



Open. Together.



OCP
SUMMIT



Networking:
Hardware

Stratum: Enabling Next-Gen SDN

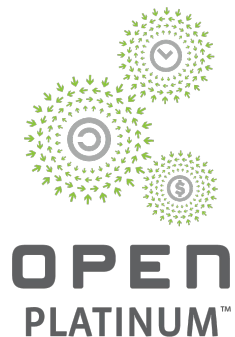
Brian O'Connor, Open Networking Foundation (ONF)

Devjit Gopalpur*, Google

Alireza Ghaffarkhah*, Google

Yi Tseng, Open Networking Foundation (ONF)

*On behalf of many at Google (Waqar Mohsin, Shashank Neelam, Jim Wanderer, Lorenzo Vicisano, Amin Vahdat, ...)



Open. Together.

Google's History

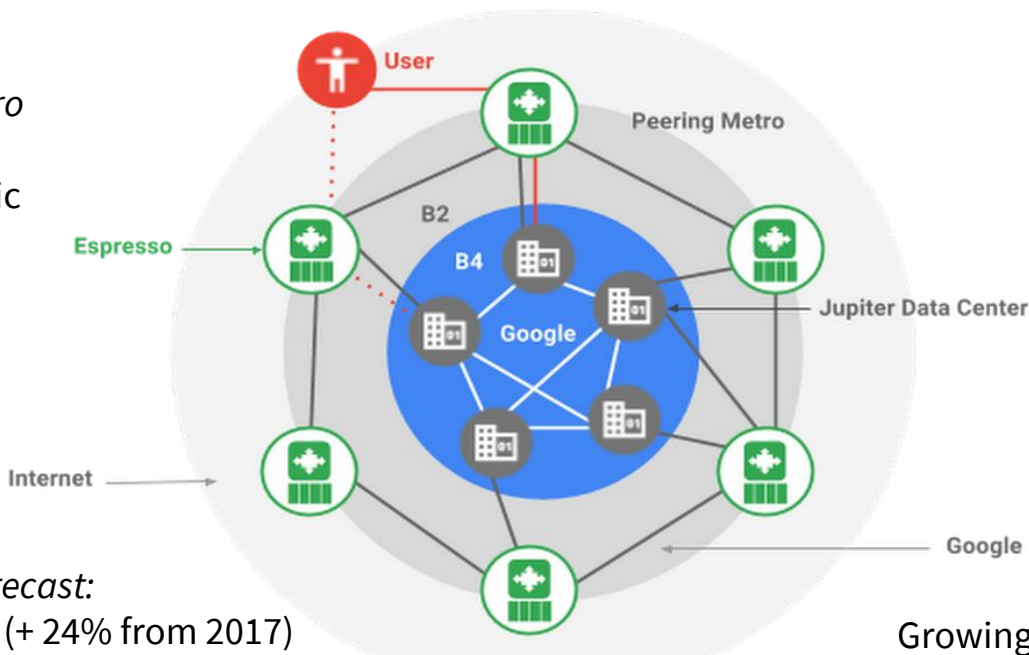


NETWORKING

Google runs SDN networks at scale

Espresso

SDN Peering Edge / Metro
70 metro sites
25% of all Internet traffic



Jupiter

SDN Data Center
1.3 Pbps
100,000+ servers/site

B4

SDN WAN
Inter-datacenter traffic
Growing faster than Internet traffic

Cisco Global Internet Forecast:

~150 EB/month in 2018 (+ 24% from 2017)

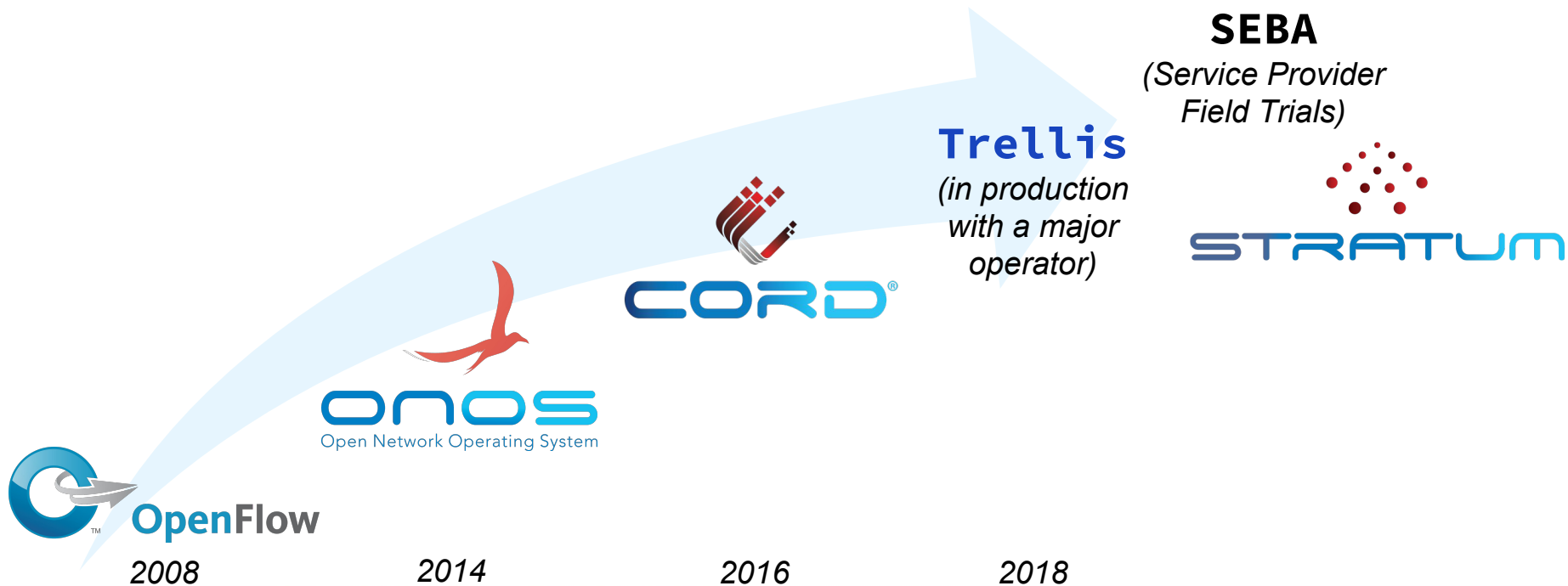
<https://www.blog.google/topics/google-cloud/making-google-cloud-faster-more-available-and-cost-effective-extending-sdn-public-internet-espresso/>
<https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf>



Open. Together.

