
- *Scale UP*
- Vendor Lock-In

Lower TCO

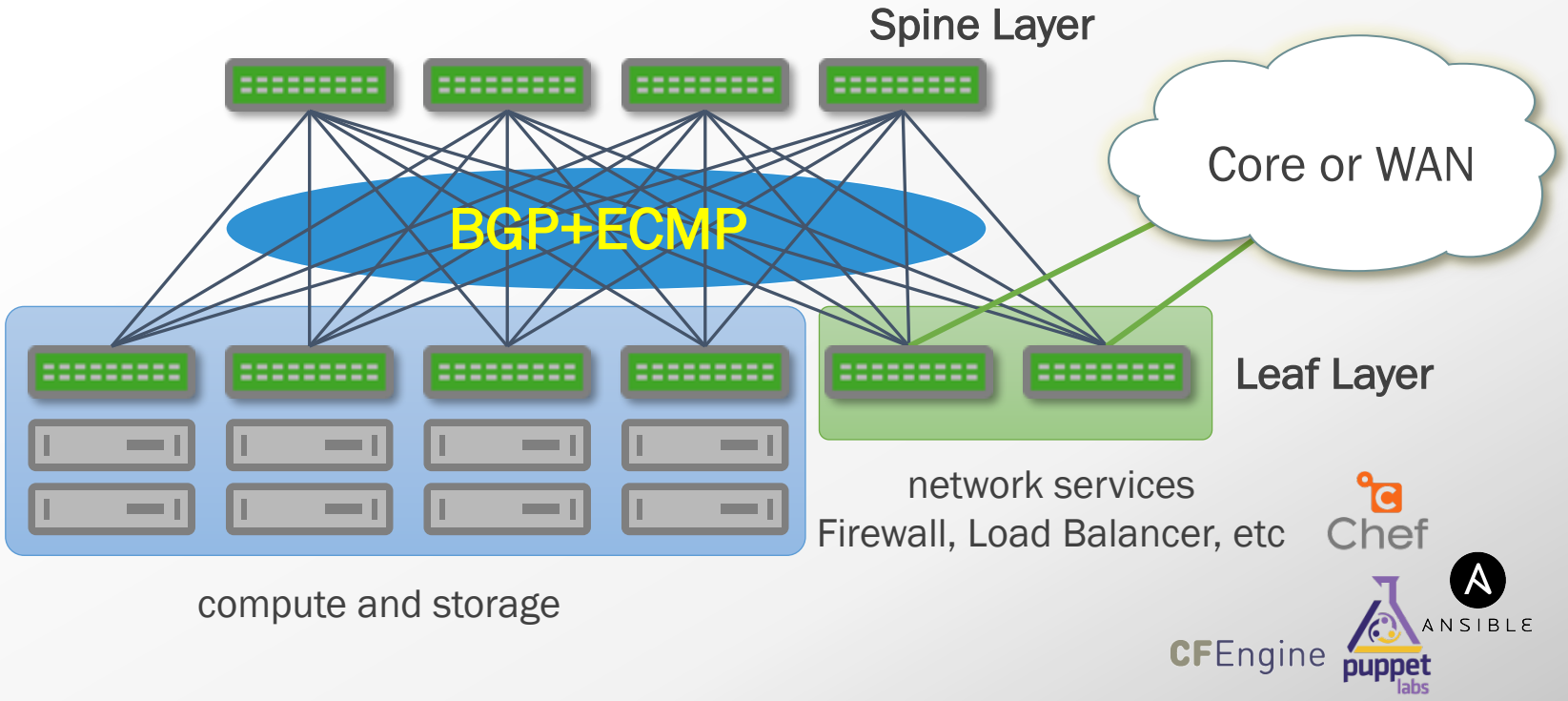


Leaf & Spine / Clos Fabric for Scale-out

CLOS/Leaf & Spine Design Benefits

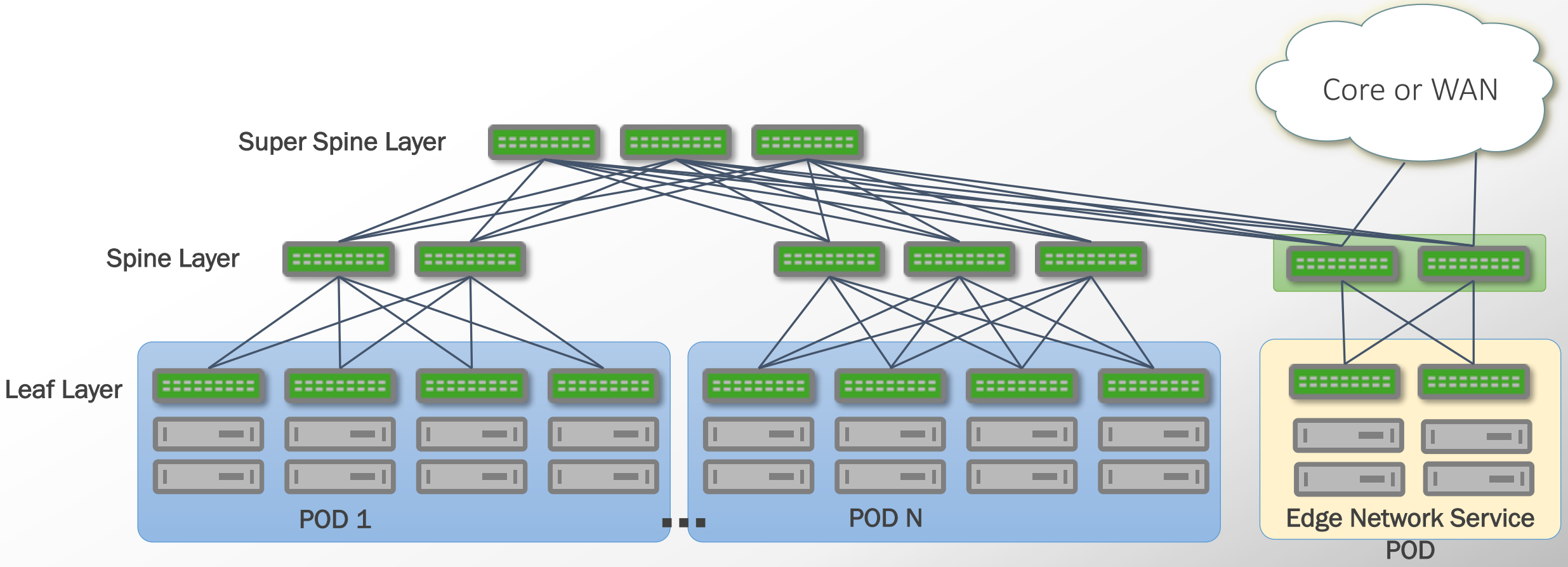
2 tier (3-stage) CLOS Folded Architecture

