OpenFlow™ Data Plane Abstraction (OF-DPA):

Abstract Switch Specification

Version 1.0

February 21, 2014

Broadcom Corporation
3151 Zanker Road
San Jose, CA 95134

This document represents the initial release of the complete specification for OF-DPA 1.0. It is intended for external distribution to solicit feedback and comments. As such, it is subject to change based on feedback received.

Table of Contents

1 **Introduction** **11**

2 **OF-DPA Components** **12**

3 **The OF-DPA Abstract Switch** **14**

3.1 Abstract Switch Overview 14

3.2 Flow Tables 15
3.2.1 Ingress Port Flow Table 16
3.2.1.1 Flow Entry Types and Match Fields 17
3.2.1.2 Instruction Types 18
3.2.1.3 Action Set 18
3.2.1.4 Counters and Flow Expiry 18
3.2.2 VLAN Flow Table 19
3.2.2.1 Flow Entry Types and Match Fields 19
3.2.2.2 Instruction Types 20
3.2.2.3 Action Set 21
3.2.2.4 Counters and Flow Expiry 21
3.2.3 Termination MAC Flow Table 22
3.2.3.1 Flow Entry Types and Match Fields 22
3.2.3.2 Instruction Types 23
3.2.3.3 Counters and Flow Expiry 23
3.2.4 Bridging Flow Table 24
3.2.4.1 Flow Entry Types and Match Fields 25
3.2.4.2 Instruction Types 25
3.2.4.3 Action Set 26
3.2.4.4 Counters and Flow Expiration 26
3.2.5 Unicast Routing Flow Table 27
3.2.5.1 Flow Entry Types and Match Fields 27
3.2.5.2 Instruction Types 28
3.2.5.3 Action Set 28
3.2.5.4 Counters and Flow Expiration 29
3.2.6 Multicast Routing Flow Table 29
3.2.6.1 Flow Entry Types and Match Fields 30
3.2.6.2 Instruction Types 30
3.2.6.3 Action Set 31
3.2.6.4 Counters and Flow Expiration 31
3.2.7 Policy ACL Flow Table 32
3.2.7.1 Flow Entry Types and Match Fields 32
3.2.7.2 Instruction Types 35
3.2.7.3 Action Set 36
3.2.7.4 Counters and Flow Expiration 37

3.3 Group Table 38
3.3.1 OF-DPA L2 Interface Group Entries 39
3.3.1.1 Naming Convention 39
3.3.1.2 Action Buckets 39
3.3.1.3 Counters 40

3.3.2 OF-DPA L2 Rewrite Group Entries 40
3.3.2.1 Naming Convention 41
3.3.2.2 Action Buckets 41
3.3.2.3 Counters 42
3.3.3 OF-DPA L3 Unicast Group Entries 42
3.3.3.1 Naming Convention 43
3.3.3.2 Action Buckets 43
3.3.3.3 Counters 44
3.3.4 OF-DPA L2 Multicast Group Entries 44
3.3.4.1 Naming Convention 45
3.3.4.2 Action Buckets 45
3.3.4.3 Counters 46
3.3.5 OF-DPA L2 Flood Group Entries 46
3.3.5.1 Naming Convention 47
3.3.5.2 Action Buckets 47
3.3.5.3 Counters 48
3.3.6 OF-DPA L3 Interface Group Entries 48
3.3.6.1 Naming Convention 49
3.3.6.2 Action Buckets 49
3.3.6.3 Counters 50
3.3.7 OF-DPA L3 Multicast Group Entries 50
3.3.7.1 Naming Convention 51
3.3.7.2 Action Buckets 51
3.3.7.3 Counters 51
3.3.8 OF-DPA L3 ECMP Group Entries 52
3.3.8.1 Naming Convention 53
3.3.8.2 Action Buckets 53
3.3.8.3 Counters 53
3.3.9 OF-DPA L2 Overlay Group Entries 54
3.3.9.1 OF-DPA L2 Overlay Group Subtypes 54
3.3.9.2 Naming Convention 56
3.3.9.3 Action Buckets 56
3.3.9.4 Counters 57
3.3.10 Fast Failover Group Entries 57

3.4 Meters 57

3.5 Ports 57
3.5.1 Physical Ports 57
3.5.1.1 Features 58
3.5.1.2 Counters 60
3.5.2 Reserved Ports 61
3.5.3 Logical Ports 62
3.5.3.1 Overlay Tunnels 62
3.5.3.2 VXLAN Tunnel Configuration 63
3.5.3.3 Logical Port Counters 67

3.6 Queues 68
3.6.1 Configuration 68
3.6.2 Counters 68

3.7 Vendor Extension Features 69
3.7.1 Source MAC Learning 69
3.7.1.1 Controller Managed Learning 69
3.7.1.2 Configuration 70
3.7.2 Group Properties 70

3.7.3 MTU Check....70
3.8 Table Numbering 70

APPENDIX A : References....72

List of Figures

Figure 1: OF-DPA Component Layering 12
Figure 2: OpenFlow OF-DPA Abstract Switch Pipeline......................................16
Figure 3: OF-DPA L2 Interface Group Entry 39
Figure 4: OF-DPA L2 Rewrite Group Entry Usage...41
Figure 5: OF-DPA L3 Unicast Group Entry Usage 43
Figure 6: OF-DPA L2 Multicast Group Entry Usage..45
Figure 7: OF-DPA L2 Flood Group Entry Usage 47
Figure 8: OF-DPA L3 Interface Group Entry 48
Figure 9: OF-DPA L3 Multicast Group Entry Usage..50
Figure 10: OF-DPA L3 ECMP Group Entry Usage..52
Figure 11: OF-DPA L2 Overlay Flood over Unicast Tunnels...............................54
Figure 12: OF-DPA L2 Overlay Flood over Multicast Tunnels 55
Figure 13: OF-DPA L2 Overlay Multicast over Unicast Tunnels 55
Figure 14: OF-DPA L2 Overlay Multicast over Multicast Tunnels 56
Figure 15: Port Properties UML Diagram 58
Figure 16: OpenFlow Feature Sub-Classes 58
Figure 17: Tunnel Logical Port UML...63
Figure 18: VXLAN Tunnel Configuration UML 65

List of Tables

Table 1: Ingress Port Flow Table Entry Types 17
Table 2: Ingress Port Flow Table Match Fields 18
Table 3: Ingress Port Flow Table Instructions 18
Table 4: Ingress Port Flow Table Counters 18
Table 5: Ingress Port Flow Table Expiry....19
Table 6: VLAN Flow Table Flow Entry Types....20
Table 7: VLAN Flow Table Match Fields 20
Table 8: VLAN Flow Table Instructions 21
Table 9: VLAN Flow Table Action List....21
Table 10: VLAN Flow Table Counters....21
Table 11: VLAN Flow Table Expiry 22
Table 12: Termination MAC Flow Table Entry Types....22
Table 13: Termination MAC Flow Table Match Fields....23
Table 14: Termination MAC Flow Table Instruction Set 23
Table 15: Termination MAC Flow Table Counters 23
Table 16: Termination MAC Flow Table Expiry 24
Table 17: Bridging Flow Table Flow Entry Types 24
Table 18: Bridging Flow Table Match Fields 25
Table 19: Bridging Flow Table Instructions 25
Table 20: Bridging Flow Table Actions by Flow Entry Type 26
Table 21: Bridging Flow Table Counters 26
Table 22: Bridging Flow Table Expiry....27
Table 23: Unicast Routing Flow Table Entry Types....27
Table 24: Unicast Routing Flow Table IPv4 Header Match Fields 27
Table 25: Unicast Routing Flow Table IPv6 Header Match Fields 28
Table 26: Unicast Routing Flow Table Instructions 28
Table 27: Unicast Routing Flow Table Actions....28
Table 28: Unicast Routing Flow Table Counters 29
Table 29: Unicast Routing Flow Table Flow Expiry 29
Table 30: Multicast Routing Flow Table Entry Types 29
Table 31: Multicast Routing Flow Table IPv4 Match Fields 30
Table 32: Multicast Routing Flow Table IPv6 Match Fields 30
Table 33: Multicast Routing Flow Table Instructions 30
Table 34: Multicast Routing Flow Table Actions....31
Table 35: Multicast Routing Flow Table Counters....31
Table 36: Multicast Routing Flow Table Flow Expiry 31
Table 37: Policy ACL Flow Table Flow Entry Types....32
Table 38: Policy ACL Flow Table IPv4 Match Fields....33
Table 39: Policy ACL Flow Table IPv6 Match Fields....34
Table 40: Policy ACL Flow Table Instruction Set 36
Table 41: Policy ACL Flow Table VLAN Flow Entry Action Set....36
Table 42: Policy ACL Flow Table Tenant Flow Entry Action Set....37
Table 43: Policy ACL Flow Table Counters....37
Table 44: Policy ACL Flow Table Expiry 38
Table 45: OF-DPA Group Table Entry Identifier Naming Convention 38

Table 46: OF-DPA L2 Interface Group Entry Type Naming Convention 39
Table 47: OF-DPA L2 Interface Group Entry Bucket Actions.............................40
Table 48: OF-DPA L2 Interface Group Entry Counters......................................40
Table 49: OF-DPA L2 Rewrite Group Entry Type Naming Convention 41
Table 50: OF-DPA L2 Rewrite Group Entry Bucket Actions...............................42
Table 51: OF-DPA L2 Rewrite Group Entry Counters..42
Table 52: OF-DPA L3 Unicast Group Entry Naming Conventioin 43
Table 53: OF-DPA L3 Unicast Bucket Actions 44
Table 54: OF-DPA L3 Unicast Group Entry Counters 44
Table 55: OF-DPA L2 Multicast Group Entry Type Naming Convention 45
Table 56: OF-DPA L2 Multicast Bucket Actions 46
Table 57: OF-DPA L2 Multicast Group Entry Counters......................................46
Table 58: OF-DPA L2 Flood Group Entry Naming Convention 47
Table 59: OF-DPA L2 Flood Bucket Actions 48
Table 60: OF-DPA L2 Flood Group Entry Counters 48
Table 61: OF-DPA L3 Interface Group Entry Type Naming Convention 49
Table 62: OF-DPA L3 Interface Group Entry Bucket Actions.............................49
Table 63: OF-DPA L3 Interface Group Entry Counters......................................50
Table 64: OF-DPA L3 Multicast Group Entry Naming Convention 51
Table 65: OF-DPA L3 Multicast Bucket Actions 51
Table 66: OF-DPA L3 Multicast Group Entry Counters......................................52
Table 67: OF-DPA L3 ECMP Group Entry Naming Convention.........................53
Table 68: OF-DPA L3 ECMP Group Entry Bucket Actions 53
Table 69: OF-DPA L3 ECMP Group Entry Counters...53
Table 70: OF-DPA L2 Overlay Flood Group Entry Naming Convention.............56
Table 71: OF-DPA L2 Overlay Group Sub-Type Entry Bucket Actions 56
Table 72: OF-DPA L2 Overlay Flood Group Sub-Type Entry Counters 57
Table 73: OF-DPA Port Features...59
Table 74: Port Features Bitmap 60
Table 75: OF-DPA Physical Port Counters 61
Table 76: OF-DPA Reserved Ports..61
Table 77: VXLAN Tunnel Endpoint Logical Port Configuration Parameters........66
Table 78: VXLAN Access Logical Port Configuration Parameters 66
Table 79: VXLAN Tenant Configuration Parameters..67
Table 80: VXLAN Next Hop Configuration Parameters......................................67
Table 81: Logical Port Counters...67
Table 82: OF-DPA Queue Configuration Parameters 68
Table 83: OF-DPA Queue Counters 68
Table 84: Source MAC Learning Feature Configuration 70
Table 85: Flow Table Number Assignments 70

Acronyms and Abbreviations

In most cases, acronyms and abbreviations are defined on first use and are also defined in the table below.

Term	Acronym	Description
Capabilities		Switch features as understood by controllers.
Flow		Sequence of packets with the same selection of header field values. Flows are unidirectional.
Flow Table		OpenFlow™ flow table as defined in the OpenFlow 1.3.1 specification.
Flow Entry		Entry in an OpenFlow flow table with its match fields and instructions.
Hybrid Switch		Control plane that has OpenFlow programmability in addition to a legacy control plane.
Group Table		The OpenFlow group table.
Group Table Entry		An OpenFlow group table entry.
Meter Table		The OpenFlow meter table.
Software-Defined Networking	SDN	User programmable control plane. In this document, SDN refers to the user ability to customize packet processing and forwarding.
Unit		A member switch within a chassis or switch stack.

1 INTRODUCTION

This document specifies an OpenFlow abstract switch model – called OpenFlow Data Plane Abstraction (OF-DPA) for Broadcom® Ethernet switch devices. The primary goal of this open specification is to enable Broadcom-based devices to be programmable using OpenFlow, and foster further growth of the ecosystem of open source and commercial OpenFlow agents and controllers that can be utilized to enable wider OpenFlow-based network infrastructure deployments. This abstract switch model is based on OpenFlow 1.3.1 [4] and utilizes its provisions to provide access to multiple tables implemented in Broadcom switch Application-Specific Integrated Circuits (ASICs). The intent is to facilitate general availability of production-quality OpenFlow switches from product vendors as well as provide a development platform for use in academic and industrial research networks.

This document represents the first feature-complete specification for OF-DPA version 1.0. It is published openly and meant to be used alongside the *Open Flow Data Plan Abstraction (OF-DPA) API Guide and Reference Manual*[1] for developing OpenFlow 1.3.1 agents and controllers. While the specification is deemed complete for features supported in OF-DPA v1.0, Broadcom solicits feedback and comments at all times to further improve the specification. As such, it may be subject to change based on feedback received from interested parties.

This document assumes familiarity with OpenFlow 1.3.1 and related Software Defined Networking (SDN) technologies.

[1] Available as an HTML document in the OF-DPA v1.0 software release package

2 OF-DPA COMPONENTS

OF-DPA is a software component that provides a hardware adaption layer between OpenFlow and Broadcom switch ASICs. It is layered above the Broadcom switch software development switch (SDK) that, in turn, provides the driver for configuring, programming, and controlling the Broadcom switch ASICs.

The OF-DPA API, as defined in the *Open Flow Data Plan Abstraction (OF-DPA) API Guide and Reference Manual*, enables programming of Broadcom ASICs using OpenFlow abstractions. However, it does not process OpenFlow protocol messages. To create a complete OpenFlow switch using OF-DPA, an OpenFlow agent is required. To complete the picture, an OpenFlow controller is required to field an OpenFlow deployment using OF-DPA-enabled switches. Figure 1 illustrates the relationship of OF-DPA with the other OpenFlow system components.

