OCP의 소프트웨어 프로젝트

2020. 11.



안 종 석 james@jslab.kr **JS Lab**

목차

- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- **6. Software Projects**
- 7. SONIC

- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- 6. Software Projects
- 7. SONIC

1. 개요

1) News

- AT&T Taps DriveNets for Core Routing Software Deployment
- Distributed disaggregated chassis (DCC) white box router to the Open Compute Project (OCP)

AT&T Taps DriveNets for Core Routing Software Deployment



Matt Kapko | Editor September 29, 2020 8:01 AM Share this article:











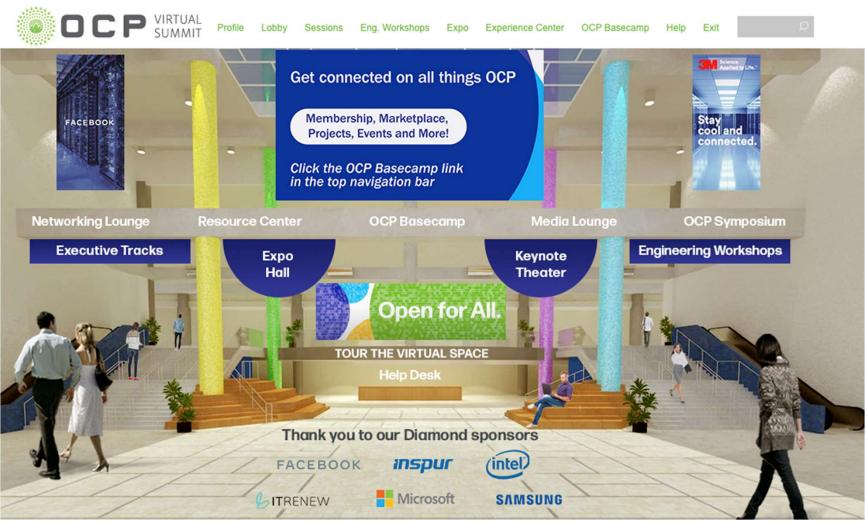


https://www.sdxcentral.com/articles/news/att-taps-drivenets-for-core-routing-software-deployment/2020/09/

1. 개요

2) Virtual Summit

OCP Virtual Summit



1. 개요

3) Community in Korea

OCP Korea Community @ Facebook



- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- 6. Software Projects
- 7. SONIC

1) OCP의 시작

- OCP의 시작 배경
 - 페이스북에서 컴퓨팅 인프라 확장을 위해 가장 효율적이고 경제적으로 가능한 방법을 찾던 팀에서 시작
 - Clean Slate 로 직접 제작한 서버, 전원공급, 랙, 배터리 백업을 제작
 - 적용한 Prineville data center에 38% 에너지 절감과 24% 비용 절감







1) OCP의 시작

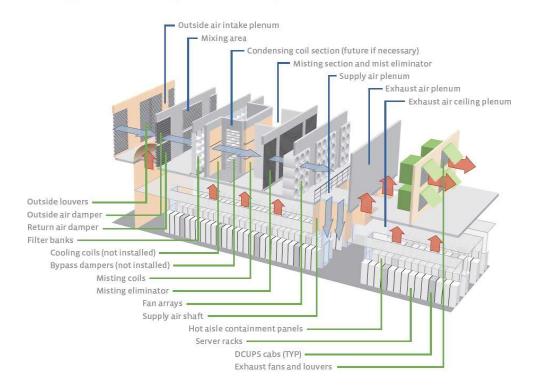
- OCP의 시작 배경
 - 전기 분배 시스템의 480-volt 사용
 - 서버의 효율에 기여 하지 않는 것의 제거
 - 겨울에 통로의 더워진 공기를 사무실 난방에 재사용, 외부 공기를 데이터센터 내에 흐르게 함
 - 중앙 무중단 전기 공급의 필요성을 제거



2) Facebook data centers

All Facebook data centers are 100% OCP

Open Compute Project Data Center





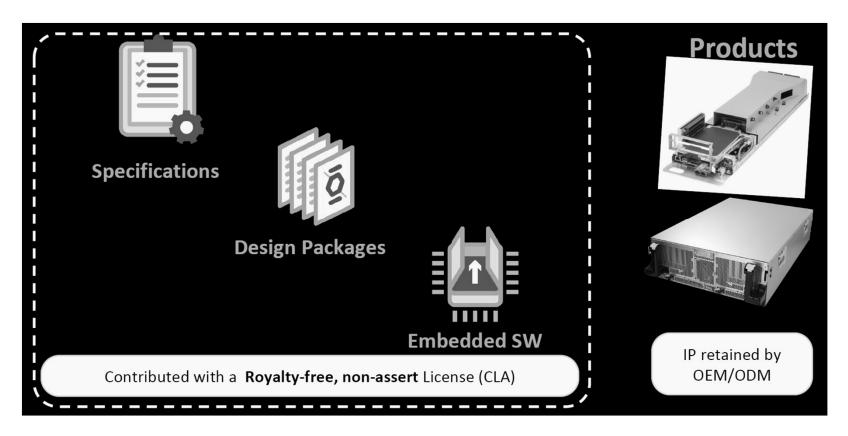
Facebook OCP Prineville Data Center
PUE = 1.06



Typical Data Center
PUE > 1.4

3) OCP의 발전

- OCP의 발전
 - OEM/ODM은 지적 자산을 소유하나, 로열티나 라이선스 없이 기술사양/설계/내장소프트웨어를 기부



4) Hardware

- Open Hardware from OCP
 - 스위치는 처음으로 2015년 Wedge를 발표하였고, 이후 6-pack과 Backpack등을 발표



- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- 6. Software Projects
- 7. SONIC