- Scale-Out architecture that ideal for East-West traffic
- Scale-Out in smaller increments – Pay-As-You-Grow
- Simpler Network – Fewer Protocols to build larger network
- One-hop away predictable latency
- Multiple paths for rich redundancy
- ECMP can be leveraged due to multiple paths

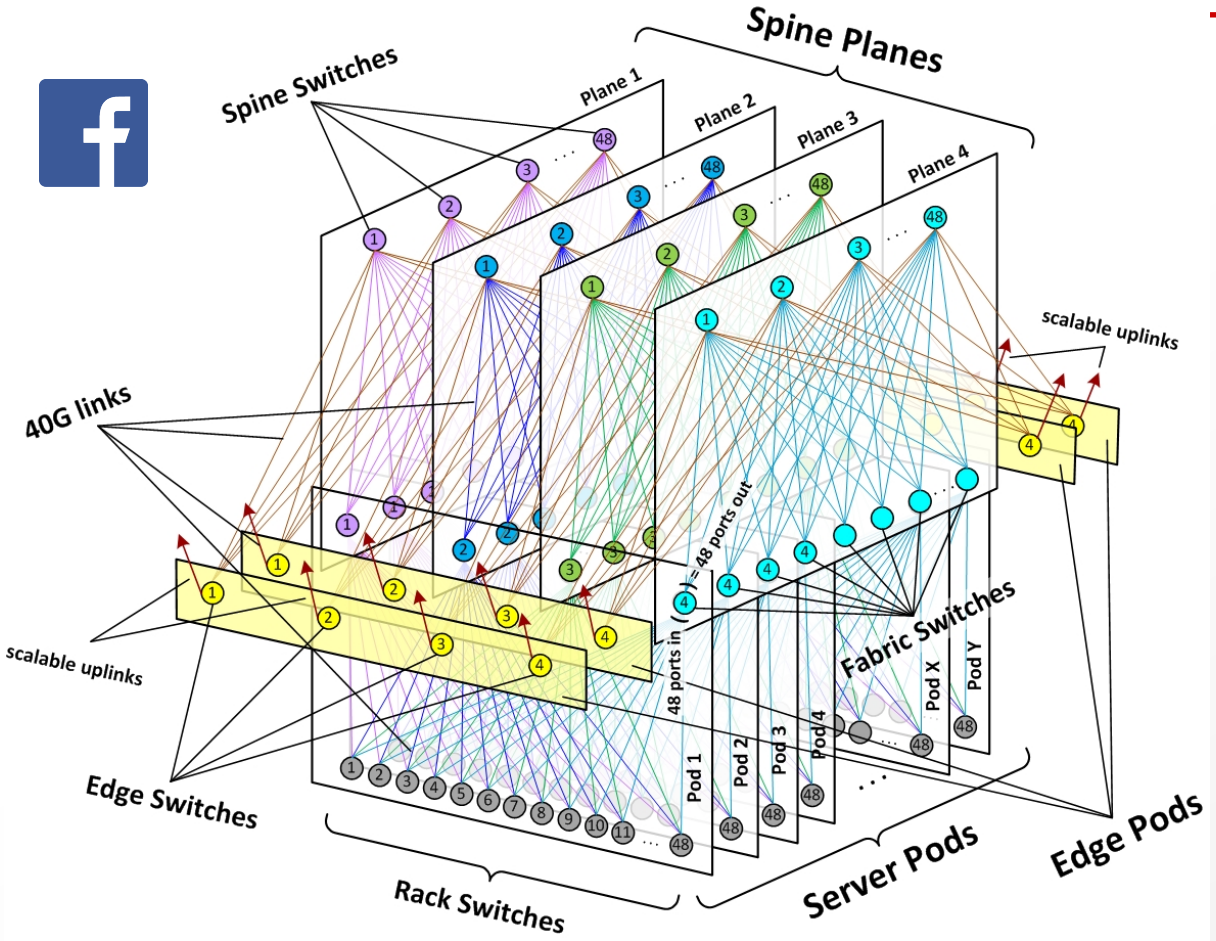


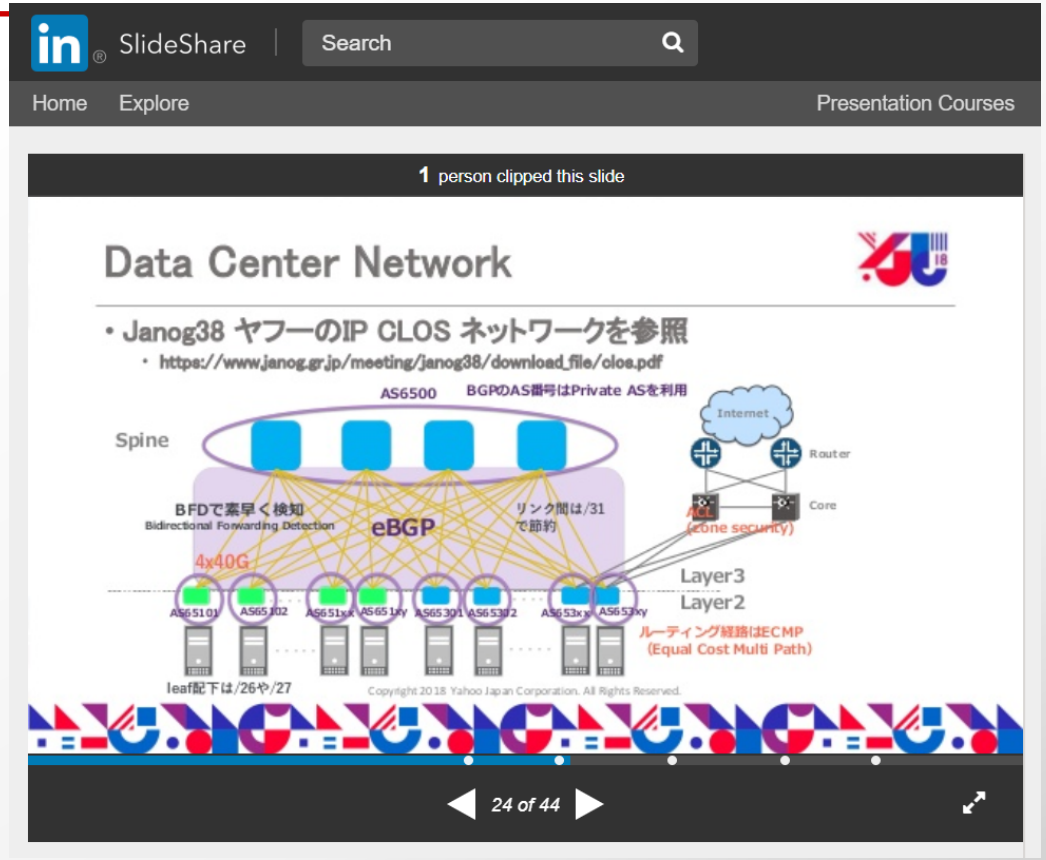
CLOS/Leaf & Spine Design Benefits (Cont'd)

