OpenFlow Data Plane Abstraction (OF-DPA™): Abstract Switch Specification

Version 2.01

January 26, 2016

Last Edit: February 9, 2016 4:24 PM

File Name: OF-DPA 2.01 Specification - 26Jan 2016.docx

Broadcom Corporation
3151 Zanker Road
San Jose, CA 95134

Revision History

Revision Number	Date	Change
1.0	15 Nov 2013	Initial release -1.0 draft
2.0	20 Nov 2014	Initial 2.0 draft
2.01	26 Jan 2016	Updates for 2.01.

This document represents an interim release of the specification for OF-DPA 2.01. It is intended for limited external distribution only to solicit feedback and comments. As such, it is subject to change based on feedback received.

Table of Contents

1 Introduction 19

2 OF-DPA Components 20

3 The OF-DPA Abstract Switch 23

3.1 Abstract Switch Overview 24
3.2 Bridging and Routing 25
3.3 Data Center Overlay Tunnels 26
3.4 MPLS-TP Customer Edge Device 28
3.4.1 VPWS 28
3.4.2 MPLS-TP L2 VPN Groups 30
3.5 MPLS Label Edge Router (LER) 32
3.6 MPLS Label Switch Router (LSR) 34
3.7 MPLS Protection Switching 35
3.8 Quality of Service (QoS) 35
3.8.1 QoS Concepts – Traffic Class and Color 36
3.8.2 Meter Table Operation 36
3.8.3 Bridging and Routing QoS 41
3.8.4 Overlay Tunnel QoS 42
3.8.5 MPLS QoS 43
3.9 Operation, Administration, and Maintenance (OAM) 45
3.9.1 OAM Concepts – Maintenance Points 45
3.9.2 Network Protection Apps 47
3.9.3 MPLS-TP OAM 48
3.9.3.1 Ethernet Service OAM 49
3.9.3.2 G.8113.1 OAM for MPLS-TP 52
3.10 Protection Switching 54

4 OF-DPA Object Details 57

4.1 Flow Tables 58
4.1.1 Ingress Port Flow Table 58
4.1.1.1 Flow Entry Types and Match Fields 58
4.1.1.2 Instruction Types 60
4.1.1.3 Actions 60
4.1.1.4 Counters and Flow Expiry 61
4.1.2 Injected OAM Flow Table 61
4.1.2.1 Flow Entry Types and Match Fields 62
4.1.2.2 Instruction Types 62
4.1.2.3 Actions 63
4.1.2.4 Counters and Flow Expiry 64
4.1.3 VLAN Flow Table 64
4.1.3.1 Flow Entry Types and Match Fields 65

4.1.3.2 Instruction Types 68
4.1.3.3 Actions 68
4.1.3.4 Counters and Flow Expiry 69
4.1.4 VLAN 1 Flow Table 70
4.1.4.1 Flow Entry Types and Match Fields 70
4.1.4.2 Instruction Types 71
4.1.4.3 Actions 71
4.1.4.4 Counters and Flow Expiry 72
4.1.5 Ingress Maintenance Point Flow Table 72
4.1.5.1 Flow Entry Types and Match Fields 73
4.1.5.2 Instruction Types 74
4.1.5.3 Actions 74
4.1.5.4 Counters and Flow Expiry 75
4.1.6 MPLS L2 Port Flow Table 75
4.1.6.1 Flow Entry Types and Match Fields 76
4.1.6.2 Instruction Types 76
4.1.6.3 Actions 77
4.1.6.4 Counters and Flow Expiry 77
4.1.7 QoS DSCP and PCP Trust Flow Tables 78
4.1.7.1 Flow Entry Types and Match Fields 78
4.1.7.2 Instruction Types 79
4.1.7.3 Actions 79
4.1.7.4 Counters and Flow Expiry 80
4.1.8 L2 Policer Flow Table 80
4.1.8.1 Flow Entry Types and Match Fields 80
4.1.8.2 Instructions 81
4.1.8.3 Action List Actions 81
4.1.8.4 Counters and Flow Expiry 81
4.1.9 L2 Policer Actions Flow Table 82
4.1.9.1 Flow Entry Types and Match Fields 82
4.1.9.2 Instructions 83
4.1.9.3 Actions 83
4.1.9.4 Counters and Flow Expiry 84
4.1.10 Termination MAC Flow Table 84
4.1.10.1 Flow Entry Types and Match Fields 84
4.1.10.2 Instruction Types 86
4.1.10.3 Counters and Flow Expiry 86
4.1.11 L3 Type Flow Table 87
4.1.11.1 Flow Entry Types and Match Fields 87
4.1.11.2 Instruction Types 88
4.1.11.3 Actions 88
4.1.11.4 Counters and Flow Expiry 88
4.1.12 MPLS Flow Tables 88
4.1.12.1 Flow Entry Types and Match Fields 89
4.1.12.2 Instruction Types 93
4.1.12.3 Actions 93
4.1.12.4 Counters and Flow Expiry 96
4.1.13 MPLS L3 Type Flow Table 96
4.1.13.1 Flow Entry Types and Match Fields 96
4.1.13.2 Instruction Types 97
4.1.13.3 Actions 98
4.1.13.4 Counters and Flow Expiry 98
4.1.14 MPLS Label Trust Flow Table 99
4.1.14.1 Flow Entry Types and Match Fields 99

4.1.14.2 Instruction Types 99
4.1.14.3 Actions 99
4.1.14.4 Counters and Flow Expiry 100
4.1.15 MPLS Type Flow Table 100
4.1.15.1 Flow Entry Types and Match Fields 100
4.1.15.2 Instructions 101
4.1.15.3 Counters and Flow Expiry 101
4.1.16 Bridging Flow Table 101
4.1.16.1 Flow Entry Types and Match Fields 102
4.1.16.2 Instruction Types 103
4.1.16.3 Action Set 104
4.1.16.4 Counters and Flow Expiration 104
4.1.17 Unicast Routing Flow Table 105
4.1.17.1 Flow Entry Types and Match Fields 105
4.1.17.2 Instruction Types 106
4.1.17.3 Action Set 107
4.1.17.4 Counters and Flow Expiration 107
4.1.18 Multicast Routing Flow Table 108
4.1.18.1 Flow Entry Types and Match Fields 108
4.1.18.2 Instruction Types 109
4.1.18.3 Action Set 110
4.1.18.4 Counters and Flow Expiration 110
4.1.19 MPLS-TP Maintenance Point Flow Table 111
4.1.19.1 Flow Entry Types and Match Fields 111
4.1.19.2 Instruction Types 112
4.1.19.3 Actions 112
4.1.19.4 Counters and Flow Expiry 112
4.1.20 Policy ACL Flow Table 113
4.1.20.1 Flow Entry Types and Match Fields 113
4.1.20.2 Instruction Types 118
4.1.20.3 Action List Actions 118
4.1.20.4 Action Set Actions 118
4.1.20.5 Counters and Flow Expiration 120
4.1.21 Color Based Actions Flow Table 120
4.1.21.1 Flow Entry Types and Match Fields 121
4.1.21.2 Instruction Types 121
4.1.21.3 Actions 122
4.1.21.4 Counters and Flow Expiry 122

4.2 Egress Flow Tables 123
4.2.1 Egress VLAN Flow Table 123
4.2.1.1 Flow Entry Types and Match Fields 123
4.2.1.2 Instruction Types 124
4.2.1.3 Actions 124
4.2.1.4 Counters and Flow Expiry 125
4.2.2 Egress VLAN 1 Flow Table 125
4.2.2.1 Flow Entry Types and Match Fields 125
4.2.2.2 Instruction Types 126
4.2.2.3 Actions 127
4.2.2.4 Counters and Flow Expiry 127
4.2.3 Egress Maintenance Point Flow Table 127
4.2.3.1 Flow Entry Types and Match Fields 128
4.2.3.2 Instruction Types 129
4.2.3.3 Actions 129

4.2.3.4 Counters and Flow Expiry 130
4.2.4 Egress DSCP PCP Remark Flow Table 130
4.2.4.1 Flow Entry Types and Match Fields 130
4.2.4.2 Instruction Types 131
4.2.4.3 Actions 131
4.2.4.4 Counters and Flow Expiry 132
4.2.5 Egress TPID Flow Table 132
4.2.5.1 Flow Entry Types and Match Fields 132
4.2.5.2 Instruction Types 133
4.2.5.3 Actions 133
4.2.5.4 Counters and Flow Expiry 134
4.3 Group Table 134
4.3.1 OF-DPA L2 Interface Group Entries 135
4.3.1.1 Naming Convention 136
4.3.1.2 Action Buckets 136
4.3.1.3 Counters 136
4.3.2 OF-DPA L2 Unfiltered Interface Group Entries 137
4.3.2.1 Naming Convention 137
4.3.2.2 Action Buckets 137
4.3.2.3 Counters 138
4.3.3 OF-DPA L2 Rewrite Group Entries 138
4.3.3.1 Naming Convention 138
4.3.3.2 Action Buckets 139
4.3.3.3 Counters 139
4.3.4 OF-DPA L3 Unicast Group Entries 139
4.3.4.1 Naming Convention 140
4.3.4.2 Action Buckets 140
4.3.4.3 Counters 140
4.3.5 OF-DPA L2 Multicast Group Entries 141
4.3.5.1 Naming Convention 141
4.3.5.2 Action Buckets 141
4.3.5.3 Counters 142
4.3.6 OF-DPA L2 Flood Group Entries 142
4.3.6.1 Naming Convention 142
4.3.6.2 Action Buckets 143
4.3.6.3 Counters 143
4.3.7 OF-DPA L3 Interface Group Entries 143
4.3.7.1 Naming Convention 144
4.3.7.2 Action Buckets 144
4.3.7.3 Counters 144
4.3.8 OF-DPA L3 Multicast Group Entries 145
4.3.8.1 Naming Convention 145
4.3.8.2 Action Buckets 145
4.3.8.3 Counters 146
4.3.9 OF-DPA L3 ECMP Group Entries 146
4.3.9.1 Naming Convention 147
4.3.9.2 Action Buckets 147
4.3.9.3 Counters 147
4.3.10 OF-DPA L2 Overlay Group Entries 148
4.3.10.1 OF-DPA L2 Overlay Group Sub-Types 148
4.3.10.2 Naming Convention 148
4.3.10.3 Action Buckets 149
4.3.10.4 Counters 149

4.3.11 OF-DPA MPLS Interface Group Entry 150
4.3.11.1 Naming Convention 150
4.3.11.2 Action Buckets 150
4.3.11.3 Counters 150
4.3.12 OF-DPA MPLS Label Group Entries 151
4.3.12.1 Naming Convention 151
4.3.12.2 MPLS VPN Label Action Buckets 151
4.3.12.3 MPLS Tunnel Label 1 Action Buckets 153
4.3.12.4 MPLS Tunnel Label 2 Action Buckets 154
4.3.12.5 MPLS Swap Label Action Buckets 155
4.3.12.6 Counters 156
4.3.13 OF-DPA MPLS Fast Failover Group Entry 156
4.3.13.1 Naming Convention 156
4.3.13.2 Action Buckets 157
4.3.13.3 Counters 157
4.3.14 OF-DPA MPLS ECMP Group Entry 157
4.3.14.1 Naming Convention 157
4.3.14.2 Action Buckets 158
4.3.14.3 Counters 158

4.4 Meters 159
4.4.1 Meter Table Entries 159
4.4.2 Meter Bands 160

5 Configuration 162

5.1 Ports 162
5.1.1 Physical Ports 163
5.1.1.1 Features 163
5.1.1.2 Counters 166
5.1.2 Reserved Ports 167
5.1.3 Logical Ports 168
5.1.3.1 Overlay Tunnels 168
5.1.3.2 VXLAN Tunnel Logical Port Configuration 169
5.1.3.3 OAM Protection Liveness Logical Ports 173

5.2 Queues 173
5.2.1 Configuration 174
5.2.2 Counters 174

5.3 OAM Message Processing 174
5.3.1 MPLS-TP Ethernet OAM Configuration 175
5.3.1.1 Fault Management 175
5.3.1.2 Performance Monitoring 183
5.3.2 MPLS-TP G.8113.1 OAM Configuration 190
5.3.2.1 Fault Management 190
5.3.2.2 Performance Monitoring 196

5.4 Protection 203
5.4.1 MPLS-TP Linear Protection 203

6 Vendor Extension Features 208

6.1 Source MAC Learning 208
6.1.1 Controller Managed Learning 209
6.1.1.1 Configuration 209

6.2 Additional Group Properties 209
6.3 MTU Check 209
6.4 Table Numbering 210
6.4.1 Egress Tables 212
6.5 Experimenter Features 212
6.5.1 Action Tables 213
6.5.1.1 MPLS Label Remark Action Tables 213
6.5.1.2 Color Set Meter Band 216
6.5.2 Actions 216
6.5.3 Match Fields 219
6.5.4 Extension Fields 223

APPENDIX A : References 225

List of Figures

Figure 1 OF-DPA Component Layering 20
Figure 2 Abstract Switch Objects Used for Bridging and Routing 25
Figure 3 Abstract Switch Objects Used for Overlay Tunnels 27
Figure 4 Abstract Switch Objects for MPLS-TP Initiation (VPWS) 28
Figure 5 Abstract Switch Objects for MPLS-TP Termination (VPWS) 29
Figure 6 MPLS L2 VPN Groups 31
Figure 7 Abstract Switch Objects for MPLS L3 VPN Initiation 32
Figure 8 Abstract Switch Objects for MPLS L3 VPN Termination 33
Figure 9 MPLS L3 Label Groups 33
Figure 10 Abstract Switch Objects Used for an MPLS LSR 34
Figure 11 MPLS LSR Groups 34
Figure 12 MPLS-TP 1:1 Protection 35
Figure 13 Token Bucket Operation 37
Figure 14 TrTCM Meter Operation 37
Figure 15 SrTCM Meter Operation 39
Figure 16 Modified TrTCM Meter Operation 40
Figure 17 Bridging and Routing Abstract Switch Showing QoS Objects 42
Figure 18 Overlay Tunnel Abstract Switch Showing QoS Objects 43
Figure 19 MPLS-TP VPWS Initiation with QoS Objects 43
Figure 20 MPLS-TP VPWS Tunnel Termination with QoS Objects 44
Figure 21 MPLS L3 VPN Initiation with QoS Objects 44
Figure 22 MPLS L3 VPN Termination with QoS Objects 45
Figure 23 MPLS LSR with QoS Objects 45
Figure 24 OAM MEP and MIP Examples 46
Figure 25 OAM and Protection Overview 47
Figure 26 MPLS-TP Service OAM Examples 48
Figure 27 Ethernet Service OAM over MPLS-TP Packet Format 49
Figure 28 MPLS-TP Initiation - Ethernet Service OAM over MPLS-TP Data Frames 50
Figure 29 MPLS-TP Termination - Ethernet Service OAM over MPLS-TP Data Frames 50
Figure 30 MPLS-TP Initiation - Ethernet over MPLS-TP OAM PDU 51
Figure 31 MPLS-TP Termination - Ethernet over MPLS-TP OAM PDU 52
Figure 32 OAM MPLS-TP G.8113.1 Packet Formats 52
Figure 33 MPLS-TP Initiation - G.8113.1 OAM Data Frame 53
Figure 34 MPLS-TP Termination - G.8113.1 OAM Data Frame 53
Figure 35 MPLS-TP - G.8113.1 OAM PDU 54
Figure 36 MPLS-TP - G.8113.1 OAM PDU Termination 54
Figure 37 Protection Switching Process 55
Figure 38 OF-DPA L3 ECMP Group Entry Usage 146
Figure 39 Meter Entry Example (TrTCM) 159
Figure 40 Port Properties Configuration 163
Figure 41 OpenFlow Feature Sub-Classes 164
Figure 42 Tunnel Logical Port Configuration 169
Figure 43 VXLAN Tunnel Configuration 171
Figure 44 OAM Ethernet Fault Management Configuration 175
Figure 45 OAM Ethernet Performance Monitoring Configuration 184
Figure 46 OAM G.8113.1 Fault Management Configuration 191
Figure 47 OAM G.8113.1 Performance Monitoring Configuration 197
Figure 48 MPLS-TP Linear Protection Configuration 204