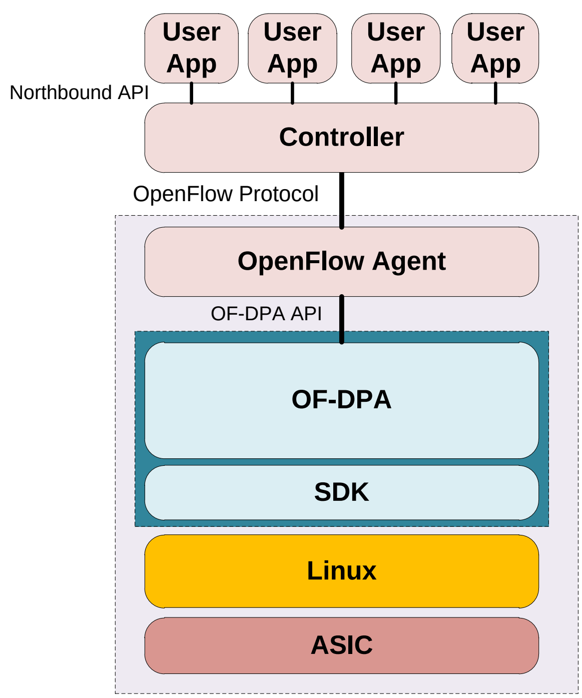

Figure 1: OF-DPA Component Layering

In Figure 1, user applications obtain services from an OpenFlow controller via a Northbound API. The Northbound API enables applications to communicate with and control one or more OpenFlow switches. In addition, the controller would likely provide advanced services such as discovery and enumeration of OpenFlow switches, along with a network-wide database of network resources, including internal and external interfaces.

The controller maintains a (secure) channel with each OpenFlow switch, over which it exchanges OpenFlow protocol messages. At the switches, OpenFlow agents maintain their end of the (secure) channel, processing received OpenFlow protocol messages and sending OpenFlow messages in response to local events.

Controllers are available from multiple sources. Any controller should be usable as long as it supports the OpenFlow 1.3.1 features defined by the Open Networking Foundation (ONF) specification and required by OF-DPA. In addition to commercial products, there are a number of readily available open source controllers, including Ryu, OpenDaylight, and CPqD.

The OF-DPA API represents hardware objects to the agent in terms of OpenFlow objects such as flow tables, group table entries, and ports. These objects are defined in the ONF OpenFlow 1.3.1 specification and instantiated by the OF-DPA Abstract Switch. As OF-DPA maintains all of the state that matches the hardware to OpenFlow, the agent is expected to do a relatively straightforward and minimal translation of OpenFlow messages into OF-DPA API calls and vice-versa, while maintaining a minimal amount of state.

The application writer programs to the description of the OF-DPA Abstract Switch in this document. Once the application gains access to an OF-DPA enabled network device, it can orchestrate and implement packet processing functions by adding flow entries to OpenFlow flow tables and defining action sets for packet editing and forwarding in terms of OpenFlow group entries. It can interrogate status of OpenFlow ports and queues, and receive events such as port state changes or flow expiration through services of the controller via the OpenFlow agent on the switch.

The next section of this document describes the OF-DPA Abstract Switch in detail. The companion document, *Open Flow Data Plan Abstraction (OF-DPA) API Guide and Reference Manual*, describes the details of the OF-DPA API.

3 THE OF-DPA ABSTRACT SWITCH

The OF-DPA abstract switch is a specialization of the OpenFlow 1.3.1 abstract switch. This section describes the OF-DPA abstract switch in terms of OpenFlow abstract objects as visible to the OpenFlow controller.

The OF-DPA OpenFlow abstract objects can be thought of as programming points for Broadcom ASICs. These include flow tables with action sets, group table entries, logical and physical ports, and queues. The OF-DPA objects support OpenFlow specific state, for example, statistics counters. They also manage hardware resources on behalf of OpenFlow.

Supporting OpenFlow in switch hardware involves some tradeoffs. As has been noted, the generality promised by OpenFlow can come at a cost of latency, as well as cost and power inefficiencies. The OF-DPA abstract switch is optimized to support single pass, full bandwidth packet processing performance that makes efficient use of the hardware and available table memory resources for bridging, routing, and data center tunnel gateway use cases, trading off unrestricted generality in favor of latency, performance and cost, while enabling a logically centralized control plane with programming flexibility.

The Abstract Switch described in this section represents the first version of OF-DPA. As such, it allows programming a subset of the functionality available in Broadcom ASICs. Future versions are expected to support additional features and packet flow use cases.

3.1 Abstract Switch Overview

OF-DPA flow tables accommodate specific types of flow entries with associated semantic rules, including constraints such as: which match fields are available, which instructions and actions are supported, how priorities can be assigned to flow entries, which next table(s) flow entries can go to, and so forth.

The flow tables fully conform to the OpenFlow 1.3.1 specification [4]. OF-DPA provides API calls to support interrogating tables for capabilities. These capabilities can include supported match fields, actions, instructions, etc. They also include status properties such as current resource usage.

Similar to flow tables, OF-DPA defines a set of group table entries. These have specific naming conventions, properties, and supported action buckets. The group table entries also conform to the

OpenFlow 1.3.1 specification, with a vendor extension. In order to support routing, a vendor extension is needed to override the default source removal and allow routing back to IN_PORT. This is fundamentally required to support a one-arm router, for example, a common use case.

Remember that OF-DPA tables are programming abstractions and do not necessary directly correspond to hardware tables.

Users must program flow tables and group entries according to the allowed entry types. The OF-DPA API validates calls and returns errors if constraints and/or conventions are violated. The OpenFlow agent that interfaces to OF-DPA may also do some argument validation and execute local iterative procedures.

Most actions are programmed using action buckets in group table entries. This not only proves to be a very efficient and modular programming approach, in that the controller can optimize hardware resources better than the switch, but also the controller intrinsically has more CPU power and memory than the control processor on a typical switch. The controller also understands what the application is trying to do, especially when programming requires updating multiple tables. However, when compared with OpenFlow 1.0 programming, it involves more controller-to-switch messages, since more objects need to be programmed. It also potentially requires the controller to keep track of more switch states, although this state can be interrogated as needed.

Some functionality must be configured using logical ports as a vendor extension. In general, this is to handle packet processing functions that the OpenFlow 1.3.1 specification currently does not handle, such as processing encapsulation headers. VXLAN data center overlay tunnels are handled by specialized configuration of logical ports rather than by directly programming flow and group entries.[2]

3.2 Flow Tables

OF-DPA presents the application writer with the flow tables shown in Figure 2. Table **goto** instructions require packets to traverse tables in a specified order. Packets accumulate actions in an action set. Actions are applied prior to the packet being forwarded when there is no next table to go to. Forwarding actions are, for the most part, provided by different types of group entries. Apply actions instructions and action lists are only used in specific constrained ways, mostly to send packets to the controller.

In the OpenFlow general case, packets pass from table to table and can be arbitrarily modified between tables. To take advantage of this generality, each table stage needs to be able to reparse packets. In OF-DPA this kind of packet flow is conceptual - packets are parsed early in the pipeline, and header fields are extracted. After that it is only these values that are used for matching in subsequent tables, or that are modified by **apply actions** instructions. It is not expected that this distinction will matter to applications.

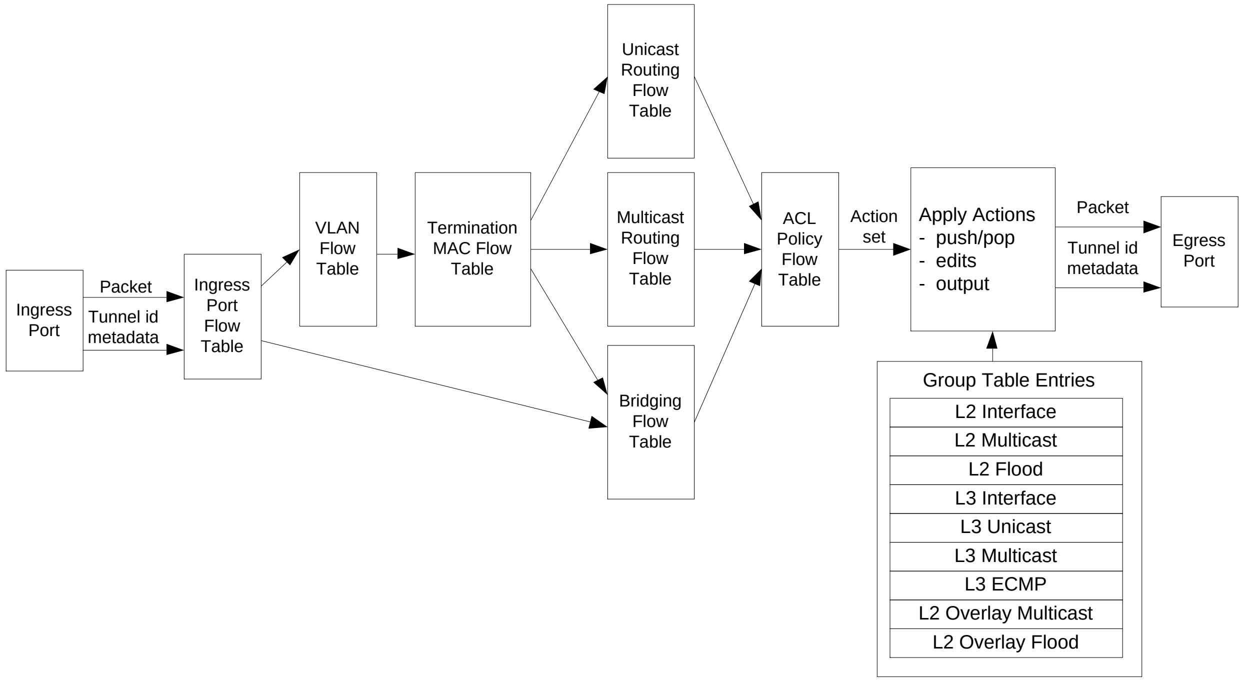

Figure 2: OpenFlow OF-DPA Abstract Switch Pipeline

Figure 2 shows the tables and group table entries in the OF-DPA abstract pipeline. Section 4.2 describes each of these tables in terms of its supported match fields, flow entry rule types, instructions, actions, expiration provisions, and statistics counters. Default miss actions are also specified for each table as applicable. Section 4.3 describes the group table entry types and action set constraints.

Ingress packets can have associated optional tunnel ID metadata. For packets from physical ports, this value is always zero. For packets from data center overlay tunnel logical ports, the tunnel ID metadata is required to identify the tenant forwarding domain, and hence cannot be zero. This version of OF-DPA only supports bridging overlay tunnel traffic in isolated tenant forwarding domains.

Note: When tenant packets are forwarded, the tunnel ID is automatically supplied to the egress logical port. OF-DPA treats the tunnel ID as read-only and does not permit it to be modified using a Set Field action[3].

3.2.1 Ingress Port Flow Table

The Ingress Port Flow Table is the first table in the pipeline and, by convention, is numbered zero. 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. To facilitate directing packets, OF-DPA encodes the port type in the ifNum port identifier so that it can be matched using specific rule types. Packets from certain logical port types can be processed differently than packets from physical ports or other logical ports by port type. OpenFlow uses a 32-bit value for ifNums, which allows OF_DPA to divide

[3] Note that OpenFlow 1.3.2 allows Set Field modification of OXM_OF_TUNNEL_ID.

it into ranges and encode the port type in the high order 16 bits. In this version of OF-DPA, the high order 16 bits are zero for physical ports and one for overlay tunnel logical ports.

The Ingress Port Flow Table presents what is essentially a demultiplexing logic function as an OpenFlow table so that it can be "programmed" from the controller. By default, packets from physical ports go to the VLAN Flow Table. For logical ports, this table must have an entry that admits ingress packets by matching the ingress ifNum against a match field specifying a range of IN_PORTs and has an appropriate Goto-Table instruction specifying the next table to process that type of packet. There is a software constraint that prevents certain types of rules from being added to other tables unless there is appropriate flow entry in the Ingress Port Flow Table. Only a small number of rule types need to be supported, one for each port type as a masked range of ifNums.

Note: The actual ifNum port range depends on the number of available interfaces and is platform dependent.

The default on miss is for packets from physical ports to go to the VLAN Flow Table. There is no default rule for packets from logical ports, which are dropped on miss.

3.2.1.1 Flow Entry Types and Match Fields

The Ingress Port Flow Table supports the flow entry types listed in Table 1. This table would typically have one rule enabling ingress packets from each port type.

Table 1: Ingress Port Flow Table Entry Types

Type	Description
Normal Ethernet Frames	Matches packets from local physical ports. Single wildcard rule with a masked IN_PORT value in the physical port range (mask=0xffff0000, value=0x00000000) and a Goto-Table instruction specifying the VLAN Flow Table.
Overlay Tunnel Frames	Matches packets from data center overlay tunnel logical ports. Single wildcard rule with a masked IN_PORT value in the physical port range (mask=0xffff0000, value=0x00010000) and a Goto-Table instruction specifying the Bridging Flow Table as the next table. The controller must add this rule before adding rules in subsequent tables that match tunneled packets.

Note: Future versions of OF-DPA will support other flow entry types for other ranges of logical ports.

The Ingress Port Flow Table uses the match fields listed in Table 2.

Table 2: Ingress Port Flow Table Match Fields

Field	Bits	Maskable	Description
IN_PORT	32	Yes	Ingress port. Maskable to support ranges of ifNum values.

3.2.1.2 Instruction Types

The Ingress Port Flow Table supports the single Goto-Table instruction listed in Table 3.

Table 3: Ingress Port Flow Table Instructions

Name	Argument	Description
Goto-Table	Table	Next table. For this version of OF-DPA, must be the VLAN Flow Table for physical ports, and the Bridging Flow Table for data center overlay tunnel logical ports.

3.2.1.3 Action Set

The Ingress Port Flow Table does not assign any actions.

3.2.1.4 Counters and Flow Expiry

The Ingress Port Flow Table supports the basic table and flow entry counters listed in Table 4.

Table 4: Ingress Port Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table.
Duration (seconds)	32	Per entry	Seconds since this flow entry was installed.

Only hard interval time-out ageing per entry is supported, as indicated in Table 5.

Table 5: Ingress Port Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.

3.2.2 VLAN Flow Table

The VLAN Flow Table is used for IEEE 801.Q compliant VLAN assignment and filtering.[4] All packets must have an associated VLAN ID in order to be processed by subsequent tables. Packets that do not match any entry in the VLAN table are filtered, that is, dropped by default.