- 3 tier (5-stage) Folded CLOS Architecture



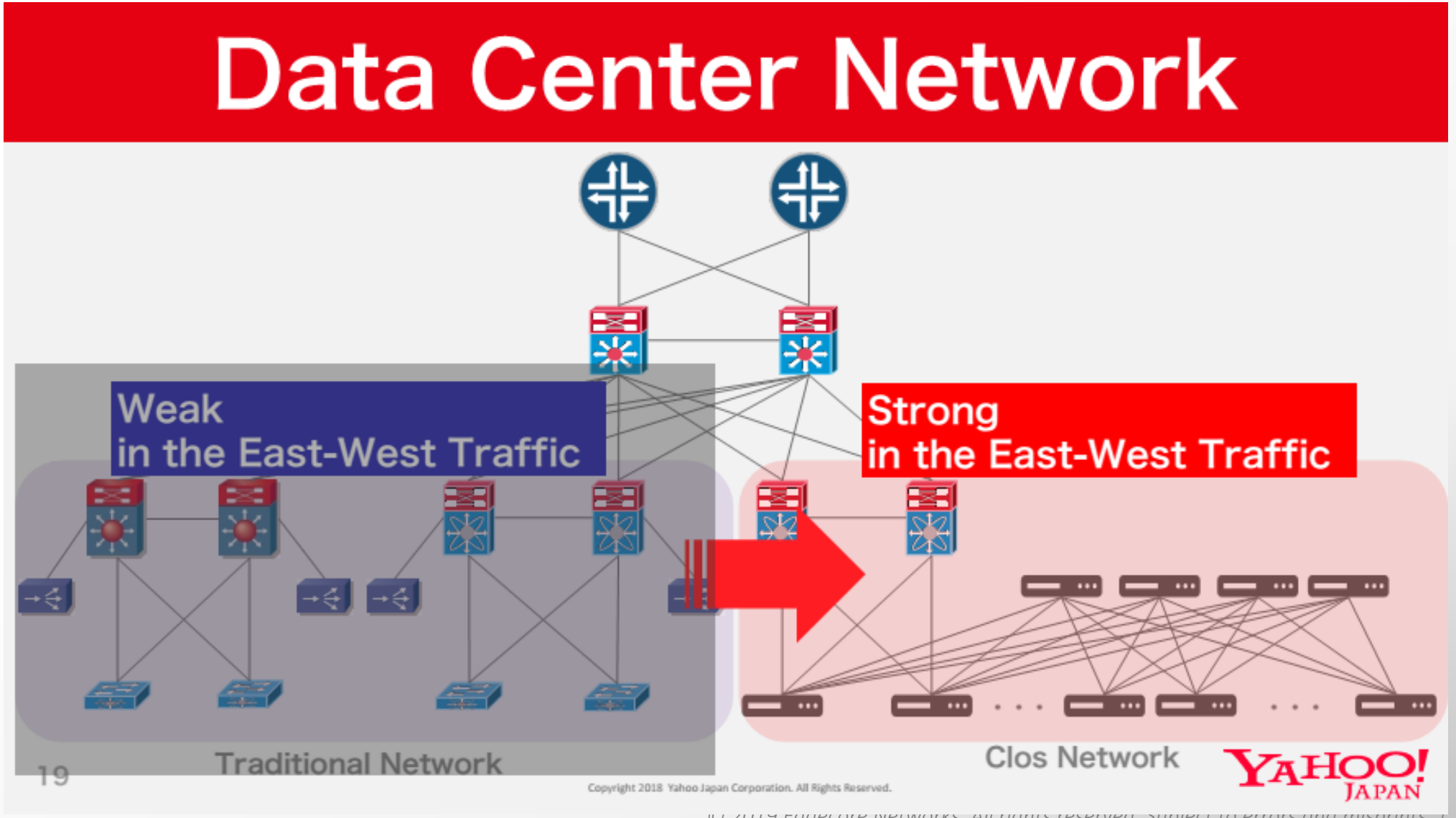
The World-wide Best Practice for Server/Storage Networks