List of Tables

Table 1 OpenFlow 1.3.4 Features Required by OF-DPA v2.0....21
Table 2 TrTCM Color-Aware Operation....38
Table 3 SrTCM Color-Aware Operation....39
Table 4 Modified TrTCM Color-Aware Operation....40
Table 5 Ingress Port Flow Table Entry Types....58
Table 6 Ingress Port Flow Table Match Fields....59
Table 7 Ingress Port Flow Table Instructions....60
Table 8 Ingress Port Flow Table QoS Next Table by Entry Type....60
Table 9 Ingress Port Flow Table Action List....61
Table 10 Ingress Port Flow Table Counters....61
Table 11 Ingress Port Flow Table Expiry....61
Table 12 Injected OAM Flow Table Entry Types....62
Table 13 Injected OAM Flow Table Match Fields....62
Table 14 Injected OAM Flow Table Instructions....63
Table 15 Injected OAM Flow Table Action List....63
Table 16 Injected OAM Flow Table Counters....64
Table 17 Injected OAM Flow Table Expiry....64
Table 18 VLAN Flow Table Flow Entry Types....65
Table 19 VLAN Flow Table Match Fields....67
Table 20 VLAN Flow Table Instructions....68
Table 21 VLAN Flow Table Action List Actions....68
Table 22 VLAN Flow Table Counters....69
Table 23 VLAN Flow Table Expiry....70
Table 24 VLAN 1 Flow Table Flow Entry Types....70
Table 25 VLAN 1 Flow Table Match Fields....70
Table 26 VLAN 1 Flow Table Instructions....71
Table 27 VLAN 1 Flow Table Action List Actions....71
Table 28 VLAN 1 Flow Table Counters....72
Table 29 VLAN 1 Flow Table Expiry....72
Table 30 Ingress Maintenance Point Flow Table Entry Types....73
Table 31 Ingress Maintenance Point Flow Table Match Fields....73
Table 32 Ingress Maintenance Point Flow Table Instructions....74
Table 33 Ingress Maintenance Point Flow Table Action List....74
Table 34 Ingress Maintenance Point Flow Table Action Set....74
Table 35 Ingress Maintenance Point Flow Table Counters....75
Table 36 Ingress Maintenance Point Flow Table Expiry....75
Table 37 MPLS L2 Port Metadata Naming Convention....75
Table 38 MPLS L2 Port Flow Table Flow Entry Types....76
Table 39 MPLS L2 Port Flow Table Match Fields....76
Table 40 MPLS L2 Port Flow Table Instructions....76
Table 41 MPLS L2 Port Flow Table Action List....77
Table 42 MPLS L2 Port Flow Table Counters....77
Table 43 MPLS L2 Port Flow Table Expiry....77
Table 44 QoS DSCP Trust Flow Table Match Fields....78
Table 45 QoS PCP Trust Flow Table Match Fields....78
Table 46 QoS DSCP and PCP Trust Flow Table Instructions....79
Table 47 QoS DSCP and PCP Trust Flow Table Action List....79
Table 48 QoS DSCP and PCP Trust Flow Table Counters....80
Table 49 QoS DSCP and PCP Trust Flow Table Expiry....80
Table 50 L2 Policer Flow Table Entry Types....80

Table 51 L2 Policer Flow Table Match Fields 81
Table 52 L2 Policer Flow Table Instructions 81
Table 53 L2 Policer Flow Table Actions 81
Table 54 L2 Policer Flow Table Counters 81
Table 55 L2 Policer Flow Table Expiry 82
Table 56 L2 Policer Actions Flow Table Entry Types 82
Table 57 L2 Policer Actions Flow Table Match Fields 82
Table 58 L2 Policer Actions Flow Table Instructions 83
Table 59 L2 Policer Actions Flow Table Action List Actions 83
Table 60 L2 Policer Actions Flow Table Action Set Actions 83
Table 61 L2 Policer Actions Flow Table Counters 84
Table 62 L2 Policer Actions Flow Table Expiry 84
Table 63 Termination MAC Flow Table Entry Types 85
Table 64 Termination MAC Flow Table Match Fields 85
Table 65 Termination MAC Flow Table Instruction Set 86
Table 66 Termination MAC Flow Table Counters 86
Table 67 Termination MAC Flow Table Expiry 87
Table 68 L3 Type Flow Table Entry Types 87
Table 69 L3 Type Flow Table Match Fields 87
Table 70 L3 Type Flow Table Counters 88
Table 71 MPLS Flow Table 0 Flow Entry Types 89
Table 72 MPLS Flow Table 1 and 2 Flow Table Entry Types 90
Table 73 MPLS Flow Table Match Fields 92
Table 74 MPLS Flow Table Instructions 93
Table 75 MPLS Flow Table Action List Actions 94
Table 76 MPLS Flow Table Action Set Actions 95
Table 77 MPLS Flow Table Counters 96
Table 78 MPLS Flow Table Expiry 96
Table 79 MPLS L3 Type Flow Table Entry Types 96
Table 80 MPLS L3 Type Flow Table Match Fields 97
Table 81 MPLS L3 Type Flow Table Instructions 97
Table 82 MPLS L3 Type Flow Table Action List 98
Table 83 MPLS L3 Type Flow Table Counters 98
Table 84 MPLS L3 Type Flow Table Expiry 99
Table 85 MPLS Label Trust Flow Table Match Fields 99
Table 86 MPLS Label Trust Flow Table Instructions 99
Table 87 MPLS Label Trust Flow Table Action List 100
Table 88 MPLS Label Trust Flow Table Counters 100
Table 89 MPLS Type Flow Table Entry Types 100
Table 90 MPLS Type Flow Table Match Fields 101
Table 91 MPLS Type Flow Table Instructions 101
Table 92 Tunnel Id Metadata Naming Convention 101
Table 93 Bridging Flow Table Flow Entry Types 102
Table 94 Bridging Flow Table Match Fields 103
Table 95 Bridging Flow Table Instructions 103
Table 96 Bridging Flow Table Action Set 104
Table 97 Bridging Flow Table Counters 104
Table 98 Bridging Flow Table Expiry 105
Table 99 Unicast Routing Flow Table Entry Types 105
Table 100 Unicast Routing Flow Table IPv4 Header Match Fields 106
Table 101 Unicast Routing Flow Table IPv6 Header Match Fields 106
Table 102 Unicast Routing Flow Table Instructions 106
Table 103 Unicast Routing Flow Table Action Set 107
Table 104 Unicast Routing Flow Table Counters 107

Table 105 Unicast Routing Flow Table Expiry 108
Table 106 Multicast Routing Flow Table Entry Types 108
Table 107 Multicast Routing Flow Table IPv4 Match Fields 108
Table 108 Multicast Routing Flow Table IPv6 Match Fields 109
Table 109 Multicast Routing Flow Table Instructions 110
Table 110 Multicast Routing Flow Table Action Set 110
Table 111 Multicast Routing Flow Table Counters 110
Table 112 Multicast Routing Flow Table Expiry 111
Table 113 MPLS-TP Maintenance Point Flow Table Entry Types 111
Table 114 MPLS-TP Maintenance Point Flow Table Match Fields 111
Table 115 MPLS-TP Maintenance Point Flow Table Instructions 112
Table 116 MPLS-TP Maintenance Point Flow Table Actions 112
Table 117 MPLS-TP Maintenance Point Flow Table Counters 112
Table 118 MPLS-TP Maintenance Point Flow Table Expiry 113
Table 119 Policy ACL Flow Table Flow Entry Types 113
Table 120 Policy ACL Flow Table IPv4 Match Fields 114
Table 121 Policy ACL Flow Table IPv6 Match Fields 116
Table 122 Policy ACL Flow Table Instructions 118
Table 123 Policy ACL Flow Table Action List Actions 118
Table 124 Policy ACL Flow Table VLAN Flow Entry Action Set 119
Table 125 Policy ACL Flow Table Tunnel Flow Entry Action Set 119
Table 126 Policy ACL Flow Table Counters 120
Table 127 Policy ACL Flow Table Expiry 120
Table 128 Color Based Actions Flow Table Entry Types 121
Table 129 Color Based Actions Flow Table Match Fields 121
Table 130 Color Based Actions Flow Table Instructions 121
Table 131 Color Based Actions Flow Table Action List Actions 122
Table 132 Color Based Actions Flow Table Action Set Actions 122
Table 133 Color Based Actions Flow Table Counters 122
Table 134 Color Based Actions Flow Table Expiry 122
Table 135 Egress VLAN Flow Table Flow Entry Types 123
Table 136 Egress VLAN Flow Table Match Fields 124
Table 137 Egress VLAN Flow Table Instructions 124
Table 138 Egress VLAN Flow Table Action List 124
Table 139 Egress VLAN Flow Table Counters 125
Table 140 Egress VLAN Flow Table Expiry 125
Table 141 Egress VLAN 1 Flow Table Flow Entry Types 126
Table 142 Egress VLAN 1 Flow Table Match Fields 126
Table 143 Egress VLAN 1 Flow Table Instructions 126
Table 144 Egress VLAN 1 Flow Table Action List 127
Table 145 Egress VLAN 1 Flow Table Counters 127
Table 146 Egress VLAN 1 Flow Table Expiry 127
Table 147 Egress Maintenance Point Flow Table Entry Types 128
Table 148 Egress Maintenance Point Flow Table Match Fields 128
Table 149 Egress Maintenance Point Flow Table Instructions 129
Table 150 Egress Maintenance Point Flow Table Action Set Actions 129
Table 151 Egress Maintenance Point Flow Table Action List Actions 129
Table 152 Egress Maintenance Point Flow Table Counters 130
Table 153 Egress Maintenance Point Flow Table Expiry 130
Table 154 Egress DSCP PCP Remark Flow Table Entry Types 130
Table 155 Egress DSCP PCP Remark Flow Table Match Fields 131
Table 156 Egress DSCP PCP Remark Flow Table Instructions 131
Table 157 Egress DSCP PCP Remark Flow Table Action List Actions 131
Table 158 Egress DSCP PCP Remark Flow Table Counters 132

Table 159 Egress DSCP PCP Remark Flow Table Expiry 132
Table 160 Egress TPID Flow Table Entry Types 132
Table 161 Egress TPID Flow Table Match Fields 133
Table 162 Egress TPID Flow Table Instructions 133
Table 163 Egress TPIPD Flow Table Action List Actions 133
Table 164 Egress TPID Flow Table Counters 134
Table 165 Egress TPID Flow Table Expiry 134
Table 166 OF-DPA Group Table Entry Identifier Naming Convention 135
Table 167 OF-DPA L2 Interface Group Entry Type Naming Convention 136
Table 168 OF-DPA L2 Interface Group Entry Bucket Actions 136
Table 169 OF-DPA L2 Interface Group Entry Counters 136
Table 170 OF-DPA L2 Unfiltered Interface Group Naming Convention 137
Table 171 OF-DPA L2 Unfiltered Interface Group Bucket Actions 137
Table 172 OF-DPA L2 Unfiltered Interface Group Entry Counters 138
Table 173 OF-DPA L2 Rewrite Group Entry Type Naming Convention 138
Table 174 OF-DPA L2 Rewrite Group Entry Bucket Actions 139
Table 175 OF-DPA L2 Rewrite Group Entry Counters 139
Table 176 OF-DPA L3 Unicast Group Entry Naming Conventioin 140
Table 177 OF-DPA L3 Unicast Bucket Actions 140
Table 178 OF-DPA L3 Unicast Group Entry Counters 141
Table 179 OF-DPA L2 Multicast Group Entry Type Naming Convention 141
Table 180 OF-DPA L2 Multicast Bucket Actions 141
Table 181 OF-DPA L2 Multicast Group Entry Counters 142
Table 182 OF-DPA L2 Flood Group Entry Naming Convention 142
Table 183 OF-DPA L2 Flood Bucket Actions 143
Table 184 OF-DPA L2 Flood Group Entry Counters 143
Table 185 OF-DPA L3 Interface Group Entry Type Naming Convention 144
Table 186 OF-DPA L3 Interface Group Entry Bucket Actions 144
Table 187 OF-DPA L3 Interface Group Entry Counters 145
Table 188 OF-DPA L3 Multicast Group Entry Naming Convention 145
Table 189 OF-DPA L3 Multicast Bucket Actions 146
Table 190 OF-DPA L3 Multicast Group Entry Counters 146
Table 191 OF-DPA L3 ECMP Group Entry Naming Convention 147
Table 192 OF-DPA L3 ECMP Group Entry Bucket Actions 147
Table 193 OF-DPA L3 ECMP Group Entry Counters 147
Table 194 OF-DPA L2 Overlay Group Entry Naming Convention 148
Table 195 OF-DPA L2 Overlay Group Sub-Type Entry Bucket Actions 149
Table 196 OF-DPA L2 Overlay Group Sub-Type Entry Counters 149
Table 197 OF-DPA MPLS Interface Group Entry Naming Convention 150
Table 198 OF-DPA MPLS Interface Group Entry Bucket Actions 150
Table 199 OF-DPA MPLS Interface Group Type Entry Counters 151
Table 200 OF-DPA MPLS Label Group Entry Naming Convention 151
Table 201 OF-DPA MPLS VPN Label Group Bucket Actions 152
Table 202 OF-DPA MPLS Tunnel Label 1 Group Bucket Actions 153
Table 203 OF-DPA MPLS Tunnel Label 2 Actions 154
Table 204 OF-DPA MPLS Swap Label Actions 155
Table 205 OF-DPA MPLS Label Group Sub-Type Entry Counters 156
Table 206 OF-DPA MPLS Fast Failover Group Entry Naming Convention 156
Table 207 OF-DPA MPLS Fast Failover Group Entry Bucket Actions 157
Table 208 OF-DPA MPLS Fast Failover Group Entry Counters 157
Table 209 OF-DPA MPLS ECMP Group Entry Naming Convention 158
Table 210 OF-DPA MPLS ECMP Group Entry Bucket Actions 158
Table 211 OF-DPA MPLS ECMP Group Entry Counters 158
Table 212 OF-DPA Meter Table Entry Parameters 160

