Open DCN networking

Andrew Qu

2019.01.15
Open Networking @ Tencent

- Cloud Computing
- Software Defined Networking
- Network Function Virtualization
- Open Network Devices
- Open Compute Hardware
- Intention based Processes
- Automation
- Open source Software
Open Networking Ecosystems @ Tencent

- Open source Software
  - Open/R
  - GoBGP
  - FRRouting
  - BIRD
- Software Defined Networking
- Intention based Processes Automation
- Open Network Devices
  - Ansible
  - influxdb
  - Grafana
  - Kafka
  - OpenDaylight
  - Batfish
  - MiniNet
  - StarlingX
  - Co-development
Tencent open DCN networking is in making

Proactive Validation
Comprehensive correctness guaranteed

Network Design (Changes) → Review Emulating Lab → Deploy

Monitor

Open source Software

Open Network Devices
Software Defined Networking

Continuous Validation
Promptly reports active and potential violations

Intention based Processes
Automation
Automation

Tencent DCN Open Switches

- Front panel 48*25G+8*100G;
- 3+1 Fans, 1+1 PSUs;
- ASIC: BCM56873 (TD3)
- Selected Cage, PSU, FAN;
- Selected CPU: IntelD-1527;
- Dual: BIOS, Dual BMC image;
- Optical Module Init Process;
- I2C architecture;
- UART customization;

- 4U, Horizontal, Pluggable;
- 100G or 400G Line cards;
- 100G and 400G LCs;
- 5+1PSUs, 2+2 Fans;
- ASIC: BCM56880 (TH3)
- Trident4 considered;
- Selected Cage, PSU, FAN;
- CPU: Intel D-1527;
- Dual BIOS, Dual BMC image
- I2C architecture;
- UART customization
Tencent DCN Switch roadmap

- LA/LC: Initial deployment in 2019, more afterwards
Tencent DCN: Open Optical module

Certification Process/Lab setup

- Background checking
  - Stability in resource
  - Historical shipment record
  - Capacity in production
  - Shipments
  - System issue record
- Module device analysis
  - DPA analysis
  - Module performance test
  - Reliability
  - Long fiber test
- Interoperability
  - Between Vendors
  - Long fiber test

Systems Test → EFT → MP

Interoperability

Risk in compatibility of high speed electrical signal: Between ASIC PHY and the Optical module, SERDES signal degradation via PCB could be severe, which causes jitter, hence the error in decoding.

Other problems are like: System flaw in shipments, Switch hardware failure in shipments, Software patch caused issue, more specifically, these issues are hard to find in early testing cycle.

Reliability

Tencent现网故障统计

- Laser: 36%
- Others: 80%
**Tencent DCN: Open Optical module**

--- Next Generation consideration

**Architecture evolution**

- CUF
  - 4*25G NRZ CWDM4-5M
  - 4*100G PAM4 FR4-2km
  - 4*100G PAM4 DR4-2km
- LC
  - 4*25G NRZ SR4-MM
  - 8*50G PAM4 MM SR8
  - 4*100G PAM4 5M DR4
- LA
  - 25G NRZ
  - 25G AOC
  - 50G PAM4
  - 100G ACC

**25G Based**

**100G Based**

**100G/400G ODM Tencent spec Modules**

- Low cost 100G ACC technology:

<table>
<thead>
<tr>
<th>Technology</th>
<th>100G PAM4</th>
<th>25G NRZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>2.8m</td>
<td>7m</td>
</tr>
</tbody>
</table>

- Silicon Photonic 400G eDR4:

**Cost**

- 100G SiP
  - 500m
  - 2km

**Distance**

- 100G EML
THANKS
Tencent open DCN networking in making

Network Emulator

Configuration

Network Policy Checker

Domain Network Controller

Certifications:
- All devices are password protected

Violations:
- Subnets of Leaf-1 and Leaf-3 cannot communicate
- rtr-y failure reduces availability