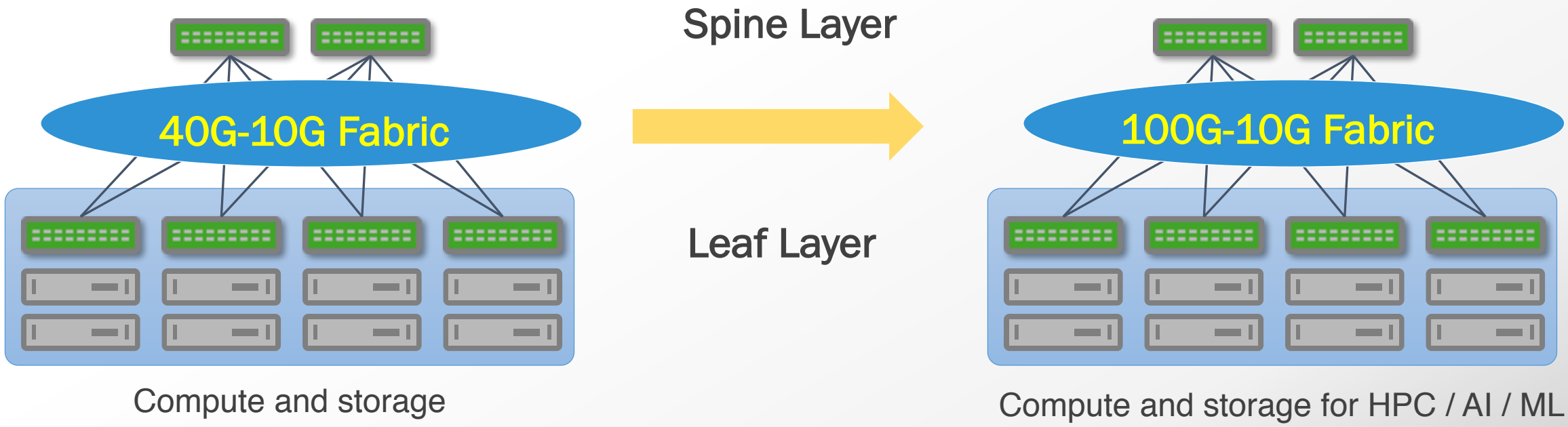


The **"IP Fabric"** Leaf-Spine Network Design (RFC 7938)
Started by Facebook and now deployed by enterprises and organizations around the world