Table 213 Meter Entry Counters 160
Table 214 Meter Band Configuration Parameters 161
Table 215 Meter Band Counters 161
Table 216 Port Type Numbering Conventions 162
Table 217 OF-DPA Port Features 164
Table 218 Port Features Bitmap 165
Table 219 OF-DPA Physical Port Counters 166
Table 220 OF-DPA Reserved Ports 167
Table 221 VXLAN Tunnel Endpoint Logical Port Configuration Parameters 172
Table 222 VXLAN Access Logical Port Configuration Parameters 172
Table 223 VXLAN Tenant Configuration Parameters 173
Table 224 VXLAN Next Hop Configuration Parameters 173
Table 225 OF-DPA Queue Configuration Parameters 174
Table 226 OF-DPA Queue Counters 174
Table 227 Ethernet OAM MEG 176
Table 228 Ethernet OAM Local MIP 176
Table 229 Ethernet OAM Local MEP 176
Table 230 Ethernet OAM Remote MEP 177
Table 231 Ethernet OAM Remote MIP 177
Table 232 Ethernet OAM CCM 178
Table 233 Ethernet OAM MEP CCM Database 178
Table 234 Ethernet OAM Loopback 179
Table 235 Ethernet OAM LCK 180
Table 236 Ethernet OAM AIS 181
Table 237 Ethernet OAM CSF 181
Table 238 Ethernet OAM Link Trace 181
Table 239 Ethernet OAM Link Trace Reply 182
Table 240 Ethernet OAM Proactive LM 184
Table 241 Ethernet OAM Proactive LM PM 15min/24h Counters 185
Table 242 Ethernet OAM On-Demand LM 185
Table 243 Ethernet OAM On Demand LM Snapshot 186
Table 244 Ethernet OAM Proactive DM 187
Table 245 Ethernet OAM Proactive DM PM 15min/24h Counters 188
Table 246 Ethernet OAM On-Demand DM 189
Table 247 Ethernet OAM On Demand DM Snapshot 189
Table 248 G.8113.1 OAM MEG 191
Table 249 G.8113.1 OAM Local MEP 192
Table 250 G.8113.1 OAM Remote MEP 192
Table 251 G.8113.1 OAM Remote MIP 192
Table 252 G.8113.1 OAM Local MIP 193
Table 253 G.8113.1 OAM CCM 193
Table 254 G.8113.1 OAM MEP CCM Database 193
Table 255 G.8113.1 OAM Loopback 194
Table 256 G.8113.1 OAM LCK 195
Table 257 G.8113.1 OAM AIS 195
Table 258 G.8113.1 OAM CSF 196
Table 259 G.8113.1 OAM Proactive LM 197
Table 260 G.8113.1 OAM Proactive LM PM 15min/24h Counters 198
Table 261 G.8113.1 OAM On-Demand LM 198
Table 262 G.8113.1 OAM On Demand LM Snapshot 199
Table 263 G.8113.1 OAM Proactive DM 200
Table 264 G.8113.1 OAM Proactive DM PM 15min/24h Counters 201
Table 265 G.8113.1 OAM On-Demand DM 202
Table 266 G.8113.1 OAM On Demand DM Snapshot 202

Table 267 MLP Protection Group 204
Table 268 MLP Protection Group Status..... 205
Table 269 MLP Config 206
Table 270 MLP Head End Config..... 206
Table 271 MLP Tail End Config..... 207
Table 272 Source MAC Learning Feature Configuration..... 209
Table 273 Flow Table Number Assignments..... 210
Table 274 MPLS Label Remark Table Entry Fields..... 213
Table 275 OF-DPA Experimenter Actions 216
Table 276 OF-DPA Experimenter Match Fields..... 219

Terminology

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 1.3.4 specification
Flow Entry		Entry in an OpenFlow flow table with its match fields and instructions.
Hybrid Switch		Switch that has OpenFlow programmability in addition to legacy control plane features. There is no standard hybrid switch model although different approaches have been identified, ranging from integrated to completely independent pipelines.
Group Table		The OpenFlow group table, consisting of group table entries.
Meter Table		The OpenFlow meter table, consisting of meter table entries.
OpenFlow Logical Switch	OFLS	A set of OpenFlow resources that can be managed as a single entity, includes a datapath and its control channel.
Open Networking Foundation	ONF	Open Networking Foundation (ONF) is a user-driven organization dedicated to the promotion and adoption of SDN through open standards development, primarily OpenFlow.
Software Defined Networking	SDN	The principles of SDN as defined by ONF are: separation of control and forwarding functions; logically centralized intelligence; programmable; open-standards based; and vendor neutral programming interfaces [1].
Table Type Pattern	TTP	Formal description of an OpenFlow 1.3.4 abstract switch in terms of programmable pipeline objects.
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 the OpenFlow protocol as an OpenFlow Logical Switch, and in so doing 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 model is based on the long term stable OpenFlow 1.3.4 specification [8] and utilizes its provisions to provide access to multiple tables implemented in Broadcom switch Application-Specific Integrated Circuits (ASICs). It is formally defined as a Table Type Pattern (TTP) using the notation specified in the ONF OpenFlow Table Type Patterns specification [12]. 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 2.0. OF-DPA 2.0 incorporates and adds functionality to OF-DPA version 1.0 [13]. It is published openly and meant to be used alongside the *OpenFlow Data Plane Abstraction (OF-DPA) API Guide and Reference Manual*[1] for developing OpenFlow 1.3.4 agents and controllers. While the specification is deemed complete for features supported in OF-DPA 2.0, 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 1.3.4 with extensions packages [9][11], and the goals of related Software Defined Networking (SDN) technologies.

[1] Available as an HTML document in the OF-DPA v2.0 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 kit (SDK) that, in turn, provides the driver for configuring, programming, and controlling the Broadcom switch ASICs.

The OF-DPA API, as defined in the *OpenFlow Data Plane Abstraction (OF-DPA) API Guide and Reference Manual*, presents a specialized hardware abstraction layer (HAL) that allows programming 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. In addition, an OpenFlow Controller is required to field an OpenFlow network deployment using OF-DPA enabled switches. Figure 1 illustrates the relationship of OF-DPA with the other OpenFlow system components.

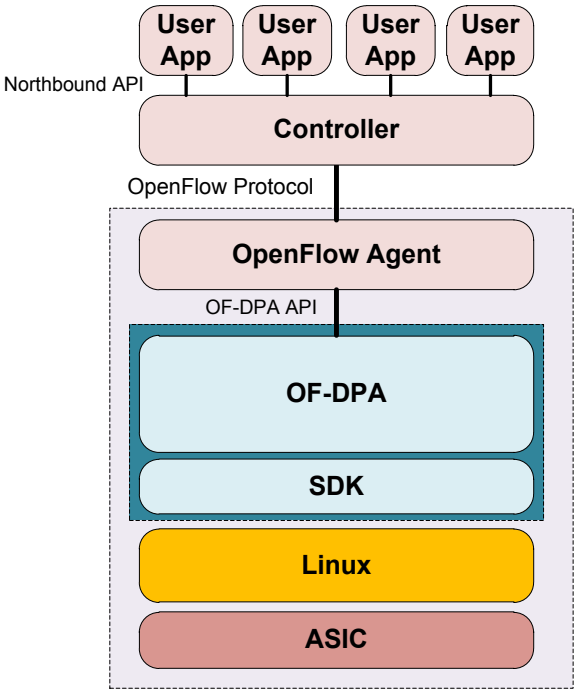

Figure 1 OF-DPA Component Layering

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.

In Figure 1, the OpenFlow Controller communicates with an OpenFlow switch using OpenFlow as the Southbound API. 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 1.3.4 features defined by the Open Networking Foundation (ONF) specification and required by OF-DPA. These required features are listed in Table 1[2]. In addition to commercial products, there are a number of readily available open source controllers that support OpenFlow 1.3.4 and later, including Ryu [23] and OpenDaylight [24].

Table 1 OpenFlow 1.3.4 Features Required by OF-DPA v2.0

Feature	Description
Pipeline Match Fields	Metadata fields that accompany the packet during pipeline processing but are not parsed from packet headers.
Experimenter Protocol Extensions	The Experimenter facility provides a standard way to extend the OpenFlow protocol to support additional functionality. OF-DPA 2.0 defines new Experimenter symmetric messages, multipart messages, flow table match fields, actions, and Meter bands.
Select and Fast Failover Group Types	These are optional types in OpenFlow 1.3.4 but required.
LOCAL Reserved Port	This is optional in OpenFlow 1.3.4 but required.
Packet Registers	Defined in the OpenFlow Extensions Package (EXT-244).

The OF-DPA API represents hardware objects to the agent in terms of objects such as flow tables, group table entries, queues, and ports that can be programmed using the protocol described in the OpenFlow 1.3.4 specification. The OF-DPA Abstract Switch provides instances of many of the standard OpenFlow

[2] OF-DPA also has configuration APIs that require deployments to define a configuration approach.

Logical switch (OFLS) objects defined in the OpenFlow 1.3.4 specification. As OF-DPA maintains the state that maps OpenFlow to the hardware, an agent is expected to do a relatively straightforward translation of OpenFlow messages into OF-DPA API calls and vice-versa, while maintaining a minimal amount of state.

In addition to standard objects, OF-DPA 2.0 defines new ancillary table objects needed to support use cases such as MPLS-TP, OAM, and QoS. These enhance the standard OFLS with features such as per-flow loss management counters and per-flow packet dropping. As with the OpenFlow 1.3 extensions [9][11], they are managed using new protocol message types, but invoked in action lists or action sets using new actions. To support certain OAM functions, OF-DPA 2.0 makes use of the Egress Tables planned for OpenFlow 1.5. Egress tables permit match action processing after groups are applied and the output port is decided. All OF-DPA 2.0 extensions are programmable using existing OpenFlow 1.3 protocol facilities such as experimenter message types[3]. Refer to Section 6 "Vendor Extension Features" on page 208 for detailed descriptions of these extensions.

The OF-DPA 2.0 Abstract Switch is formally defined using the Table Type Pattern (TTP) framework [12] developed by the ONF Forwarding Abstractions Working Group. A TTP is an OpenFlow abstract switch model that describes the forwarding behaviors that can be programmed by a controller. In particular, the TTP framework permits the application writer to express switch requirements, and a controller and switch to agree on supported features at run time.

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 with action lists and/or action sets for packet editing and forwarding. Most packet forwarding uses OpenFlow group entries. An application can interrogate the 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 provides an overview of the OF-DPA Abstract Switch using diagrams to illustrate the objects and the relationships between them for particular packet flows. For a detailed description of the objects, flows, and relationships consult the TTP[4]. For details of the OF-DPA API, consult the *OpenFlow Data Plane Abstraction (OF-DPA) API Guide and Reference Manual*.

[3] Some of these are forward-looking in that they anticipate features proposed for OpenFlow 1.5.
[4] The TTP is included as a text document in the OF-DPA v2.0 software release package.

3 THE OF-DPA ABSTRACT SWITCH

The OF-DPA Abstract Switch is a specialization of the OpenFlow 1.3.4 OFLS with some extensions. This section describes the OF-DPA Abstract Switch in terms of OpenFlow abstract objects as visible to the OpenFlow controller.

The OF-DPA Abstract Switch 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 adaptation layer provides support for OpenFlow specific state, for example, statistics counters. It also maps OpenFlow objects to hardware and manages hardware resources.

Supporting OpenFlow in switch hardware involves some tradeoffs. As has been noted elsewhere, the generality promised by OpenFlow can come at a cost of latency, as well as cost and power inefficiencies. In addition, to effectively use this generality a specific multi-table pipeline needs to be designed and configured first. The OF-DPA Abstract Switch may be viewed as coming preconfigured and optimized to support single pass, full bandwidth packet processing performance that makes efficient use of the hardware and available table memory resources, trading off unrestricted generality in favor of latency, performance and cost, while enabling a logically centralized control plane with programming flexibility [1].

The OF-DPA Abstract Switch includes functionality to support: bridging; routing; data center tunnel gateways; MPLS provider edge and label switch routing; and QoS use case packet flows. Although all are available simultaneously in the same switch, different flows utilize different sequences of tables and group entries.

This section describes the OF-DPA Abstract Switch packet flows supported in OF-DPA 2.0. These should be taken as informative as the flow diagrams typically show only the relevant objects for that use case. The TTP file should be taken as normative. Note that the flows supported by OF-DPA expose a proper subset of the functionality available in Broadcom ASICs. Future versions of OF-DPA may be 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 conform to the OpenFlow 1.3.4 specifications[5]. In addition to normal flows, OF-DPA supports the following types of special flow entries:

- **Built-in**. Built-in flow entries come preinstalled in specific tables, following the models in the ONF TTP specification [12]. They are visible to the controller but cannot be modified or deleted. Built-in entries have pre-assigned match fields, priority, and actions. Table miss rules are all done using built-in functions.

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.

In addition to flow tables, OF-DPA defines a set of group table entry types. The OpenFlow 1.3.4 specification defines four types of groups: indirect, all, select, and fast failover. OF-DPA further refines these group entry types according to how they can be referenced or chained in packet flows. This is done using specific naming conventions, properties, and supported action buckets. All OF-DPA group table entry types can be programmed using OpenFlow 1.3.4 as long as group mods respect the typing conventions.

One motivation for group typing is supporting fundamental differences in use case requirements. For example, in order to support "one-arm" routing using group table entries there needed to be a way to override OpenFlow's default source removal and allow routing back to IN_PORT. This was accomplished by defining L3 group entry types with different properties from L2 groups. Group entry typing is also useful to enforce constraints on group entry chains and for VLAN configuration on physical ports.

