Demystifying Open Source Network Operating Systems

Senthil Kumar Ganesan
Dell Technologies
@skg_net
https://www.linkedin.com/in/skgnet/
Disaggregation

Traditional Networking  →  Open Networking

Hardware & Software  →  Closed Software  →  Open Hardware

Integrated  →  Disaggregated H/W & S/W  →  Disaggregated H/W & Modular S/W

- Rich Feature Set
- Pre Integrated & Tested
- Single Point Support

- H/W Independence
- Power of choice
- Cost Efficiency

- Innovation
- Development Speed
- Cost Efficiency
- Modern Operating Paradigms

Open Networking with Open Source Software

[Open / Closed] Protocol / Application
Open Network Abstraction
Open Platform Abstraction
Open Hardware

Open Networking

Open Hardware

Closed Software
Open Networking – Network Operating Systems

Closed Software

Open Hardware

* All product names logos and brands are property of their respective owners
Anatomy of a Network Switch and OS

- **HW**
  - CPU
  - Peripherals (FAN, LED, Sensors etc)
  - NPU

- **Peripheral Layer**
  - Platforms Drivers
  - NPU Drivers

- **Operating System**
  - Drivers
  - Platform Abstraction
  - NPU Abstraction

- **Abstraction Layer**
  - Protocol / Application Layer
  - CLI
  - SNMP
  - Netconf
  - REST
  - GNMI
  - BGP
  - STP
  - LACP
  - ACL
  - QoS

- **Management Layer**
  - Network Infrastructure Management (IPC, logging, etc)

- **Bootloader**
Open Networking Components

ONIE
ONL
SAI
switchdev
OpenConfig
Open Network Install Environment

- Modern, Efficient Network Installer
- Open Compute Subproject, initially contributed by Cumulus
- Defacto standard boot loader, its a small operating system, pre-installed as firmware on bare metal network switches.
- Provides an environment for automated operating system provisioning
- ~161 devices currently supported.
Open Networking Components

Open Networking Linux

- A Linux Distribution for Open Networking Switches
- Reference NOS for the Open Compute Project (OCP)
- ~120 H/W Platforms supported, becoming the defacto standard.
- Uses Debian and stock LTS Linux kernel
- Provides Platform Abstraction, via ONLP API
ONIE

• Modern, Efficient Network Installer
• Open Compute Subproject, initially contributed by Cumulus
• Defacto standard boot loader, its a small operating system, pre-installed as firmware on bare metal network switches.
• Provides an environment for automated operating system provisioning
• ~161 devices currently supported.

ONL

• A Linux Distribution for Open Networking Switches
• Reference NOS for the Open Compute Project (OCP)
• ~120 H/W Platforms supported, becoming the defacto standard.
• Uses Debian and stock LTS Linux kernel
• Provides Platform Abstraction, via ONLP API

SAI

• SAI defines API to provide a vendor-independent way of controlling forwarding elements, such as a switching ASIC, an NPU or a software switch in a uniform manner.
• SAI helps easily port new ASIC by running the same application stack on all the hardware, enabled by a simple, consistent programming interface.
• CRUD operation over extensible Entity/Attribute/Value data model
• Provides Network Abstraction for all major ASIC vendors (Barefoot, Boardcom, Cavium, Innovium, Mellanox)
Open Networking Components

ONIE

Open Network Install Environment
- Modern, Efficient Network Installer
- Open Compute Subproject, initially contributed by Cumulus
- Defacto standard boot loader, its a small operating system, pre-installed as firmware on bare metal network switches.
- Provides an environment for automated operating system provisioning
- ~161 devices currently supported.

ONL

Open Networking Linux
- A Linux Distribution for Open Networking Switches
- Reference NOS for the Open Compute Project (OCP)
- ~120 H/W Platforms supported, becoming the defacto standard.
- Uses Debian and stock LTS Linux kernel
- Provides Platform Abstraction, via ONLP API

SAI

Switch Abstraction Interface
- SAI defines API to provide a vendor-independent way of controlling forwarding elements, such as a switching ASIC, an NPU or a software switch in a uniform manner.
- SAI helps easily port new ASIC by running the same application stack on all the hardware, enabled by a simple, consistent programming interface.
- CRUD operation over extensible Entity/Attribute/Value data model
- Provides Network Abstraction for all major ASIC vendors (Barefoot, Broadcom, Cavium, Innovium, Mellanox)

Switchdev

- The Ethernet switch device driver model (switchdev) is an in-kernel driver model for switch devices which offload the forwarding (data) plane from the kernel.
- Provides Network Abstraction – currently only Mellanox is supported
- Offloads L2 & L3 from Linux Kernel and aims to re-use the same set of linux network tool set for switches
Open Networking Components

ONIE
• Modern, Efficient Network Installer
• Open Compute Subproject, initially contributed by Cumulus
• Defacto standard boot loader, its a small operating system, pre-installed as firmware on bare metal network switches.
• Provides an environment for automated operating system provisioning
• ~161 devices currently supported.

