Build Reliable Cloud Networks with SONiC and ONE

Wei Bai 白巍
Microsoft Research Asia

OCP China Technology Day, Shenzhen, China
Two Open Source Cornerstones for High Reliability

Networking OS: SONiC

Network Verification: ONE
Networking OS: SONiC
A Solution to Unblock Hardware Innovation

Monitoring, Management, Deployment Tools, Cutting Edge SDN

Switch Abstraction Interface (SAI)

Merchant Silicon
Switch Abstraction Interface (SAI)

Simple, consistent, and stable network application stack

Help consume the underlying complex, heterogeneous hardware easily and faster

https://github.com/opencomputeproject/SAI
SONiC High-Level Architecture

[Diagram showing the high-level architecture of SONiC, including layers such as Switch Hardware, Linux Kernel, Linux User Space, and Application, with components like Hardware Sensor App, LAG App, LLDP App, BGP App, SNMP App, Object Library, Switch State Service, SAI API, SAI Implementation, Sysfs, Netdev, Platform Drivers, Networking Drivers, ASIC Drivers, Fan/Power/Led, Transceivers (Media), and ASIC.]
Switch State Service (SWSS)

- APP DB: persist App objects
- SAI DB: persist SAI objects
- Orchestration Agent: translation between apps and SAI objects, resolution of dependency and conflict
- SyncD: sync SAI objects between software and hardware

**Key Goal:** Evolve components independently
**SONiC Containerization**

---

*Software for Open Networking in the Cloud*

Unauthorized access and/or use are prohibited. All access and/or use are subject to monitoring.


---

```
CONTAINER ID   IMAGE                         COMMAND                     CREATED       STATUS          PORTS      NAMES
8a3abdd1a8cc  docker-snmp-sv2:latest       "/usr/bin/supervisord"       4 days ago    Up   2 days       snmp
580ad5ee729a  docker-syncd-brcm:latest    "/usr/bin/supervisord"       4 days ago    Up   2 days       synced
2bd4ee0d793b5  docker-orchagent-brcm:latest  "/usr/bin/supervisord"       4 days ago    Up   2 days       swss
6320785de98e  docker-dhcp-relay:latest    "/usr/bin/docker_in_."       4 days ago    Up   2 days       dhcp relay
79234c4d010e  docker-fpm-guagua:latest    "/usr/bin/supervisord"       4 days ago    Up   2 days       bgp
89b76361d4d85d  docker-platform-monitor:latest  "/usr/bin/supervisord"     4 days ago    Up   2 days       pmon
3499815fd2c3  docker-lldp-sv2:latest      "/usr/bin/supervisord"       4 days ago    Up   2 days       lldp
0ce0b837def2  docker-teamd:latest         "/usr/bin/supervisord"       4 days ago    Up   2 days       teamd
144eda8ba6bf  docker-database:latest      "/usr/bin/supervisord"       4 days ago    Up   3 days       database
```
SONiC Containerization

• Components developed in different environments

• Source code may not be available

• Enables choices on a per-component basis
SONiC – Powering Microsoft At Cloud Scale

Features and Roadmap

**Current:** BGP, ECMP, ECN, WRED, LAG, SNMP, SYSLOG, ACL, LLDP, NTP, VLAN, DHCP, AD, TACACS+, VLAN Trunk, CoPP, IPv6, Everflow, Fast reboot, RDMA, PFC WD, QoS, Telemetry, Warm reboot, OOM, VxLAN, VRF

**Roadmap:** FRR, L3 MLAG, sFLOW, BGP EVPN, NAT
Inviting Contributions in All Areas

• New ideas on white/open network devices
• New features, applications and tools
• Download it, test it and use it!

• Website:  https://azure.github.io/SONiC/
• Mailing list: sonicproject@googlegroups.com
• GitHub:  https://github.com/Azure/SONiC/
• Wiki:  https://github.com/Azure/SONiC/wiki/
Network Verification: ONE
Astronauts use high-fidelity emulators to practice complex, high-risk missions.
Azure engineers use **Open Network Emulator (ONE)** to practice complex, high-risk network operations.
Open Network Emulator

**Fast**
network with 1000s of devices created in minutes

**Seamless**
push-button deployment

**High fidelity**
devices work exactly as production
support from multiple vendors
ONE typical usage scenario

Network engineer describes desired change

An emulated replica is created

Health verified by Z3 theorem prover

Pass/Fail (with feedback, including counter-examples)
More Details

• Mailing list: crystalnet-dev@microsoft.com

• Publication

CrystalNet: Faithfully Emulating Large Production Networks

Hongqiang Harry Liu*, Yibo Zhu*, Jitu Padhye, Jiaxin Cao, Sri Tallapragada, Nuno P. Lopes, Andrey Rybalchenko, Guohan Lu, Lihua Yuan
Microsoft
{harliu,yibzh,padhye,jiacao,srita,nlopes,rybal,gulv,lyuan}@microsoft.com

ABSTRACT
Network reliability is critical for large clouds and online service providers like Microsoft. Our network is large, heterogeneous, complex and undergoes constant churns. In such an environment even small issues triggered by device failures, buggy device software, configuration errors, unproven management tools and unavoidable human errors can quickly cause large outages. A promising way to minimize such network outages is to proactively validate all network operations in a high-fidelity network emulator, before they are carried out in production. To this end, we present CrystalNet, a cloud-scale, high-fidelity network emulator. It runs real network device firmwares in a network of containers and virtual machines, loaded with production configurations. Network engineers can use the same management tools and methods to interact with the emulated network as they do

1 INTRODUCTION
CrystalNet is a high-fidelity, cloud-scale network emulator in daily use at Microsoft. We built CrystalNet to help our engineers in their quest to improve the overall reliability of our networking infrastructure. A reliable and performant networking fabric is critical to meet the availability SLAs we promise to our customers.
It is notoriously challenging to run large networks like ours in a reliable manner [11, 13, 15, 31]. Our network consists of tens of thousands of devices, sourced from numerous vendors, and deployed across the globe. These devices run complex (and hence bug-prone) routing software, controlled by complex (and hence bug-prone) configurations. Furthermore, churn is ever-present in our network: apart from occasional hardware failures, upgrades, new deployments and other changes are always ongoing.
Acknowledgements

• Xin Liu
• Ze Gan
• Guohan Lu
• Yongqiang Xiong
• Lihua Yuan
Thanks!
Cloud Priorities

- Fast
- Supportable
- Reliable
- Secure

Borrow from Albert Greenberg
SONiC: Software for Open Networking in the Cloud

- Switch Abstraction Interface (SAI)
  - Cross-ASIC portability

- Modular Design with Switch State Service (SwSS)
  - Decoupling software components
  - Consistent application development model

- Containerization of SONiC
  - Serviceability
  - Cross-platform portability
SONiC: Software for Open Networking in the Cloud

- Switch Abstraction Interface (SAI)
  - Cross-ASIC portability

- Modular Design with Switch State Service (SwSS)
  - Decoupling software components
  - Consistent application development model

- Containerization of SONiC
  - Serviceability
  - Cross-platform portability
SONiC High-Level Architecture
How Routing Works in SONiC
SONiC: Software for Open Networking in the Cloud

- Switch Abstraction Interface (SAI)
  - Cross-ASIC portability
- Modular Design with Switch State Service (SwSS)
  - Decoupling software components
  - Consistent application development model
- Containerization of SONiC
  - Serviceability
  - Cross-platform portability
Demo: SONiC + ONE
Topology