ONF's History

The ONF has a lot of experience building SDN and NFV solutions



SDN Provides Many Benefits

- **Fine-grained control** enables support for more complex QoS and load balancing policies
- **Control plane optimizations** difficult to achieve using traditional networking
- **Enhanced network visibility** for troubleshooting, monitoring, and auditing
- **New features** can be added **by operators** at **software time scale**, a boost for innovation
- ... and the list goes on

How do we deliver SDN on Open hardware?

- New control interface
 - Common control plane abstraction defines pipeline capability and behavior
 - Programmability and extensibility for different types of switching chips
- Common models for configuration and monitoring
- Common interfaces for operations
 - Diagnostics, Security, Software upgrade
- Common platform abstraction (e.g. Open Network Linux Platform API)
- **Open source switch stack**

What does the new switch stack give us?

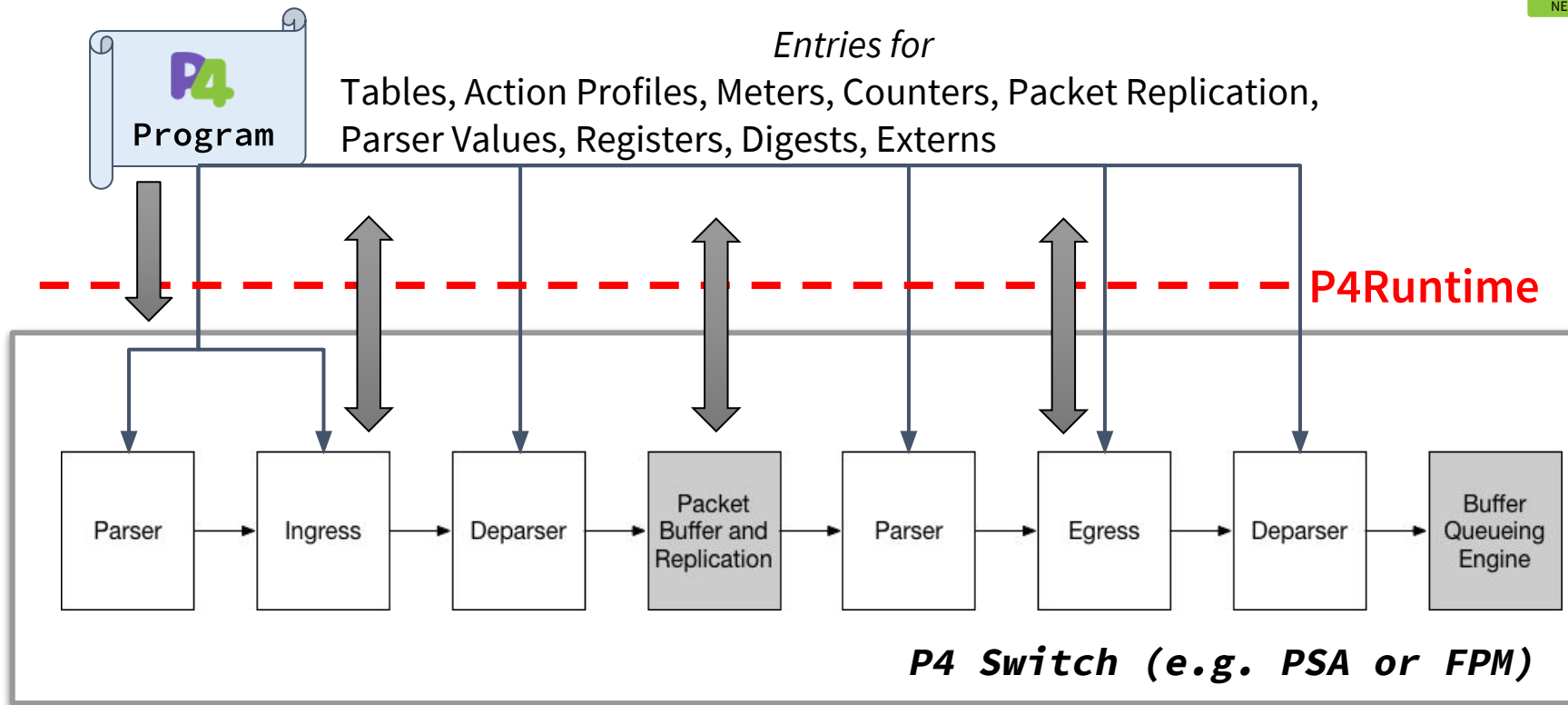


- Support for **vendor-neutral** control applications
 - Control plane is written once, compiled for multiple backends, i.e. hardware.
 - Contract provides extensibility. New use cases and network roles do not require modification of APIs or switch software.
- Support for **programmable hardware**
 - Even more flexibility - backend faithfully mimics software intent.
 - Pushes hardware abstraction up the stack.
 - Uniform runtime interface for heterogeneous silicon as well as network intent.
- Support for a **uniform network model**
 - Vendor-agnostic model of topology.
 - Simplifies operability of a multi-vendor network.

... and hence ...

- Enhanced deployment **velocity** at **scale**
 - Introduction of new functionality, hardware, etc. using common workflows.
 - Incremental support for new equipment.
 - Rapid prototyping by operators and vendors using a well-defined contract.
- Simplified **migration** of services
 - From traditional devices to programmable devices.
 - Between heterogeneous device blocks.
- Unified device **management**
 - Operators use common tools to deploy, configure, monitor and troubleshoot devices from multiple vendors.