Remember that OF-DPA tables are programming abstractions and do not necessary directly correspond one-to-one with hardware tables. However they are designed to faithfully capture both use case requirements and the hardware packet flow semantics, while being straightforward to program from standard controllers.

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. This includes the requirement that objects must exist before they can be referenced from other objects. The OpenFlow agent that interfaces to OF-DPA may also do some argument validation and execute local iterative procedures.

[5] Some features use extensions defined by ONF as well as some implemented in OpenFlow 1.5. Experimenter fields, actions, and messages are defined so that an OpenFlow 1.3.4 controller can program features.

Many forwarding and editing actions for encapsulation/push and field modify are programmed using one or more 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 the controller intrinsically has more CPU power and memory than the control processor on a typical switch for this task. 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 may in some cases require more messages between the controller and switches, since more objects need to be programmed[6]. It also potentially requires the controller to keep track of more switch state, although this state can be interrogated as needed.

Some functionality must be configured using logical ports. In general, this is to handle packet processing functions that OpenFlow has historically not supported, such as adding and removing encapsulation headers. VXLAN data center overlay tunnels are handled by specialized configuration of logical ports.

Note: OF-DPA does not support reassembly. The switch drops fragments by default.

The following sections are organized by use case packet flow.

3.2 Bridging and Routing

The OF-DPA Abstract Switch objects that can be programmed for bridging and routing are shown in Figure 2. This packet flow adds egress tables but can be programmed the same as OF-DPA 1.0 [13].

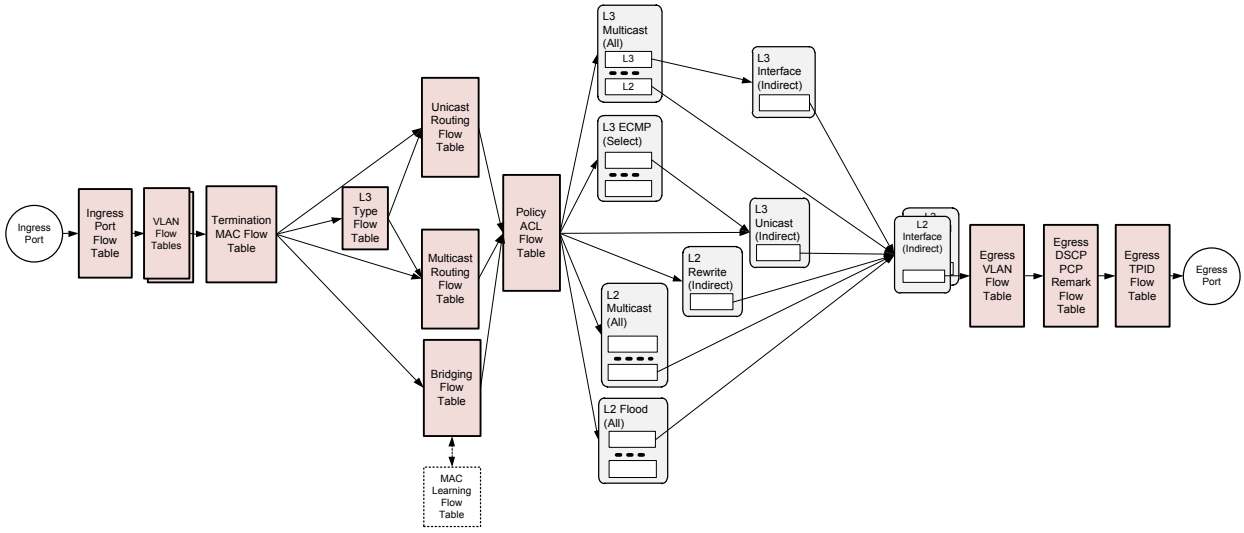

Figure 2 Abstract Switch Objects Used for Bridging and Routing

For Bridging and Routing, packets enter and exit the pipeline on physical ports local to the switch. The Ingress Port Flow Table (table 0) is always the first table to process a packet. Flow entries in this table can

[6] In other cases fewer messages are required, since factoring tables by function can significantly reduce the number of flow entries that need to be programmed.

distinguish normal from tunnel traffic by matching associated Tunnel Id metadata. Normal Bridging and Routing packets from physical ports always have a Tunnel Id value of 0. To simplify programming this table provides a built-in default rule that passes through packets with Tunnel Id 0 that do not match any higher priority rules.

All packets in the Bridging and Routing flow must have a VLAN. The VLAN Flow Table can do VLAN filtering for tagged packets, VLAN translation, and VLAN assignment for untagged packets. The VLAN tables can process up to two VLAN tags. If the forwarded packet has more than one VLAN tag, the outermost VLAN Id is used.

The Termination MAC Flow Table matches destination MAC addresses to determine whether to bridge or route the packet and, if routing, whether it has a unicast or multicast destination MAC address[7]. MAC learning is supported using a "virtual" flow table that is logically synchronized with the Bridging Flow Table.

When MAC learning is enabled, OF-DPA does a lookup in the Bridging Flow Table using the source MAC, outermost VLAN Id, and IN_PORT. A miss is reported to the controller using a Packet-In message. This lookup is done for all packets, whether bridged or routed. The MAC Learning Flow Table cannot be directly read or written by the controller.[8]

The ACL Policy Flow Table can perform multi-field wildcard matches, analogous to the function of an ACL in a conventional switch.

OF-DPA makes extensive use of OpenFlow Group entries, and most forwarding and packet edit actions are applied based on OpenFlow group entry buckets. Groups support capabilities that are awkward or inefficient to program in OpenFlow 1.0, such as multi-path and multicast forwarding, while taking advantage of functionality built into the hardware.

3.3 Data Center Overlay Tunnels

OF-DPA objects for Data Center Overlay tunnels are shown in Figure 3. These are unchanged from OF-DPA 1.0.

Data Center Overlay Tunnel processing is able to separate the traffic for different virtual networks (sometimes referred to as tenants) into isolated forwarding domains. Packets are forwarded based on a non-zero Tunnel-id value that identifies a particular such forwarding domain. A rule must be programmed in the Ingress Port Flow table to admit Data Center Overlay tunnel packets for a particular

[7] The L3 Type Flow Table permits processing packets with a Unicast DMAC but Multicast DIP to be processed as multicast packets. Support depends on the platform.

[8] The MAC Learning Flow Table has a "virtual" table number which is reported to the Controller in a table miss Packet-In message. It does not appear as part of the pipeline since its table number assignment would violate the OpenFlow requirement for packets to traverse tables in monotonically increasing order.

tenant. In addition, flow entries in the Bridging Flow Table and ACL Policy Flow Table match tunnel traffic by Tunnel Id instead of by VLAN Id[9].

OF-DPA uses a naming convention for Tunnel-id metadata where the high order 16 bits identify the tunnel type. This permits flow entries to potentially distinguish among different types of tunnel traffic.

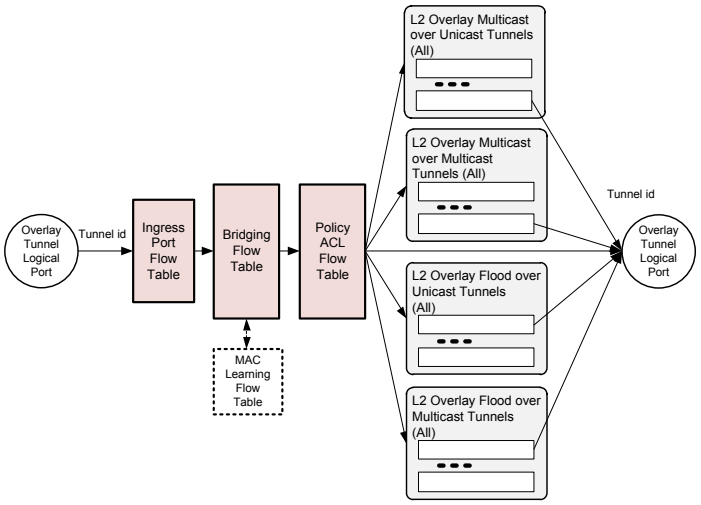

Figure 3 Abstract Switch Objects Used for Overlay Tunnels

In OF-DPA 1.0 Data Center Overlay tunneling is implemented using a combination of configuration APIs, logical ports, and flow tables. In particular the OF-DPA pipeline receives inner packets after encapsulation headers have been removed. These enter from logical ports accompanied by Tunnel Id metadata. OF-DPA forwards packets to logical ports with Tunnel Id metadata.

OF-DPA 1.0 supplied an information model for tunnel configuration intended to be used as an extension to OF-Config. The configuration differs from that in OF-Config 1.2 [14] in two ways. One is that it separately configures the VTEP to a "network" logical port and the tenant VNI to a Tunnel Id, rather than combine the two. This has the dual advantages of permitting tenants to be multiplexed on a VTEP, while doing MAC learning on the ingress logical port. The other difference is that local server facing "access" ports are also configured as logical ports. Multiple tenants can be multiplexed on a physical port, with tenant VLANs mapped to Tunnel Ids.

Another issue is how to route the frames once they are encapsulated with IP headers. OF-DPA overlays are designed to be used in top of rack switches. Since a control plane is not available for routing, the configuration permits static routes with ECMP to be configured using OF-Config.

The OF-DPA 1.0 information model is described in terms of UML diagrams. These can be used to produce YANG models that augment the OF-Config YANG model. They could also be used to generate extensions to OVSDB [16].

[9] This follows the VXLAN convention of treating the VNI as a 24-bit "VLAN extension".

Configuring tunnels this way depends on having a suitable way to do configuration. In general this would require using a configuration protocol and associated configuration agent in addition to requiring an OpenFlow agent.

3.4 MPLS-TP Customer Edge Device

This section describes the OF-DPA objects for MPLS-TP L2 Customer Edge (CE). This functionality is new in OF-DPA 2.0.

The OpenFlow 1.3.4 OFLS does not intrinsically provide all of the functionality needed for MPLS-TP. As a result OF-DPA 2.0 introduces a number of extensions, most of which are in the form of additional pipeline metadata match fields and related actions, but some involve new object types. In contrast to the approach used for Data Center Overlay tunnels, OF-DPA implements MPLS-TP tunneling directly in the pipeline. To support this, OF-DPA 2.0 introduces new actions to push and pop Ethernet headers.

Note: These extensions are modeled after extensions proposed for a later version of OpenFlow. The experimenter programming code points are described in Section 6[10]

3.4.1 VPWS

Figure 4 shows the OF-DPA Abstract Switch objects used in VPWS initiation packet flows. The MPLS-TP L2 VPN groups used for label processing are discussed in more detail in Section 3.4.2.

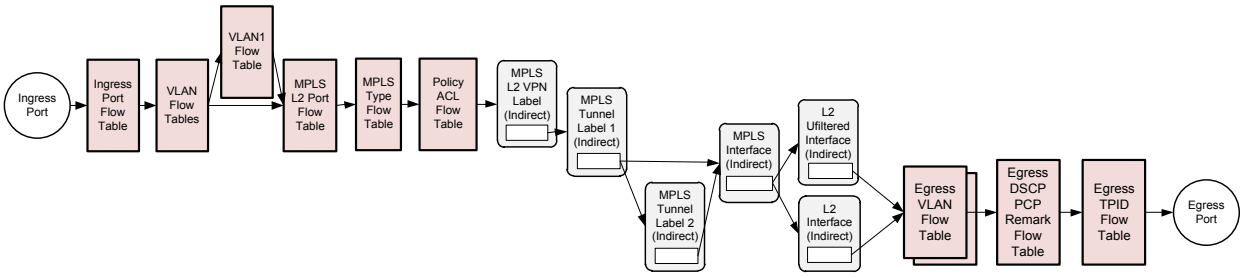

Figure 4 Abstract Switch Objects for MPLS-TP Initiation (VPWS)

VPWS is an E-Line service that provides a point-to-point service. The forwarding decision is fixed; the pseudo-wire directly determines how to forward the packet without the need for a bridging lookup. VPWS does not require learning, flooding, or multicast support.

For VPWS initiation, packets are classified to a particular customer pseudo-wire based on some combination of ingress port and packet VLAN header fields. To accommodate IEEE 802.1Q VLAN stacking, user packets may need to be classified based on both outer and inner VLAN tags. Since OpenFlow 1.3.4 only supports flows that match the outermost VLAN tag, OF-DPA uses a second flow

[10] At this writing OpenFlow 1.5 is a work in progress. As a result, there may be changes to OF-DPA to align OpenFlow 1.5 as features become better defined. If history is any indication, OpenFlow 1.5 support in agents and controllers may not be stable for some time.

table, the VLAN 1 Flow Table, to match both outer and inner tags. For stacked VLAN tags, a flow entry in the VLAN Flow table is first used to match the outermost VLAN, with a Set-Field action in the action list to set a new pipeline match field (OVID) to temporarily hold the matched outermost VLAN Id. The action list then does a Pop-VLAN to expose the inner tag, and a Goto-Table instruction specifies the VLAN 1 Flow Table, where a flow entry can then match both VLAN values together. [11]

To facilitate features that will be introduced later, such as QoS classification, additional pipeline match fields, MPLS L2 Port and MPLS Type, are defined. An MPLS L2 Port can represent either a local attachment (UNI) or network (NNI) logical ingress interface. OF-DPA uses a naming convention for MPLS L2 Port identifiers that partitions it into UNI and NNI ranges in order to distinguish the packet source in flow entries.

In addition, a Tunnel Id metadata value must be assigned to the flow. Analogous to the case with Data Center Overlay Tunnel logical ports, the MPLS forwarding pipeline implements an isolated forwarding domain for a particular customer pseudo-wire, with packets identified by Tunnel-id. A separate Tunnel-id range is defined in order for flow entries to distinguish MPLS-TP packets from data center overlay tunnel packets.

Both MPLS L2 Port and Tunnel-id must be assigned to MPLS-TP flows. Together these are used to represent packets and direction for a particular customer flow.

Figure 5 shows the OF-DPA Abstract Switch objects that are used for VPWS termination packet flows. For simplicity, the two VLAN flow tables and L2 Interface group entries are shown stacked. For VPWS termination, flows typically require no more than one VLAN tag to be matched.