1) Board

Board Members



Mark Roenigk Chairman & President Facebook



Rocky Bullock
Secretary
OCP Foundation



Andy Bechtolsheim Board Member Individual



Jim Hawkins Board Member Rackspace



Partha Ranganathan Board Member Google



Kushagra Vaid Board Member Microsoft

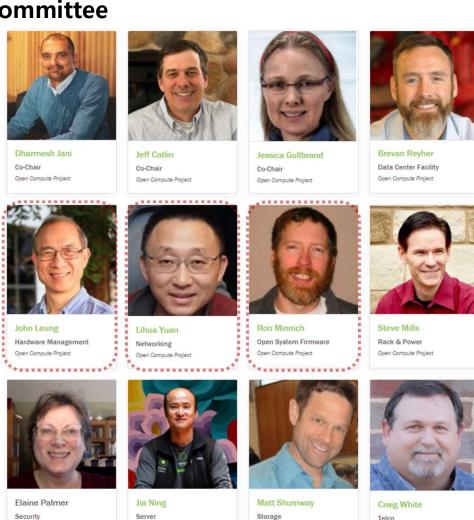


Rebecca Weekly
Board Member

2) Incubation

Incubation Committee

Open Compute Project



Open Compute Project

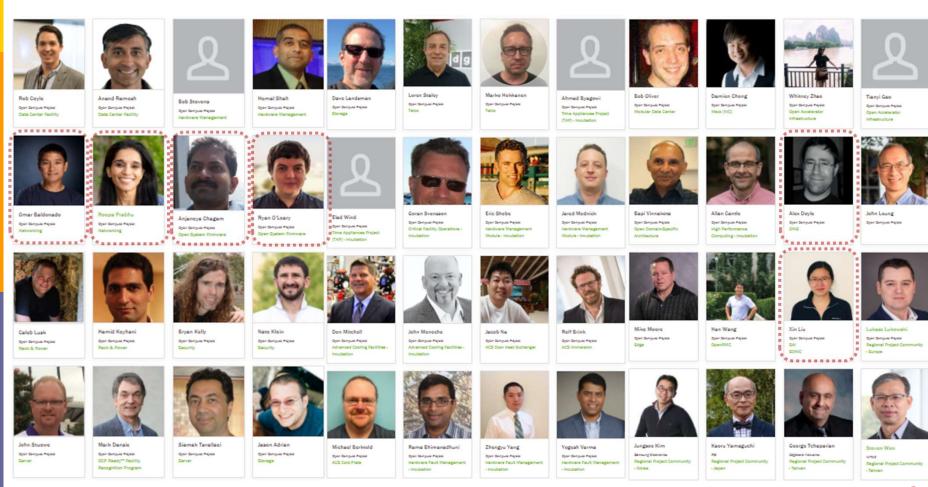
Open Compute Project

Telco

Open Compute Project

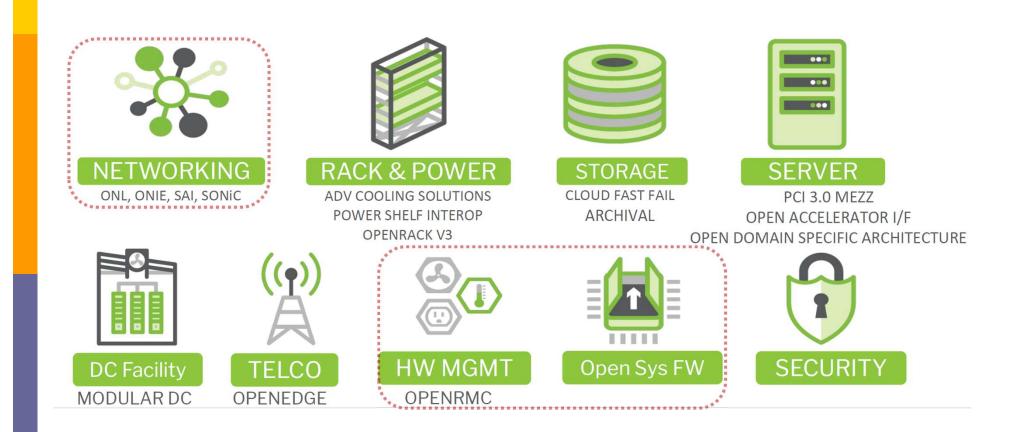
4) Projects

- Project Leads
- https://www.opencompute.org/about/project-leads



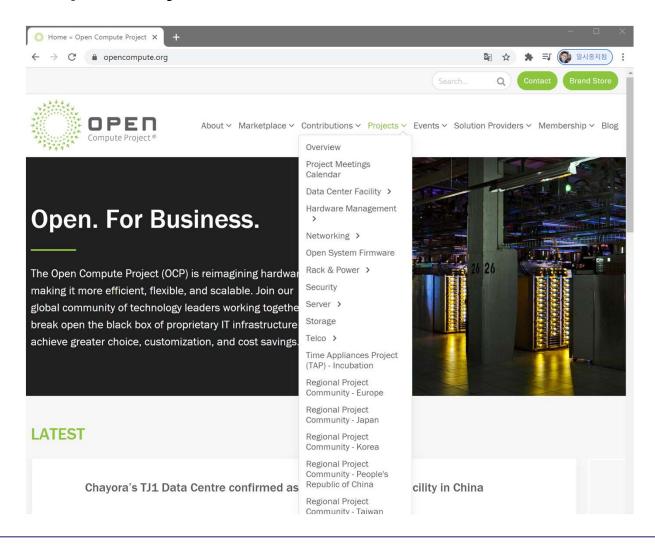
5) Projects & Sub-Projects

OCP Community: Projects & Sub-Projects



5) Projects & Sub-Projects

Open Compute Projects



6) Share

Shared via a Royalty Free Patent Non-Assert CLA(Contributor License Agreement), any OSI license or Copyright License























- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- 6. Software Projects
- 7. SONIC

1) PRODUCT RECOGNITION PROGRAM

OCP PRODUCT RECOGNITION PROGRAM

- Products that carry the OCP Accepted[™] recognition comply 100% with an OCP approved specification and the design files are open sourced and available.
- Products that carry the OCP Inspired[™] recognition comply 100% with an OCP approved specification and are available from a Gold, Silver or Platinum member of OCP.





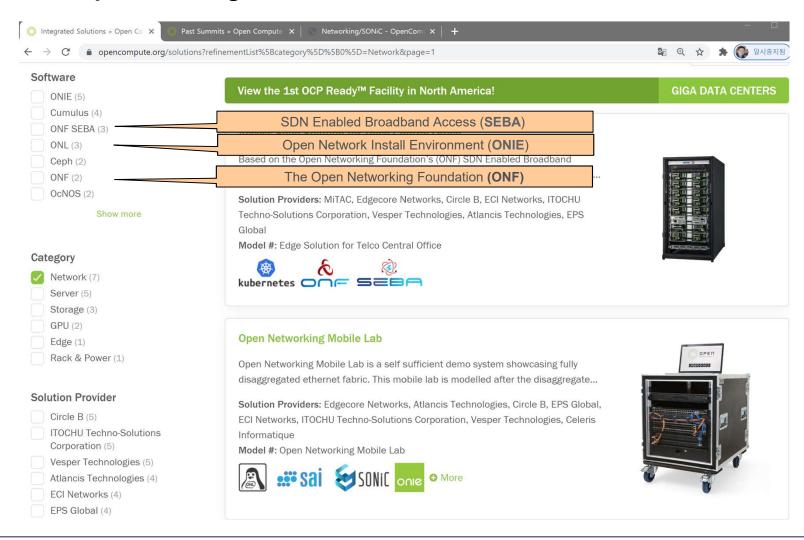
2) Products and Data centers

Marketplace – Products and Data centers



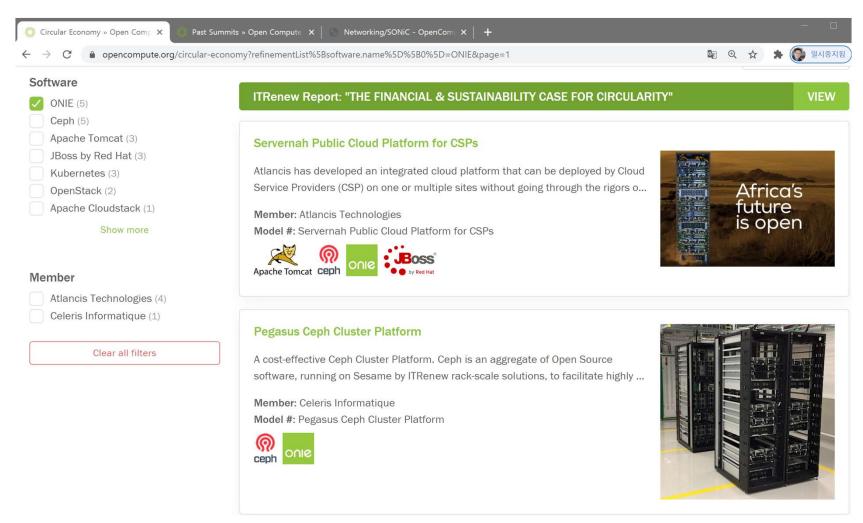
3) Integrated Solutions

Marketplace – Integrated Solutions



4) Circular Economy

Marketplace – Circular Economy



5) 6-pack, Backpack

- 6팩 (6-pack): 페이스북이 2015년 2월12일 공개
- Backpack: 100G Datacenter용



6) Minipack

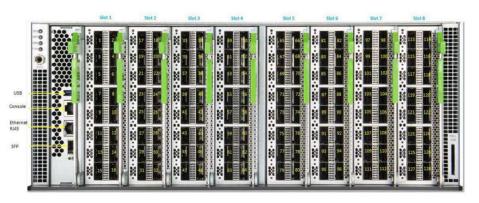
Minipack

- 20%+ switches in data centers are white box + open source software
- First OCP 400G port white box switch
- A new building block switch based on Broadcom Tomahawk 3 ASIC
- Support 2 types of port interface modules (PIM), 100G and 400G
- PIM card plug and play hardware enabled
- Smaller size
- Lower power