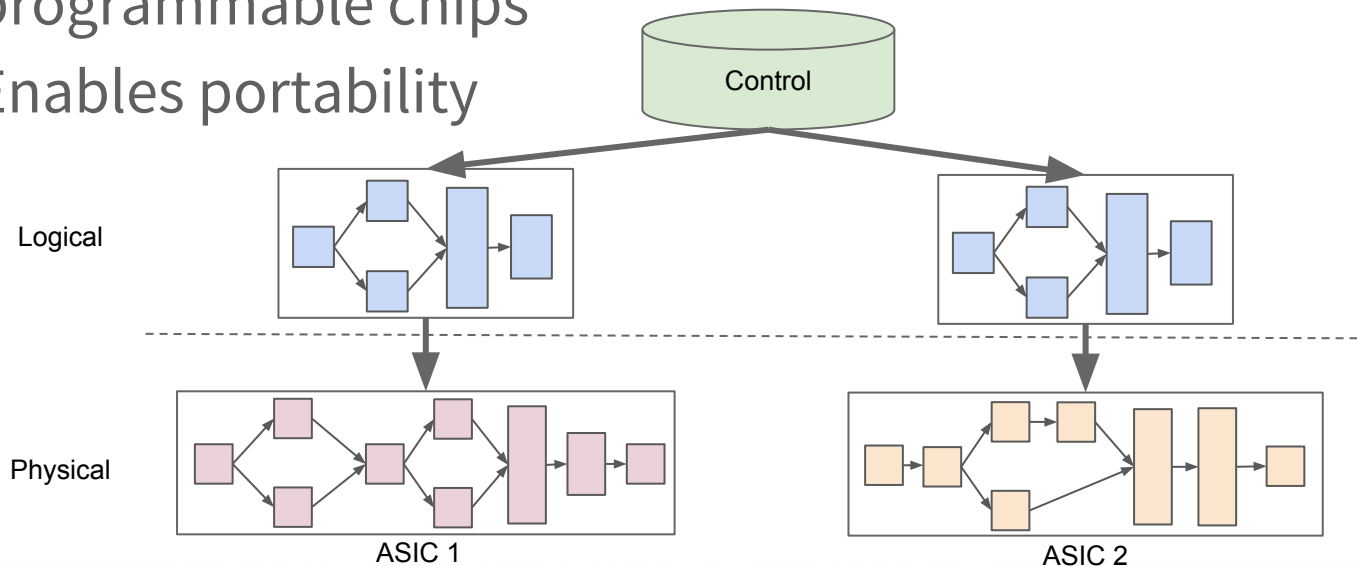
Control Interface: P4Runtime



Slide adapted from P4.org

Role of P4

- Provide clear pipeline definition using P4 tailored to role
- Useful for fixed-function/traditional ASICs as well as programmable chips
- Enables portability