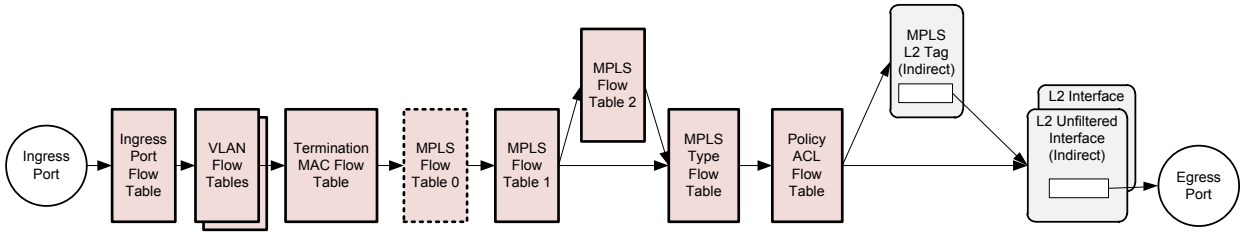

Figure 5 Abstract Switch Objects for MPLS-TP Termination (VPWS)

The Termination MAC Flow Table identifies MPLS frames that require MPLS tunnel termination processing using flow entries that match destination MAC address and EtherType, and that have a Goto-Table instruction that specifies an MPLS Flow Table.

For the same reason that two VLAN tables were used to enable matching two VLAN tags, multiple MPLS Flow Tables are used to enable matching up to three MPLS labels. Each table is used to match an outermost MPLS shim header. The first table, MPLS 0, can be used to match and pop an outermost LSP

[11] This approach was used rather than defining new VLAN header match fields to avoid issues with enabling OpenFlow 1.3.4 to differentiate between tags by TPID.

label[12]. MPLS Flow Tables 1 and 2 can be used to match an LSP label or a pseudo-wire label. In the bottom-of-stack case OF-DPA provides additional new match fields and actions to identify and pop a control word if one is expected, and to remove the outermost L2 header[13]. The pseudo-wire label matching rule also assigns the Tunnel-id, and supplies a group action for forwarding the packet.

MPLS Flow Table 0 only supports a subset of the actions supported by MPLS Flow Tables 1 and 2. All flow entries in MPLS Flow Table 0 should also be installed in MPLS Flow Tables 1 and 2. MPLS Flow Tables 1 and 2 are synchronized and contain the same rules, so updating MPLS Flow Table 1 also updates MPLS Flow Table 2.

Note: The OF-DPA 2.0 API does not permit MPLS Flow Table 2 to be modified directly.

Flexible policies can be applied to the resulting packet using the Policy ACL Flow Table.

If the frame has a service delimiting tag [38], an MPLS L2 Tag group entry can optionally be used to process the tag. The resulting frame is forwarded to an egress interface using either an L2 Interface group entry or an L2 Unfiltered Interface group entry. As in the Bridging and Routing flows, these groups are used to represent the tagging properties of the egress port.

Note: In diagrams overlaid L2 Interface group entries are shown as in Figure 5. This should be interpreted as representing either an L2 Interface or an L2 Unfiltered group. L2 Unfiltered group entries are described in Section 4.3.2.

3.4.2 MPLS-TP L2 VPN Groups

Figure 6 shows the L2 VPN MPLS groups used in VPWS flows and includes positions where MPLS Fast Failover groups can be inserted. OF-DPA 2.0 supports MPLS-TP 1:1 linear protection. Figure 6 shows the MPLS Fast Failover and MPLS L2 VPN group entry chaining relationships for two level 1:1 protection.

[12] The dotted lines indicate that MPLS 0 is optional and may not be supported on all platforms.
[13] OpenFlow expects the MPLS pop action to specify the EtherType of the result packet, but in this case the result is "Ethernet in Ethernet." For consistency it could be programmed to Transparent Ethernet Bridging (0x6558) [33].

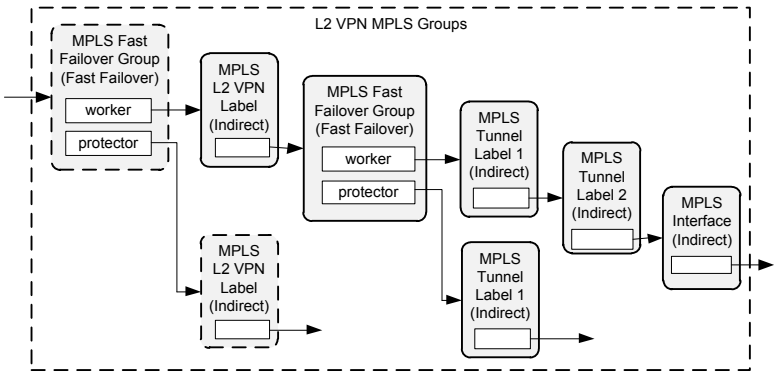

Figure 6 MPLS L2 VPN Groups

MPLS Fast Failover group entries have two buckets. The primary (worker) bucket is for the normal working path, and the second (protector) bucket is for the protection path. An external network protection switching process decides which to use and when to switch over from one to the other. The operation of the network protection switching process and its relation to OAM objects is described in Section 3.9.

Note: The number of levels of protection available is platform dependent.

The MPLS L2 VPN Label group entry bucket actions can include:

- Push L2 Header (initial values are zero; to be populated by actions in the MPLS Interface group)
- Push CW (new action to push control word)
- Push MPLS shim header (for PW label; must set the EtherType in the L2 header to 0x8847)
- Set-Field MPLS Label
- Set-Field BOS (bottom of stack)
- Set-Field TC (EXP) – explicit value or from table based on Traffic Class and Color. Optional.
- Set-Field TTL
- Set-Field PCP (in outermost L2 header) – explicit value or from table. Optional.
- Group (MPLS Tunnel Label)

The MPLS Tunnel Label 1 group entry bucket actions can include:

- Push MPLS shim header (for LSP label; must set the EtherType to 0x8847)
- Set-Field MPLS Label
- Set-Field TC (EXP) – explicit value, copy out, or set from table. Optional
- Set-Field TTL (value or copy out)
- Set-Field PCP (in L2 header) – explicit value or from table (Optional)
- Group (MPLS Tunnel Label 2 or MPLS Interface)

The MPLS Tunnel Label 2 Group entry may be used to push a second outer tunnel label and has the same actions as MPLS Tunnel Label 1 except the chained group can only be MPLS Interface.

The MPLS Interface group entry sets the outer MAC-DST, MAC-SRC, and VLAN Id. The outer L2 header is assumed to have a VLAN field.

3.5 MPLS Label Edge Router (LER)

The OF-DPA objects for MPLS L3 VPN are shown in Figure 7 for initiation and Figure 8 for termination.

The MPLS Label Edge Router packet flow supports routing into and out of MPLS L3 VPN tunnels. An LER is both an IP router and an MPLS tunnel endpoint. The LER can support multiple VPNs for different customers.

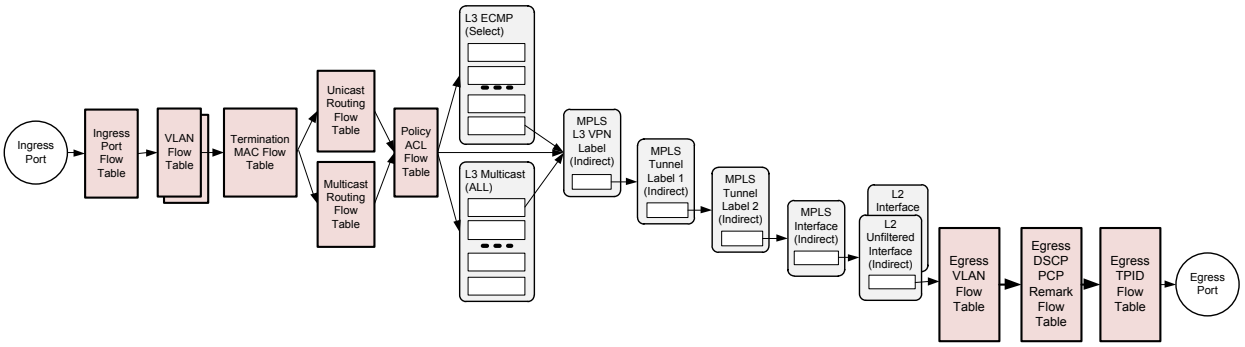

Figure 7 Abstract Switch Objects for MPLS L3 VPN Initiation

For tunnel initiation IP packets are routed to an MPLS L3 VPN. For isolation multiple virtual routing tables (VRF) are supported, selected using VRF pipeline metadata. OF-DPA defines a pipeline match field for VRF.

Customer ingress traffic is differentiated based on VLAN, and VLAN Flow Table rules can optionally assign a VRF for a customer's traffic. IP packets can be forwarded either directly or through L3 Multicast or L3 ECMP group entries to MPLS Label Group Entries that push MPLS labels and update fields in the Ethernet header. In these groups the buckets must chain to the appropriate L3 group entry types.

The MPLS L3 VPN packet flows also support protection switching for MPLS frames using MPLS Fast Failover Group Entries to select between primary and backup paths based on the liveness property of a referenced port or group.

Note: Support for L3 VPN automatic protection switching is platform dependent. OF-DPA does not support OAM for L3 VPN packet flows.

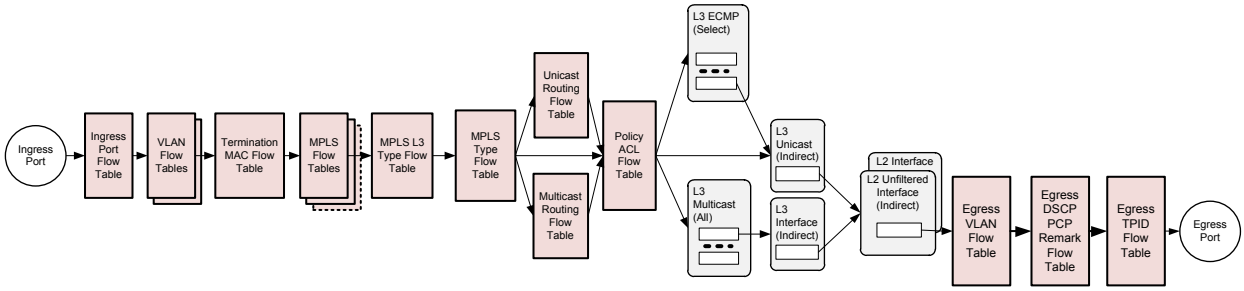

Figure 8 Abstract Switch Objects for MPLS L3 VPN Termination

For L3 VPN tunnel termination, MPLS frames are identified by destination MAC and VLAN in the Termination MAC Flow Table. MPLS shim headers are processed as previously indicated by the MPLS Flow Tables.

Popping the shim headers turns MPLS L3 VPN these frames into IP packets that can be routed using the Routing Flow Tables or directly forwarded to L3 Unicast, L3 Multicast, or L3 ECMP Group Entries. Packets to be routed that have multicast group destinations are forwarded using the Multicast Routing Flow Table rather than the Unicast Routing Flow Table.

The MPLS L3 Type Flow table detects the IP version and whether the packet has a multicast destination. This is a fixed table containing only built-in flow entries. It cannot be modified using OpenFlow, and its flow entries never expire.

The MPLS bottom of stack label should assign VRF pipeline metadata for the VPN. VPN traffic is isolated by being forwarded to the VLAN associated with that VPN's VRF.

Figure 9 shows the MPLS L3 VPN group entry types.

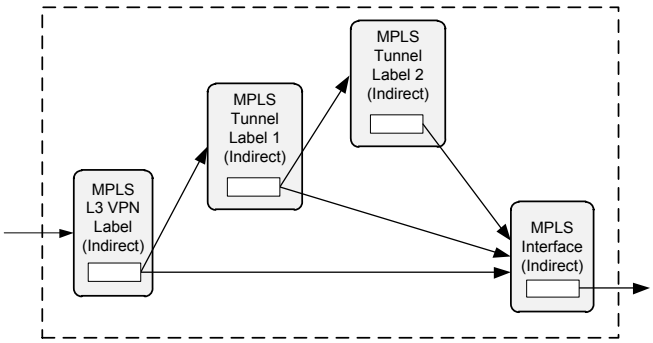

Figure 9 MPLS L3 Label Groups

The MPLS L3 VPN Label group entry bucket actions include:

- Push MPLS shim header (for VPN label; this sets the EtherType to 0x8847)
- Set-Field MPLS Label

- Set-Field BOS (bottom of stack)
- Set-Field TC (EXP) – explicit value or from table based on Traffic Class and Color. (Optional)
- Set-Field TTL
- Set-Field PCP (in L2 header) – explicit value or from table. (Optional)
- Group (MPLS Tunnel Label)

The MPLS Tunnel Label group entry bucket actions are the same as for the L2 VPN case.

3.6 MPLS Label Switch Router (LSR)

The OF-DPA objects for an MPLS Label Switch Router (MPLS-TP P node) are shown in Figure 10. Except for the MPLS SWAP Label group, the flow uses many of the same objects as previously described.

An LSR forwards MPLS frames by optionally popping one or more labels and then swapping a label. For OF-DPA these are programmed as actions in the MPLS Flow Tables. A swap action can apply either to a tunnel (LSP) label that is not bottom of stack, and also to a PW label that is bottom of stack for PW stitching scenarios.

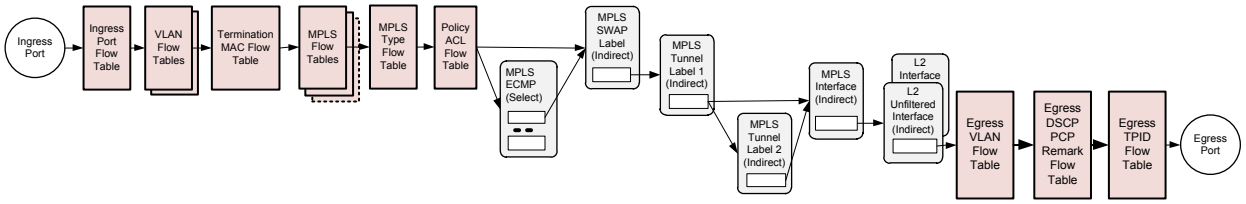

Figure 10 Abstract Switch Objects Used for an MPLS LSR

OF-DPA supports optional forwarding to an MPLS ECMP group for multipath load balancing. This is only used for MPLS L3 VPN since it is not supported in the MPLS-TP specifications. The path selection function is implemented in the hardware platform. Note that OF-DPA 2.0 does not provide support for MPLS entropy labels.

The LSR group entries including Fast Failover are shown in Figure 11.