OCP Adoption Example: Yahoo Japan

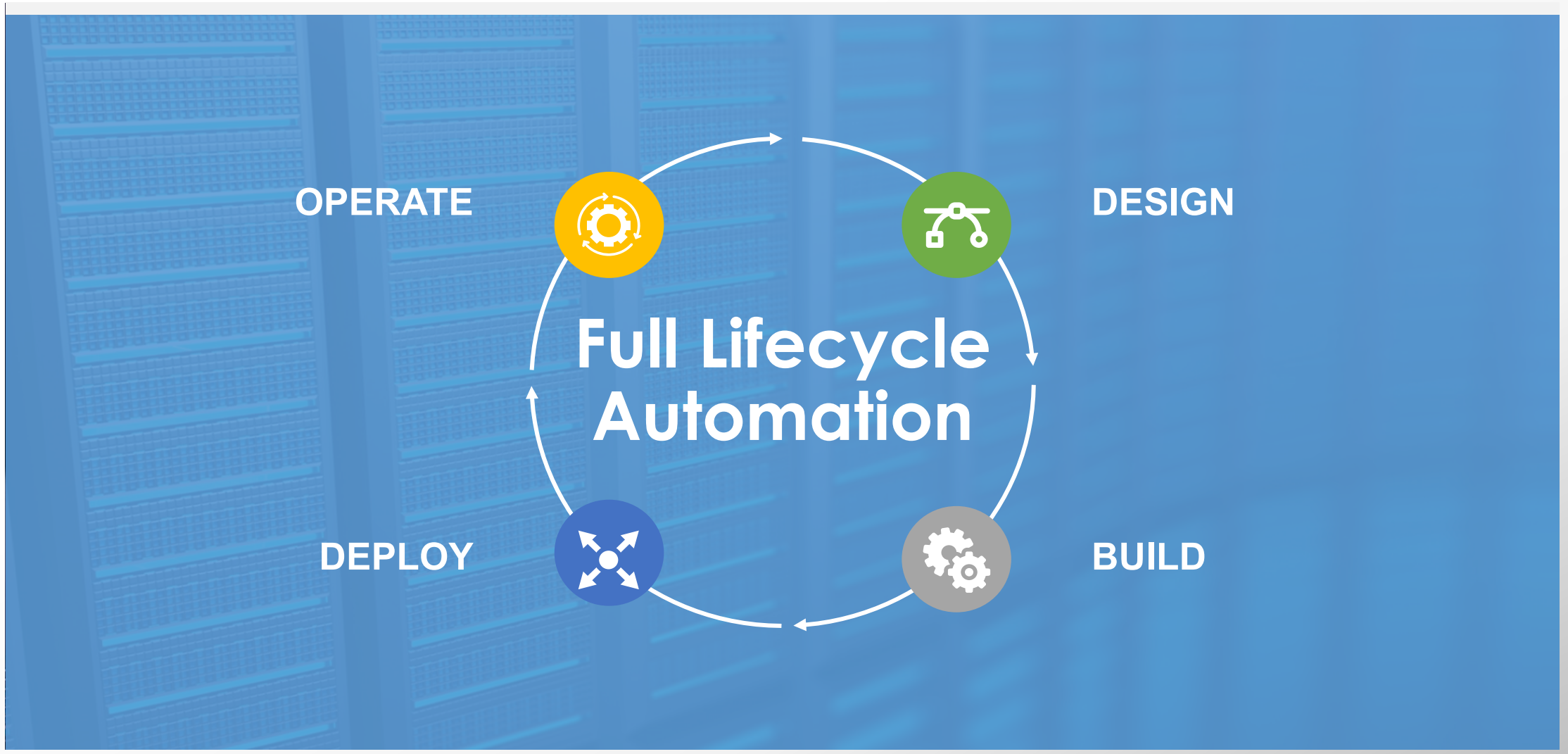


Industry Moving to Larger Bandwidth Fabric



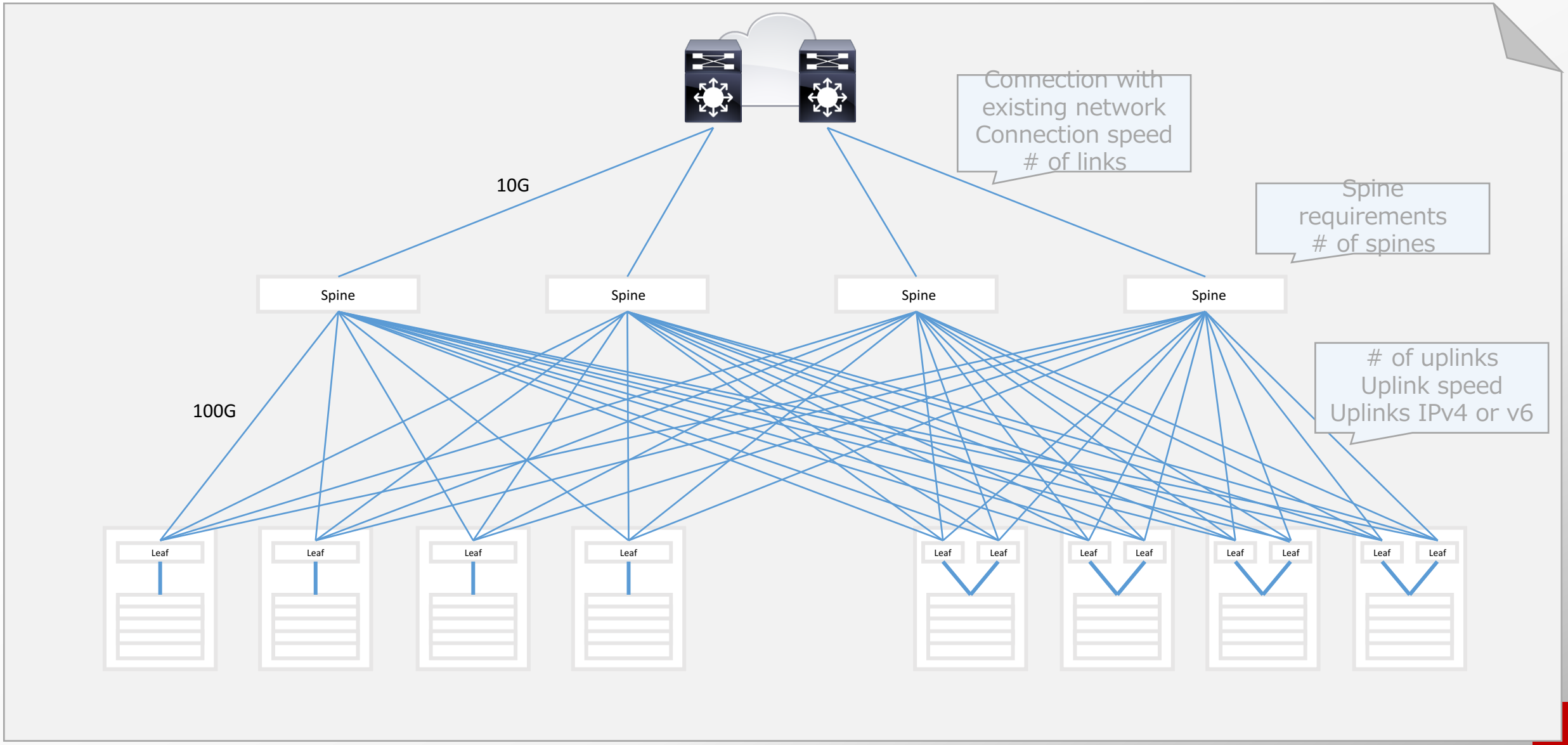
Full Life Cycle Automation

Full Life Cycle Automation



Design & Build Stage

Design & Build -- Template

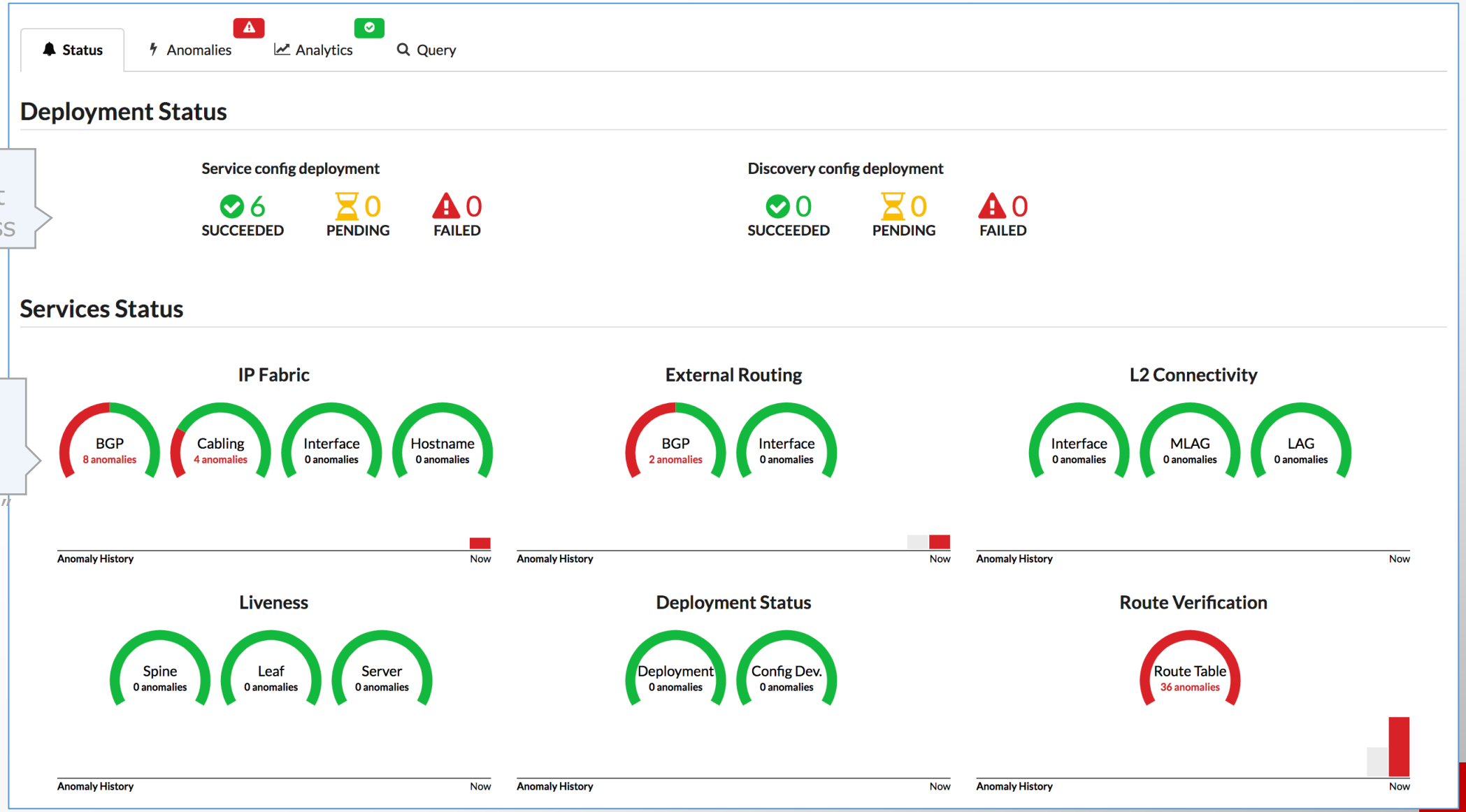


Deploy Stage

Deploy – Blueprint – Telemetry