Note: IEEE-defined BPDUs are always received untagged.[5]

3.2.2.1 Flow Entry Types and Match Fields

The VLAN Flow Table supports the Flow Entry Types listed in Table 6. Flow entries are differentiated based on IN_PORT, MAC-DST, regardless of whether the packet was tagged, and, if so, the VLAN ID in the tag.

OpenFlow uses a 16-bit field for VLAN ID, but only the low order 12 bits are needed to express a VLAN ID. OpenFlow defines special values to indicate tagged and untagged packets. In particular, the VLAN ID 0x0000 (OFPVID_NONE) is used to represent an untagged packet, and 0x1000 (OFPVID_PRESENT) for a priority tagged packet. All tagged packets are represented by VLAN ID values between 0x1001 and 0x1FFE[6] (OFPVID_PRESENT | VLAN ID value).

Note: OF-DPA does not support the special VLAN tag values defined in OpenFlow 1.3.1 that test for the presence of a VLAN tag. Therefore matching packets, based on whether or not they have a VLAN tag, are not supported. Refer to Table 12 in the *OpenFlow 1.3.1 Specification*, item [4] in Appendix A.

Entries in the OF-DPA VLAN Flow table are mutually exclusive. Any explicit rule priority assignments are ignored.

[4] The VLAN Flow Table presents the hardware port and VLAN configuration bitmaps to the OpenFlow controller as a flow table.
[5] There are vendor-specific BPDUs that are VLAN tagged.
[6] 0x1FFF represents an illegal VLAN id.

Table 6: VLAN Flow Table Flow Entry Types

Type	Description
VLAN Filtering	Exact match on IN_PORT and VLAN_VID parsed from the packet. For tagged packets with a VLAN tag containing a VLAN_VID greater than zero. Cannot be masked. VLAN_VID cannot be used in a Port VLAN Assignment rule for untagged packets. The only instruction is Goto-Table and must specify the Termination MAC Flow Table. Tagged packets that miss are considered VLAN_VIDs that are not allowed on the port and are dropped.
Untagged Packet Port VLAN Assignment	Exact match on IN_PORT and VLAN ID == 0 (lower 12 bits of match field) value using a mask value of 0x0fff (masks off OFPVID_PRESENT). Action set must assign a VLAN_VID. The VLAN_VID cannot be one used in a VLAN Filtering rule. If the packet does not have a VLAN tag, one will be pushed, if necessary, at packet egress. Rule must have a Goto-Table instruction specifying the Termination MAC Flow Table. Untagged packets are dropped if there is no port VLAN assignment rule.

Note: The untagged packet rule applies to both untagged packets, which match VLAN_VID = 0x1000, and IEEE 802.1P priority tagged packets, which match VLAN_VID = 0x0000. However the VLAN-PCP match field will be set from the value in the VLAN tag rather than default to zero in the case of a packet without a VLAN tag.

Note: A VLAN Flow Table rule cannot specify an IN_PORT and VLAN_VID combination that is used in a Tunnel Access Logical Port configuration. Conversely, it must include a rule to permit an IN_PORT and VLAN_VID combination used in a Tunnel Next Hop configuration.

The VLAN Flow Table matches the fields listed in Table 7.

Table 7: VLAN Flow Table Match Fields

Field	Bits	Maskable	Description
IN_PORT	32	No	Ingress port. Must be a value in the physical port range (less than 1024).
VLAN_VID	16	Yes	Outer VLAN id. The mask value must be 0x1ffff for VLAN filtering rules and 0x0fff for untagged packet rules.

3.2.2.2 Instruction Types

The VLAN table supports the instruction types listed in Table 8.

Table 8: VLAN Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	The VLAN Flow Table supports the actions specified in Table 9.
Goto-Table	Table	For VLAN filtering or Port VLAN assignment the next table can only be the Termination MAC Flow Table. A packet is dropped if it matches an entry that has no next table specified.

3.2.2.3 Action Set

The VLAN table uses Apply Actions for port VLAN tagging and assignment. The action list can have at most one of each action type.

Table 9: VLAN Flow Table Action List

Name	Argument	Description
Set Field	VLAN_VID, must be between 1 and 4094.	Sets the VLAN ID for an untagged port VLAN assignment rule. If the packet does not have a VLAN tag, then one is pushed with the specified VLAN ID and priority zero. If the VLAN tag exists, then the VLAN ID will be replaced with the specified value.

Note: The untagged packet action is the same as in the OpenFlow 1.0 specification. The implicit addition of a tag to an untagged packet is unspecified in the OpenFlow 1.3.1 specification.

3.2.2.4 Counters and Flow Expiry

The VLAN Flow Table supports the table and flow entry counters listed in Table 10.

Table 10: VLAN Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed

Only hard interval time-out ageing per entry is supported, as indicated in Table 11.

Table 11: VLAN Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.

3.2.3 Termination MAC Flow Table

The Termination MAC Flow Table determines whether to do bridging or routing on a packet. It identifies destination MAC, VLAN, and Ethertype for routed packets. Routed packet rule types use a goto instruction to indicate that the next table is one of the routing tables. The default on a miss is to go to the Bridging Flow Table.

3.2.3.1 Flow Entry Types and Match Fields

The Termination MAC Flow Table implements the flow entry types listed in Table 12.

Table 12: Termination MAC Flow Table Entry Types

Name	Description
Unicast MAC	Used to identify an IPv4 or IPv6 router MAC. Priority must be assigned so as to be lower than a multicast MAC rule, if one exists.
IPv4 Multicast MAC	Wildcard rule that recognizes all IPv4 multicast MAC addresses specified in RFC 1112. If specified, this must be ETH_DST = 01-00-5e-00-00-00 with mask ff-ff-ff-80-00-00. There can only be one flow entry of this type.
IPv6 Multicast MAC	Wildcard rule that recognizes all IPv6 MAC addresses specified in RFC 2464. If specified, this must be ETH_DST = 33-33-00-00-00-00 with mask ff-ff-00-00-00-00. There can only be one flow entry of this type.

The Termination MAC Flow Table match fields are listed in Table 13. Strict rule priority must be assigned by the controller so that every flow entry has a unique priority.

Table 13: Termination MAC Flow Table Match Fields

Field	Bits	Maskable	Description
IN_PORT	32	Yes	Physical (local) input port. Field maskable only.
ETH-TYPE	16	No	The only allowed values are 0x0800 or 0x86dd.
ETH_DST	48	Yes	Ethernet destination MAC. Prefix maskable for only the specific multicast IP flow entries in Table 12. Must be field masked for unicast destination MACs.
VLAN_VID	16	Yes	Matches against the Outer VLAN ID. Field maskable only.

3.2.3.2 Instruction Types

The Termination MAC Flow Table can have the instructions shown in Table 14.

Table 14: Termination MAC Flow Table Instruction Set

Name	Argument	Description
Goto-Table	Table	Unicast MAC rules can only specify the Routing Flow Table. Multicast MAC rules can only specify the Multicast Routing Flow Table. Packet is dropped if rule matches and there is no next table.
Apply Actions	Action List	Optional. If supplied, can only contain one action - output a copy to CONTROLLER.

3.2.3.3 Counters and Flow Expiry

The Termination MAC Flow Table counters are listed in Table 15.

Table 15: Termination MAC Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Number of active flow entries in the table
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed

Termination MAC Flow Table expiry provisions are as indicated in Table 16.

Table 16: Termination MAC Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.

3.2.4 Bridging Flow Table

The Bridging Flow Table supports Ethernet packet switching for potentially large numbers of flow entries using the hardware L2 tables. The default on a miss is to go to the Policy ACL Flow Table.

Note: The Policy ACL Flow Table is preferred for matching BPDUs.

The Bridging Flow Table forwards based on either VLAN (normal switched packets) or Tunnel ID (isolated tenant forwarding domain packets) using the flow entry types in Table 17.

Table 17: Bridging Flow Table Flow Entry Types

Type	Description
Unicast VLAN Bridging	Matches switched unicast Ethernet frames by VLAN ID and MAC_DST. MAC_DST must be unicast and cannot be masked. VLAN ID must be nonzero. Tunnel ID must be zero.
Multicast VLAN Bridging	Matches switched multicast Ethernet frames by VLAN ID and MAC_DST. MAC_DST must be multicast and cannot be masked. VLAN ID must be nonzero. Tunnel ID must be zero.
DLF VLAN Bridging	Matches switched Ethernet frames by VLAN ID only. MAC_DST is optional and must be field masked if specified. Must have lower priority than any unicast or multicast flow entries that specify this VLAN. VLAN ID must be non-zero. Tunnel ID must be zero.
Unicast Tenant Bridging	Matches switched unicast Ethernet frames by tunnel ID and MAC_DST. MAC_DST must be unicast and cannot be masked. Tunnel ID must be nonzero. VLAN ID must be zero.
Multicast Tenant Bridging	Matches switched multicast Ethernet frames by tunnel ID and MAC_DST. MAC_DST must be multicast and cannot be masked. Tunnel ID must be nonzero. VLAN ID must be zero.

Type	Description
DLF Tenant Bridging	Matches switched Ethernet frames by tunnel ID only. MAC_DST is optional and must be field masked if specified. Must have lower priority than any unicast or multicast flow entries that specify this tunnel ID. Tunnel ID must be nonzero. VLAN ID must be zero.

Note: Exact match rules must be given higher priority assignments than any wildcard rules. In any event, exact match rules are evaluated before any wildcard rules.

3.2.4.1 Flow Entry Types and Match Fields

Match fields for flow entry types are described in the following tables.

Table 18: Bridging Flow Table Match Fields

Field	Bits	Maskable	Description
ETH_DST	48	Yes	Ethernet destination MAC. Field maskable only; allowed values depend on flow entry type.
VLAN_VID	16	No	VLAN ID.
TUNNEL ID	32	No	Identifies isolated forwarding domain for tenant traffic.

3.2.4.2 Instruction Types

Default next table if no match is the ACL Policy Flow Table.

Table 19: Bridging Flow Table Instructions

Name	Argument	Description
Write-Actions	Action set	Only the actions in Section 3.2.4.3 can be specified.
Apply-Actions	Action list	Optional. If specified, can contain only a single output action to send a copy to CONTROLLER
Goto-Table	Table	Must be the ACL Policy Flow Table if specified. If packet matches and no next table is specified then the packet is dropped.

3.2.4.3 Action Set

The Bridging Flow Table supports the actions in Table 20 by flow entry type. The OF-DPA API validates consistency of flow entry type and OF-DPA group entry type references.

Table 20: Bridging Flow Table Actions by Flow Entry Type

Type	Argument	Description
Unicast VLAN Bridging	Group ID	Must be an OF-DPA L2 Interface group entry for the forwarding VLAN
Multicast VLAN Bridging	Group ID	Must be an OF-DPA L2 Multicast group entry for the forwarding VLAN
DLF VLAN Bridging	Group ID	Must be an OF-DPA L2 Flood group entry for the forwarding VLAN
Unicast Tenant Bridging	Output	Must be an overlay tunnel logical port associated with the tenant forwarding domain (tunnel ID)
Multicast Tenant Bridging	Group ID	Must be an OF-DPA L2 Overlay Multicast sub-type group entry for the tenant forwarding domain (tunnel ID)
DLF Tenant Bridging	Group ID	Must be an OF-DPA L2 Overlay Flood subtype group entry for the tenant forwarding domain (tunnel ID)

3.2.4.4 Counters and Flow Expiration

The Bridging Flow Table counters are listed in Table 21.

Table 21: Bridging Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Number of active entries in the table.
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed.

Bridging Flow Table expiry provisions are shown in Table 22.

Table 22: Bridging Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.
Idle Timeout	32	Number of seconds of inactivity - after which a flow entry is removed. Optional, flow entry does not age out if unspecified or zero.

3.2.5 Unicast Routing Flow Table

The Unicast Routing Flow Table supports routing for potentially large numbers of IPv4 and IPv6 flow entries using the hardware L3 tables.

The Unicast Routing Flow Table is a single table but organized as two mutually exclusive logical subtables by IP protocol, and supports the flow entry types listed in Table 23. A single table number is used for both logical tables.

Table 23: Unicast Routing Flow Table Entry Types

Type	Table	Prerequisite(s)	Description
IPv4 Unicast	Table 24	Ethertype=0x0800	Matches routed unicast IPv4 packets.
IPv6 Unicast	Table 25	Ethertype=0x86dd	Matches routed unicast IPv6 packets.

3.2.5.1 Flow Entry Types and Match Fields

Match fields for flow entry types are described in the following tables.

Table 24: Unicast Routing Flow Table IPv4 Header Match Fields

Field	Bits	Maskable	Description
ETH_TYPE	16	No	Must be 0x0800.
IPv4 DST	32	Yes	Must be a unicast IPv4 address. Bit-maskable, but must be prefix-masked.

Table 25: Unicast Routing Flow Table IPv6 Header Match Fields

Field	Bits	Maskable	Description
ETH_TYPE	16	No	Must be 0x86dd.
IPV6_DST	128	Yes	Must be a unicast IPv6 address.

Note: Exact match rules must be given higher priority assignments than any LPM prefix match rules. In any event, the hardware evaluates exact match rules before any wildcard rules.

3.2.5.2 Instruction Types

Default next table on a miss is the ACL Policy Flow Table.

Table 26: Unicast Routing Flow Table Instructions

Name	Argument	Description
Write-Actions	Action set	Only the actions in Table 27 can be specified.
Goto-Table	Table	Must be ACL Policy if specified. If not specified, the packet is dropped.

Other instruction types, specifically Apply Actions, are not supported.

3.2.5.3 Action Set

The actions in Table 27 are supported.

Table 27: Unicast Routing Flow Table Actions

Name	Argument	Description
Group	Group ID	Must be an OF-DPA L3 Unicast Group Entry.

The group entry includes the decrement TTL and MTU check actions, so these need not be explicitly specified in the action set.

3.2.5.4 Counters and Flow Expiration

The Routing Flow Table counters are listed in Table 28.

Table 28: Unicast Routing Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table.
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed.

Unicast Routing Flow Table expiry provisions are shown in Table 29.

Table 29: Unicast Routing Flow Table Flow Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.
Idle Timeout	32	Number of seconds of inactivity, after which a flow entry is removed. Optional, entry does not age out if zero or not specified.

3.2.6 Multicast Routing Flow Table

The Multicast Routing Flow Table supports routing for IPv4 and IPv6 multicast packets.

The Multicast Routing Flow Table is also organized as two mutually exclusive logical subtables by IP protocol, and supports the flow entry types listed in Table 30.