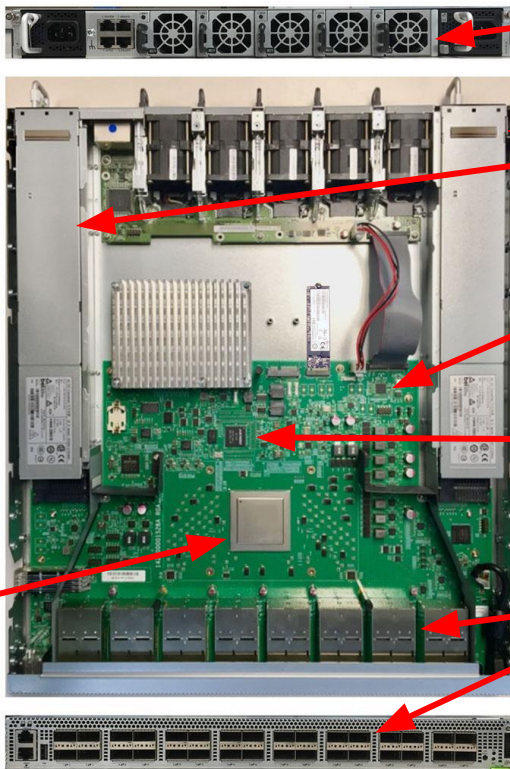
OAM Interfaces: gNMI and gNOI

- gNMI for:
 - Configuration
 - Monitoring
 - Telemetry
- gNOI for Operations

Switch Chip Configuration

QoS Queues and Scheduling
 Serialization / Deserialization
 Port Channelization

Management Network



Fan Speed

Power supplies

Monitor Sensors
 e.g. temperature

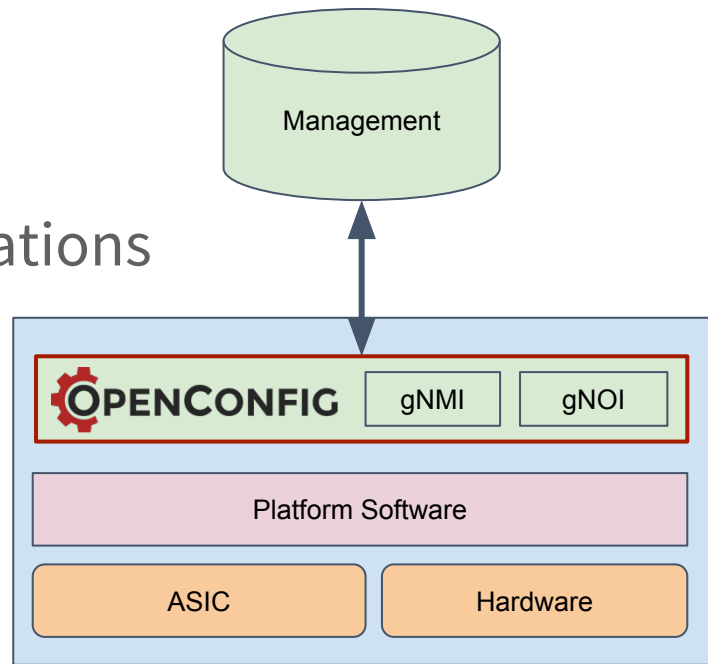
Software Deployment and Upgrade

Port State and Mapping
 LED Control

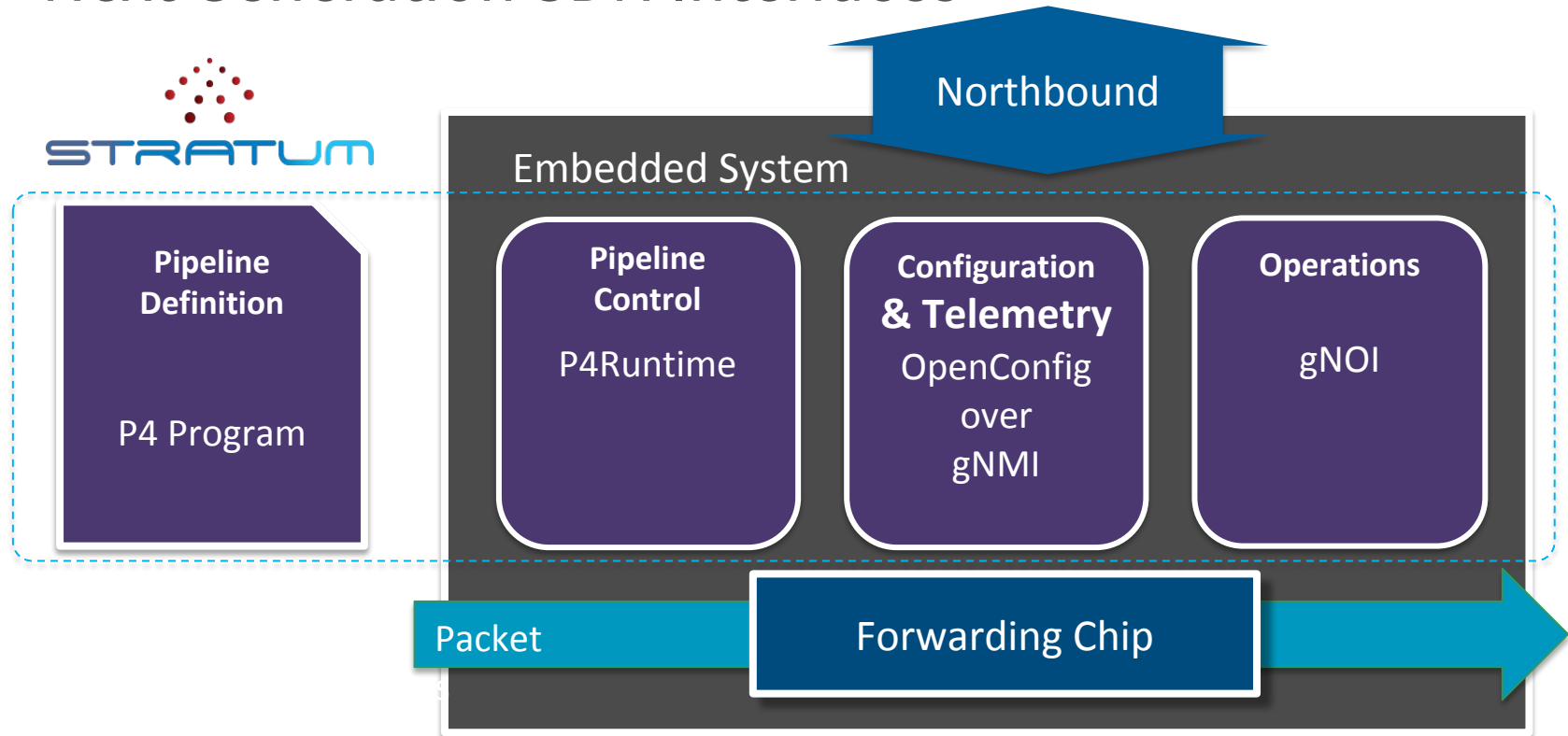
... and the list goes on.