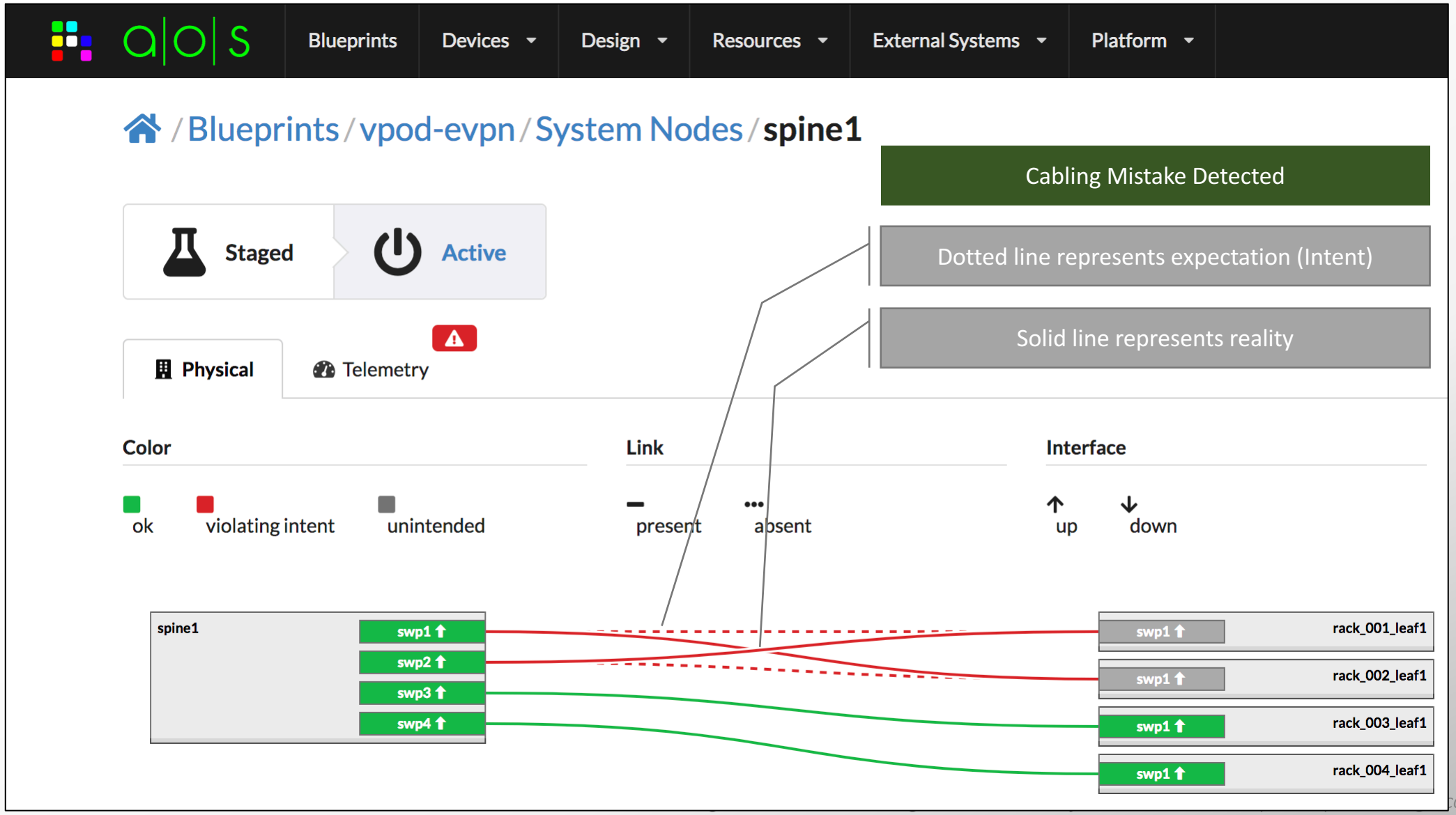
Config deployment task progress

Monitors the health of the overall network system by comparing “Intent” and “Actual”



Operate

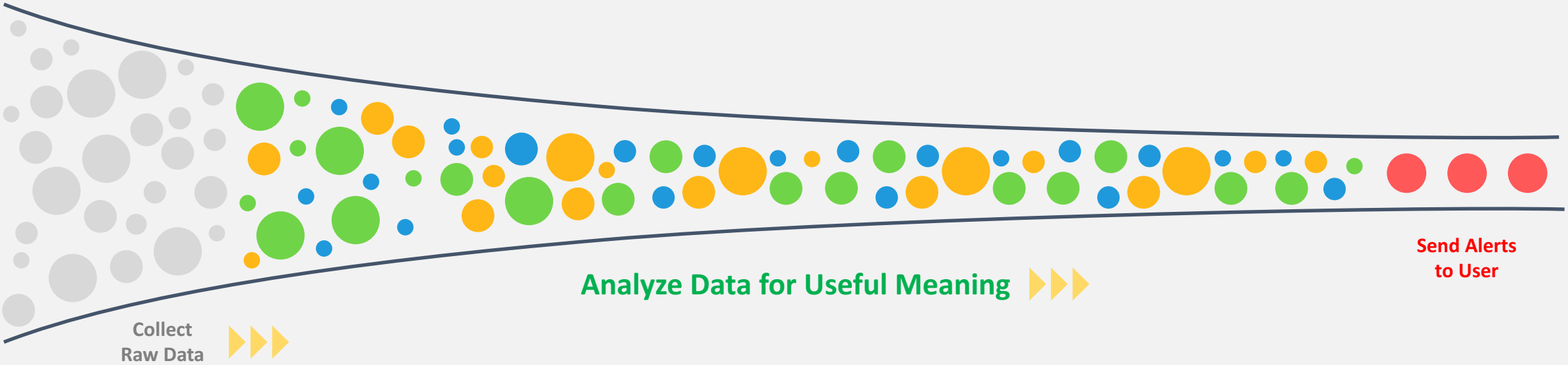
Deploy – Blueprint – Telemetry



Real-time Monitoring

Real-time Monitoring

- Streaming Telemetry - Next Generation Monitoring
- Real-Time Analyze Data for Useful Meaning
- No More SNMP