Table 30: Multicast Routing Flow Table Entry Types

Type	Table	Prerequisite(s)	Description
IPv4 Multicast	Table 31	Ethertype=0x0800	Matches routed multicast IPv4 packets.
IPv6 Multicast	Table 32	Ethertype=0x86dd	Matches routed multicast IPv6 packets.

3.2.6.1 Flow Entry Types and Match Fields

Match fields for flow entry types are described in the following tables.

Table 31: Multicast Routing Flow Table IPv4 Match Fields

Field	Bits	Maskable	Description
ETH_TYPE	16	No	Must be 0x0800. Required pre-requisite.
VLAN_VID	16	No	VLAN ID.
IPV4_SRC	32	Yes	Field maskable only.
IPV4_DST	32	No	Must be an IPv4 multicast group address.

Table 32: Multicast Routing Flow Table IPv6 Match Fields

Field	Bits	Maskable	Description
ETH_TYPE	16	No	Must be 0x86dd. Required prerequisite.
VLAN_VID	16	No	VLAN ID.
IPV6_SRC	128	Yes	Field maskable only.
IPV6_DST	128	No	Must be an IPv6 multicast group address.

3.2.6.2 Instruction Types

Default next table on miss is the ACL Policy Flow Table.

Table 33: Multicast Routing Flow Table Instructions

Name	Argument	Description
Write Actions	Action set	Only the actions in Table 33 can be specified.
Goto-Table	Table	Must be the Policy ACL Flow Table. In the event that there is no group entry referenced and no next table specified, the packet will be dropped.

Other instruction types, specifically Apply Actions, are not supported.

3.2.6.3 Action Set

The Multicast Routing Table supports the actions in Table 34.

Table 34: Multicast Routing Flow Table Actions

Name	Argument	Description
Group	Group ID	Must be an OF-DPA L3 Multicast group entry with the forwarding VLAN ID as a name component.

Note: The group entry includes the decrement TTL and MTU check actions.

3.2.6.4 Counters and Flow Expiration

The Multicast Routing Flow Table counters are as shown in Table 35.

Table 35: Multicast Routing Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table.
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed.

Multicast Routing Flow Table expiry provisions are shown in Table 36.

Table 36: Multicast Routing Flow Table Flow Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.
Idle Timeout	32	Number of seconds of inactivity after which a flow entry is removed. Optional: entry does not age out if zero or not specified.

3.2.7 Policy ACL Flow Table

The Policy ACL Flow Table supports wide, multifield matching. Most fields can be wildcard matched, and explicit priority must be included in all flow entry modification API calls. This is the preferred table for matching BPDU and ARP packets. It is also the only table where QoS actions are available.

The Policy ACL Flow Table is organized as mutually exclusive logical subtables. Flow entries in the IPv6 logical tables match only IPv6 packets by either VLAN ID or Tunnel ID, but not both. The non-IPv6 logical table matches any packet except for IPv6 packets by either VLAN ID or Tunnel ID, but not both. By OpenFlow single-entry match semantics, since the Policy ACL Flow Table is considered a single table, a packet can match, at most, one rule in the entire table.

Note: The Ethertype prerequisite must be explicitly provided and cannot be masked.

The Policy ACL Flow Table supports the flow entry types listed in Table 37.

Table 37: Policy ACL Flow Table Flow Entry Types

Type	Table	Prerequisite	Description
IPv4 VLAN	Table 38	Ethertype != 0x86dd, IN_PORT is a physical port.	Matches packets by VLAN ID except for IPv6. VLAN ID is optional but must be nonzero if supplied. Tunnel ID is optional but must be zero if supplied.
IPv6 VLAN	Table 39	Ethertype != 0x86dd, IN_PORT is a physical port.	Matches only IPv6 packets by VLAN ID. VLAN ID is optional but must be nonzero if supplied. Tunnel ID is optional but must be zero if supplied.
IPv4 Tenant	Table 38	Ethertype=0x86dd, IN_PORT is a physical port.	Matches packets by tunnel ID except for IPv6. A nonzero Tunnel ID is required and is not maskable. VLAN ID must be field masked if supplied.
IPv6 Tenant	Table 39	Ethertype=0x86dd, IN_PORT is a physical port.	Matches only IPv6 packets by tunnel ID. A nonzero Tunnel ID is required and is not maskable. VLAN ID must be field masked if supplied.

3.2.7.1 Flow Entry Types and Match Fields

The available match fields for Policy ACL Flow Table flow entry types are as described in the following tables.

Table 38: Policy ACL Flow Table IPv4 Match Fields

Field	Bits	Maskable	Description
IN_PORT	32	Yes	Physical or logical ingress port. For physical ports, the only mask value permitted is 0xffff0000, corresponding to any physical port. Logical ports cannot be masked.
ETH_SRC	48	Yes	Ethernet source MAC.
ETH_DST	48	Yes	Ethernet destination MAC.
ETH_TYPE	16	No	Any value except 0x86dd.
VLAN_VID	16	Yes	VLAN ID. Cannot be masked for a VLAN bridging redirect rule that specifies an L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	Yes	Priority from VLAN tag. Always has a value; will be zero if packet was untagged.
TUNNEL ID	32	No	Tenant forwarding domain. Only applicable to tenant flow entry types.
IPV4_SRC	32	Yes	Matches SIP if Ethertype = 0x0800.
ARP_SPA	32	Yes	Matches ARP source protocol address if Ethertype = 0x0806.
IPV4_DST	32	Yes	Matches DIP if Ethertype = 0x0800.
IP_PROTO	8	Yes	IP protocol field from IP header if Ethertype = 0x0800.
IP_DSCP	6	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474.
IP_ECN	2	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168.
TCP_SRC	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 6.
UDP_SRC	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 17.
SCTP_SRC	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 132.

Field	Bits	Maskable	Description
ICMPV4_TYPE	8	Yes	If Ethertype = 0x0800 and IP_PROTO = 1.
TCP_DST	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 6.
UDP_DST	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 17.
SCTP_DST	16	Yes	If Ethertype = 0x0800 and IP_PROTO = 132.
ICMPv4_CODE	8	Yes	If Ethertype = 0x0800 and IP_PROTO = 1.

Table 39: Policy ACL Flow Table IPv6 Match Fields

Field	Bits	Maskable	Description
IN_PORT	32	Yes	Physical or logical ingress port. For physical ports, the only mask value permitted is 0xffff0000, corresponding to any physical port. Logical ports cannot be masked.
ETH_SRC	48	Yes	Ethernet source MAC.
ETH_DST	48	Yes	Ethernet destination MAC.
ETH_TYPE	16	No	Must be 0x86dd.
VLAN_VID	16	Yes	VLAN ID. Cannot be masked for a VLAN bridging redirect rule that specifies an L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	Yes	Priority from VLAN tag. Always has a value; will be zero if packet was untagged.
TUNNEL ID	32	No	Tenant forwarding domain. Only applicable to tenant flow entry types.
IPV6_SRC	128	Yes	Matches IPv6 SIP.
IPV6_DST	128	Yes	Matches IPv6 DIP.
IP_PROTO	8	Yes	Matches IPv6 Next header.

Field	Bits	Maskable	Description
IPV6_FLABEL	20	Yes	Matches IPv6 flow label.
IP_DSCP	6	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474.
IP_ECN	2	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168.
TCP_SRC	16	Yes	If Ethertype = 0x86dd and IP_PROTO = 6.
UDP_SRC	16	Yes	If Ethertype = 0x86dd and IP_PROTO = 17.
SCTP_SRC	16	Yes	If Ethertype = 0x86dd and IP_PROTO = 132.
ICMPV6_TYPE	8	Yes	If Ethertype = 0x86dd and IP_PROTO = 58.
TCP_DST	16	Yes	If Ethertype = 0x86dd 00 and IP_PROTO = 6.
UDP_DST	16	Yes	If Ethertype = 0x86dd and IP_PROTO = 17.
SCTP_DST	16	Yes	If Ethertype = 0x86dd and IP_PROTO = 132.
ICMPv6_CODE	8	Yes	If Ethertype = 0x86dd and IP_PROTO = 58.

Notes:

1. IPv6 Neighbor Discovery field matching is not supported in this version of OF-DPA.
2. Not all IPv6 match fields are supported on all platforms.
3. OF-DPA permits bit masking L4 source and destination ports, as well as ICMP code. The OpenFlow 1.3.1 specification does not require these to be maskable.

3.2.7.2 Instruction Types

The only instruction is write actions. Since there is no next table, there can be no Goto-Table or Write Metadata instructions.

Table 40: Policy ACL Flow Table Instruction Set

Name	Argument	Description
Apply Actions	Action list	Optional. If supplied, can only contain one action - output a copy to CONTROLLER.
Clear Actions	–	Used to clear any group entry reference from the action set for dropping the packet. Cannot be combined with write actions.
Write Actions	Action set	Only the actions in Table 41 can be specified.

The packet is dropped if there is no group action that specifies output ports, since there is no next table.

Note: Apply-actions to CONTROLLER would be used in order to output the packet to the CONTROLLER reserved port, rather than an output action in the write-actions action set.

3.2.7.3 Action Set

The Policy ACL Flow Table supports the actions listed in Table 41 for VLAN match rule types, and the actions in Table 42 for tenant match rule types.

Table 41: Policy ACL Flow Table VLAN Flow Entry Action Set

Name	Argument	Description
Group	Group	Sets output group entry for processing the packet after this table. Group must exist, be consistent with the type of rule and packet, and can be any of the following: L2 Interface, L2 Rewrite, L2 Multicast, L3 Unicast, L3 Multicast, or L3 ECMP respectful of VLAN ID naming conventions. In particular, if the output is an L2 Rewrite group that does not set the VLAN ID, the L2 Interface group it references must be consistent with the VLAN ID in the matched flow entry.
Set-Queue	Queue-id	Determines queue to be used when packet is finally forwarded. Zero indicates the default queue.
Set-Field	VLAN PCP	New outer VLAN priority marking.
Set-Field	IP_DSCP	New IP DSCP marking. Applicable to IP packets only.

As with Unicast and Multicast Routing Flow Table actions, the decrement TTL and MTU checks are encoded by referencing an L3 Unicast or Multicast group entry.

Note: If the group entry type is L2 Interface, L2 Rewrite, or L2 Multicast, then these checks will not be done.

Table 42: Policy ACL Flow Table Tenant Flow Entry Action Set

Name	Argument	Description
Group	Group	Sets output group entry for multicast forwarding. Group entry must exist and be an OF-DPA L2 Overlay Multicast subtype in the tenant forwarding domain.
Output	ifNum	Sets output port for unicast forwarding. Must be a tunnel logical port in the tenant forwarding domain.
Set-Field	IP_DSCP	New IP DSCP marking. Applicable to IP packets only.

3.2.7.4 Counters and Flow Expiration

The Policy ACL Flow Table counters are listed in Table 43. These are applicable to both VLAN and Tenant flow entries.

Table 43: Policy ACL Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table.
Duration (sec.)	32	Per entry	Seconds since this flow entry was installed.
Received Packets	64	Per entry	Number of packets to which this rule applies.
Received Bytes	64	Per entry	Number of bytes to which this rule applies.

Policy ACL Flow Table expiry provisions are shown in Table 44. Each flow entry can have its own time-out values.

Table 44: Policy ACL Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.
Idle Timeout	32	Number of seconds of inactivity, after which a flow entry is removed. Optional: entry does not age out if zero or not specified.

3.3 Group Table

Most forwarding actions are embodied in group table entries. OF-DPA supports a defined set of group table entry types, effectively partitioning the group table into logical subtables.

Each group entry has an identifier, type, counters, and one or more action buckets. OpenFlow has a single monolithic group table, but OF-DPA differentiates among types of group entries. For this purpose, OF-DPA encodes the group entry type in a group entry identifier field. The basic naming convention followed is illustrated in Table 45.

Table 45: OF-DPA Group Table Entry Identifier Naming Convention

Field	Bits	Description
Index	[27:0]	28-bit field, used to uniquely identify a group entry of the indicated type. May be used to further encode properties of the group entry, such as VLAN_VID.
Type	[31:28]	4-bit field that encodes the entry type, one of: 0: OF-DPA L2 Interface 1: OF-DPA L2 Rewrite 2: OF-DPA L3 Unicast 3: OF-DPA L2 Multicast 4: OF-DPA L2 Flood 5: OF-DPA L3 Interface 6: OF-DPA L3 Multicast 7: OF-DPA L3 ECMP 8: OF-DPA L2 Overlay

The OF-DPA API performs consistency checks on the group entry type when a group action is used in a flow entry.

The index scheme varies by OF-DPA group entry type and is described in the following sections.

3.3.1 OF-DPA L2 Interface Group Entries

L2 Interface Group entries are of OpenFlow indirect type, with a single action bucket.

OF-DPA L2 Interface group entries are used for egress VLAN filtering and tagging. The identifier scheme is diagramed in Figure 3 and detailed in Table 45. If a VLAN is allowed on a port, an appropriate group entry must be defined for that VLAN and port combination. Only predefined group entries can be referenced in flow entry or group entry action sets.

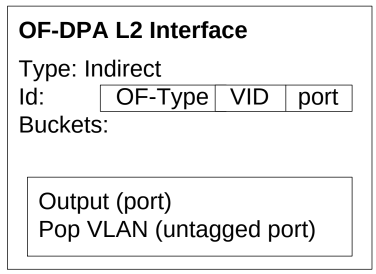

Figure 3: OF-DPA L2 Interface Group Entry

3.3.1.1 Naming Convention

Table 46 details the OF-DPA L2 Interface group entry identifier subfields that encode combinations of egress port and VLAN ID.

Table 46: OF-DPA L2 Interface Group Entry Type Naming Convention

Field	Bits	Description
Port identifier	[15:0]	Identifies a physical port (ifNum).
VLAN Id	[27:16]	VLAN ID.
Type	[31:28]	0 (L2 Unicast).

3.3.1.2 Action Buckets

The single action bucket specifies the output port, and whether or not the packet is egressed tagged. Although the pop action is a NOP if the packet has no VLAN tag, packets should always have a VLAN tag when the actions in the output group table are applied.

Note: If the packet came in untagged and a port VLAN was assigned, a VLAN tag was pushed as a VLAN Flow Table action.

Table 47: OF-DPA L2 Interface Group Entry Bucket Actions

Field	Argument	Description
Output	Port	Physical output port.
Pop VLAN	None	Pop the VLAN tag before sending the packet.

Clearly OF-DPA L2 Interface group entries must be defined before being used. OF-DPA maintains reference counts for used entries, and an entry cannot be deleted if it is referenced by a flow entry or another group.