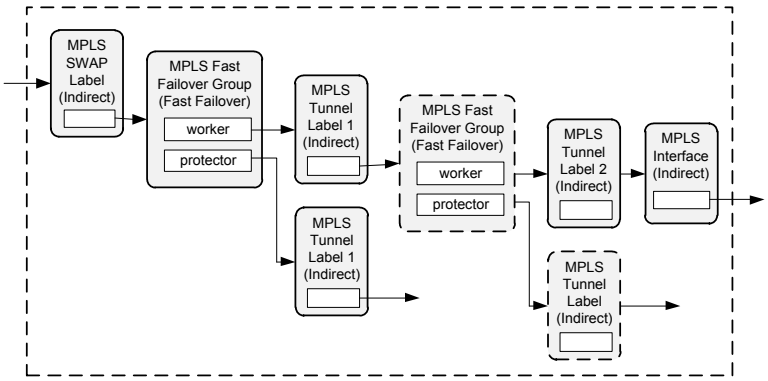

Figure 11 MPLS LSR Groups

The MPLS SWAP Label group entry bucket actions include:

- Push MPLS Label (if LSP label popped)
- Set-Field MPLS Label (set or swap label)
- Set-Field TC (EXP, optional)
- Decrement and check TTL
- Set-Field TTL (optional)
- Set-Field PCP (in outermost L2 header) – explicit value or from table. Optional.
- Group (MPLS Tunnel Label)

The MPLS Tunnel Label group entry bucket actions are the same as for the earlier cases.

The MPLS Interface group entry sets the outer MAC-DST, MAC-SRC, and VLAN Id for forwarding. The outer L2 header is assumed to have a VLAN field.

3.7 MPLS Protection Switching

OF-DPA 2.0 supports MPLS-TP 1:1 linear protection. The MPLS Fast Failover group entry for 1:1 protection is shown in Figure 12.

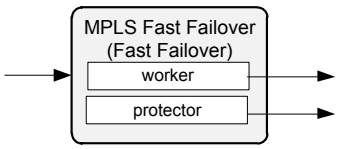

Figure 12 MPLS-TP 1:1 Protection

MPLS Fast Failover group entries have two buckets. The primary bucket is for the normal working path, and the second bucket is for the protection path. An external network protection switching process decides which to use and when to switch over from one to the other.

MPLS Fast Failover buckets must be programmed with a watch_port that specifies "watching" an OAM Protection Liveness logical port. The Protection Switching process can change the state of an OAM Protection Liveness logical port in order to switchover to the protection path. The OAM Protection Liveness logical port must be included in the MPLS Linear Protection configuration as shown in Figure 48. The operation of the network protection switching process and its relation to OAM objects is described in Section 3.9.

Note: The number of levels of protection available is platform dependent. Note that in most instances one or two levels are sufficient.

3.8 Quality of Service (QoS)

This section describes the OF-DPA Abstract Switch objects supporting QoS for Bridging and Routing and for MPLS. Note that OpenFlow 1.3.4 does not fully support the metering and marking capabilities

required by standards such as DiffServ [28]. As a result, some concepts need to be introduced before describing the abstract switch packet flows.

3.8.1 QoS Concepts – Traffic Class and Color

QoS packet processing following the DiffServ model [28] comprises the following:

- Classification - assigning a Traffic Class and Color, or drop precedence. Trust policies are associated with ingress ports and can determine whether or not to assign Traffic Class and Color based on incoming packet header fields;
- Metering – policing, which can change the Color based on flow properties such as packet and byte rate;
- Marking - setting QoS fields in the packet headers based on Traffic Class and Color; and
- Shaping – queuing or dropping the packet based on Traffic Class and Color, respectively. Queues are serviced based on the scheduling discipline. Typically queues implement a color-based admission scheme such that they stop accepting packets with higher drop precedence after the number of entries in the queue exceeds some threshold[14].

OF-DPA 2.0 defines new pipeline match fields for Traffic Class and Color. Traffic Class can take one of eight values corresponding to DiffServ per-hop behavior (PHB) groups: BE (0), AF1 (1), AF2 (2), AF3 (3), AF4 (4), EF (5), CS6 (6), and CS7 (7)[15]. Color can be one of Green (0), Yellow (1), or Red (2).

OF-DPA supports multiple methods for doing classification and marking. In general, the approach used can depend on the use case and the particular packet flow. OF-DPA defines new table types for classification and marking that differ from standard OpenFlow Match Action Tables. These classification and marking tables can be configured using OF-DPA APIs similar to the ones for flow tables.

The L2 Policer and the Policy ACL Flow Tables provide Meter instructions. These instructions reference a Meter Table entry to police a flow. The L2 Policer Actions and the Policy ACL Color Based Actions Flow Tables can be used for applying color based actions, respectively.

3.8.2 Meter Table Operation

OF-DPA uses a type of Meter Table entry with two meter bands. Each meter band can indicate actions to apply immediately to the packet. In OpenFlow 1.3.4, the meter instruction must be evaluated before other instructions.[16]

[14] Shaping properties and configuration are platform dependent.
[15] Traffic Class is encoded using a 4-bit field, with values 8-15 interpreted as CS7.
[16] OpenFlow 1.5 deprecates the meter instruction in favor of a meter action. This does not materially change the operation of OF-DPA 2.0 meters, but upgrading the OpenFlow version would change how meters were programmed from the controller.

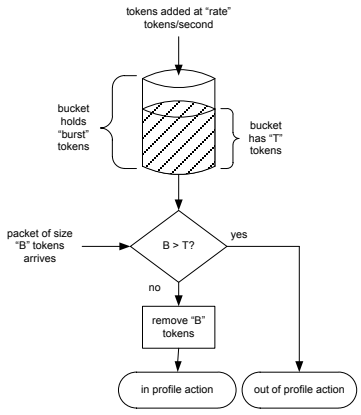

Figure 13 Token Bucket Operation

As opposed to OpenFlow 1.3.4 Meters, OF-DPA Meters are implemented using the standard token bucket algorithms, and hence the configuration implies not only the rate and burst limits but also how rate and burst are measured. For the most part, rate and burst are configured in terms of standard OpenFlow Meter Band parameters. Figure 13 illustrates the operation of the token bucket algorithm. OpenFlow Meter Band rate and burst parameters are used to configure the rate at which tokens are added and the bucket size in tokens, as well as whether the tokens represent kilobits or packets.[17] However, OF-DPA meters also support three modes (two rate three color, single rate three color, or modified two rate three color) and two forms of coloring (whether color aware or color blind).

Operation of the two rate three color mode (TrTCM), illustrated in Figure 14, is as specified in RFC 2698 [30].

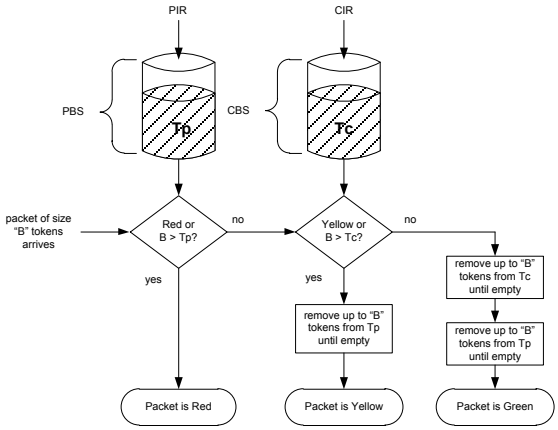

Figure 14 TrTCM Meter Operation

TrTCM meters can be color-aware or color-blind. TrTCM color-aware operation in terms of actions and incoming packet color is as shown in Table 2. In color-blind mode the incoming packet is always Green.

[17] OpenFlow meter bands specify packet rates and burst sizes independent of how measured.

Table 2 TrTCM Color-Aware Operation

Incoming Color	B <= Tc?	B <= Tp?	Result Color	Tc	Tp
Green	Yes	Yes	Green	Tc -= B	Tp -= B
	Yes	No	Red	-	-
	No	Yes	Yellow	-	Tp -= B
	No	No	Red	-	-
Yellow	Yes	Yes	Yellow	-	Tp -= B
	Yes	No	Red	-	-
	No	Yes	Yellow	-	Tp -= B
	No	No	Red	-	-
Red	Yes	Yes	Red	-	-
	Yes	No	Red	-	-
	No	Yes	Red	-	-
	No	No	Red	-	-

The single rate three color mode (SrTCM), illustrated in Figure 15, operates as specified in RFC 2697 [29].

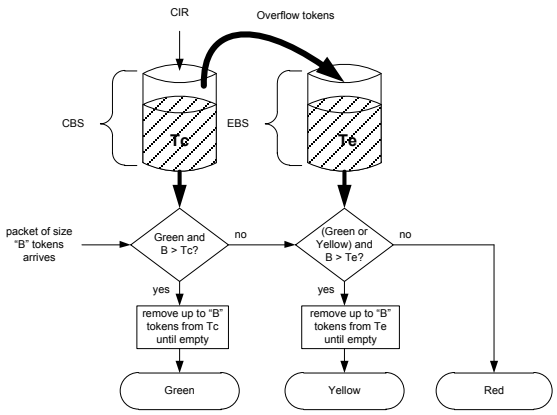

Figure 15 SrTCM Meter Operation

As with TrTCM, SrTCM meters can be color-aware or color-blind. SrTCM color-aware operation is as shown in Table 3. In color-blind mode the incoming packet is always Green.

Table 3 SrTCM Color-Aware Operation

Incoming Color	B <= Tc?	B <= Te?	Result Color	Tc	Te
Green	Yes	Yes	Green	Tc -= B	
	Yes	No	Green	- Tc -= B	-
	No	Yes	Yellow	-	Te -= B
	No	No	Red	-	-
Yellow	Yes	Yes	Yellow	-	Te -= B
	Yes	No	Red	-	-
	No	Yes	Yellow	-	Te -= B
	No	No	Red	-	-
Red	Yes	Yes	Red	-	-
	Yes	No	Red	-	-
	No	Yes	Red	-	-

Incoming Color	B <= Tc?	B <= Te?	Result Color	Tc	Te
	No	No	Red	-	-

The modified or "efficient" two rate three color mode, illustrated in Figure 15, operates as specified in RFC 4115 [31]. Its operation, shown in Table 3, is similar to SrTCM except for the way in which tokens are added, EIR is configured independently. In addition OF-DPA optionally supports the MEF modified TrTCM coupling mode [32], where Te is increased by both EIR tokens and overflow tokens from Tc as in the SrTCM. This is indicated in Figure 16 using the dotted overflow arrow.

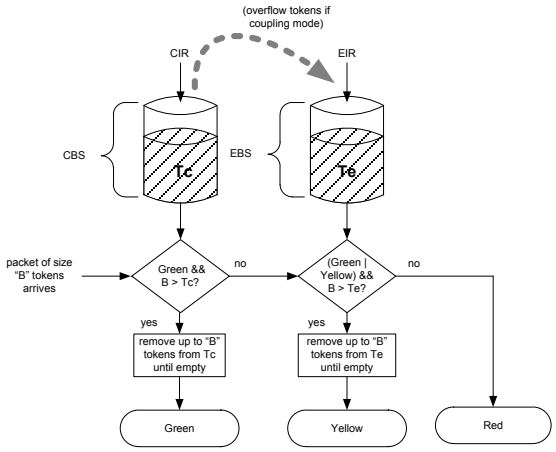

Figure 16 Modified TrTCM Meter Operation

As before meters can be color-aware or color-blind. In color-blind mode the incoming packet is always considered Green.

Table 4 Modified TrTCM Color-Aware Operation

Incoming Color	B <= Tc?	B <= Te?	Result Color	Tc	Te
Green	Yes	Yes	Green	Tc -= B	
	Yes	No	Green	- Tc -= B	-
	No	Yes	Yellow	-	Te -= B
	No	No	Red	-	-
Yellow	Yes	Yes	Yellow	-	Te -= B

Incoming Color	B <= Tc?	B <= Te?	Result Color	Tc	Te
	Yes	No	Red	-	-
	No	Yes	Yellow	-	Te -= B
	No	No	Red	-	-
Red	Yes	Yes	Red	-	-
	Yes	No	Red	-	-
	No	Yes	Red	-	-
	No	No	Red	-	-

In OF-DPA 2.0 the packet Traffic Class value is used to directly assign the queue. As a consequence OF-DPA does not support the Set Queue action.

Queues can be configured in terms of minimum and maximum rates using the OF-DPA queue configuration APIs.

Meter Table entries are configured using OF-DPA meter table APIs. An OpenFlow controller would program OF-DPA Meter Table entries using the experimenter message fields described in Section 6.

3.8.3 Bridging and Routing QoS

The OF-DPA 2.0 Bridging and Routing pipeline, including additional objects for QoS, is shown in Figure 17. Basically QoS adds Meter table instructions to the Policy ACL Flow Table and remark actions for IP and Ethernet using the Color Based Actions Flow Table. The L2 Policer table is not used.

In OF-DPA 2.0, the Ingress Port Flow Table rule GoTo-Table instructions can optionally specify the DSCP Trust Flow Table or the PCP Trust Flow Table in addition to directly to the VLAN Flow Table. These trust flow tables can implement incoming packet classification profile mappings in order to determine Traffic Class and Color assignments based on header marking. The DSCP Trust Flow Table is used to classify incoming IP packets based on the DSCP field and a QoS Trust Index[18] supplied by the Ingress Port Flow Table. Similarly, the PCP Trust Flow Table is used to classify an incoming packet based on the 802.1Q PRI and DEI fields based on a QoS Trust Index supplied by the Ingress Port Flow Table. The default port classification is used if the packet type is incompatible with the rule.

[18] Use of separate tables and QoS Index values can significantly reduce the number of required rules and controller programming burden. It also can support applying different profile mappings.

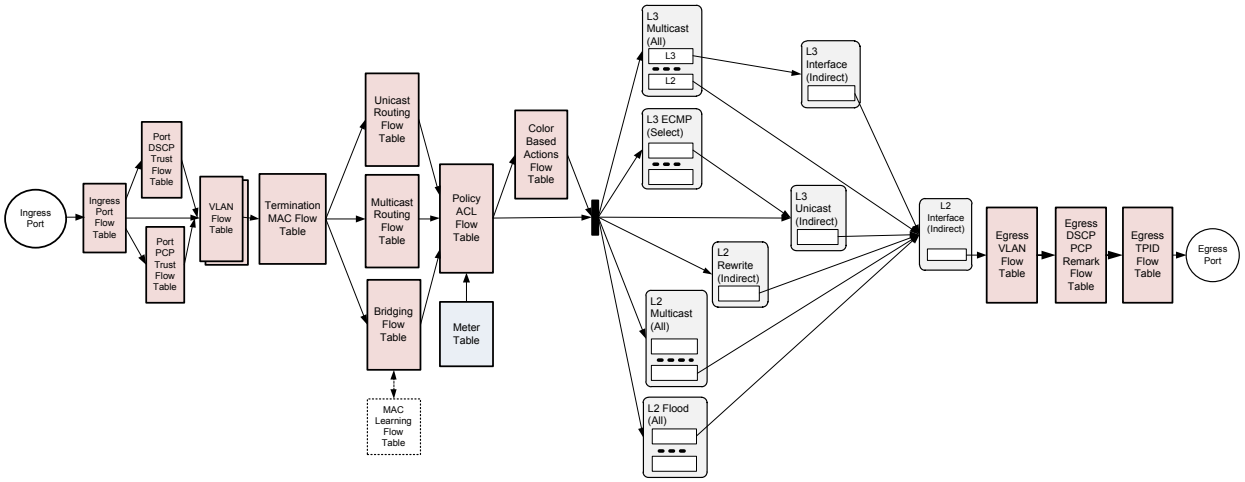

Figure 17 Bridging and Routing Abstract Switch Showing QoS Objects

Metering is used to police traffic rate according to bandwidth profiles. OF-DPA policing effectively implements three profiles: red, yellow, and green. The meter can change the value of the Color pipeline match field based on the profile.