6) Minipack

- Minipack
- 2 types of PIM cards
 - 128 X 100G, 12.8T switching bandwidth
 - 32 X 400G, 12.8T switching bandwidth



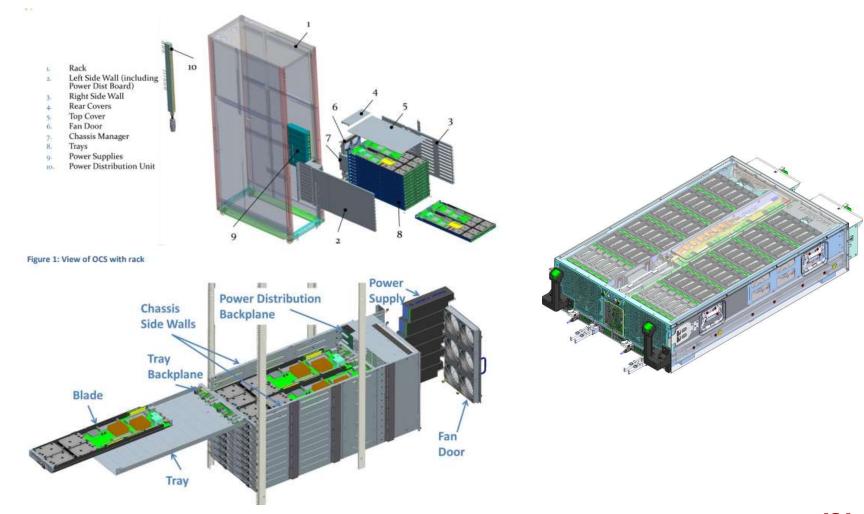


PIM-16Q 100Gb PIM-4DD 400Gb



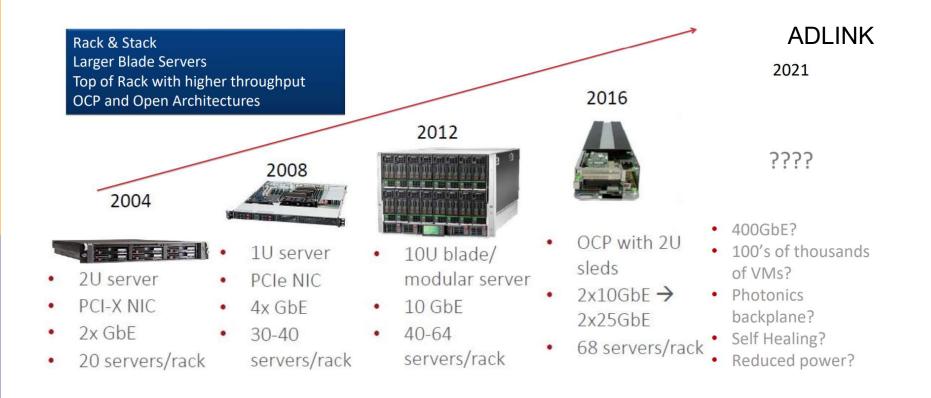
7) OCP Rack

❖ OCP Rack and Servers and Storage System



8) 하드웨어 발전

- 컴퓨팅 하드웨어 발전과 OCP
 - 서버 네트워크 연결의 발전 방향 (ADLINK)



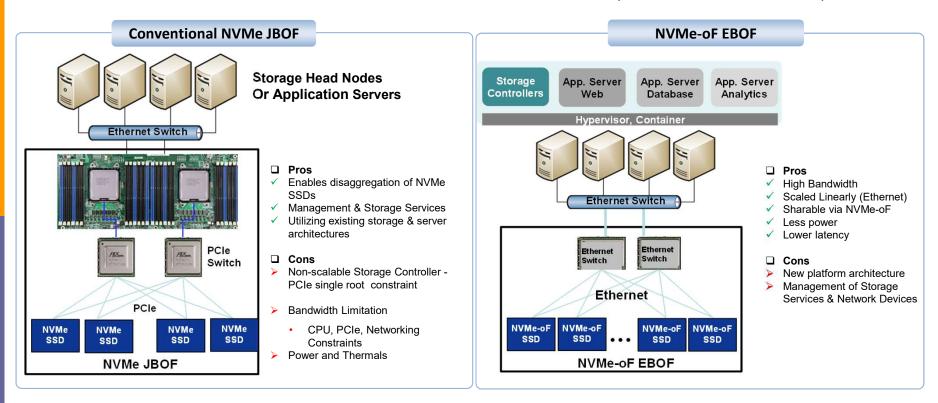
9) NVMe-oF SSD based EBOF

NVMe-oF SSD based EBOF

- NVMe-oF (NVMe over Fabrics)
- **EBOF** (Ethernet Bunch of Flash)

JBOF (just a bunch of flash)

EBOF (Ethernet Bunch of Flash)



4. OCP Marketplace10) Telecom Infra Project

- Telecom Infra Project
 - Backhaul Projects
 - Access Projects
 - Core & Management Projects

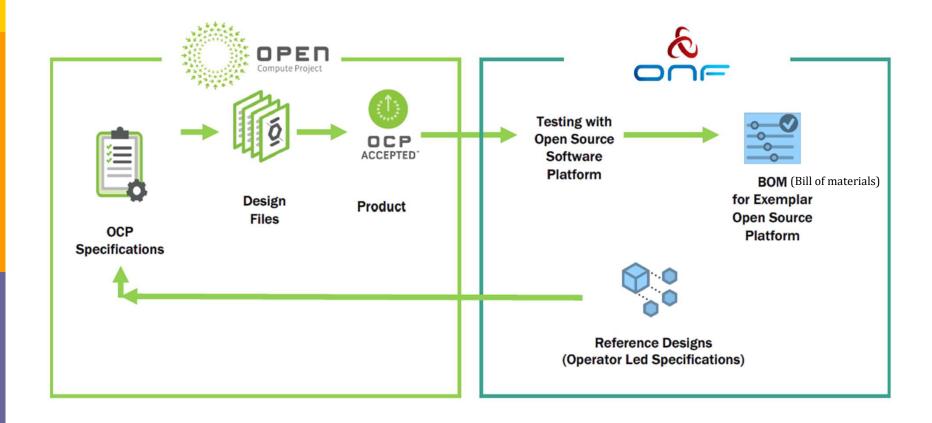


- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- **6. Software Projects**
- 7. SONIC

5. Collaborating with OCP

1) OCP and ONF

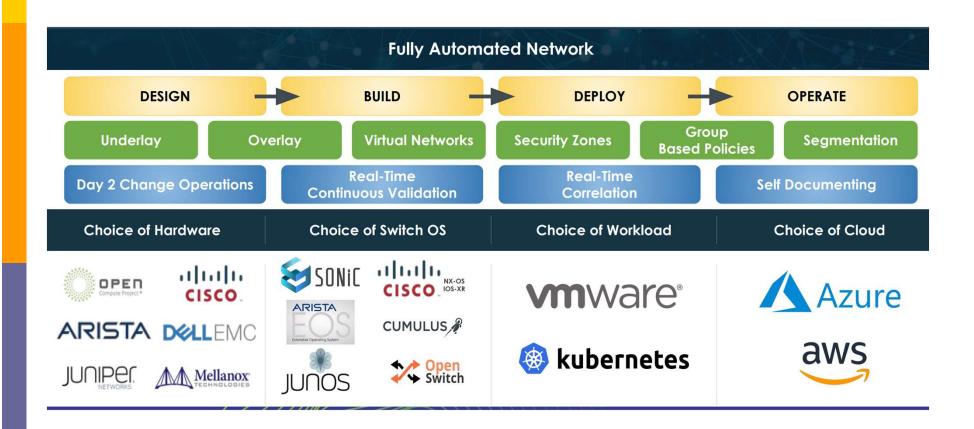
- Collaborating with OCP solutions
 - OCP and ONF



