Production OpenFlow Switches Now Available
- Building CORD Using OpenFlow Switches

CORD Build
November, 2017
Nothing That Lasts Is Built In A Day

One doesn’t know everything in the beginning

Takes time and effort to flush out all the issues

Typically takes many years for new technology to go from concept to production
OpenFlow Evolution

2009

OpenFlow 1.0:
Single table, all ternary matches
Suitable for software switch
Limited # of header match fields
Limited support for networking cases

2012

OpenFlow 1.3:
Multiple Tables, groups, meters
Extensible – new header and match fields

2014

TTPs:
Programming model for a multi-table OpenFlow switch
Formal description of tables, actions, extensions, etc. in JSON

OF-DPA
Implements a TTP to enable OpenFlow programming for line-rate hardware

Open API specification, implementations available today for multiple white box switches
OF-DPA Evolution

OF-DPA 1.0
- L2 Bridging
- L3 IPv4 and IPv6
- VxLAN

OF-DPA 2.0
- VPWS
- MPLS L3 VPN
- QoS

OF-DPA 3.x
- CORD Ready Solution

Runs on ASICs serving all segments from customer edge to core
Robust Ecosystem Available Around OF-DPA

CONTROLLERS & AGENTS
- OPEN DAYLIGHT
- Ryu
- ONOS
- Indigo
- lagopus

ODMS
- Accton
- QCT
- Inventec
- ADVANTECH
- Agema
- Celestica

✓ OPERATORS
✓ OEMS
✓ SOFTWARE VENDORS
✓ SYSTEM INTEGRATORS

Most Comprehensive OpenFlow Support in the Industry
# OF-DPA Features

## Supports Datacenter and Service Provider Use Cases

- **Layer 2** – unicast, multicast, learning, flooding
- VLAN Classification, Translation, Add/Edit/Remove
- **Layer 3** – IPv4/v6 Unicast/Multicast Routing; ECMP
- VxLAN Tunnels, Segment Routing
- VPWS, MPLS-TP VPN Groups
- MPLS LSR, LER – Initiation and Termination
- QoS – trTCM, srTCM, color blind/aware meters, MPLS Diffserve QoS
- MPLS-TP Linear Protection (G.8131), MPLS-TP OAM (G.8113.1)
- Ethernet Service OAM (G.8013/Y.1731)

Supports a wide range of use cases:
- CORD
- S-PTN
- Others
OF-DPA Use in Trellis

- XOS (Orchestration)
  - vOLT
  - Service Composition
  - Fabric Control
  - Multicast Control
  - vRouter

- ONOS Controller Cluster

- At both Spine and Leaf Access
  - MACs
  - Links

- vSG
- VNF

- White Box Hardware
- Trellis Fabric

- To controller
  - OpenFlow 1.3
  - Netconf

- Indigo OF Agent
- OF-Config Agent
- OF-DPA API
- OF-DPA
OF-DPA Software Stack

- OpenFlow Agent
  - Indigo

- OF-Config Agent
  - Netconf

- OpenFlow Data Plane Abstraction
  - Adaptation layer
  - Maps OF-DPA API to SDK API
  - SDK initializes and controls ASIC

- Open Network Linux
  - Open Compute Project (OCP)
  - ONIE

- OCP Bare Metal Hardware
  - Edgecore and others
**OF-DPA TTP Pipeline**

- **OF-DPA Features used in CORD**
  - Bridging
    - L2 Flood
    - L2 Unicast
    - L2 Multicast
    - L2 Learning
  - Routing
    - ECMP
    - IPV4 Unicast and Multicast Routing
    - IPV6 Unicast and Multicast Routing
  - MPLS L3 VPN/ L2 Pseudowire
    - Segment Routing
  - QoS Classification and Marking, Metering
  - VLAN Cross Connect

- **OF-DPA Features not used in CORD**
  - VXLAN VTEP initiation and termination
  - OAM and Protection Switching
  - Statistics counters