Packet QoS field values can optionally be remarked at egress using a number of methods:

- Using Policy ACL Color Based Actions Flow Table actions
- Using statically configured values
- Using the Egress DSCP PCP Remark Flow Table.

The Egress DSCP PCP Remark Flow Table can match Traffic Class and Color along with an index supplied by one of the L2 Interface Group entries.

3.8.4 Overlay Tunnel QoS

Figure 18 shows the analogous Tunnel Overlay Logical Port trust classification tables and meter application.

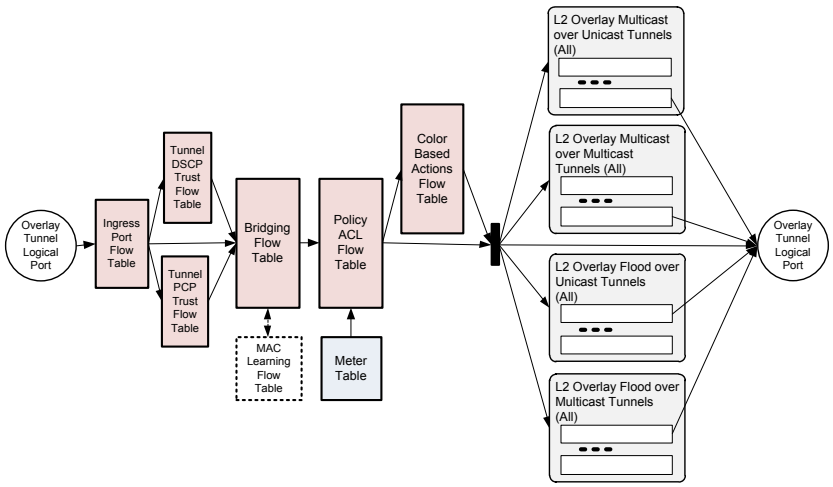

Figure 18 Overlay Tunnel Abstract Switch Showing QoS Objects

3.8.5 MPLS QoS

Figure 19 shows the OF-DPA objects added to MPLS-TP VPWS tunnel initiation for QoS. Incoming packets from local interfaces are classified to a pseudo-wire, represented as an MPLS L2 port abstraction. MPLS Trust Flow Tables can be associated with the MPLS L2 port. These tables are analogous to the L2/L3 PCP/DHCP Port Trust Tables but determined by the MPLS L2 port instead of the ingress port. The MPLS L2 Port trust tables can perform "simple" classification based on the packet QoS header fields.

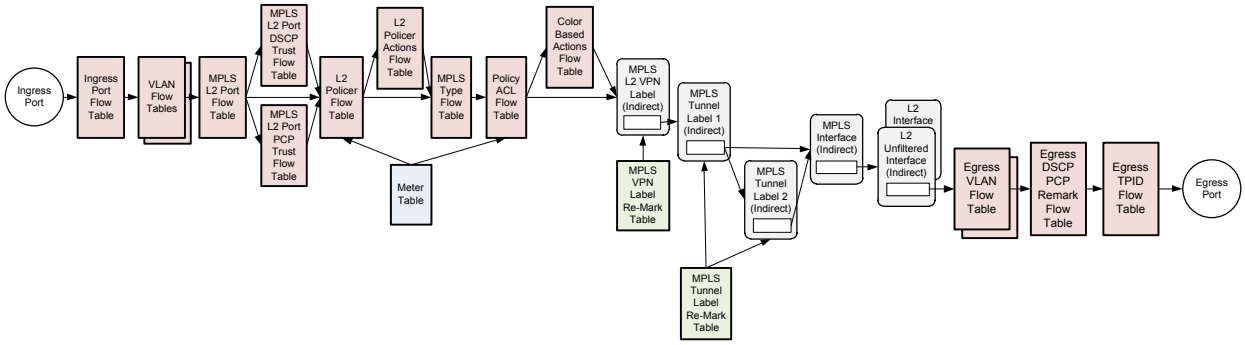

Figure 19 MPLS-TP VPWS Initiation with QoS Objects

Two metering options are supported for MPLS-TP: using the Policy ACL rules; or using the L2 Policer Flow Table. The L2 Policer table can potentially provide meters for a large number of client flows, but its availability is platform dependent.

OF-DPA 2.0 supports different options for setting the MPLS TC and outermost VLAN PCP and DEI fields in order to accommodate different use cases (e.g., pipe and uniform models for LSP and PW labels):

- Fixed values programmed into the MPLS label group entries.
- Copy out from inner label, if one exists.
- Map using a MPLS VPN or Tunnel Label Remark Table with an index supplied in the label group.
- Use the existing value unchanged.

OF-DPA defines two types of MPLS Remark tables as new ancillary table objects. The MPLS VPN Label Remark Tables are referenced by actions from OF-DPA MPLS VPN Label Group entries when originating a tunnel, and from OF-DPA MPLS Swap Label Group entries when swapping a label. The MPLS Tunnel Label Remark Tables are referenced by actions from OF-DPA MPLS Tunnel Label 1 or OF-DPA MPLS Tunnel Label 2 Group Entries. Either type can specify a new MPLS TC value as well as new 802.1p PCP and DEI values.

These new object types are used because OpenFlow 1.3.4 does not provide a way to do a data dependent profile determination (e.g., a table lookup) in the context of a group action, or to do a match action table lookup between group entries.[19]

The OF-DPA objects for MPLS-TP VPWS tunnel termination with QoS are shown in Figure 20. This adds the Meter Table and Color-Based Actions Flow Table previously described, and also supports remarking user frames using the Egress DSCP PCP Remark Table.

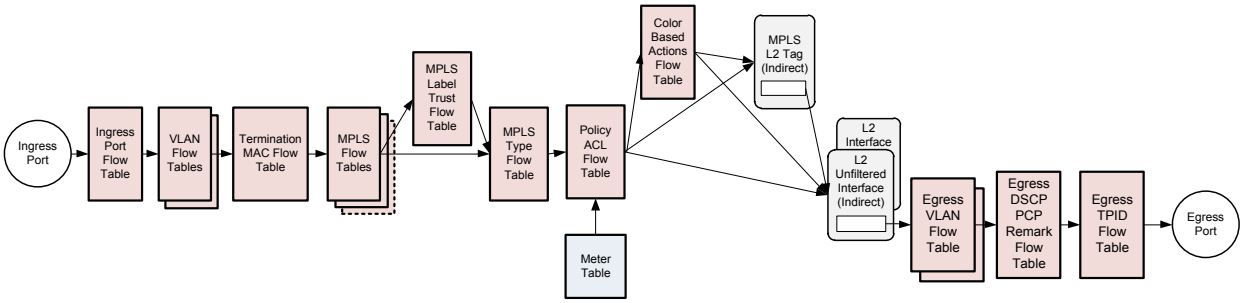

Figure 20 MPLS-TP VPWS Tunnel Termination with QoS Objects

The OF-DPA MPLS L3 VPN Initiation abstract objects with QoS are as shown in Figure 21. The corresponding MPLS L3 VPN Termination Abstract Switch objects including QoS objects for label remark is as shown in Figure 22.

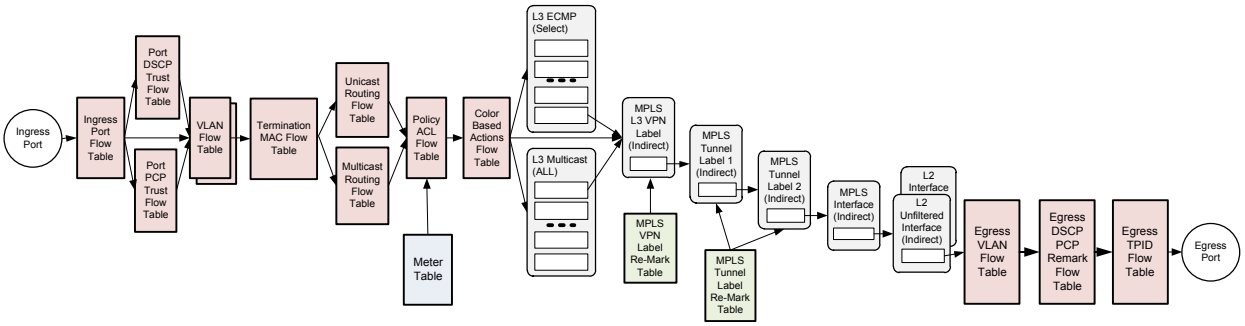

Figure 21 MPLS L3 VPN Initiation with QoS Objects

[19] Although OF-DPA does make use of Egress Tables, they are only applied at the output port immediately before packet egress and after all groups. Inner label EXP fields cannot be modified using this approach.

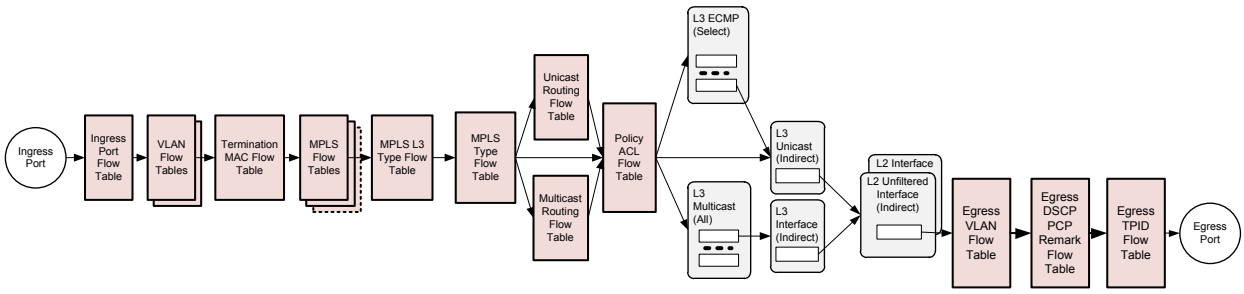

Figure 22 MPLS L3 VPN Termination with QoS Objects

For completeness, the MPLS LSR abstract objects including QoS objects is shown in Figure 23. The MPLS Label Remark tables are as described above.

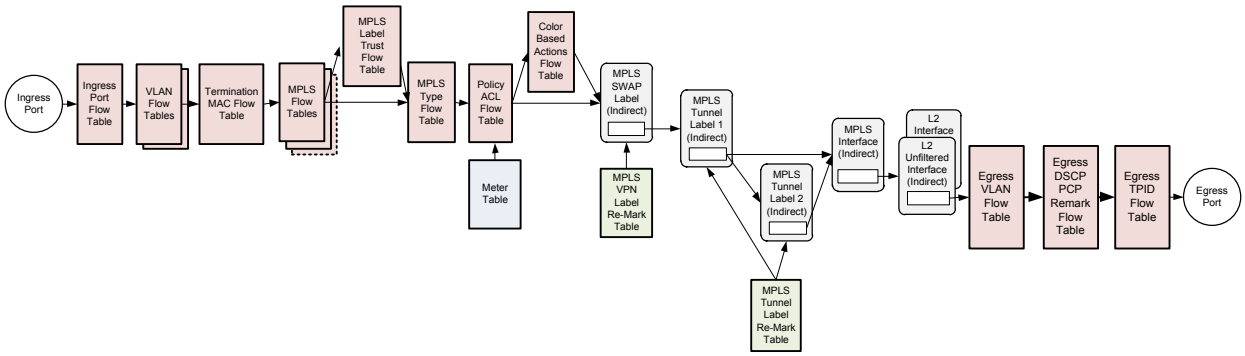

Figure 23 MPLS LSR with QoS Objects

Note that flow policing can change packet color and potentially affect queue drop precedence as well as remarking.

3.9 Operation, Administration, and Maintenance (OAM)

OAM functions are used to detect and localize loss of connectivity or degradation of performance in order to maintain service level guarantees. In particular, identifying the defect location in the network enables rapid repair based on preprovisioned redundant (protection) paths. This section provides some background context on OAM but is not intended to be a comprehensive tutorial.

Different networking scenarios use different types of OAM, and sometimes more than one can be used at the same time. OF-DPA includes support for ITU G.8113.1 and for Ethernet OAM over MPLS-TP. This section describes the OF-DPA objects used for OAM processing.

3.9.1 OAM Concepts – Maintenance Points

OAM functions are located at Maintenance Endpoints (MEP) and Maintenance Intermediate Points (MIP) belonging to a Maintenance Association (MA) for a fully connected network. MEPs are often, but not always, located at nodes that provide external access to the network. MIPs are located at strategic points

within the network. The MEPs transmit and receive OAM messages in order to detect defects on the protected path(s) between them.

OAM messages are used for fault management and performance monitoring. Depending on the type and purpose, messages might be sent periodically (proactive OAM) or on-demand. MIPs located along the path between MEPs respond to messages received from MEPs belonging to the same MA.

On a network element, MEPs and MIPs can be provisioned so that they face outward, towards the network, or inwards, towards the node. Network facing MEPs are referred to as Down MEPs and are used to test the path up to and out of an interface on the node. Inward facing MEPs are referred to as Up MEPs and are used to test the forwarding path through the node. An MA can consist of either Down MEPs or Up MEPs, but not both. Note that a MIP can have Up and Down orientations as well[20].

Maintenance Domain (MD) is an Ethernet OAM concept. An MD is a collection of nested MAs identified by Maintenance Domain Level (MDL). Nested MDs are used to verify connectivity of a path subset in order to localize faults. All Ethernet OAM messages include an MDL field.

These concepts are illustrated in Figure 24, taken from the IEEE 802.1ag specification [35].