3.3.1.3 Counters

OF-DPA L2 Interface group entry counters are as shown in Table 48.

Table 48: OF-DPA L2 Interface Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow entries or group entities currently referencing this group entry.
Duration (secs.)	32	Per entry	Seconds since this group entry was installed.

3.3.2 OF-DPA L2 Rewrite Group Entries

OF-DPA L2 Rewrite group entries are of indirect type and have a single action bucket. They are used when it is desired to modify Ethernet header fields for bridged packets. Use of an OF-DPA L2 Rewrite group entry is optional, and can only be a Policy ACL Flow Table action.

Figure 4 shows usage of an OF-DPA L2 Rewrite Group entry.

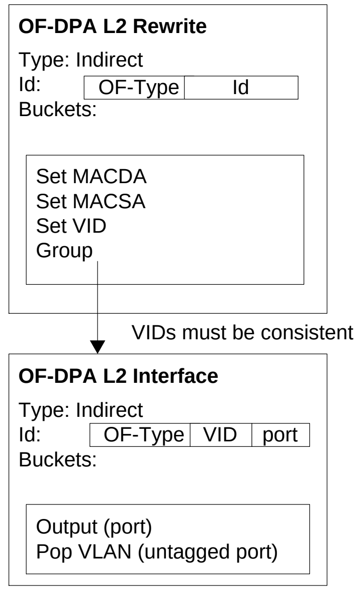

Figure 4: OF-DPA L2 Rewrite Group Entry Usage

OF-DPA L2 Rewrite actions are optional with the exception of group. This permits an OF-DPA L2 Rewrite group entry to selectively modify the source MAC, destination MAC, and/or VLAN ID.

If a Set Field action sets the VLAN ID, the VLAN ID must be the same as in the L2 Interface group entry.

Note: If the VLAN ID is not rewritten, the VLAN ID in the L2 Interface group entry must be the same as the VLAN ID matched in the Policy ACL Flow Table flow entry that forwarded to the rewrite group.

3.3.2.1 Naming Convention

Table 49 details the OF-DPA L2 Rewrite group entry identifier subfields that encode the type and VLAN ID.

Table 49: OF-DPA L2 Rewrite Group Entry Type Naming Convention

Field	Bits	Description
Id	[27:0]	Index to differentiate group entries of this type.
Type	[31:28]	1 (OF-DPA L2 Rewrite).

3.3.2.2 Action Buckets

The single action bucket specifies the output group for forwarding the packet and optional Ethernet header modifications.

Table 50: OF-DPA L2 Rewrite Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Must reference an OF-DPA L2 Interface group entry. The VLAN ID component of the referenced group entry must match the Set Field value for VLAN ID. Required.
Set Field	MAC_SRC	Rewrite the source MAC. Optional.
Set Field	MAC_DST	Rewrite the destination MAC. Optional.
Set Field	VLAN-id	Rewrite the VLAN ID. Optional.

Referenced group entries must be defined before being used. OF-DPA maintains reference counts for used entries, and an entry cannot be deleted if it is referenced by a flow entry or another group.

3.3.2.3 Counters

OF-DPA L2 Rewrite group entry counters are as shown in Table 51 for completeness.

Table 51: OF-DPA L2 Rewrite Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.3 OF-DPA L3 Unicast Group Entries

OF-DPA L3 Unicast group entries are used to supply the routing next hop and output interface for forwarding the packet.

To properly route a packet from either the Routing Flow Table or the Policy ACL Flow Table, the packet must be forwarded to an OF-DPA L3 Unicast Group entry.

Figure 5 shows an OF-DPA L3 Unicast Group entry referencing an OF-DPA L2 Unicast Group entry.

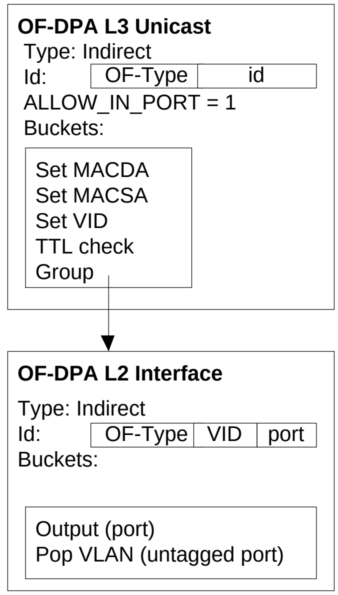

Figure 5: OF-DPA L3 Unicast Group Entry Usage

OF-DPA L3 Unicast uses the ALLOW_IN_PORT vendor extension property to allow packets to be sent out IN_PORT. This property overrides the OpenFlow default behavior, which is to not forward a packet to IN_PORT. Note that the ALLOW_IN_PORT vendor extension property applies to referenced groups as well.

The referenced OF-DPA L2 Interface Group entry must be in the same VLAN as that assigned by the OF-DPA L3 Unicast Group entry.

3.3.3.1 Naming Convention

The naming convention for OF-DPA L3 Unicast Group entries is shown in Table 52.

Table 52: OF-DPA L3 Unicast Group Entry Naming Conventioin

Field	Bits	Description
Id	[27:0]	Index to differentiate group entries of this type.
Type	[31:28]	2 (OF-DPA L3 Unicast).

3.3.3.2 Action Buckets

The single action bucket is as shown in Table 53. All actions are required.

Table 53: OF-DPA L3 Unicast Bucket Actions

Field	Argument	Description
Group	Group-id	Must reference an OF-DPA L2 Interface group entry. This group entry can output the packet to IN_PORT. The VLAN ID of the referenced group entry must match the Set Field VLAN ID value, and the port identifiers must match.
Set Field	MAC_DST	Write the next hop destination MAC.
Set Field	MAC_SRC	Write the source MAC corresponding to the L3 output interface.
Set Field	VLAN-id	Write the VLAN ID corresponding to the L3 output interface.
Decrement TTL and do MTU check	-	MTU check is a vendor extension. An invalid TTL (zero before or after decrement) is always dropped and a copy is sent to the CPU for forwarding to the CONTROLLER. Similarly, a packet that exceeds the MTU is dropped and a copy is sent to the CONTROLLER.

3.3.3.3 Counters

The OF-DPA L3 Unicast group entry counters are as shown in Table 54.

Table 54: OF-DPA L3 Unicast Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.4 OF-DPA L2 Multicast Group Entries

OF-DPA L2 multicast group entries are of OpenFlow ALL type. There can be multiple action buckets, each encoding an output port by forwarding to an OF-DPA L2 Unicast Group entry. In effect, all this group entry type does is reference one or more OF-DPA L2 Interface group entries.

Note: By default, and as per the OpenFlow 1.3.1 specification, a packet cannot be forwarded back to the IN_PORT from which it came in. An action bucket that specifies the particular packet's ingress port is not evaluated.

Figure 6 illustrates OF-DPA L2 Multicast Group Entry contents and its relationship to referenced OF-DPA L2 Unicast Group entries.

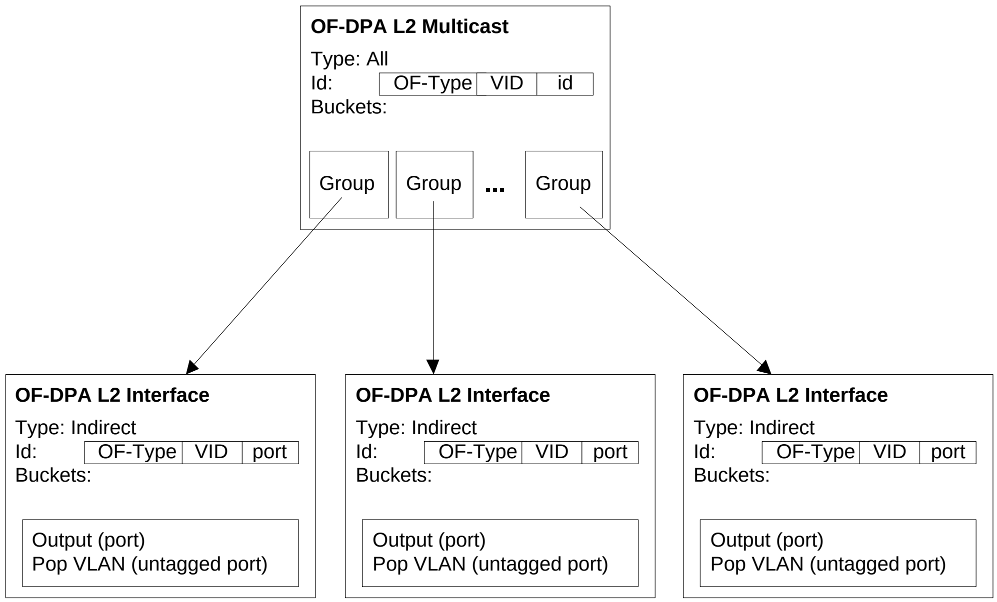

Figure 6: OF-DPA L2 Multicast Group Entry Usage

All of the OF-DPA L2 Interface Group entries referenced by the OF-DPA Multicast Group entry, and the OF-DPA Multicast Group entry itself, must be in the same VLAN.

3.3.4.1 Naming Convention

OF-DPA L2 Multicast group entries use the naming convention in Table 55.

Table 55: OF-DPA L2 Multicast Group Entry Type Naming Convention

Field	Bits	Description
Id	[15:0]	Index to differentiate group entries of this type.
VLAN Id	[27:16]	VLAN ID.
Type	[31:28]	3 (L2 Multicast).

3.3.4.2 Action Buckets

The contents of OF-DPA L2 Multicast Group entry buckets can contain only the value shown in Table 56.

Table 56: OF-DPA L2 Multicast Bucket Actions

Field	Argument	Description
Group	Group-id	Must reference an OF-DPA L2 Multicast group entry whose VLAN ID name component matches the referencing group entry.

3.3.4.3 Counters

The VL2 Multicast group entry counters are as shown in Table 57.

Table 57: OF-DPA L2 Multicast Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.5 OF-DPA L2 Flood Group Entries

The OF-DPA L2 Flood Group entries are used by VLAN Flow Table DLF rules. Like OF-DPA L2 Multicast, they are of OpenFlow *all type*. The action buckets each encode an output port. OF-DPA L2 Flood Group entry buckets each forward a replica to an output port. The main difference from OF-DPA L2 Multicast Group entries is how they are processed in the hardware.

Note: There can be only one OF-DPA L2 Flood Group Entry per VLAN.

Figure 7 illustrates an OF-DPA L2 Flood Group Entry and its usage.

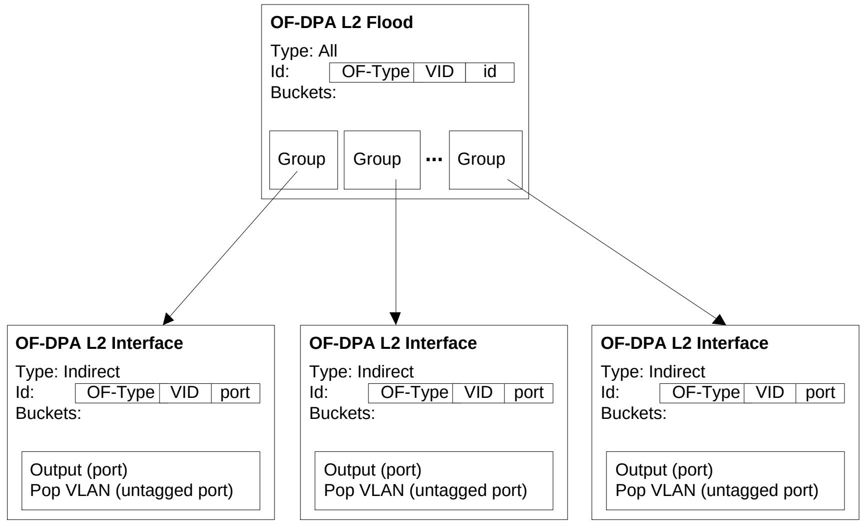

Figure 7: OF-DPA L2 Flood Group Entry Usage

All of the OF-DPA L2 Interface Group entries referenced by the OF-DPA Flood Group entry, and the OF-DPA Flood Group entry itself, must be in the same VLAN.

3.3.5.1 Naming Convention

OF-DPA L2 Flood group entries follow the naming convention shown in Table 58.

Table 58: OF-DPA L2 Flood Group Entry Naming Convention

Field	Bits	Description
Id	[15:0]	Index to differentiate group entries of this type.
VLAN Id	[27:16]	VLAN ID.
Type	[31:28]	4 (OF-DPA L2 Flood).

3.3.5.2 Action Buckets

The contents of the OF-DPA L2 Flood Group Entry action buckets can contain only the values shown in Table 59.

Table 59: OF-DPA L2 Flood Bucket Actions

Field	Argument	Description
Group	Group-id	Must reference an OF-DPA L2 Unicast group entry whose VLAN ID name component matches the VLAN ID in the referencing group entry name.

3.3.5.3 Counters

The OF-DPA L2 Multicast group entry counters are as shown in Table 60.

Table 60: OF-DPA L2 Flood Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.6 OF-DPA L3 Interface Group Entries

OF-DPA L3 interface group entries are of indirect type and have a single action bucket. They are used to supply outgoing routing interface properties for multicast forwarding. For unicast forwarding, use of OF-DPA L3 Unicast group entries is recommended.

Figure 8 provides a diagram of an OF-DPA L3 Interface Group entry.

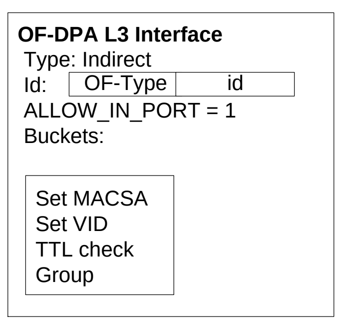

Figure 8: OF-DPA L3 Interface Group Entry

OF-DPA L3 Interface uses the ALLOW-IN-PORT vendor extension that permits packets to be sent out IN_PORT.

The VLAN ID in the name must be the same as the VLAN_VID assigned in the Set Field action.

3.3.6.1 Naming Convention

Table 61 details the OF-DPA L3 Interface group entry identifier subfields.

Table 61: OF-DPA L3 Interface Group Entry Type Naming Convention

Field	Bits	Description
Id	[27:0]	Index to differentiate group entries of this type.
Type	[31:28]	5 (OF-DPA L3 Interface).

3.3.6.2 Action Buckets

The single action bucket specifies the MAC_SRC, VLAN_VID, TTL decrement action, and an output group for forwarding the packet. All actions are required.