ONL
• A Linux Distribution for Open Networking Switches
• Reference NOS for the Open Compute Project (OCP)
• ~120 H/W Platforms supported, becoming the defacto standard.
• Uses Debian and stock LTS Linux kernel
• Provides Platform Abstraction, via ONLP API

SAI
• SAI defines API to provide a vendor-independent way of controlling forwarding elements, such as a switching ASIC, an NPU or a software switch in a uniform manner.
• SAI helps easily port new ASIC by running the same application stack on all the hardware, enabled by a simple, consistent programming interface.
• CRUD operation over extensible Entity/Attribute/Value data model
• Provides Network Abstraction for all major ASIC vendors (Barefoot, Boardcom, Cavium, Innovium, Mellanox)

Switchdev
• The Ethernet switch device driver model (switchdev) is an in-kernel driver model for switch devices which offload the forwarding (data) plane from the kernel.
• Provides Network Abstraction – currently only Mellanox is supported
• Offloads L2 & L3 from Linux Kernel and aims to re-use the same set of linux network tool set for switches

OpenConfig
• OpenConfig provides a consistent set of vendor-neutral data models (written in YANG) based on actual operational needs from use cases and requirements from multiple network operators.
• Openconfig is supported by most of the major networking vendors.
• Provides management layer abstraction – thus enabling common management or controller application to be written.
• Uses modern RPC - NETCONF, RESTCONF, GNMI - GRPC
Anatomy of a Network Switch and OS

Network Infrastructure Management (IPC, logging, etc)

Management Layer
- CLI
- SNMP
- Netconf
- REST
- GNMI

Protocol/Application Layer
- BGP
- STP
- LACP
- ACL
- QoS

Abstraction Layer
- Platform Abstraction
- NPU Abstraction

Operating System
- Drivers
- Platform Drivers
- NPU Drivers

Bootloader

HW
- CPU
- Peripherals (FAN, LED, Sensors etc)
- NPU
Anatomy of a Network Switch and OS

- **OpenConfig**
- **Management Layer**
  - CLI
  - SNMP
  - Netconf
  - REST
  - GNMI
- **Protocol / Application Layer**
  - OpenFlow
  - STP
  - LACP
  - ACL
  - QoS
- **Abstraction Layer**
  - Platform Abstraction
  - NPU Abstraction
- **Operating System**
  - Drivers
  - Platform Drivers
  - NPU Drivers
- **Infrastructure Layer**
  - Network Infrastructure Management (IPC, logging, etc)
- **HW**
  - CPU
  - Peripherals (FAN, LED, Sensors etc)
  - NPU
- **Bootloader**

Open Network Linux

Switchdev

Onie
Open Source Network Operating Systems

- Based on Debian Linux
- Build your own
- Abstraction
- Opensource protocol stacks
- Layer 3 Use cases
Open Source Network Operating Systems

Winding Down

LibreSwitch

FlexSwitch

Present

Open Network Linux

Open Switch

SONiC

Upcoming

DANOS

DENT

STRATUM

* All product names, logos, and brands are property of their respective owners.
Open Networking Linux (ONL)

- Provides Platform Abstraction
- Support Multiple NPU abstraction. OpenNSL, SAI, OF_DPA.
- Supports Multiple Control Plane.
- Thin OS with Controller (ONOS)
- Thick OS with FRR/ORC
- Becoming the de-facto standard for platform abstraction ~120 platforms (BigSwitch, Startum, Arrcus, SnapRoute etc)
- Use Cases:
  - CORD Leaf Spine Fabric (ONOS / Indigo Agent)
  - EVPN with GoBGP / Zebra / ORC / Open NSL
  - FBOSS Leaf Spine Fabric
Open Networking Linux (ONL)

- Provides Platform Abstraction
- Support Multiple NPU abstraction. OpenNSL, SAI, OF_DPA.
- Supports Multiple Control Plane.
- Thin OS with Controller (ONOS)
- Thick OS with FRR/ORC
- Becoming the de-facto standard for platform abstraction ~120 platforms (BigSwitch, Startum, Arrcus, SnapRoute etc)
- Use Cases:
  - CORD Leaf Spine Fabric (ONOS / Indigo Agent)
  - EVPN with GoBGP / Zebra / ORC / Open NSL
  - FBOSS Leaf Spine Fabric
Open Networking Linux (ONL)

- Provides Platform Abstraction
- Support Multiple NPU abstraction. OpenNSL, SAI, OF_DPA.
- Supports Multiple Control Plane.
- Thin OS with Controller (ONOS)
- Thick OS with FRR/ORC
- Becoming the de-facto standard for platform abstraction ~120 platforms (BigSwitch, Startum, Arrcus, SnapRoute etc)
- Use Cases:
  - CORD Leaf Spine Fabric (ONOS / Indigo Agent)
  - EVPN with GoBGP / Zebra / ORC / Open NSL
  - FBOSS Leaf Spine Fabric
Open vSwitch