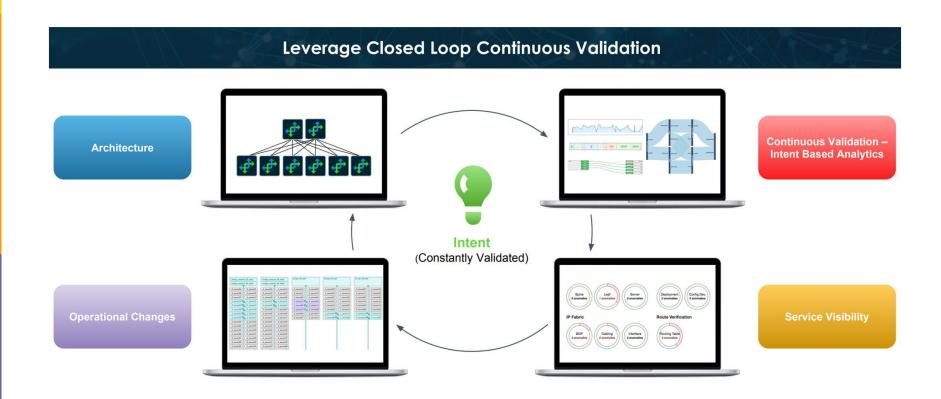
5. Collaborating with OCP

2) Intent-Based Data Center

Intent-Based Data Center Automation



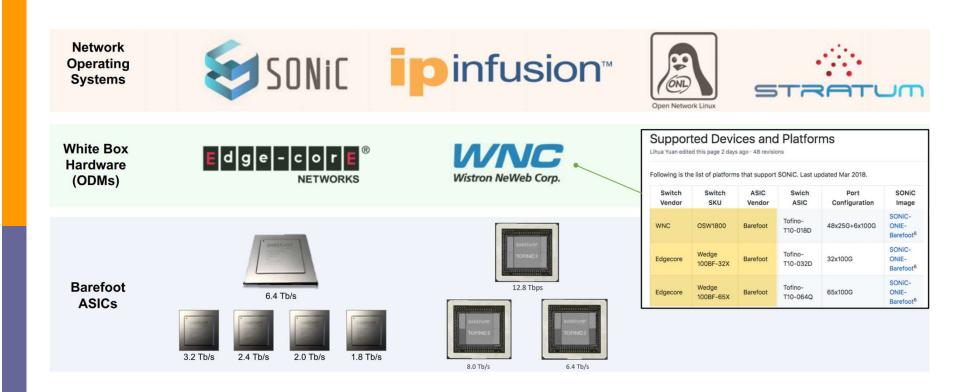
- 5. Collaborating with OCP
- 3) Intent-Based Networking
- Intent-Based Networking Analytics



5. Collaborating with OCP

4) Barefoot OCP Ecosystem

- Collaborating with OCP solutions
 - Barefoot OCP Ecosystem

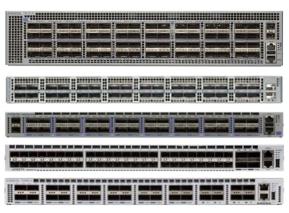


5. Collaborating with OCP 5) 아리스타

- Collaborating with OCP solutions
 - 아리스타

SONiC on Arista Hardware





2016

Arista EOS on Facebook Wedge100S











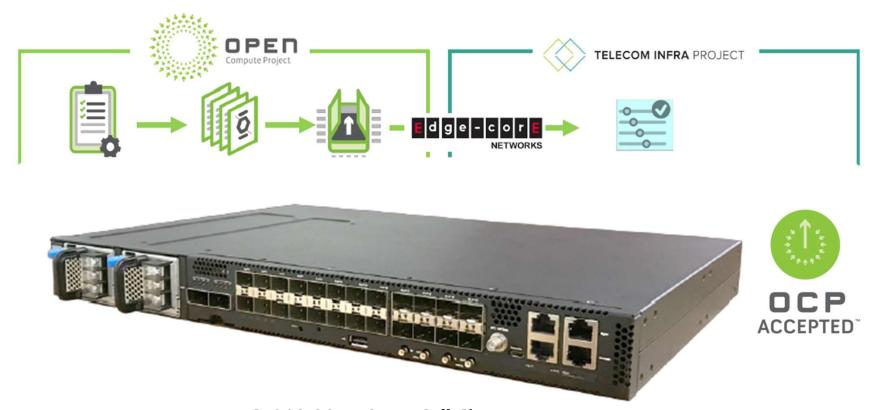




5. Collaborating with OCP

6) 에지코어

- Collaborating with OCP solutions
 - 에지코어



AS7316-26XB Open Cell Site Router

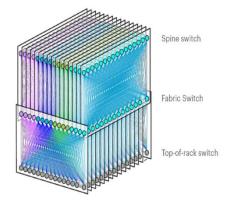
5. Collaborating with OCP7) SONiC

- Collaborating with OCP solutions
 - SONIC









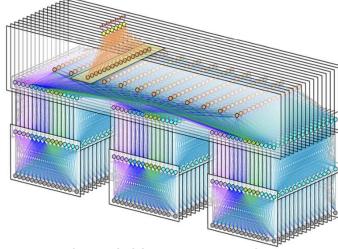
Facebook F16 data center network topology



Arista 7368X 128x100G



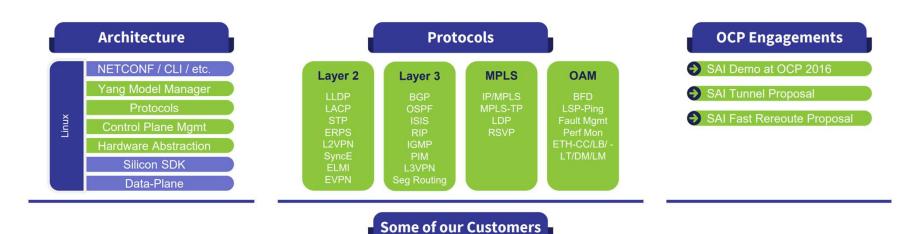
FB Minipack 128x100G



페이스북 다계층 Leaf-Spine 패브릭

5. Collaborating with OCP

- 8) Metaswitch
- Distributed Data Plane
- Metaswitch Networking Software
- https://www.metaswitch.com/products/protocol-stacks



























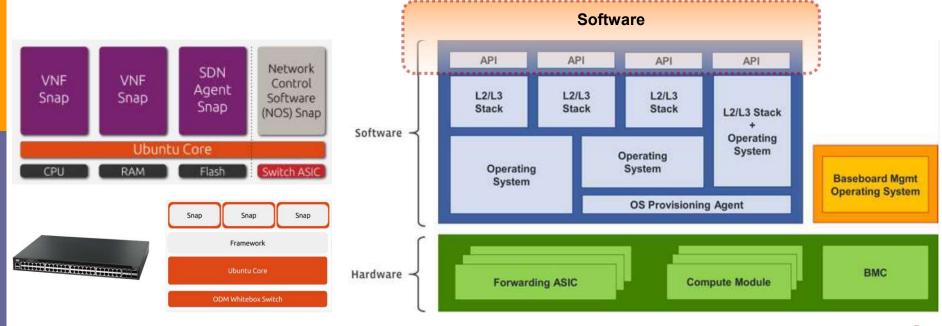


- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- **6. Software Projects**
- 7. SONIC

