Software-Defined Infrastructure at Uber Scale

Justin Dustzadeh

Head of Global Network & Software Platform

Open Networking Summit 2018
Outline
Software-Defined Infrastructure at Uber Scale

○ Who Is Uber?
○ Vision & Architectural Principles
○ Infrastructure Automation
○ Uber’s Global Network Infrastructure
Who Is Uber?
Uber by the Numbers
Continued Rapid Growth at Global Scale

- Completed 1Bth trip Dec 2015, 2Bth trip Jun 2016, 5Bth trip May 2017
- Completed 4B trips in 2017
- Completing ~15M trips daily
- Operating in 600+ cities, 78 countries, 6 continents
- 75M active monthly riders
- 3M active drivers globally
Vision & Architectural Principles
Vision & Architectural Principles
Software-Defined Infrastructure at Uber Scale

○ **Mission**: bring transportation — for everyone, everywhere

○ Highly-available global infrastructure

○ Scale and **automate** through software

○ From closed solutions to **open, standards-based** technologies

○ No single-vendor dependency
Visibility & Control

- Control
- Insights
- Real-Time Visibility
- Analytics & ML

Infrastructure
Infrastructure Automation Through Software

- Auto Provisioning
- Change Automation
- Auto Detection
- Auto Remediation

Asset Lifecycle Management

Infrastructure

- Server
- Network
- Data Center
Automated Provisioning Through Software

- Auto discovery *(network, server)*
- Network automation: zero-touch provisioning
- Server automation: imaging & configuration
- Auto-validation & deployment readiness *(network, server)*:
  - Asset tracking, health checks (device & connectivity), intended state validation

Fully-Automated Provisioning At Scale Through Software
Automated Change Management

- Vendor-agnostic, automated change management (network, server)
- Source of truth for the desired state of infrastructure
- Behavior defined in configuration data (vs through direct interactions with devices)
- Data drives device config generation
- Descriptive model for change management -- path to truly-intent-based software platform
- Infrastructure-as-code: versioned with rollbacks, reviewable, auditable, portable (on-prem/cloud)
(Near) Real-Time Visibility & Analytics

- Auto Provisioning
- Change Automation
- Auto Detection
- Auto Remediation

- Monitoring:
  - Passive and active
- Streaming telemetry (OpenConfig) with YANG-based vendor-neutral data models
- Near-real-time network visibility for intra-DC, inter-DC, cloud to DC, DC to external services
  - Reachability, round-trip & one-way latency, packet drops (silent or bursty) and black holes
- Distributed and highly-available design
- Server monitoring
Auto Mitigation & Auto Remediation

- Auto mitigation & auto remediation (network, server)
- Automated remediation workflows (RMA/troubleshooting tickets, etc.)
- Framework to add diagnostic & auto-remediation tests
- Automated server replacement
Uber’s Global Network Infrastructure

Data Center & Wide-Area Networks
Data Center Network Design Principles

Highlights

○ Resiliency
○ Simplicity
○ Scalability
○ Flexibility
○ Cost effectiveness
Data Center Network Fabric

Example 6-Plane Fabric

POP 1
POP 2

Edge Pod
Backbone router
Multi-services

DC Aggregation
Hub for inter-fabric connectivity

External Fabric Switch
Fabric egress connectivity

Fabric Switch
Inter-pod connectivity

Pod Switch
Rack aggregation

Rack Switch
Server aggregation
Data Center Network Fabric

PoP

Edge Pod

DC Agg

Fabric

PoP
Multi-Availability-Zone Design

Example Multi-Campus Metro Design

Campus 1

Zone 1

Building 1

Fabric 1

DC Agg

Edge Pod

Campus 2

Zone 2

Building 2

Fabric 1

DC Agg

Edge Pod

Building 3

Fabric 2

Campus 3

Zone 3

Building 1

Fabric 1

Fabric 2

Zone 4

Building 1

Fabric 1

DC Agg

Edge Pod

Zone 5

Building 2

Fabric 1
WAN, Regions, Availability Zones

Multi-AZ, Multi-Region

Region

PoP

Zone 1

Zone 2
Global Network Infrastructure

End-to-End View

SD-WAN
- Full line-rate crypto at 100G

DWDM
- Scalable metro connectivity

Regional & Long-Haul Transport
- Prefer a more dynamic approach, moving away from static, long-term contract models
- Open optical line systems
- FlexGrid support
- Take advantage of new transponder technology and shorter product development life cycles (DSP, modulation, baud)
- SDN → on-demand spectrum-as-a-service
- Buy ROADM ports and spectrum, not lit 100G circuits

Software-Defined
- Auto-provisioning
- Change automation
- Auto-detection
- Auto-remediation
- Server ODM
- Network whitebox

ACCESS
METRO
BACKBONE
DC
Network Disaggregation
Network Disaggregation
Data Center Switching

- Management Plane Software
- Network Protocol Stack
- Operating System (OS)
- Hardware Platform
- Switching ASIC

Network Applications & Management
Network Protocol Stack
Network OS for Bare-Metal Switches
- Open Abstraction Interfaces
- ONIE
Merchant Silicon
Bare-Metal Manufacturers (ODM)
Network Whitebox Considerations

3rd-party Network OS (NOS) Adoption

- NOS feature parity
- Hardware abstraction architecture & design
  - Kernel-agnostic NOS design: reduced operational risk for changes & new hardware support
- Code support model
  - User-managed vs vendor-managed
  - Feature enhancements & software support (patches, vulnerability management)
  - Software maturity & stability based on installed base
  - Expertise for hardware integration
- Full operational readiness