PINS

P4 Integrated Network Stack

Mythil Raman
Konstantin Weitz

May 19, 2021
Google loves SDN ... 

- We have been doing SDN for > 10 years [1]
- **Data Centers & WAN** are exclusively SDN controlled
- **Simpler algorithms driving efficiency and higher availability**
  e.g. traffic engineering in WAN [2], and WCMP in DC
- **Easier Debugging** [3]
  Entire network state is visible to the controller.
- **More Compute**
  compute on a server rack, instead of a small switch CPU
- **Merchant Silicon Switches**

[1]: Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google’s Datacenter Network
[2]: B4: Experience with a Globally-Deployed Software Defined WAN
[3]: Orion: Google’s Software-Defined Networking Control Plane
... and programmability

- long standing member of P4 community
- P4Runtime has many benefits over OpenFlow
  - extensible pipeline model
  - feature velocity
- enables rapid innovation
- provides semantic pipeline description, useful for:
  - automated verification
  - documentation
Open Source NOS on the Rise!

- **Disaggregation** of network stack and white box switches

- Switch OS landscape is quite **fragmented**
  - Stratum, SONiC, FBOSS, DANOS, DENT, ...
  - Different communities and use cases

- Open source communities are solving the **same** problems
  - Chip and peripheral configuration
  - Platform management (fans, sensors, etc.)
  - Warm boot, Non-stop Forwarding and Hitless upgrade
  - ASIC / Pipeline abstraction
    - ... and the list goes on ...

We see this as an **Opportunity**!
Motivation

Bring SDN capabilities to Open Source NOS

Enable operators to incrementally explore SDN

Make it easier for operators to customize their pipeline

Improve feature velocity
SDN Based SW Velocity

Shifting focus from Protocols → Algorithms & Microservices based controller architecture led to Orion [1] having

- **Bi-weekly** software releases
- **30** new significant capabilities
- **16x** Improvement in scale
- **50x** Jupiter availability improvement
- **100x** B4 availability increase
- **40x** Improvement in Network convergence time

[1] Orion: Google’s Software-Defined Networking Control Plane
Extending Open Source NOS

PINS

(P4 Integrated Network Stack)

Integrating SDN and Programmability
01 Architecture
Traditional NOS

- BGP
- DHCP
- LLDP
- SNMP
- ...
SDN Capabilities
APIs
02 SONiC Implementation
Remoting SAI and supporting programmability

We model SAI in P4 and expose it via P4Runtime.
Key Design Decisions

- **Open Source**: We intend to upstream our changes to the community.

- **Opt In**: Existing SONiC use-cases see no overhead/impact.

- **Mix & Match**: Mix SDN with local control.

- **Familiar Interfaces**: Reuse SAI, P4, P4Runtime, and gNMI/gNOI.

- **Remoting SAI**: P4Runtime remotes SAI, not SONiC. Low Level interfaces gives full flexibility to SDN controller.
03 Use cases
Many Use Cases for SDN & Programmability

- Software Defined WAN
- Hitless Route Sequencing
- Inline Network Functions
  - Load balancers
  - Firewalls
  - Telemetry
- Inband Network Telemetry (INT)
- Unequal Cost MultiPath (UCMP, aka WCMP)
Balanced Clos Topology

Equal Cost Multi Path (ECMP) works perfectly here.
At only 75% utilization, traffic is dropped.

BGP Link Bandwidth can help to propagate local information, but finding globally optimal solutions is still out of reach.
UCMP (or WCMP) is crucial to optimal utilization of network.

Central controller allows allocating optimal weights.
Remote Controller Performs Optimal Routing

SDN Controller learns of link failure.
Remote Controller Performs Optimal Routing

SDN Controller Computes Optimal Routing Solution
Remote Controller Performs Optimal Routing

Propagate Solution to Switches.

Modify Group Weights via P4Runtime
Remote Controller Performs Optimal Routing

[Diagram showing a network topology with spine, leaf, and host layers. Connections are labeled with bandwidths such as 20G, 40G, 30G.]
04 Summary & Questions
Key Takeaways

● Google is extending Open Source NOS (SONiC)

● Enabling the SDN ecosystem
  ○ Ability to ingest “off the shelf” switches into our DC
  ○ Aligning Open source community NOS efforts on baseline featureset.

● Enables exciting use cases
  ○ Software Defined WAN, Hitless Route Sequencing
  ○ Inline Network Functions, INT, UCMP etc.
Key Takeaways

● PINS: In incubation @ONF with a few partners
  ○ Intend to fully open source this effort in CY21
  ○ Looking for community participation.

● Watch this space for future updates.