1) Software Project 개요

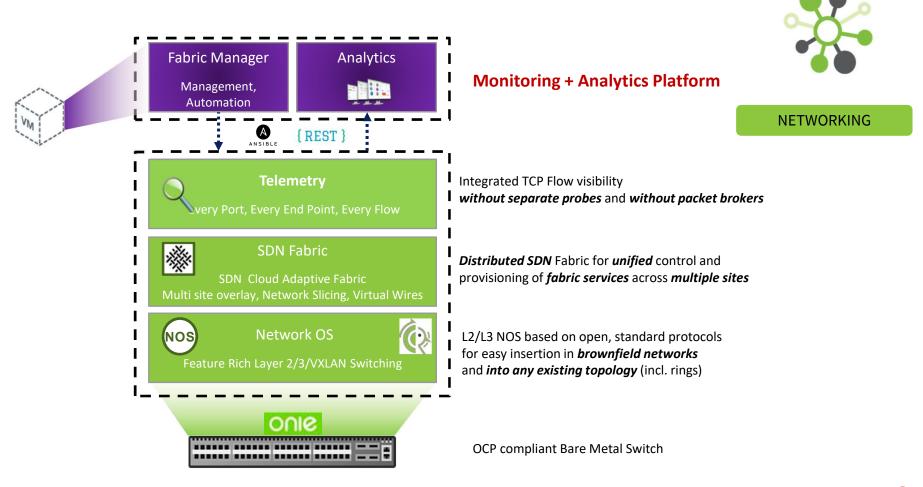
Hardware and Software





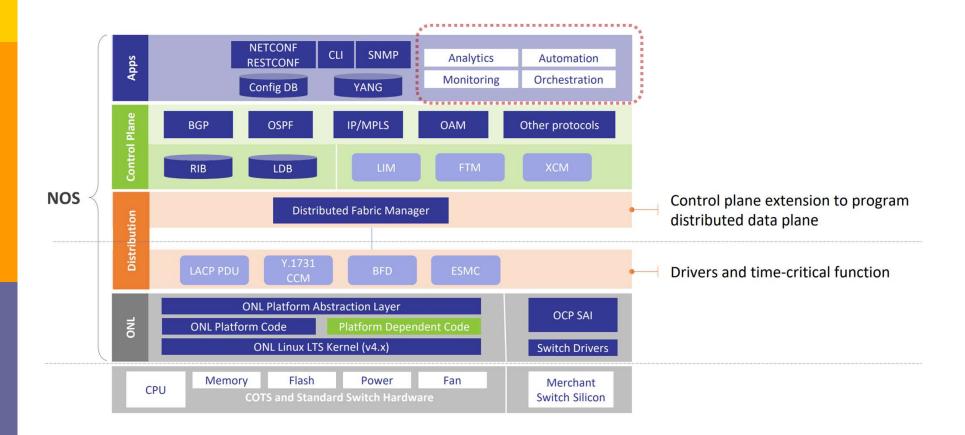
2) Monitoring and Analytics for Disaggregated Networking

Open – Disaggregated Networking



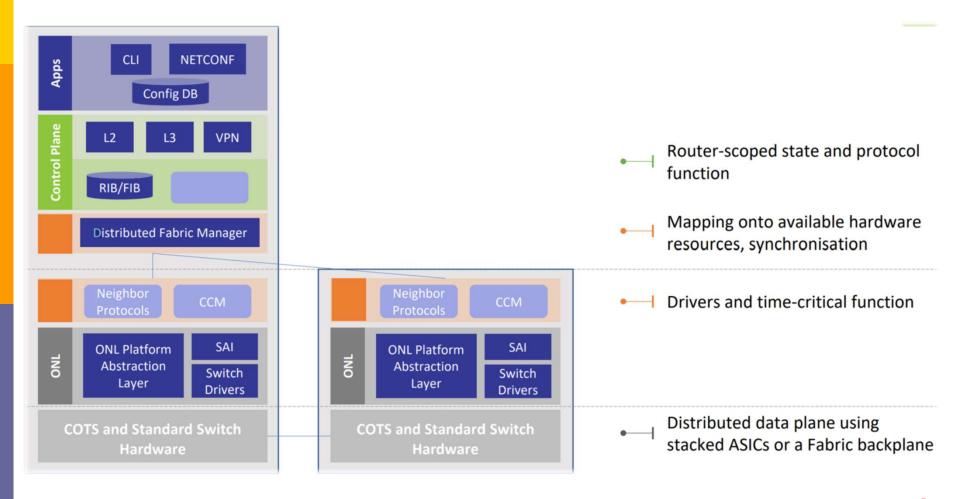
3) Management for Distributable NOS

Architecture of a Distributable NOS



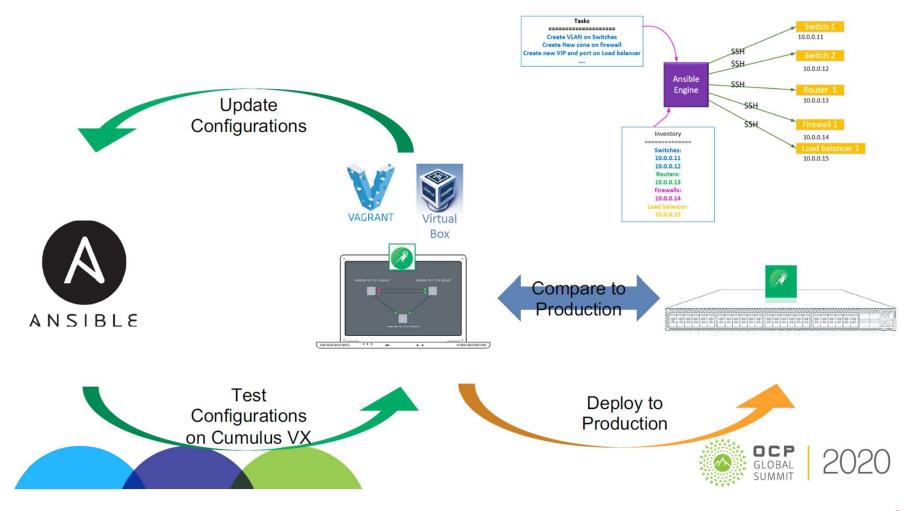
4) Management for Distributed Data Plane

Distributed Data Plane



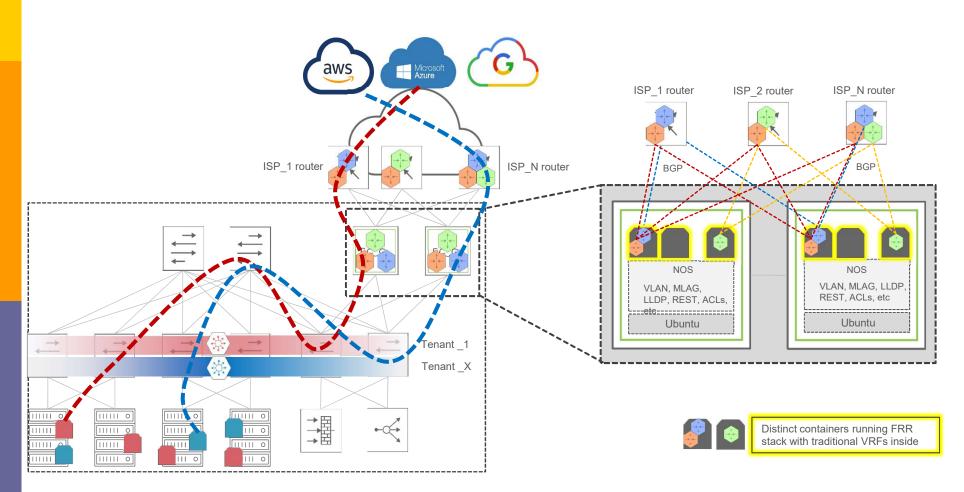
5) Automation

Ansible



6) Hardware vRouter

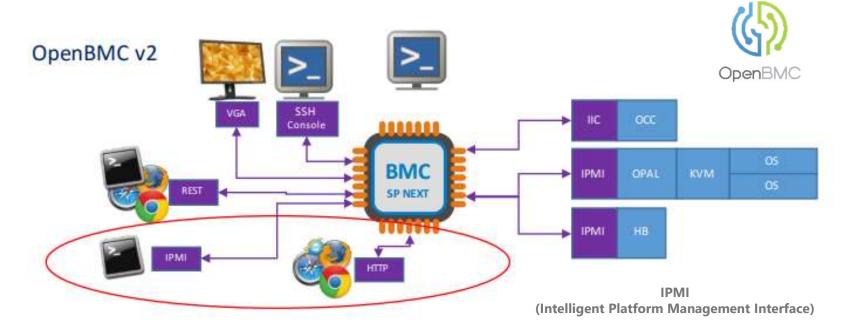
"Hardware vRouter" Based Multi-tenancy for DC Gateway Service



7) OpenBMC

The OpenBMC:

- Linux Foundation open-source project
- goal is to produce an open source implementation of the Baseboard Management Controllers (BMC) Firmware Stack.