Enhanced Configuration

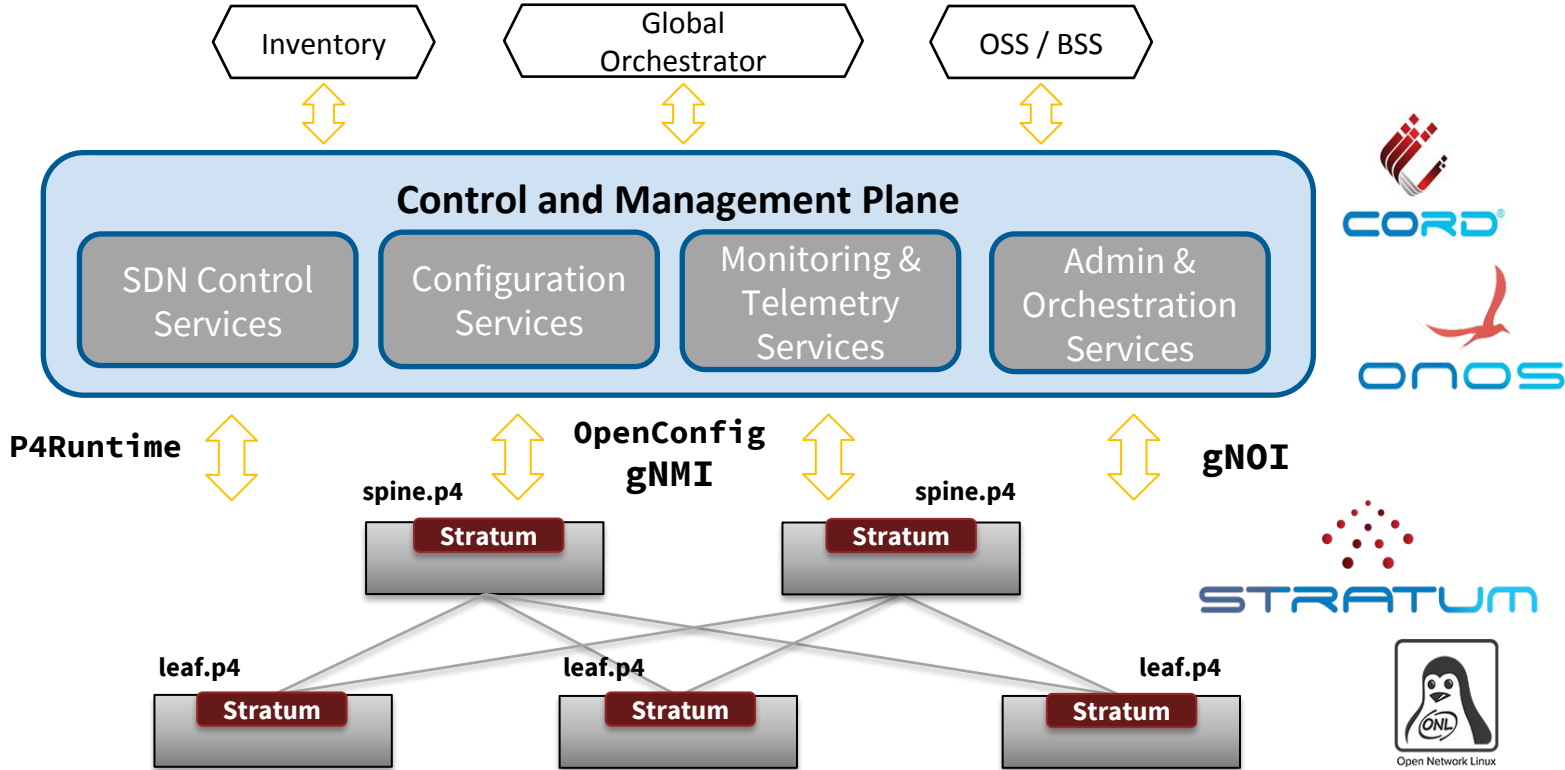
- Configuration and Management
- Declarative configuration
- Streaming telemetry
- Model-driven management and operations
 - gNMI - network management interface
 - gNOI - network operations interface
- Vendor-neutral data models



Next Generation SDN Interfaces

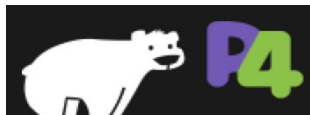


Next Generation SDN picture



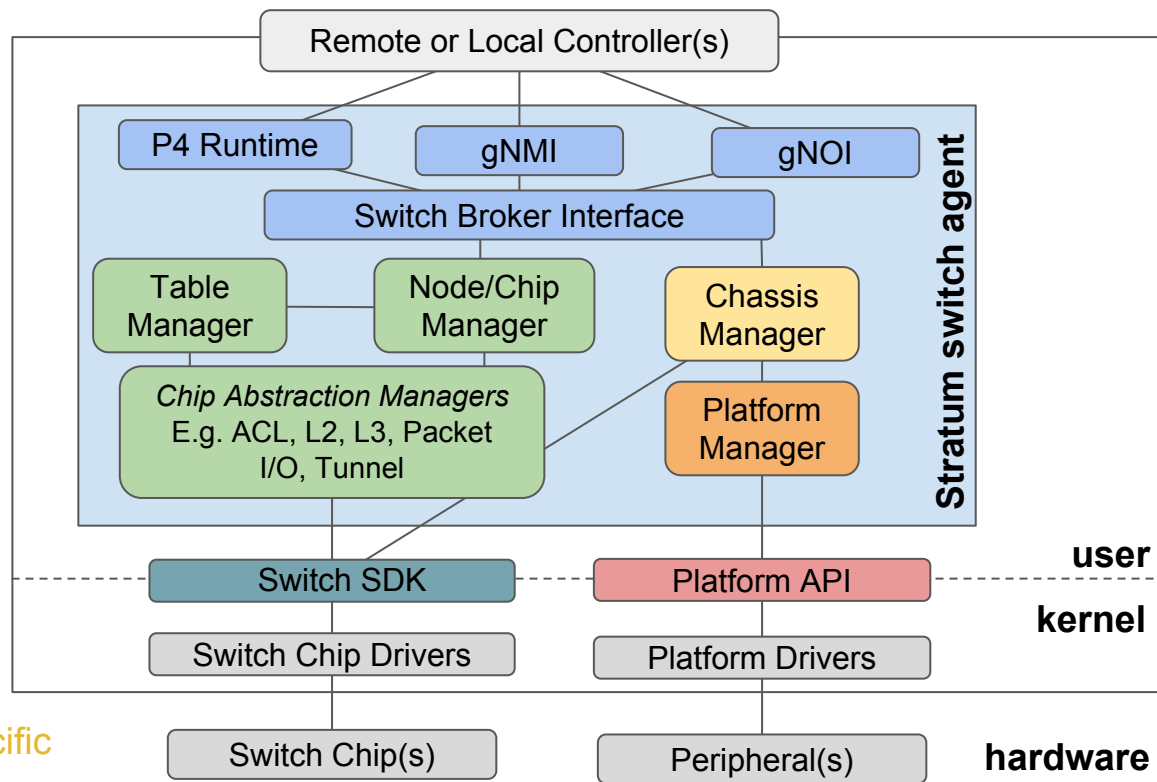
Stratum Implementation Details

- Implements **P4Runtime**, **gNMI**, and **gNOI** services
- Controlled locally or remotely using **gRPC**
- Written in **C++11**
- Runs as a **Linux** process in user space
- Can be distributed with **ONL**
- Built using **Bazel**