SNMP Weakness：

- A protocol that is over 20 years old
- Doesn't scale
- Not real-time
  - Polling Interval

- Not supported for modern workloads such as container
- Not reliable (can’t tell if traps are really being sent or delivered)
  - Polling Interval

From PULL to PUSH

	Push (Streaming Telemetry)	Pull (SNMP)
Init Check	Push all info while booting	Wait for server to pull SNMP info
Scalability	Push info to server only in need	Server need to pull info at a certain time
Latency	Low network bandwidth Real-time	High network bandwidth No real-time, Long time interval

Proactive Analytics

Proactive Analytics

- Nip it in the bud
- Intent Based Analytics

Intent Based Analytics (IBA)

- Ask your network a ‘question’:
- Are there **continuous packet drops** on any of my leaf to server interfaces?
- Is there a large **ECMP imbalance** in my leaf/spine network?
- Is my VXLAN/EVPN overlay network correctly programmed across my entire network?
- Network performance is **slow** between Server A and Storage X.
  Is there a **bottleneck** in the network somewhere?
- Are there any peer-links being actively utilized for large amounts of traffic?
- Are physical VLANs and VMware VLANs synchronized on every link between ESXi and leaf switch?
- Is there **memory leak** in any of my network devices?

Proactive Analytics

Operations – Blueprint – Intent Based Analytics (IBA)

for Proactive Monitoring

/ Blueprints / Bellevue

Dashboard | Analytics | Staged | Uncommitted | Active

Proactive Failure Detection

Find EVPN-VXLAN control plane issues

Probes / EVPN IBA Telemetry — Operational — No anomalies — EVPN

Search stages...

- EVPN Type 3 for L2 VNIs
  - EVPN Type 3 for L2 VNIs
- IPV4 EVPN Type 5 routing for virtual networks
  - IPV4 EVPN Type 5 routing for virtual networks
- IPV6 EVPN Type 5 routing for virtual networks
  - IPV6 EVPN Type 5 routing for virtual networks
- EVPN IPv4 Type 5 status for virtual networks
  - EVPN IPv4 Type 5 status for virtual networks
- EVPN IPv6 Type 5 status for virtual networks
  - EVPN IPv6 Type 5 status for virtual networks
- Expected remote VTEP count
  - Expected remote VTEP count
- Remote L2VNI VTEP status

Stage: EVPN Type 3 for L2 VNIs — Discrete State Set — Type-3_BUM

EVPN Type 3 for L2 VNI routes represents the list of VTEPS that a leaf will flood traffic to. VTEP and RD indicates that reachability between system and remote-system is impaired s

An association is made between remote VTEP IP and a deterministic route distinguisher. N

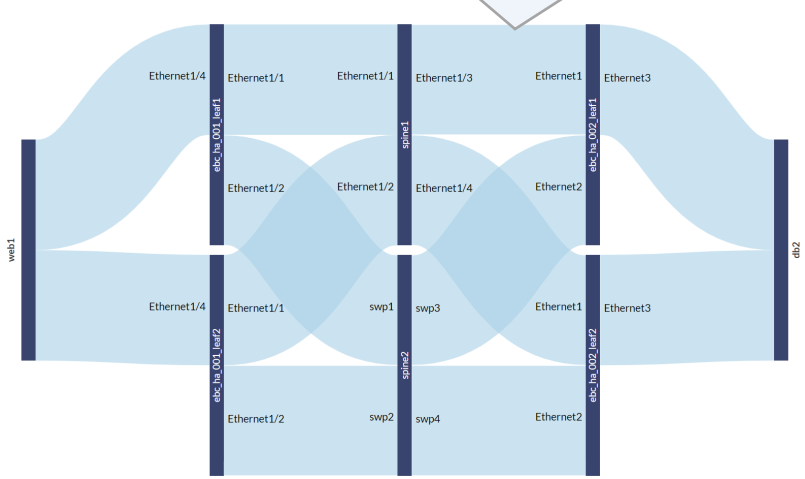
Search stage data...