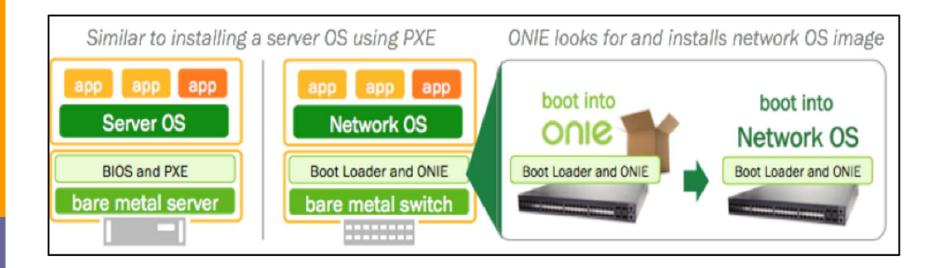


- **6. Software Projects**
- 8) Switch Abstraction Interface
- What is in SAI (Switch Abstraction Interface)

Reliability/QoS Infrastructure Monitoring **Protocol Support** SAI P4 Model TAM [Broadcom] MPLS [Mellanox] L3 Fast Reroute [Mellanox] [Metaswitch] Microburst Multi-NPU [Dell] 802.1BR [Dell] [Marvell] BFD [Dell] Capability Query Critical Resource **Segment Routing** Monitoring [MSFT] [Cavium] [MSFT] Basic L2/L3 ECN [Dell] SAI Ext API [Dell] **INT** [Barefoot] Forwarding

9) ONIE

Open Network Install Environment(ONIE)



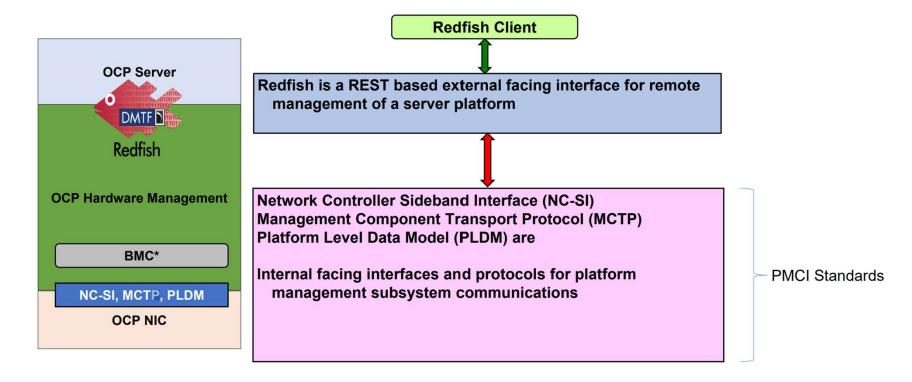
6. Software Projects10) NOS

OpenSwitch v.s. ONL (Open Network Linux)

OpenSwitch (ops) http://www.openswitch.net/	Open Network Linux (ONL) http://opennetlinux.org/
Feature	es / Functionalities
Open NOS with full L2/L3 Switching Feature. Routing / OpenFlow agents are included.	Open Platform Distribution for NOS. Routing / OpenFlow agents <u>NOT</u> included. (only samples)
Tar	get Hardware
OCP (Open Compute) swi	tch, Bare metal (White Box) Switch
C	Contributors
Hewlett Packard, Accton, Broadcom Intel, Qosmos, VMWare, Arista	Big Switch Networks (Initial Source Code Contributor), Pica8, Accton
	License
Apache License, v. 2.0	Eclipse Public License and GPL for Kernel

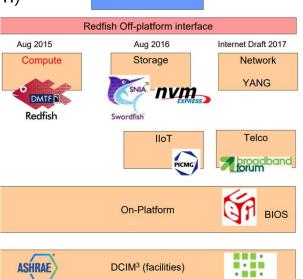
6. Software Projects11) DMTF

- DMTF Standards Applicable to OCP Platforms
 - The DMTF(Distributed Management Task Force) organization develops open manageability standards spanning diverse emerging and traditional IT infrastructures



6. Software Projects12) Redfish

- Redfish standard: SPMF(sequential pattern mining framework), DMTF
- Extending Redfish manageability
 - Domain experts from other SDO's are extending Redfish
 - Networked storage, storage services, and nonvolatile storage (SNIA, NVMExpress)
 - Ethernet Switch map YANG to Redfish
 - BIOS interface (UEFI)
 - **DC facilities infrastructure devices** (The Green Grid, ASHRAE)
 - Industrial IoT (PICMG)
 - Customer Premise Equipment (Broadband Forum)



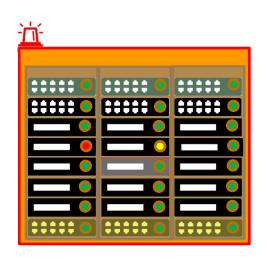
Redfish Client

the green grid

12) Rack Manager Service

Rack Manager Service

 The Rack Manager service provides the front end through a Redfishcompliant REST API for automated management and a command-line interface for manual management. It manages all devices within the rack and communicates directly with the server management system through the network.



User Management (add, remove, etc.)

PMDU Management (power state, relay state, meter alert, throttling, etc.)

Rack Manager Services (Redfish, TFTP, NFS, NTP, JTAG, etc.)

Rack Manager (rack inventory, firmware update, log services, attention LED, etc.)

System / Blade Management (power state, boot order, BIOS configuration, TPM status, mezzanine card status, remote media mount, etc.)

System / Blade Serial Sessions (session management, etc.)

Switch Management (port information, firmware update, etc.)

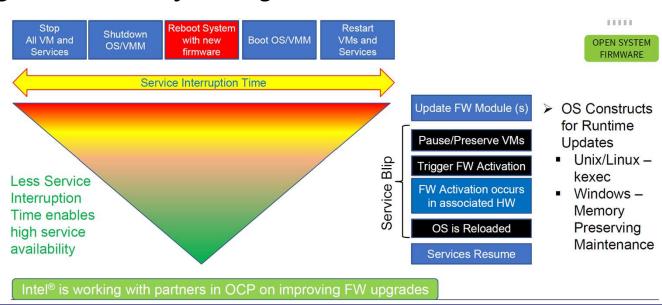
UPS Management (voltage, power, current, alarm, etc.)