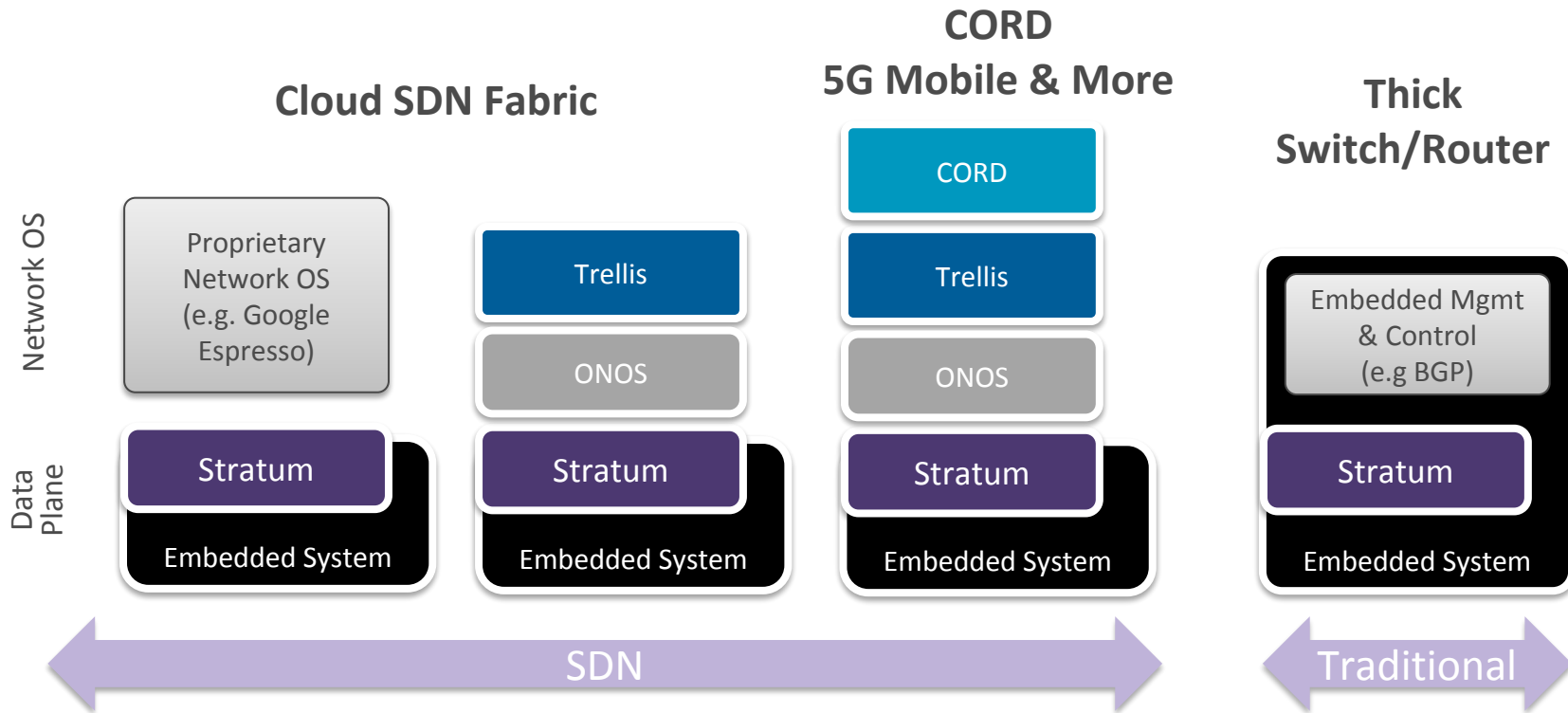


Open Network Linux

Stratum High-level Architectural Components



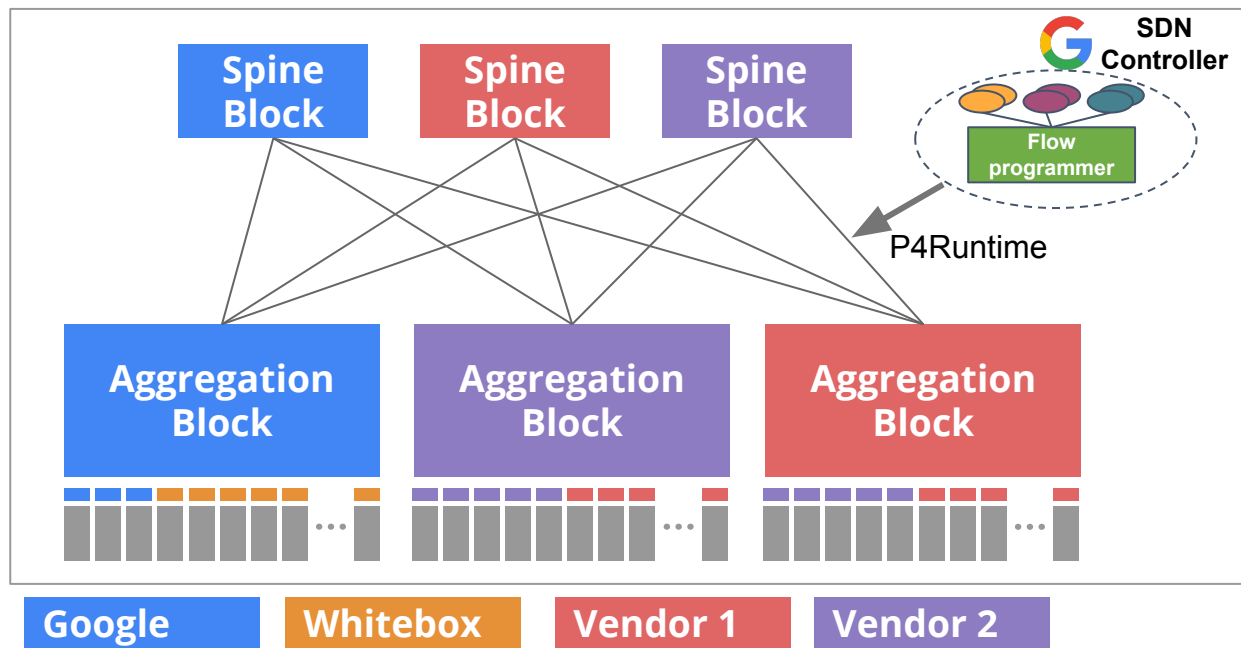
Stratum Use Cases



Google's Approach to Next-Gen Multi-Vendor SDN



- Heterogeneous network
- Single consistent API
 - P4Runtime
 - OpenConfig
- Exploit unique HW capabilities
- Leverage commercial technology / vendors