up

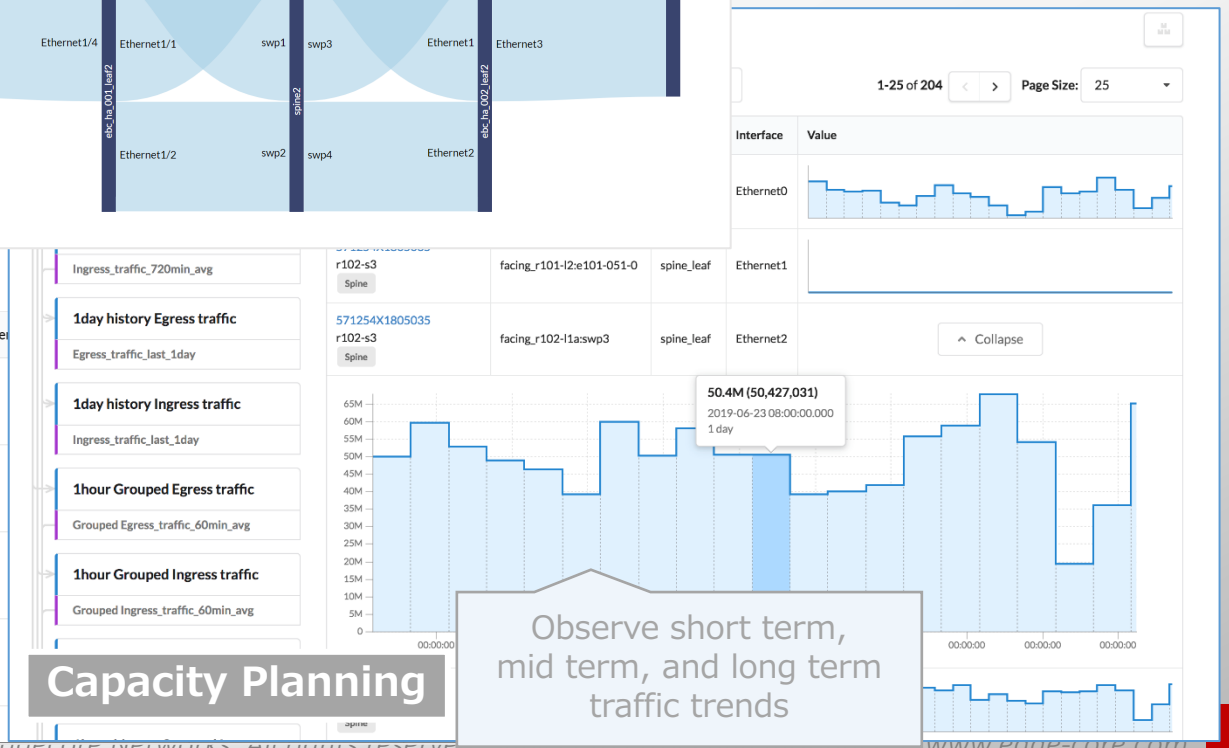
System ID	Key	Label	Local Vlan	Rd	Remote Hostname	Re
525400141590 ebc-ha-002-leaf2 Leaf	10.0.0.2:4|100.100.100.1	db	4	10.0.0.2:4	ebc-ha-001-leaf1	4
525400141590 ebc-ha-002-leaf2 Leaf	10.0.0.2:5|100.100.100.1	web	5	10.0.0.2:5	ebc-ha-001-leaf1	5
525400141590 ebc-ha-002-leaf2 Leaf	10.0.0.2:6|100.100.100.1	LB	6	10.0.0.2:6	ebc-ha-001-leaf1	6
525400141590 ebc-ha-002-leaf2 Leaf	10.0.0.3:4|100.100.100.1	db	4	10.0.0.3:4	ebc-ha-001-leaf2	4
525400141590						

Troubleshooting

Find bottlenecks in the network



Capacity Planning

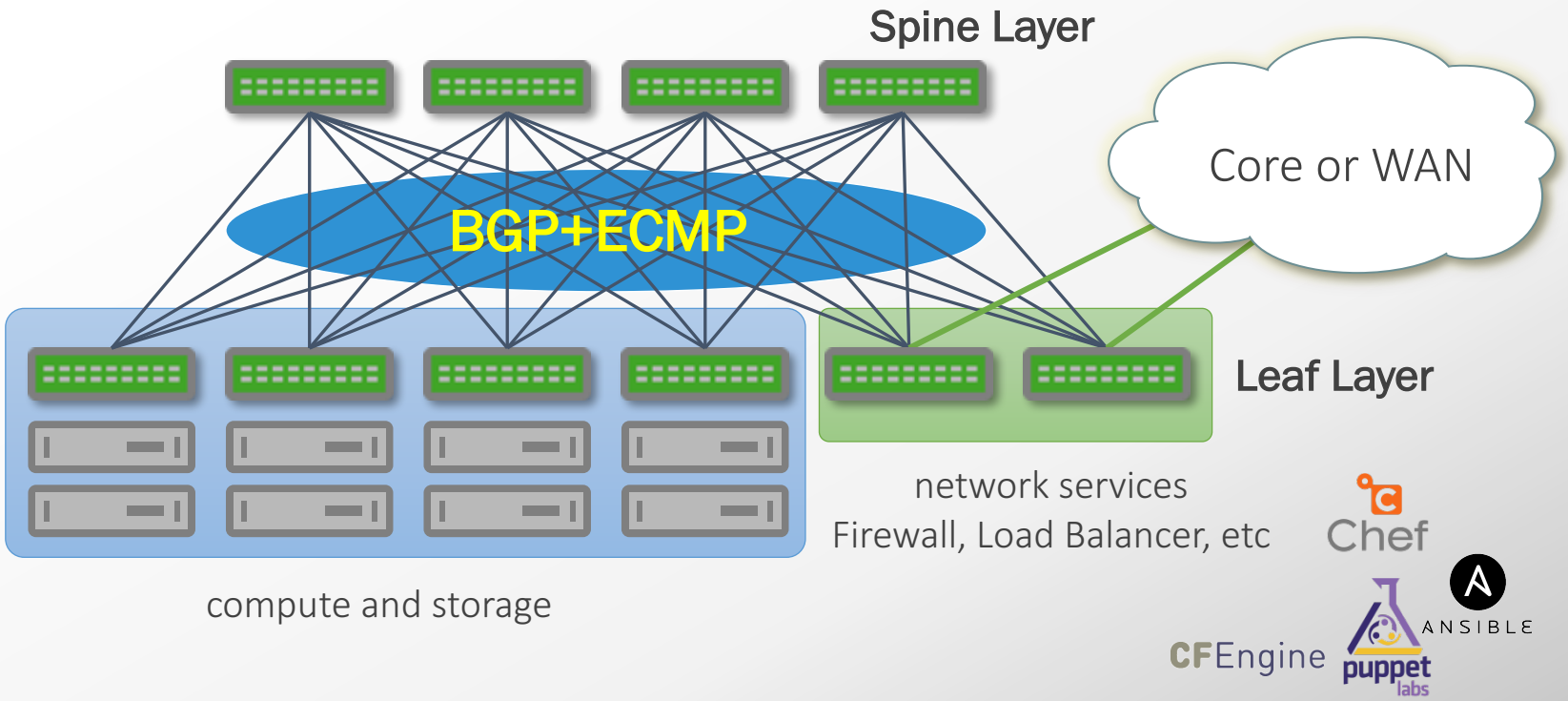


Observe short term, mid term, and long term traffic trends

Recap

Recap: Latest SDDC Design Principles

- CLOS for Scale-Out
- Vender agnostic
- Full Life Cycle Automation
- Real-time Monitoring
- Proactive Analytics



www.accton.com

www.edge-core.com

THANK YOU