High-level functionality exposed by Redfish OCP profile

13) Cloud Firmware Updates

- Supply chain integrity
- Ease of deployment at scale
- Impact less updates
- Automatic Recovery / Rollback
- Audit trails
- Root of trust
- Low boot time
- Configuration / Policy management



- 1. 개요
- 2. OCP Overview
- 3. OCP 재단 구성
- 4. OCP Marketplace
- 5. Collaborating with OCP
- 6. Software Projects
- 7. SONIC

1) SONIC 개요

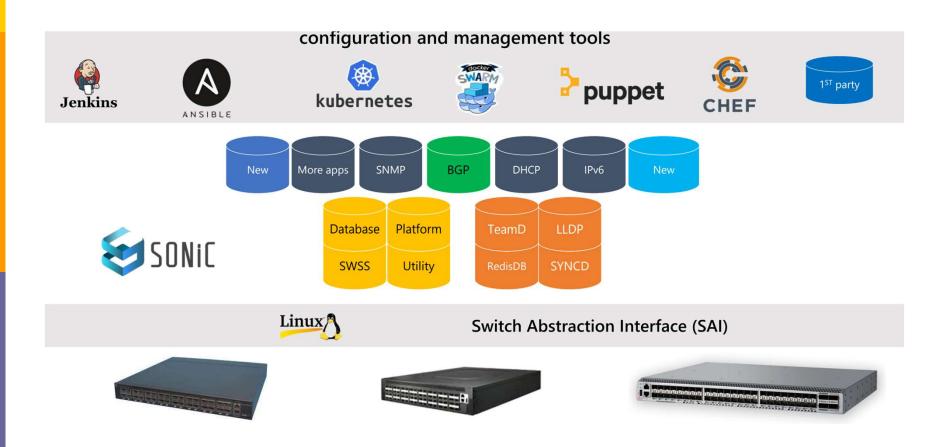
- SONiC (Software for Open Networking in the Cloud):
- redisDB with multiple tables
 - BGP Table: Stores configuration related to BGP neighbors, advertisements.
 - Port: Contains configuration related to interfaces, speeds.
 - PortChannel: Contains configuration of Link Aggregations (bonding/etherchannel, port channel).
 - VLAN: Contains VLAN IDs, member ports.
 - VLAN members: Contains configuration for individual ports for 802.1q tagging.
 - Layer 3 tables: INTERFACE, PORTCHANNEL_INTERFACE, and VLAN_INTERFACE to store IP address details of Layer 3 interfaces.
 - ACL_RULE: Contains configuration of access lists.
 - ..
 - ..

1) SONIC 개요

❖ SONiC (Software for Open Networking in the Cloud) Commands:

```
Show MAC.
admin@sonic:~$ show mac
No. Vlan MacAddress
                               Port
     1000 E2:8C:56:85:4A:CD Ethernet12
Configure VLANs and member ports:
root@sonic:/# config vlan add 1200
root@sonic:/# config vlan member add 1200 Ethernet1
root@sonic:/# config vlan member add 1200 Ethernet9 -u
                                                          (Defining Untagged)
root@sonic:/# config vlan member add 1200 Ethernet8
BGP show commands:
admin@sonic:~$ show bgp neighbors 192.168.1.161
admin@sonic:~$ show bgp neighbors 192.168.1.161 advertised-routes
admin@sonic:~$ show bgp neighbors 192.168.1.161 received-routes
admin@sonic:~$ show bgp neighbors 192.168.1.161 routes
```

1) SONIC 개요



2) SONiC history



3) SONiC +

❖ SONiC +

SONiC+ Load Balancer



- Load balancing @ switch line rate
- Can run on existing ToR switch excluding the extra load balancer servers hop
- can run in addition to all sonic functionality

SONiC+-You cannot fight what you cannot see

The Sepio security stand running on SONiC offers a unique solution that discovers ,identifies and blocks malicious physical layer devices (active and passive)

Full protection against:

- Invisible network implants
- Manipulated firmware



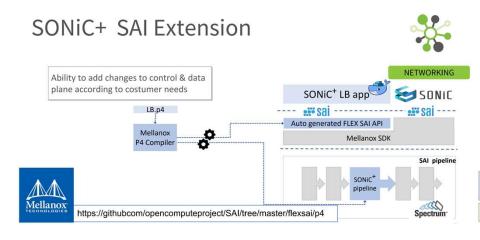
SONiC⁺ Tunnel



- Support tunnel mapping to overlay networks
- Support VRF peering to increase routing scale
- Use cases:
 - Connect bare metal servers to overlay networks
 - Map Costumer Tunnel to tenant network

3) SONiC +

❖ SONiC +

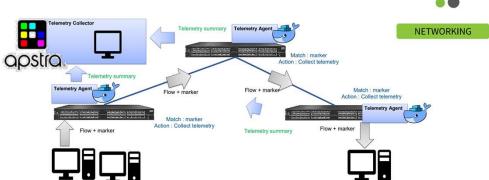


SONIC + ONL

- ONL is used by many NOS as a base OS
 - e.g) ONF Stratum, Facebook FBOSS
- building a solid base OS in the open source community benefits everyone
 - ODM only needs to port their hardware to ONL
 - SONiC gets wider hardware support



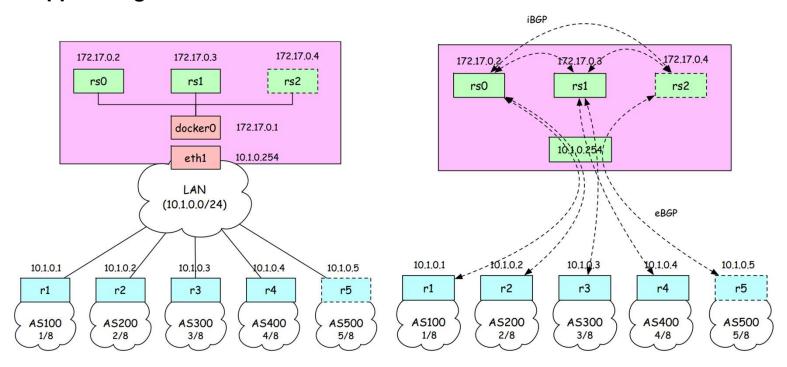
SONiC+ Telemetry



4) containerized Routing Protocol Daemon

Scale out Route Server

- Build a BGP route server that can service thousands of BGP clients
- Multiple instances of cRPD(containerized routing protocol daemon) behind a NATed bridge or load balancer
- Clients see all routes come from a single route server
- Support large number of BGP clients > 10K



5) P4 for SONiC

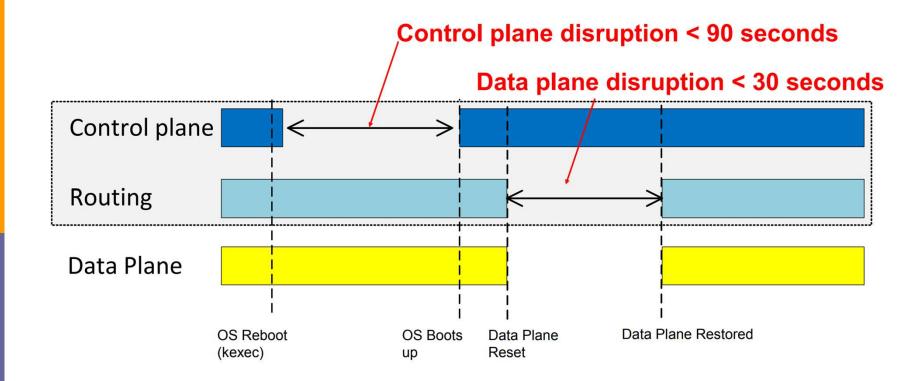
❖ Sample Switch.P4 Features for SONiC

- Ethernet switching
 - VLAN Flooding
 - MAC Learning & Aging
 - STP state
 - VLAN Translation
- IPv4 and IPv6 routing
 - Unicast Routing
 - Routed Ports & SVI
 - VRF
 - Unicast RPF Strict and Loose
 - Multicast PIM-SM/DM & PIM-Bidir
- QOS
 - QoS Classification & marking
 - Drop profiles/WRED
 - RoCE v2 / PFC
 - CoPP (Control plane policing)
 - WRED-based ECN marking