Transforming Tencent's Network: One Datacenter at a Time

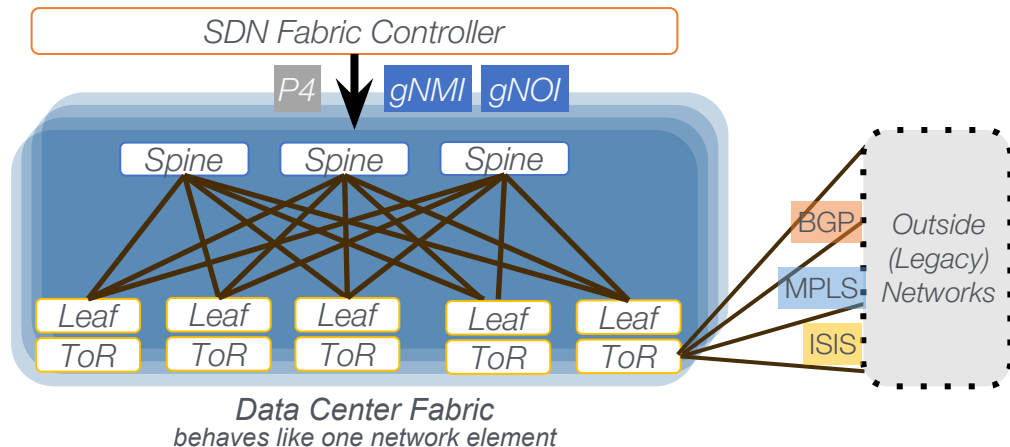


- Data center fabric as disaggregated modular switch

Switch OS → Data Center SDN Controller

Fabric Cards → Spine/Core Switches

Line Cards → Leaf/ToR Switches

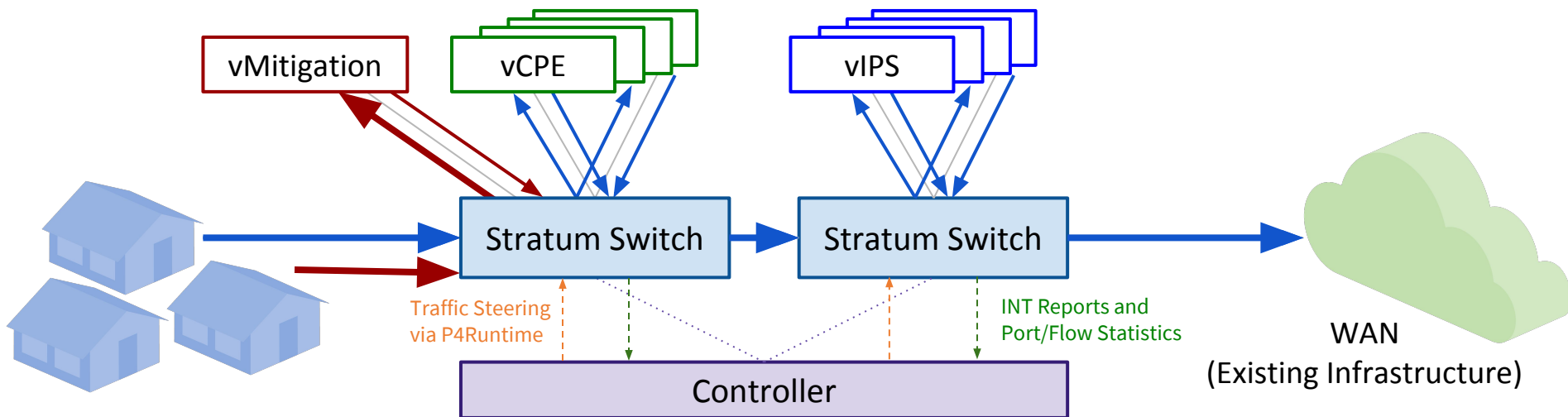


- Centralized control does not mean the entire network must have one controller.
- Rather we opt for a network of controllers, enabled by ONF CORD, Trellis and Stratum.
 - Freedom to use different protocols or RPC at outside controllers.
 - Facilitates integration with legacy networks.