- **CORD Platforms**: Trident 2, Tomahawk, Qumran-MX
OF-DPA TTP Example (in JSON)

```
"flow_tables": [
  {
    "name": "IngressPort",
    "doc": "The Ingress Port Flow Table decides, based on ingress port type, whether to forward the packet using the main pipeline or in an isolated (e.g., tenant) forwarding domain. In this version of OF-DPA, the high order 16 bits are zero for physical ports and one for overlay tunnel logical ports."
  },
  {
    "name": "overlay-Tunnel",
    "priority": 1,
    "doc": "Matches packets from data center overlay tunnel logical ports."
  },
  {
    "name": "Normal-Ethernet",
    "priority": 1,
    "doc": "Matches packets from local physical ports."
  }
],
"built_in_flow_mods": [
  {
    "name": "set_ether_type",
    "doc": "Set the Ethernet frame header type.
    "match_set": [
      {"field": "ether_type", "match_type": "range", "range": [0x0800, 0x8800]"
    ],
    "instructions": [
      {"instruction": "set_ether_type", "value": {"ether_type": 0x8800}"
    }
  }
],
"flow_mod_types": [
  {
    "name": "set_ether_type",
    "priority": 1,
    "doc": "Matches packets from data center overlay tunnel logical ports."
  }
],
"instruction_set": [
  {"instruction": "GOTO_TABLE", "table": "bridging"}
]
```
QoS and Mirroring

- Meter Table invoked as action from Policy ACL table
- OpenFlow meters defined in terms of “meter bands”
- OF-DPA defines IETF standard bands
  - RFC 2697 (SrTCM)
  - RFC 2698 (TrTCM)
  - RFC 4115 (modified TrTCM, shown)

### Meter Entry
- **Identifier**
- **Flags: Kbps, Color-aware**

#### Red Band
- Type = Color Set
- Color = R
- Rate = PIR
- Burst = PBS

#### Yellow Band
- Type = Color Set
- Color = Y
- Rate = CIR
- Burst = CBS

---

- A packet of size "B" tokens arrives
- **CIR**
  - (overflow tokens if coupling mode)
  - **EIR**

- Green
  - yes: remove up to "B" tokens from Tc until empty
  - no: Yellow

- Yellow
  - yes: remove up to "B" tokens from Te until empty
  - no: Red

---

Diagram:

- Tc
- EBS
- CBS
- Te
- Red
- Green
- Yellow
Groups and Egress Tables

- Defined in the OpenFlow 1.3 specification
  - Groups are a forwarding action abstraction
  - Applied after last ingress table
  - Used for various packet processing actions, including header editing
  - Safety property – group-id must be defined before it can be referenced

- Types of Groups
  - **Indirect** (index tables, e.g., next hop for L2 header modifications, MPLS label push)
  - **Select** (ECMP group tables – random selection, e.g., based on hash function)
  - **All** (multicast replication – usually in packet buffer engine)
  - **Fast Failover** (protection switching based on port “liveness”)

- Egress Tables
  - Applied after last group
  - Cannot change egress port (but can clone or drop packet)
  - Used for things like egress ACLs, VLAN translation, etc.
  - Optional (pass through if not programmed)
OF-DPA TTP Subset Used in Trellis

- Enables Programming Hardware Forwarding Tables Using OpenFlow 1.3
- Enables ONOS Fine-Grained SDN Control of White Box Fabric Infrastructure
Configuration Examples

- VXLAN tunneling
- Port configuration
  - Queue configuration
    - Scheduling mode (e.g., strict priority)
    - Congestion properties (e.g., WRED curves)

- Yang models and UML diagrams
- Augmentations to OF-Config 1.2
Summary: Benefits of Using OF-DPA in Trellis

- Enables OpenFlow 1.3 implementations on industry leading switch architecture
- Leverages ASIC hardware tables to deliver scale and performance
- Extends OpenFlow to expose features supported in ASIC such as QoS
- Accelerates application development with openly published OF-DPA API

OF-DPA is the leader in OpenFlow implementations on merchant silicon
Thank You