- Provides Platform Abstraction and Network Abstraction
- By default supports the Linux Protocol & Application Stacks
- Used as base operating system for Dell EMC Enterprise Operating OS10
- Can be leveraged to write custom application using Control Plane Services API.
- Production Ready, Deployed in multiple Verizon, AWINX etc.
- Part of Linux Foundation
Open vSwitch

- Provides Platform Abstraction and Network Abstraction
- By default supports the Linux Protocol & Application Stacks
- Used as base operating system for Dell EMC Enterprise Operating OS10
- Can be leveraged to write custom application using Control Plane Services API.
- Production Ready, Deployed in multiple Verizon, AWINX etc.
- Part of Linux Foundation
OPX

- Provides Platform Abstraction and Network Abstraction
- By default supports the Linux Protocol & Application Stacks
- Used as base operating system for Dell EMC Enterprise Operating OS10
- Can be leveraged to write custom application using Control Plane Services API.
- Production Ready, Deployed in multiple Verizon, AWINX etc.
- Part of Linux Foundation

Applications

- OPX Apps and Micro-Services
  - Open Source
  - Commercial

- Linux/Standard Open Source Apps
  - STP
  - LLDP
  - BGP
  - OSPF

OPX

- Platform Abstraction
- Control Plane Services (CPS)
- Network Abstraction
- Network Device Interface
- Switch Abstraction Interface
- Platform Drivers
- System Device Interface
- Linux Kernel
- Linux CPS Adapter

Choice of Hardware and NPU

Linux Foundation
- Provides a complete NOS with the modern database and micro services based architecture.

- By default supports the Linux Protocol & Application Stacks

- Production Ready & Deployed by multiple Hyperscalers and Enterprise (Microsoft, Alibaba, Tencent, LinkedIn, Comcast etc)

- Supported by Dell Technologies and Mellanox

- Deployment velocity and scale

- Use Cases:
  - Data Center Leaf & Spine
- Provides a complete NOS with the modern database and microservices-based architecture.
- By default supports the Linux Protocol & Application Stacks
- Production Ready & Deployed by multiple Hyperscalers and Enterprise (Microsoft, Alibaba, Tencent, LinkedIn, Comcast, etc)
- Supported by Dell Technologies and Mellanox
- Deployment velocity and scale
- Use Cases:
  - Data Center Leaf & Spine
- Provides a complete NOS with the modern database and microservices based architecture.
- By default supports the Linux Protocol & Application Stacks
- Production Ready & Deployed by multiple Hyperscalers and Enterprise (Microsoft, Alibaba, Tencent, LinkedIn, Comcast etc)
- Supported by Dell Technologies and Mellanox
- Deployment velocity and scale
- Use Cases:
  - Data Center Leaf & Spine
Features

- LLDP
- VRF
- Port Profiles
- STP
- LAG
- ACL
- BGP MP
- NTP
- VLAN
- OSPF
- IPv4 / IPv6
- VxLAN
- QOS
- EVPN
- CoPP
- Monitoring
- GRPC
- Fast Reboot
- Monitoring
- Mirroring
- SNMP
- ConfigDB
- FIB Acceleration
- ECMP
- DTEL
- TATACS
- Devops
- L2 Mac
- PFC
- IPv4 / IPv6
- EVPN
New, extensible control interface GNMI, GNOI, P4 Runtime

Common models and interface for configuration, management & operations

Common Platform Abstractions

Vendor-neutral control planes

Unified device management

Simplified migration

Deployment velocity and scale

Open Source Sep 2019
- New, extensible control interface GNMI, GNOI, P4
- Runtime
- Common models and interface for configuration, management & operations
- Common Platform Abstractions
- Vendor-neutral control planes
- Unified device management
- Simplified migration
- Deployment velocity and scale
- Open Source Sep 2019
- New, extensible control interface GNMI, GNOI, P4 Runtime
- Common models and interface for configuration, management & operations
- Common Platform Abstractions
- Vendor-neutral control planes
- Unified device management
- Simplified migration
- Deployment velocity and scale
- Open Source Sep 2019
  - Shared (HW agnostic)
  - Chip specific
  - Platform specific
  - Chip and Platform specific
Stratum Use Cases

Cloud SDN Fabric
- Proprietary Network OS (e.g. Google Espresso)
- ONOS
- Trellis

5G Mobile & More
- CORD
- ONOS
- Trellis

Thick Switch/Router
- Embedded Mgmt & Control (e.g. BGP)
- Stratum

SDN
- Traditional

Data plane
- Stratum
- Embedded System
Journey so far…

Hyperscalers → Service Providers → Enterprise
Challenges & Call for Action?

Support
Quality
Explore & Understand
Test Drive
Deploy

Image by Peggy und Marco Lachmann-Anke