Table 62: OF-DPA L3 Interface Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Must reference an OF-DPA L2 Interface group entry. This group entry can output the packet to IN_PORT. The VLAN ID component of the referenced group entry must match the Set Field value for VLAN ID.
Set Field	MAC_SRC	Write the source MAC corresponding to the L3 output interface.
Set Field	VLAN-id	Write the VLAN ID corresponding to the L3 output interface.
Decrement TTL and do MTU check	-	Required. MTU check is a vendor extension. An invalid TTL (zero before or after decrement action) is always dropped, and a copy is sent to the CPU for forwarding to the CONTROLLER. Similarly, a packet that exceeds the MTU is dropped, and a copy is sent to the CONTROLLER.

Referenced group entries must be defined before being used. OF-DPA maintains reference counts for used entries, and an entry cannot be deleted if it is referenced by a flow entry or another group.

3.3.6.3 Counters

OF-DPA L3 Interface group entry counters are as shown in Table 51 for completeness.

Table 63: OF-DPA L3 Interface Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.7 OF-DPA L3 Multicast Group Entries

OF-DPA L3 Multicast group entries are of OpenFlow *all type*. The action buckets describe the interfaces to which multicast packet replicas are forwarded.

Figure 9 illustrates OF-DPA L3 Multicast group entries and how they are used.

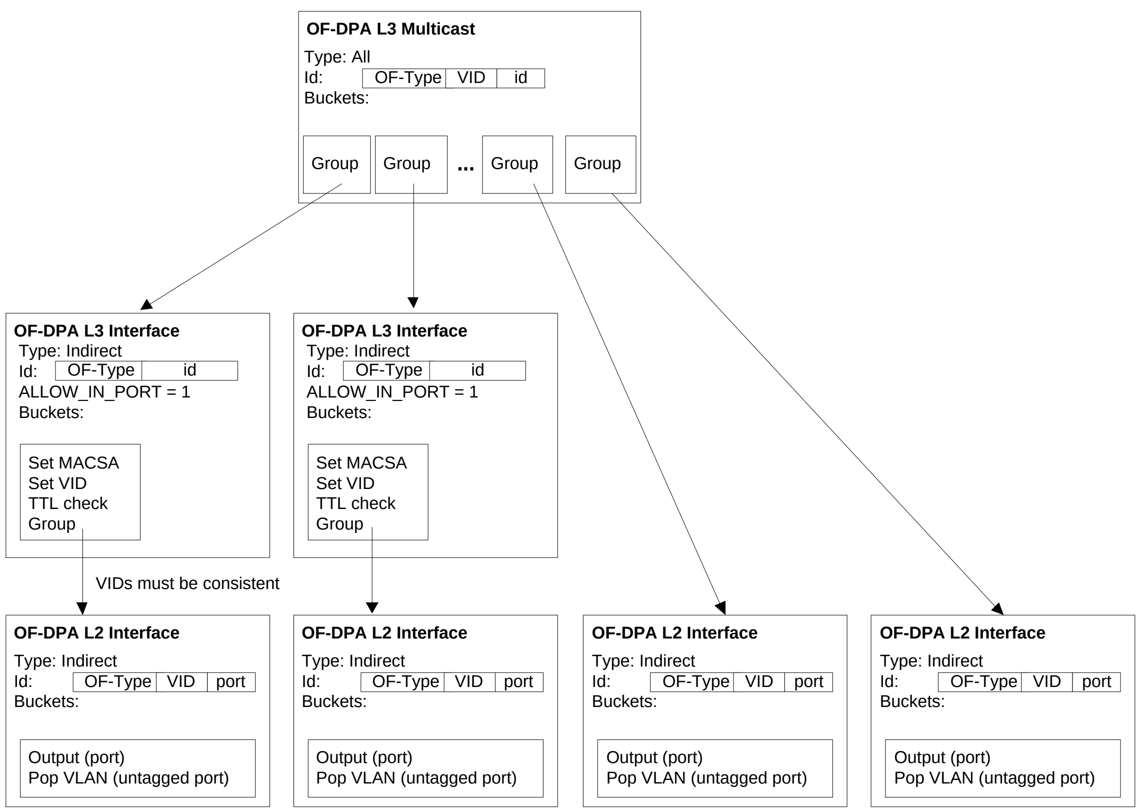

Figure 9: OF-DPA L3 Multicast Group Entry Usage

IP multicast packets are forwarded differently, depending on whether they are switched or routed. Packets must be switched in the VLAN in which they came in and cannot be output to IN_PORT. Packets that are multicast in other VLANs must be routed and must be allowed to egress via IN_PORT. This difference is reflected in the actions that are programmed in the action buckets.

Note: The referenced OF-DPA L2 Interface Group entries must be in the same VLAN as the OF-DPA L3 Multicast group entry. However the referenced OF-DPA L3 Interface Group entries must be in different VLANs from the OF-DPA L3 Multicast Group entry and from each other.

3.3.7.1 Naming Convention

The naming convention for OF-DPA L3 Multicast Group entries is shown in Table 64.

Table 64: OF-DPA L3 Multicast Group Entry Naming Convention

Field	Bits	Description
Index	[15:0]	Used to differentiate between OF-DPA L3 multicast group entries.
VLAN Id	[27:16]	VLAN ID.
Type	[31:28]	6 (OF-DPA L3 Multicast).

3.3.7.2 Action Buckets

The action buckets contain the values shown in Table 65.

Table 65: OF-DPA L3 Multicast Bucket Actions

Field	Argument	Description
Group	Group-id	OF-DPA L3 Multicast buckets can forward to multiple OF-DPA L3 Interface Group entry and OF-DPA L2 Interface Group entries. Referenced group entry names must conform to the VLAN ID requirements above.

3.3.7.3 Counters

The OF-DPA L3 Multicast group entry counters are as shown in Table 66.

Table 66: OF-DPA L3 Multicast Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (sec.)	32	Per entry	Seconds since this group entry was installed.

3.3.8 OF-DPA L3 ECMP Group Entries

OF-DPA L3 ECMP group entries are OpenFlow select groups. The action buckets reference the OF-DPA L3 Unicast group entries that are members of the multipath group for ECMP forwarding.

Figure 10 illustrates the OF-DPA L3 ECMP Group entry usage.

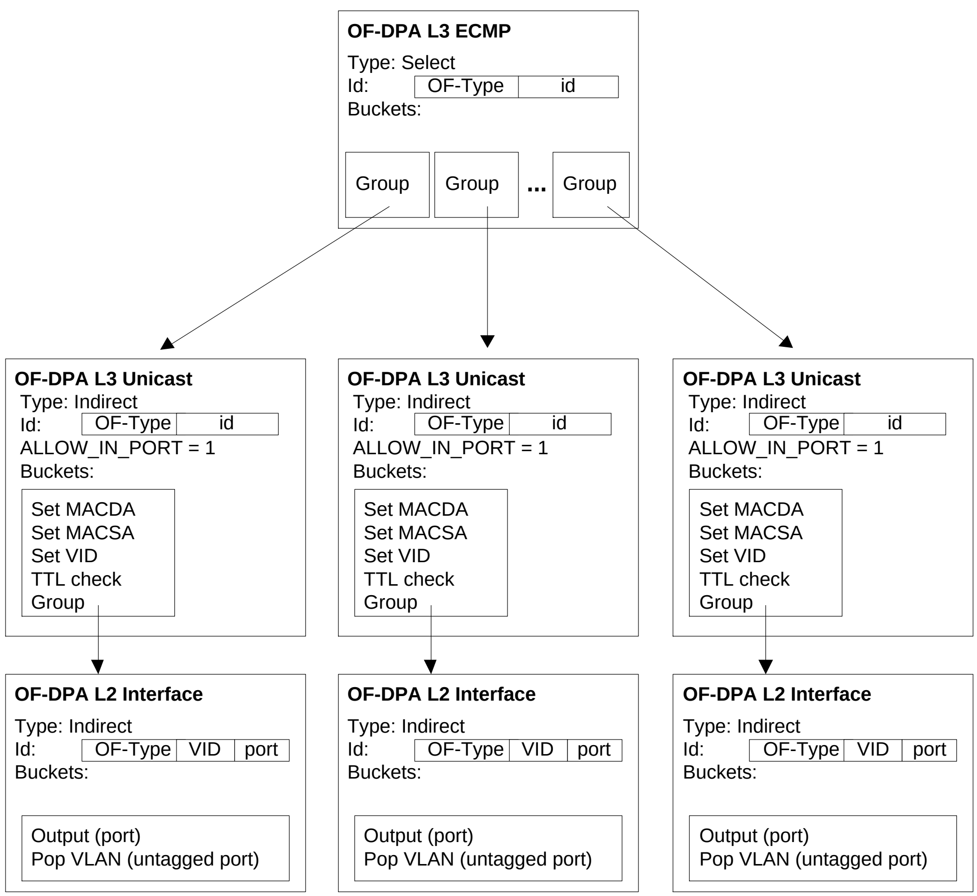

Figure 10: OF-DPA L3 ECMP Group Entry Usage

An OF-DPA L3 ECMP Group entry can be specified as a routing target instead of an OF-DPA L3 Unicast Group entry. Selection of an action bucket for forwarding a particular packet is hardware specific.

As before, the VLAN ID in the L2 Interface group entry must be consistent with the VLAN ID set in the L3 Unicast Set-Field action.

3.3.8.1 Naming Convention

The naming convention for OF-DPA L3 ECMP Group entries is as shown in Table 67.

Table 67: OF-DPA L3 ECMP Group Entry Naming Convention

Field	Bits	Description
Id	[27:0]	Used to differentiate OF-DPA L3 ECMP group entries.
Type	[31:28]	7 (OF-DPA L3 ECMP).

3.3.8.2 Action Buckets

The action buckets contain the single value listed in Table 68.

Table 68: OF-DPA L3 ECMP Group Entry Bucket Actions

Field	Argument	Description
Group	Group-id	Must reference an OF-DPA L3 Unicast group entry subject to VLAN ID naming constraints.

3.3.8.3 Counters

The OF-DPA L3 ECMP group entry counters are as shown in Table 69.

Table 69: OF-DPA L3 ECMP Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (secs.)	32	Per entry	Seconds since this group entry was installed.

3.3.9 OF-DPA L2 Overlay Group Entries

OF-DPA L2 Overlay Group Entries are of OpenFlow *all type*. The action buckets describe the tenant access logical ports and/or tunnel endpoint logical ports to which packets are to be replicated by this group. **Note:** All tenant logical ports must be for the same tenant as the tunnel ID in the group name.

Tenant access and tunnel endpoint logical port configuration is described in Section 3.5.3.

3.3.9.1 OF-DPA L2 Overlay Group Subtypes

There are four OF-DPA L2 Overlay Group subtypes. These can be considered OF-DPA group entries in their own right, but are described together here since they perform similar functions. The differences relate to usage (whether in DLF or multicast flows) and to the underlay remote tunnel endpoint type (whether unicast or multicast).

Note: Regardless of whether forwarded (overlay) packets are themselves unicast or multicast, they will be replicated using the underlay tunnel type corresponding to the OF-DPA L2 Overlay Group subtype name component.

Figure 11 shows an OF-DPA L2 Overlay Flood over Unicast Tunnels group entry. Buckets can specify multiple access and/or tunnel logical ports. OF-DPA will use unicast underlay tunnels to forward packets for the specified logical ports. OF-DPA L2 Overlay Flood over Unicast Tunnels group entries can only be referenced by tunnel DLF rule types.

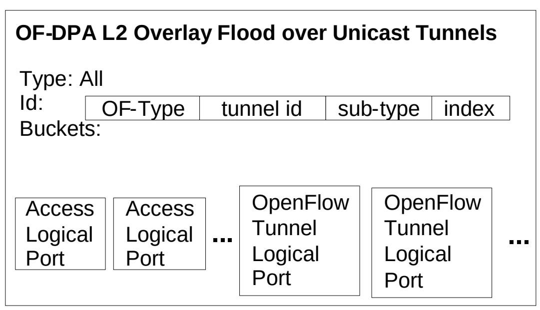

Figure 11: OF-DPA L2 Overlay Flood over Unicast Tunnels

Figure 12 illustrates an OF-DPA L2 Overlay Flood over Multicast Tunnels group entry. There can be, at most, one bucket specifying a tunnel logical port. OF-DPA will forward packets over the tenant multicast underlay tunnel configured on the tunnel logical port. A multicast IP group address must have been configured for the tenant on that logical port.

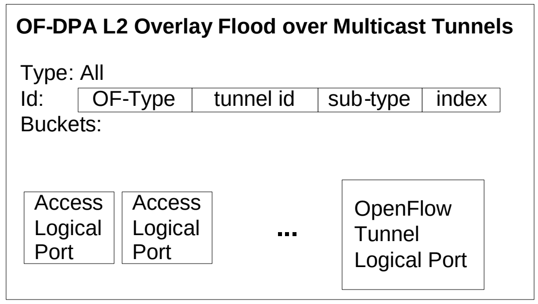

Figure 12: OF-DPA L2 Overlay Flood over Multicast Tunnels

OF-DPA L2 Overlay Flood over Multicast Tunnels group entries can only be referenced by a tunnel DLF rule.

Figure 13 shows an OF-DPA L2 Overlay Multicast over Unicast Tunnels group entry. Multiple tunnel logical port buckets can be specified. OF-DPA will use unicast underlay tunnels to forward packets for the specified logical ports. OF-DPA L2 Overlay Multicast over Unicast Tunnels group entries cannot be referenced by tunnel DLF rule types.

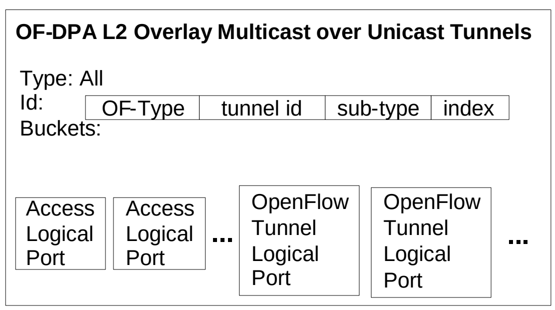

Figure 13: OF-DPA L2 Overlay Multicast over Unicast Tunnels

Figure 14 illustrates the OF-DPA L2 Overlay Multicast over Multicast Tunnels group entry. There can be, at most, one bucket specifying a tunnel logical port configured with a multicast IP group address for the tenant. OF-DPA will use unicast underlay tunnels to forward packets for the specified logical ports. OF-DPA L2 Overlay Multicast over Unicast Tunnels group entries cannot be referenced by tunnel DLF rules.