Slide adapted from Tencent

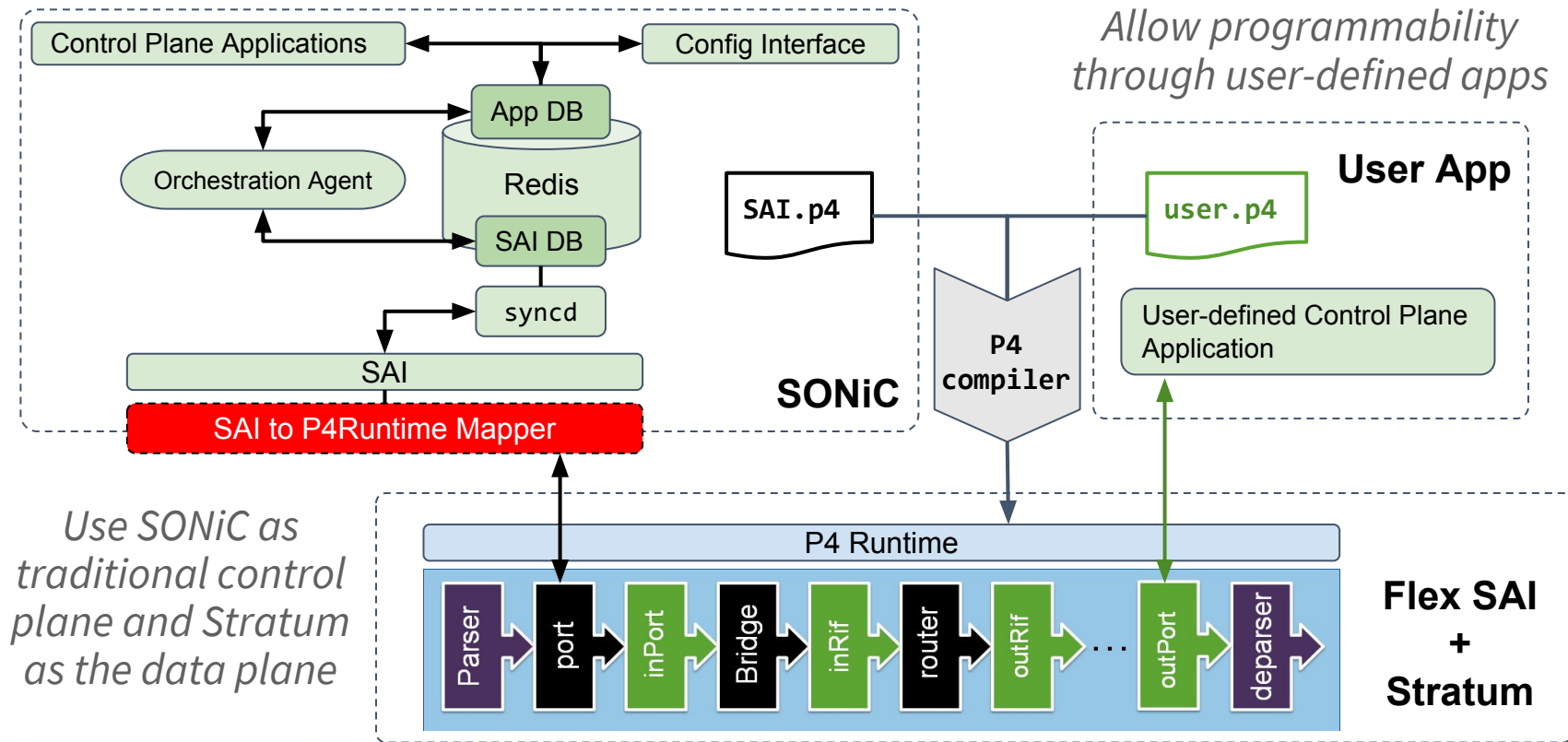
NTT's SDN-style Use Cases

- Flexible service chaining through network functions
- Auto-scaling of network functions in response to load triggers
- Detect flow bursts (e.g. DDoS) and forward throttled traffic through mitigation function



Slide adapted from NTT

Building a traditional **and** programmable router



OCP Software + Stratum



- Stratum is built on **ONL** and leverages **ONLP** as the primary platform API
- The Stratum community has been a contributor and driver for **ONLPv2** (<https://github.com/opencomputeproject/OpenNetworkLinux/tree/ONLPv2/packages>)
- **ONIE** is used to install new ONL images on Stratum switches



Open Network Linux



Targeted OCP Hardware



- Accepted Hardware

- Edgecore AS7712-32X (Broadcom Tomahawk)
- Facebook/Edgecore Wedge 100-32X (Broadcom Tomahawk)



- Hardware under Review

- Barefoot/Edgecore Wedge 100BF-32X (Barefoot Tofino)
- Barefoot/Edgecore Wedge 100BF-65X (Barefoot Tofino)

- Inspired Hardware

- Agema AG9032v1 (Broadcom Tomahawk)



Open. Together.

Stratum Community



NoviFlow



One Source Integration



Mellanox
TECHNOLOGIES



Open. Together.

Stratum Roadmap



Stratum Community Launch

Pioneer Phase

- Initial Reference Platform Support (HW & SW)
- Development Infrastructure (Build, CI, etc.)

Field Trials, Production Deployments

- Cloud and Telco networks
- ONF's CORD with major operators

Stratum Member Preview

- Expanded platform support
- Feature development
- Hackathons

Open Source Launch

Community Development

- Increase list of supported chipsets and platforms
- Synergy with open source Switch OSES and controller planes



Getting involved

<https://www.opennetworking.org/stratum/>

Contribute to the Interfaces and reference P4 programs

- Interfaces and Models: [P4Runtime](#), [gNMI](#), [gNOI](#), and the [OpenConfig models](#)
- P4 programs: [Fabric.p4](#), [Flex SAI](#), etc.

Become a Stratum Member

Join the Public Mailing List

- Periodic updates on Stratum's progress.

Related OCP Projects

- <https://www.opencompute.org/wiki/Networking/ONL>
- <https://www.opencompute.org/wiki/Networking/SONiC>
- <https://www.opencompute.org/wiki/Networking/SAI>
- <https://www.opencompute.org/wiki/Networking/ONIE>



Open. Together.

OCP Global Summit | March 14–15, 2019





Code releases

	Release 0.1 (May 2018)	Release 0.2 (Oct. 2018)	Release 0.3 (Feb. 2019)
P4Runtime	Support for pre-release	Support for 1.0.0-rc1	Support for 1.0 and minor fixes
gNMI	Basic framework	Stable support	Stable support and bug fixes
gNOI	-	Initial interfaces	4 service implementations (e.g. system, file)
Switch support	Google platforms; Partial Broadcom support	Barefoot Tofino on 3 vendors; BMv2 software sw.	Tofino platform integration; DummySwitch for testing
Platform abstraction	Basic interfaces	Support for platform mapping and DB	Add support for ONLP
Conformance Testing	-	Test framework definitions	Test framework definitions