- Tunneling: VxLAN (v4/v6), IP-in-IP, GRE
- ACL
 - Ingress MAC ACL, IPv4/v6 ACL, RACL
 - Egress MAC ACL, IPv4/v6 ACL, RACL
 - QoS ACL, System ACL, PBR
 - Port Range lookups in ACLs
- Security Features
 - Storm Control,
 - IP Source Guard
- sFlow
- PTP
- Counters
 - Route Table Entry Counters
 - VLAN/Bridge Domain Counters
 - Port/Interface Counters
 - ACL stats
- Barefoot Dataplane Telemetry

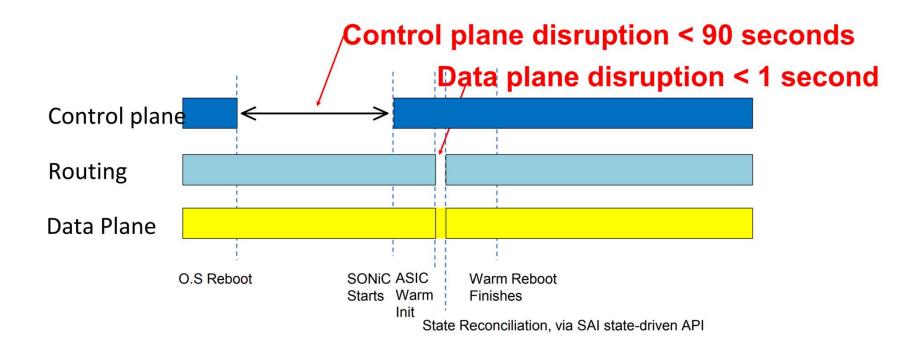
6) Fast Boot

❖ Fast Boot



7) Warm Boot

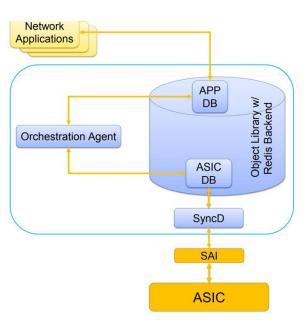
❖ Warm Boot



8) Warm Boot Architecture

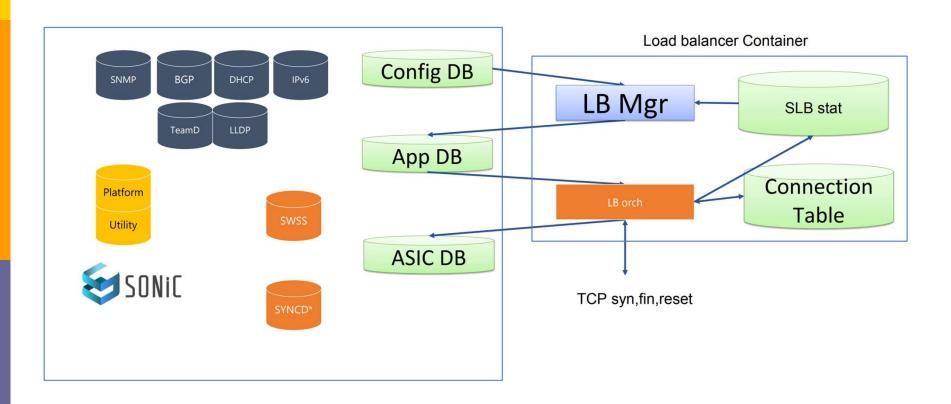
❖ Warm Boot Architecture

- 1. Warm boot script stores App/ASIC DB on disc
- 2. Redis restores App/ASIC DB after reboot
- 3. OA reads AppDB and compiles a new ASIC DB
- 4. SyncD compares old/new ASIC DB, and apply diff to the ASIC
- 5. Applications waking up in parallel
 - May staged changes to App DB
 - OA comes in as usual, updates ASIC dB
 - SyncD keeps syncing ASIC DB to hardware



9) Load Balancer

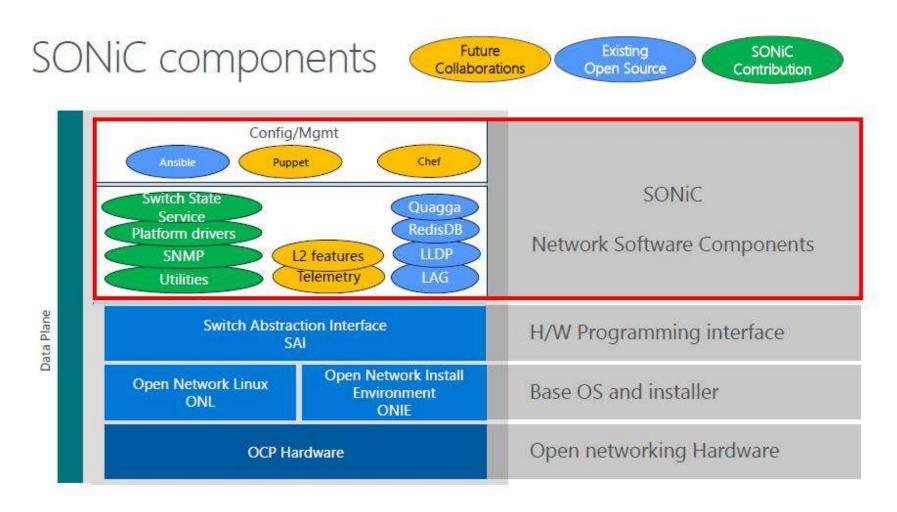
Load Balancer (Extending SwSS on SONiC)



Software Support Service (SWSS)

7. SONIC 10) SONIC 구성

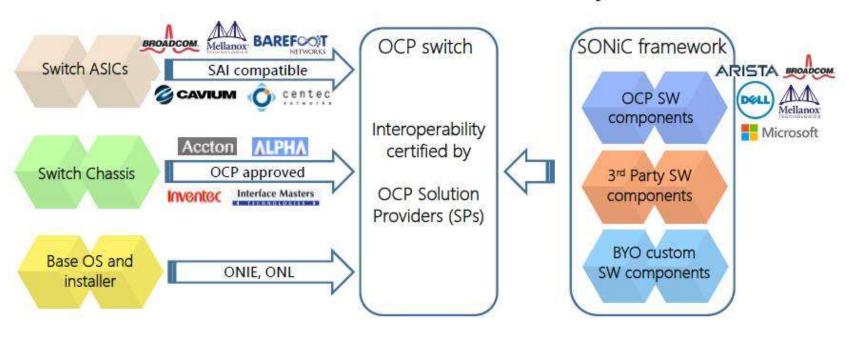
SONIC Components



7. SONIC 11) SONIC 생태계

SONiC ecosystem

SONiC vision for the OCP switch ecosystem



Fully Open Sourced switching platform - Increased choices for OCP end users

7. SONiC11) SONiC Sub-Project

SONiC Sub-Project Calls

- SONiC Kubernetes Workgroup App Ext discussion: October 23rd, 2020
- SONiC Test Subgroup: October 21st, 2020
- SONiC Test Subgroup: July 15th, 2020
- PDDF v2.0 June 16th, 2020
- Sonic-mgmt pytest introduction (II) June 3rd, 2020
- BGP un-numbered and System health monitoring HLD June 2nd, 2020
- EVPN VXLAN HLD and ZTP config update HLD May 26th, 2020
- Sonic-mgmt pytest introduction (I) May 20th, 2020
- Dynamic buffer calculation HLD May 19th, 2020
- SONiC Configuration Replace HLD May 12th, 2020
- MLAG & Firmware utils May 5th, 2020
- Fine grained ECMP & L3 kernel programming enhancement April 28th, 2020
- D-Bus Host to Container Communications HLD April 14th, 2020
- Style Python Docstrings and cEOS April 8th, 2020
- AAA HLD Discussion April 7th, 2020
- System Health and System LED Monitoring March 31st, 2020
- PyTest and Testbed Routing Design March 25th, 2020
- Add QoS Scheduler and Shaper March 17th, 2020
- Code Coverage Rate for SONiC March 11th, 2020
- D-BUS HLD Discussion January 28th, 2020