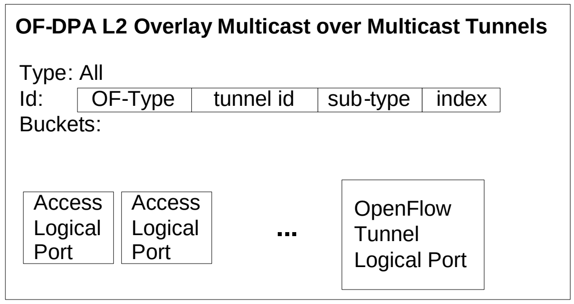

Figure 14: OF-DPA L2 Overlay Multicast over Multicast Tunnels

3.3.9.2 Naming Convention

The naming convention for OF-DPA L2 Overlay Flood group entries is shown in Table 70.

Table 70: OF-DPA L2 Overlay Flood Group Entry Naming Convention

Field	Bits	Description
Index	[9:0]	Used to differentiate L2 Overlay group entries of the same subtype.
Subtype	[11:10]	Identifies the type of underlay tunnel used: 0: OF-DPA L2 Overlay Flood Group over Unicast Tunnels 1: OF-DPA L2 Overlay Flood Group over Multicast Tunnels 2: OF-DPA L2 Overlay Multicast Group over Unicast Tunnels 3: L2 Overlay Multicast Group over Multicast Tunnels
Tunnel Id	[27:12]	Low order 16 bits of the 32-bit tenant forwarding domain identifier. Must uniquely identify the tenant.
Type	[31:28]	8 (OF-DPA L2 Overlay Flood).

3.3.9.3 Action Buckets

The action buckets for all OF-DPA L2 Overlay Group Entry subtypes contain the values shown in Table 71.

Table 71: OF-DPA L2 Overlay Group Sub-Type Entry Bucket Actions

Field	Argument	Description
Output	Logical port	Must be a logical port in the tenant forwarding domain. Can be either an access or tunnel logical port.

3.3.9.4 Counters

The OF-DPA L2 Overlay Flood group entry counters are as shown in Table 72. These counters are individually maintained by subtype.

Table 72: OF-DPA L2 Overlay Flood Group Sub-Type Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per entry	Number of flow or group entities currently referencing this group entry.
Duration (seconds)	32	Per entry	Seconds since this group entry was installed.

3.3.10 Fast Failover Group Entries

This version of OF-DPA does not support fast failover group entries.

3.4 Meters

This version of OF-DPA does not support meter bands.

3.5 Ports

This section lists the OF-DPA supported properties for physical, reserved, and logical ports.

3.5.1 Physical Ports

OF-DPA supports physical ports that are available on specific target platforms.

Physical ports are front panel ports on the abstract switch. Figure 15, from OF-Config 1.1.1, shows the UML port configuration data model.

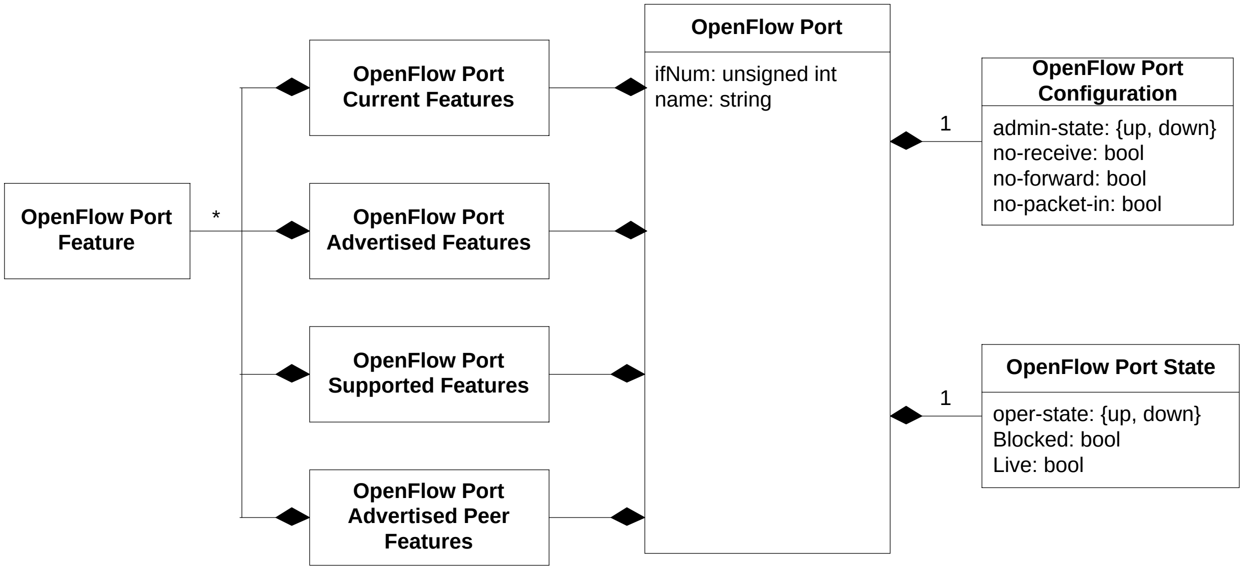

Figure 15: Port Properties UML Diagram

3.5.1.1 Features

OpenFlow Port Features are further modeled in terms of the subclasses shown in Figure 16, also from OF-Config 1.1.1.

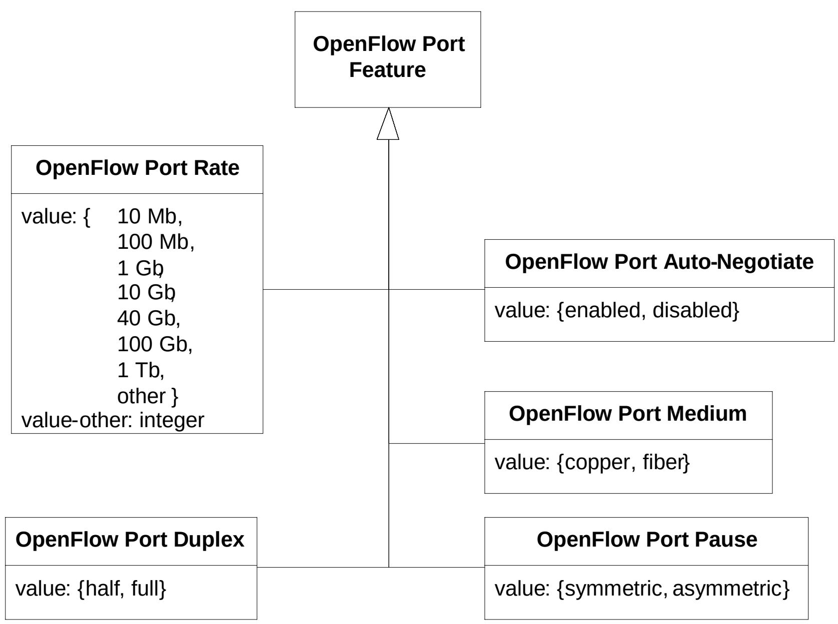

Figure 16: OpenFlow Feature Sub-Classes

For OF-Config, which is based on NetConf and YANG, UML models such as these are used to directly derive XML style sheets for representing configuration protocol messages.

OF-DPA supports the physical port features listed in Table 73.

Table 73: OF-DPA Port Features

Name	Bits	Configurable?	Description
Number	32	No	ifNum (should be the same as in interface MIB).
Hardware Address	48	No	MAC address assigned to port.
Name	128	Yes	16-byte string name (should be the same as in interface MIB).
Configured State	32	Yes	Port is administratively up (0) or down (1).
Current State	32	No	Port link state is up (0) or down (1).
Current Features	32	No	OF-DPA supports the feature bitmap in Table 74. A one indicates the feature is currently active.
Advertised Features	32	No	OF-DPA supports the feature bitmap in Table 74. A zero bit indicates the feature is not available.
Supported Features	32	No	OF-DPA supports the features in Table 74. A zero bit indicates the feature is not supported.
Peer Features	32	No	Bitmap indicating capabilities advertised by the peer from Table 74.
Current Speed	32	No	Current port bitrate in kbps.
Max Speed	32	No	Maximum port bitrate in kbps.

Note: Not all of the above may be applicable to the LOCAL reserved port.

Table 74 shows the port features bitmap referenced from the table above and the OpenFlow Port Features subclasses in Figure 16.

Table 74: Port Features Bitmap

Feature	Bit	Description
10 Mbps HD	0	10 Mbps half-duplex
10 Mbps FD	1	10 Mbps full-duplex
100 Mbps HD	2	100 Mbps half-duplex
100 Mbps FD	3	100 Mbps full-duplex
1GB HD	4	1 Gbps half-duplex
1GB FD	5	1 Gbps full-duplex
10GB FD	6	10 Gbps full-duplex
40GB FD	7	40 Gbps full-duplex
100GB FD	8	100 Gbps full-duplex
1TB FD	9	1 Tbps full-duplex
Other	10	Other rate, not in the above list
Copper	11	Copper medium
Fiber	12	Fiber medium
Autoneg	13	Auto-negotiation
Pause	14	Pause enabled
Pause_Asym	15	Asymmetric pause supported

3.5.1.2 Counters

OF-DPA supports the port counters listed in Table 75.

Table 75: OF-DPA Physical Port Counters

Name	Bits	Description
Received Packets	64	Total packets received.
Transmitted Packets	64	Total packets transmitted.
Received Bytes	64	Total bytes received.
Transmitted Bytes	64	Total bytes transmitted.
Receive Drops	64	Received packets dropped for any reason.
Transmit Drops	64	Transmitted packets dropped for any reason.
Receive Errors	64	Received packet errors.
Transmit Errors	64	Transmit packets errors.
Receive Frame Alignment Errors	64	Received packets with frame alignment errors.
Receive Overrun Errors	64	Received packet overruns.
Receive CRC Errors	64	Received packet CRC errors.
Collisions	64	Collisions.
Duration (sec.)	32	Time in seconds since configured.

3.5.2 Reserved Ports

OF-DPA supports the reserved ports listed in Table 76.

Table 76: OF-DPA Reserved Ports

Name	Description	Use	Supported?
ALL	Not supported in OF-DPA.	Output	No

Name	Description	Use	Supported?
IN_PORT	L3 Unicast Group entries obviate the need for this port.	Output	No
CONTROLLER	The OpenFlow controller. Packet is sent to CPU.	Input or output	Yes
TABLE	Used in PACKET_OUT messages to indicate that the packet must be recirculated through the pipeline. Must always be the first table in the pipeline, if specified.	Output	Yes
ANY	Use specific L2 and L3 Multicast Group entries.	-	No
LOCAL	Used to send packets to the local CPU.	Input or output	Yes
NORMAL	Use LOCAL instead.	Output	No
FLOOD	Use L2 Flood Group Entries instead.	Output	No

3.5.3 Logical Ports

Logical Ports are used to model functionality as external to the OpenFlow pipeline. The abstract switch receives packets from an ingress logical port after they have been processed by an external function and forwards packets to a destination logical port for processing by an external function.

Logical port external function parameters are configured on the Logical Port.

OF-DPA uses Logical ports for overlay tunnels.

3.5.3.1 Overlay Tunnels

Tunnel packets enter the OF-DPA data path from Tunnel Logical Ports, along with tunnel ID metadata. The tunnel ID identifies the tenant forwarding domain. Tunnel Logical Ports are modeled according the UML data model in Figure 17, which shows Tunnel Logical Ports as abstract classes.

Note: There are two Tunnel Logical Port subtypes shown. Access ports connect local servers in the tenant forwarding domain. Tunnel Endpoints connect to remote switches.

Tunnel Logical Ports must have a specified protocol in order to be instantiated. This version of OF-DPA supports VXLAN [10] overlays.

Note: VXLAN support is switch hardware platform-dependent.

The Tunnel Endpoint abstract class provides the necessary configuration parameters common to different protocol subclasses. This includes the local and remote endpoint addresses, the TTL for packet origination, and the multipath properties for forwarding tunnel initiation packets.

Similarly, the Tunnel Access Port abstract class provides necessary parameters for locally attached servers. Three methods are supported: all traffic on a port, all traffic with a particular VLAN ID on a port, or packets tagged with an IEEE 802.1BR [9] port extension tag (ETAG) on that port.

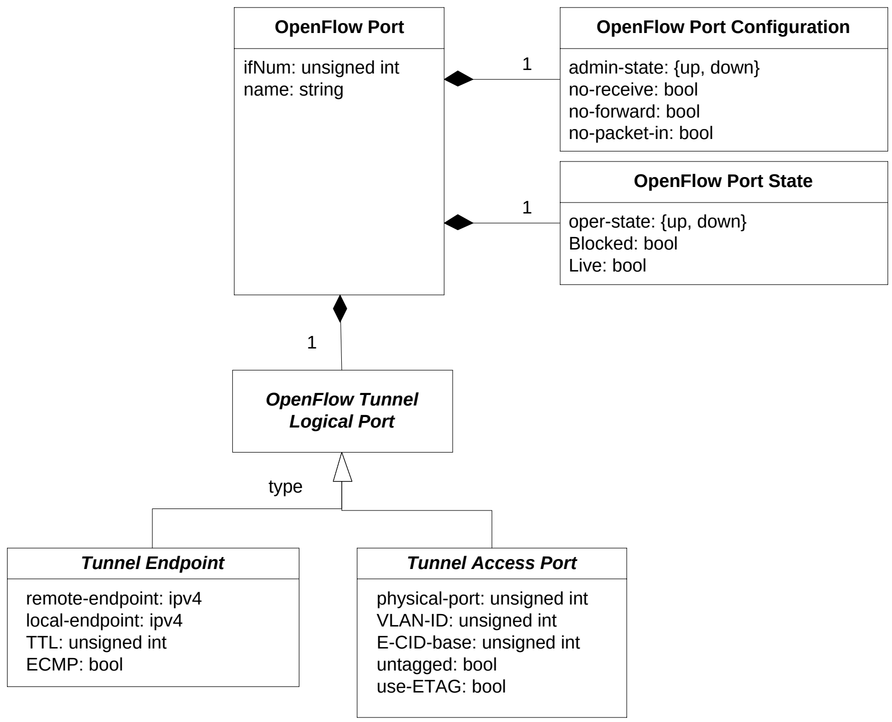

Figure 17: Tunnel Logical Port UML

3.5.3.2 VXLAN Tunnel Configuration

VXLAN Logical Port configuration is shown in the UML diagram in Figure 18. Two types of VXLAN Logical Port can be configured: VXLAN Tunnel Endpoints and VXLAN Access Ports.

VXLAN Tunnel Endpoint Logical Ports are used to forward packets to a remote tunnel endpoint, or VTEP. The VXLAN Tunnel Endpoint class is configured with protocol specific header properties as well as tunnel initiation forwarding properties. The use-entropy flag indicates that a hash value is to be inserted, instead of the configured udp-src-port-if-no-entropy setting.

Note: This version of OF-DPA supports hardware with a single system-wide configuration for the terminator-udp-dest-port and use-entropy settings. As a result, all configured VXLAN Tunnel Endpoints must specify the same values for these parameters.

VXLAN Tunnel Endpoints must be configured with forwarding state for tunnel initiation packets. This can be specified in terms of a VXLAN Unicast Tunnel Next Hop or an ECMP VXLAN Next Hop Group multipath object, if the ECMP flag is set. The ECMP class aggregates one or more VXLAN Unicast Tunnel Next Hop objects.

Traffic for multiple tenants can be multiplexed over a VXLAN Tunnel Endpoint. The VXLAN Tenant class provides the protocol header information (VN_ID) for distinguishing a particular Tenant's traffic. It also identifies the Tenant's isolated forwarding domain for ingress and egress packets. The VXLAN Tunnel class can also provide an IP multicast group address for this Tenant's traffic.

Note: Multiple VXLAN Tenants can share an IP multicast address.

The VXLAN Access Port class configures a logical port for a VXLAN Tenant's locally attached endpoint. The Access Port configuration specifies how traffic is classified to a particular VXLAN Tenant isolated forwarding domain. This can be one of: all traffic on a port; traffic on a port with a particular VLAN ID; and traffic with a particular E-Tag [9].

Note: There is some interaction between the overlay tunnel configuration, VLAN Flow Table entries, and L2 Interface Groups. For Access Ports, configuration must be mutually exclusive in order to isolate overlay tenant traffic. This means that a VLAN Flow Table entry must not specify filtering for local tenant traffic configured via an Access Port, and an L2 Interface Group must not call out the same port and VLAN properties as for local tenant traffic. Tunnel Endpoint operation, however, depends upon L2 Interface Group settings for forwarding underlay VXLAN packets initiated by a Tunnel Next Hop, and upon VLAN Flow Table entries to permit receiving underlay terminated tunnel packets.

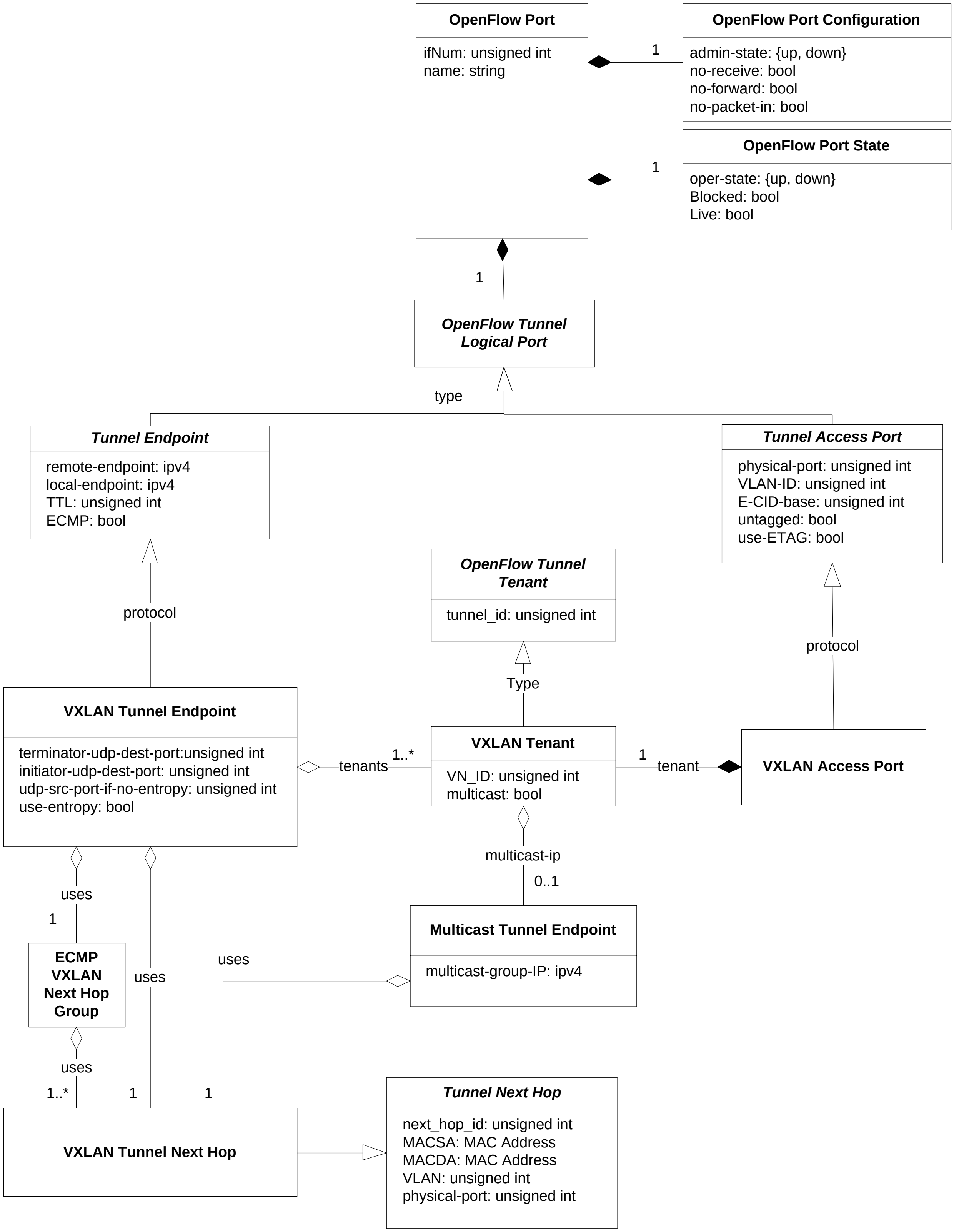

Figure 18: VXLAN Tunnel Configuration UML

OF-DPA provides configuration APIs for VXLAN Tunnel Endpoints, VXLAN Tenants, VXLAN Access Ports, ECMP VXLAN Next Hop Groups, and VXLAN Unicast Tunnel Next Hops. Configuration parameters are listed in Table 77 through Table 80.

Note: OF-DPA derives header fields for multicast tunnel origination packets from related configuration values. For example, if multiple local endpoint IP addresses are configured, OF-DPA will select one.

Table 77: VXLAN Tunnel Endpoint Logical Port Configuration Parameters

Name	Bits	Description
Remote Endpoint	32	IPv4 SIP in termination packets or DIP for unicast origination packets.
Local Endpoint	32	IPv4 DIP for termination packets or SIP for origination packets.
TTL	8	TTL value for use in origination packets.
ECMP	1	Use multipath forwarding for origination packets.
Terminator UDP Dest Port	16	Destination UDP port for recognizing termination VXLAN frames.
Initiator UDP Dest Port	16	Destination UDP port to put in originating VXLAN frames.
UDP Source Port	16	Default source port to use if entropy option is not used.
Use Entropy	1	Insert hash value in place of UDP source port.

Table 78: VXLAN Access Logical Port Configuration Parameters

Name	Bits	Description
Port	32	Local port.
VLAN id	16	VLAN ID to match or use if VLAN tagged.
E-CID	16	E-CID value to match or use if IEEE 802.1BR tagged.
Untagged	1	All traffic on port is for the same tenant.
Use ETAG	1	Use IEEE 802.1BR tagging rather than VLAN ID.

Table 79: VXLAN Tenant Configuration Parameters

Name	Bits	Description
Tunnel id	32	Value to associate with packets for this tenant. Identifies the tenant forwarding domain.
VN_ID	24	Segment identifier in the VXLAN header that identifies this tenant.
Multicast IP	32	Multicast group IP address associated with this tenant.

Table 80: VXLAN Next Hop Configuration Parameters

Name	Bits	Description
Next Hop Id	32	Identifier used to reference next hop objects.
MACSA	64	Underlay source MAC address.
MACDA	64	Underlay destination MAC address.
VLAN id	16	VLAN ID to use if tagged.
Port	32	Egress port for forwarding.

3.5.3.3 Logical Port Counters

OF-DPA supports the counters shown in Table 81 on logical ports.

Table 81: Logical Port Counters

Name	Bits	Description
Received Packets	64	Total packets received.
Transmitted Packets	64	Total packets transmitted.
Received Bytes	64	Total bytes received.
Transmitted Bytes	64	Total bytes transmitted.

Name	Bits	Description
Duration (sec.)	32	Time in seconds since configured.

3.6 Queues

OF-DPA supports eight queues per standard port.

3.6.1 Configuration

OF-DPA queue configuration parameters are listed in Table 82. Queue_Id is always relative to the port to which the queue is attached. Queue_Id values must be a value between zero and seven.

Table 82: OF-DPA Queue Configuration Parameters

Name	Bits	Description
Queue_Id	32	Identifier for this specific queue. Must be a value between 0 and 7.
Port	32	Port to which this queue is attached.
Max Rate	16	Maximum rate in terms of a percentage of the port rate, specified in increments of .1%. A value of 1000 means no maximum rate.
Min Rate	16	Minimum rate in terms of a percentage of the port rate, specified in increments of 0.1%. A value of 1000 means no minimum rate.

3.6.2 Counters

OF-DPA queues counters are as shown in Table 83.

Table 83: OF-DPA Queue Counters

Name	Bits	Description
Transmit Packets	64	Total packets transmitted.
Transmit Bytes	64	Total bytes transmitted.
Duration (sec.)	64	Duration in seconds.

3.7 Vendor Extension Features

OF-DPA provides vendor extensions for source MAC learning and L3 forwarding IN_PORT control. These are intended to be programmed using Experimental Protocol Elements as described in the OpenFlow 1.3.1 specification and are described in this section.

3.7.1 Source MAC Learning

One recognized issue with the OpenFlow 1.3.1 specification is the inadequate support for a flow table that maintains a single database that can be accessed using different lookup keys for different purposes.[7]

The Bridging Flow Table contains MAC forwarding entries and is looked up by MAC_DST and VLAN. An exact match in the table provides the output port in the form of an L2 Interface Group entry. However, if there is no exact match, a wildcard flow entry that matches just the VLAN can provide an L2 Flood Group entry for destination location forwarding (DLF).

Source MAC learning is typically used to discover the MAC-to-port binding for populating the MAC table. A second lookup is done in the same table using the MAC_SRC and VLAN. If there is a match, the output port is compared against the IN_PORT. If there is a miss, an entry for this MAC and VLAN is added to the table along with the port. If there is a match, but the port values are different, it means the end station has moved and the port is updated.[8]

OF-DPA implements optional logic for identifying when a MAC-to-port binding needs to be learned as a vendor extension. This function looks up all packets, regardless of whether they will be processed using the Bridging Flow Table or the Routing Flow Table. If the MAC_SRC and VLAN miss, or if the port has changed, the logic does one of two things, depending on the configuration.

Note: Network Virtualization SDN use cases, especially in data center and enterprise networks, centrally manage L2 forwarding and VLAN tables based on network discovery and do not rely on learning and flooding.

3.7.1.1 Controller Managed Learning

If all learning is to be managed by the OpenFlow controller, OF-DPA will send a PACKET_IN message to the controller with a reason code (no match), special table ID that indicates the learning lookup, and the MAC_SRC, VLAN, and IN_PORT match fields that missed. The controller can then send a flow mod message to the switch to add an appropriate entry in the Bridging Flow Table or update Group Entries. Since a PACKET_OUT is not expected as well, there is no need to buffer the miss packet, which would have already been forwarded normally.

To prevent multiple PACKET_IN learning messages, OF-DPA adds a pending (disabled) entry in the Bridging Flow Table. This entry will be removed after a configured interval if the controller does not come back with a flow mod.

[7] This is addressed to some degree by the table synchronization features in the OpenFlow 1.4 specification.
[8] This is subject to security policy.

For controller managed learning the feature would be configured with destination CONTROLLER.

3.7.1.2 Configuration

The configuration options are listed in Table 84.

Table 84: Source MAC Learning Feature Configuration

Name	Description
Enable	Enable the source MAC learning feature.
Destination	CONTROLLER.
Duration	If the destination is CONTROLLER, indicates the time interval after which the pending entry is removed if the controller does not issue a Flow Mod to keep it.

3.7.2 Group Properties

OF-DPA adds the vendor extension property **ALLOW_IN_PORT** to all group entries. This property applies to the group entry and to any referenced group entries unless the override is explicitly overridden. Since the default, defined in the OpenFlow 1.3.1 specification, is to not forward packets to IN_PORT, this property need not be explicitly specified unless it is desired to override the default behavior.

Only the L3 Interface Group Entries use this feature in OF-DPA.

3.7.3 MTU Check

OF-DPA adds an MTU check with the TTL check described in the OpenFlow 1.3.1 specification. The same error code is used for both TTL and MTU check.

MTU check is required in order to implement an IP router and enable it to send the correct ICMP destination unreachable messages.

3.8 Table Numbering

OF-DPA table number assignments are shown in Table 85.

Table 85: Flow Table Number Assignments

Flow Table Name	Number
Ingress Port Flow Table	0

Flow Table Name	Number
VLAN Flow Table	10
Termination MAC Flow Table	20
Unicast Routing Flow Table	30
Multicast Routing Flow Table	40
Bridging Flow Table	50
Policy ACL Flow Table	60
Source MAC Learning Flow Table	254

APPENDIX A :REFERENCES

[1] OpenFlow 1.0 Specification, https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.0.0.pdf

[2] OpenFlow 1.1 Specification, https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.1.0.pdf

[3] OpenFlow 1.2 Specification, https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.2.pdf

[4] OpenFlow 1.3 Specification, https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.3.0.pdf

[5] OpenFlow 1.3.1 Specification, https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.3.1.pdf

[6] OF-Config 1.1 Specification, https://www.opennetworking.org/images/stories/downloads/of-config/of-config-1.1.pdf

[7] A Framework for Enabling Security Controls in OpenFlow Networks, in Proceedings of the HotSDN Workshop at SIGCOMM, Helsinki, Finland, August 2012, http://www.openflowsec.org/FortNOX_Sigcomm_HotSDN_2012.pdf

[8] Open vSwitch, http://openvswitch.org/

[9] IEEE Standard for Local and metropolitan area networks – Virtual Bridged Local Area Networks – Bridge Port Extension, IEEE Std 802.1BR™-2012

[10] VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks - http://datatracker.ietf.org/doc/draft-mahalingam-dutt-dcops-vxlan