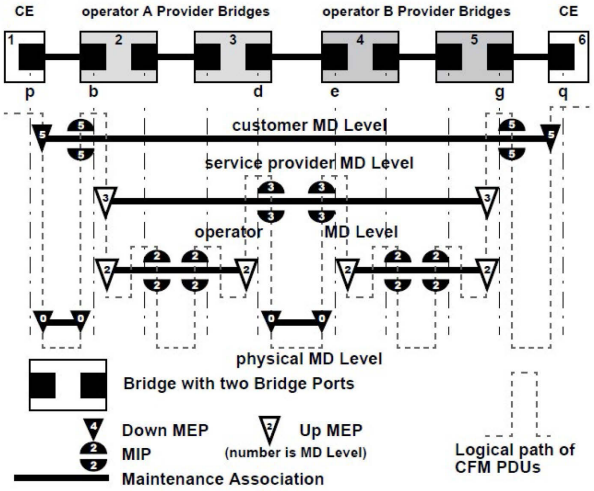

Figure 24 OAM MEP and MIP Examples

Note: The term Maintenance Association is defined in IEEE 802.1ag for Ethernet OAM. The equivalent term defined in ITU G.8113/Y.1731 is Maintenance Entity Group (MEG).

OF-DPA OAM provides operations for provisioning and supporting MEPs and MIPs. These include using OpenFlow match action tables to identify frames that require processing at a MEP or MIP, and then

[20]Actually, MIPs are both, with Up and Down MHFs (MIP Half Functions), represented with semi-circles.

performing OAM actions on these frames. On control frames (OAM messages), the actions can be to do OAM message processing. On data frames, the actions can include performance monitoring instrumentation, for example, updating transmit and receive counters for loss measurement (LM).

A particular Maintenance Point in the MEG is identified using a unique LMEP Id[21]. OF-DPA defines an LMEP Id pipeline metadata match field for this purpose. OF-DPA uses LMEP Ids to bridge between the OAM configuration and packet forwarding domains.

3.9.2 Network Protection Apps

To meet OAM proactive fault management and performance monitoring requirements, OF-DPA supports offloading OAM processing functions to one or more Local Engines hosting local Network Protection Apps. There are multiple implementation options for a Local Engine, such as: software on the local CPU; external hardware or software device such as an FPGA or sidecar processor; internal processing engine dedicated to OAM[22]. In all cases, the local Network Protection App can perform time-sensitive OAM functions such as transmitting and receiving continuity check (CCM) frames.

OF-DPA uses an OpenFlow LOCAL Reserved port to interface between the data plane and a local Network Protection App, as shown in Figure 25. Network Protection Apps receive and transmit OAM frames with parameters similar to those in Packet_In and Packet_Out messages.

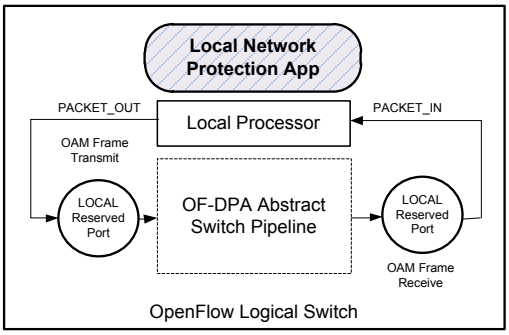

Figure 25 OAM and Protection Overview

OF-DPA Abstract Switch objects send frames to the Network Protection App using an output action to LOCAL, accompanied by OAM pipeline metadata fields. Conversely, a Network Protection App can inject OAM frames into the OF-DPA pipeline that ingress via the LOCAL port, accompanied by packet metadata.

OF-DPA flow entries execute an output action to the LOCAL Reserved Port and provide associated metadata fields. These pipeline match fields can include the packet IN_PORT, LMEP Id, transmit and receive counters, and packet arrival timestamp.

[21] LMEP is used generically and applies to both MEPs and MIPs
[22] Available on certain Broadcom platforms

Note: The mechanism for associating metadata with the LOCAL Reserved Port is implementation dependent and outside of the scope of this specification.

Note: The level of accuracy in loss measurements is affected by the latency between reading counters and inserting counts into injected OAM frames. In general, a Local Network Protection App should be used to provide better accuracy. It is further recommended that a local interface be provided for accessing counter values, although such a facility is outside of the scope of this specification.

3.9.3 MPLS-TP OAM

MPLS-TP support uses MPLS Fast Failover Groups to implement protection for network paths. Fast Failover Groups rely on OAM path fault detection for liveness. When OAM detects a fault on the worker path, the OAM connectivity verification process state machine will notify the network protection state machine, which in turn will invalidate the liveness property being monitored by the worker bucket. This then causes the Fast Failover Group to switch over to the protector path. The MPLS Fast Failover Group bucket "watch_port" parameter is configured to watch the operational status of an OAM Network Protection Logical Port defined for this purpose.[23]

The requirements for MPLS-TP OAM are described in RFC 5860 [34]. Examples are illustrated in Figure 26.

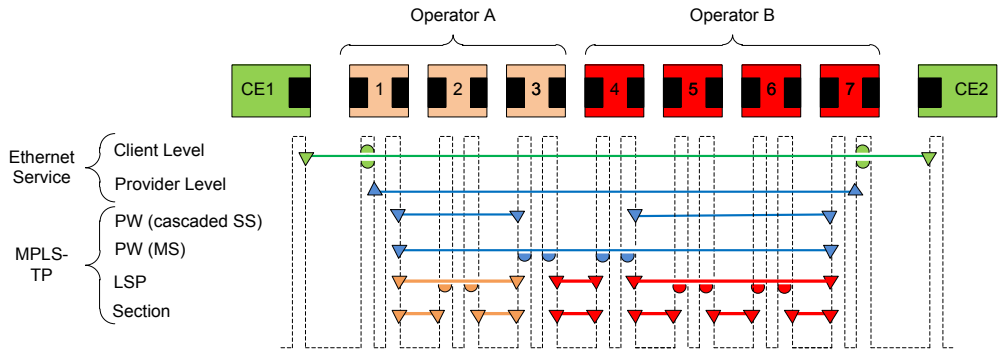

Figure 26 MPLS-TP Service OAM Examples

As shown in Figure 26, MPLS-TP requires support for the following different types of OAM:

- Client Service OAM is end-end between client points of service (UNI-C). Since MPLS-TP provides an Ethernet service, client service OAM exchanges Ethernet OAM frames through the MPLS tunnel between Down MEPs on customer equipment (CE). Client Service OAM can also query MIPs at the local and remote provider edge devices.
- Provider Service OAM is end-end between service provider points of service (UNI-N). This also exchanges Ethernet OAM frames through the MPLS tunnel between Up MEPs in LERs.

[23] In OpenFlow 1.3.4, Fast Failover Group buckets are configured to "watch" either port or group liveness, with group liveness determined (at run time) by port liveness. OF-DPA MPLS Fast Failover Group buckets "watch" a logical port, relying on the Network Protection App to control bucket liveness by changing the state of the watched logical port.

- MPLS-TP Channel OAM is at the pseudo-wire (PW) level and may be end-end (multi-segment, MS-PW), or stitched across multiple operator domains (single-operator, cascaded SS-PW). Two operator domains are shown in Figure 26. Multi-domain end-end at the PW level is also shown. Channel OAM is only between Down MEPs. MS-PW is shown with client MIPs on nodes with server level MEPs.
- MPLS-TP Path OAM is at the LSP level and is used to protect the path between network elements. Path OAM MEs are always within a single operator domain and only between Down MEPs. Path OAM has MIPs on nodes with server level MEPs. While one level of LSP is shown above, there may be segments with nested LSPs.
- MPLS-TP Section OAM is on the single hop between two network elements. Section OAM is between Down MEPs on adjacent nodes.

Each requires identifying one or more packet formats. Packet formats are described in the following sections.

The following sections provide details of OF-DPA MPLS-TP OAM processing for Ethernet and G.8113.1. Configuration for message processing and liveness determination can be found in Section 5.3 on page 174.

3.9.3.1 Ethernet Service OAM

Ethernet OAM methods are described in ITU-T Y.1731 [22] and IEEE 802.1ag [35]. Ethernet OAM frames are encapsulated within MPLS. These frames are identified by having ETH-TYPE=0x8902.

The packet format is shown in Figure 27.

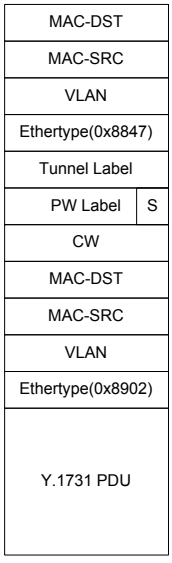

Figure 27 Ethernet Service OAM over MPLS-TP Packet Format

For tunnel initiation at the Maintenance Point, the OF-DPA data path requirement is to count transmitted user data frames if LM is enabled. This would be the case for a link level Client Service Down MEP, for example. User frames are recognized in the VLAN Flow Table as for the general initiation case. A new

OAM Data Plane Counter Table is defined along with a new action to increment based on LMEP Id and Traffic Class for the affected flow.

In addition, an action to query the Drop Status table is used to drop client frames during a Lock condition. The flow tables involved are shown in the pipeline fragment in Figure 28.

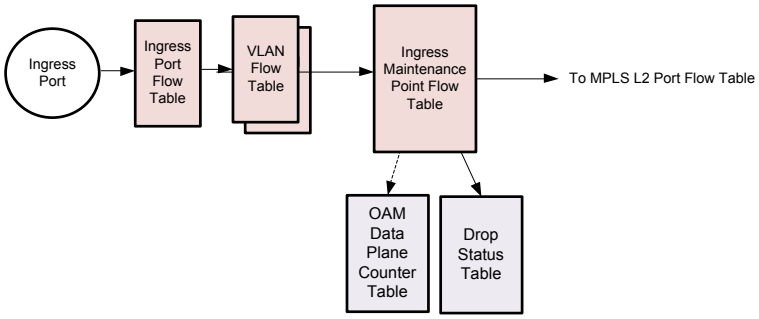

Figure 28 MPLS-TP Initiation - Ethernet Service OAM over MPLS-TP Data Frames

The connection to the OAM Data Plane Counter table is shown as a dotted line to highlight the fact that the count action is performed using the action set. Counters are not actually incremented until the action set is applied after the last match action table in the pipeline. This accommodates the fact that the packet may be dropped before that point due to a policing action.

Figure 29 shows the corresponding pipeline fragment for user data frame tunnel termination at a Maintenance Point. Note that this requires the use of egress tables. Egress tables permit match action processing in the context of an egress port on the payload frame after encapsulation headers are removed.

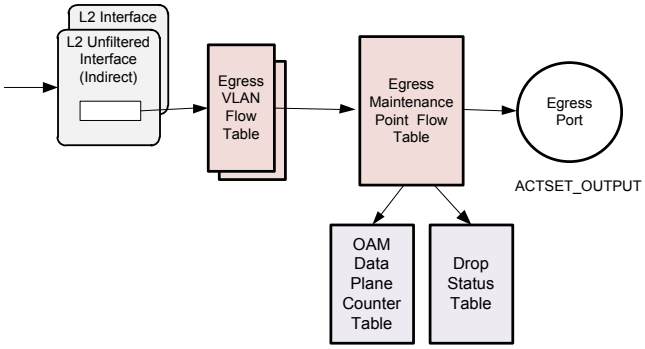

Figure 29 MPLS-TP Termination - Ethernet Service OAM over MPLS-TP Data Frames

Figure 30 shows the objects that handle Client Service MIP, Client Service link Down MEP, and Provider Service Up MEP OAM PDUs for MPLS-TP initiation.

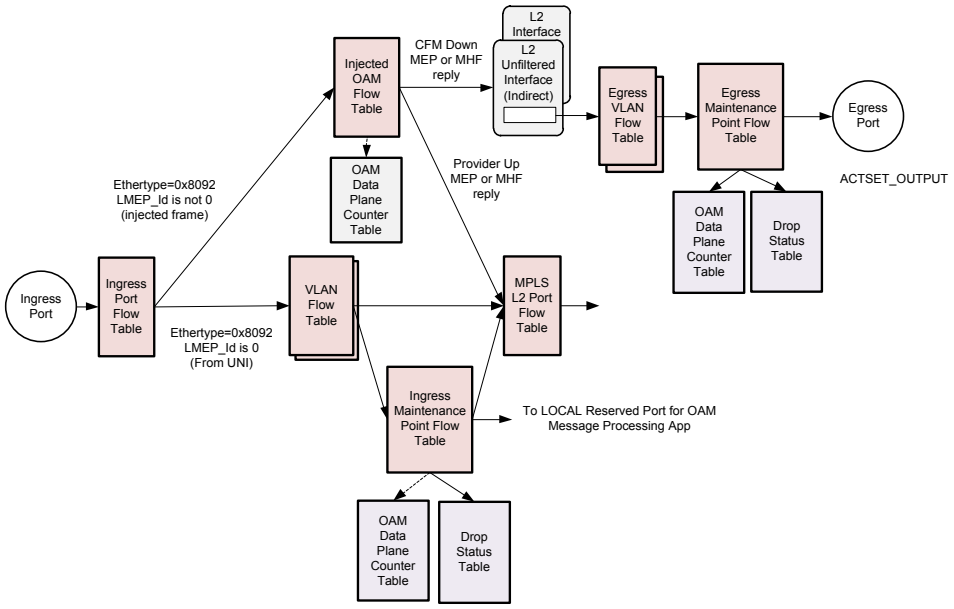

Figure 30 MPLS-TP Initiation - Ethernet over MPLS-TP OAM PDU

Client Service OAM PDU frames from the CE enter the pipeline without an LMEP Id assignment and are classified by the VLAN Flow Table as though customer data frames. If a MIP is configured then the matching rule sets the LMEP Id and forwards the frame to the Ingress Maintenance Point Flow Table for maintenance point processing. This table provides rules that match on the opcode and MDL from the frame and the LMEP Id to decide whether to output the frame for OAM message processing, count and forward frames normally as data, or drop.

The OAM processing engine injects OAM PDU frames via a reserved port along with LMEP Id, Traffic Class, and Color pipeline metadata. Ingress Port Flow Table rules identify these frames by matching ETH_TYPE=0x8902 and non-zero LMEP Id, with a Goto-Table instruction specifying the Injected OAM Flow Table. Rules in the Injected OAM Flow Table can match the LMEP Id and classify these to a customer flow (Up MHF replies) or output to the CE (CFM Down MEP PDUs).

Note: An arrival timestamp value is associated with OAM PDUs. This timestamp is a pipeline metadata value. Its maintenance is a local implementation matter.

A new Set-Counter-Fields action is defined for setting the counter pipeline metadata fields by reading current values in the OAM Data Plane Counter Table for a specified LMEP Id and Traffic Class, and for setting the timestamp metadata from the arrival timestamp value.

Note: Whether implementations update OAM PDUs directly with timestamp and counter values in the data plane is a local matter and not programmed using OpenFlow.

Figure 31 shows the objects that handle CFM Down MEP, Provider Service Up MEP, and Client Service UP MHF OAM PDUs for the MPLS-TP termination case. The processing approach is analogous to the initiation case using the corresponding egress tables.

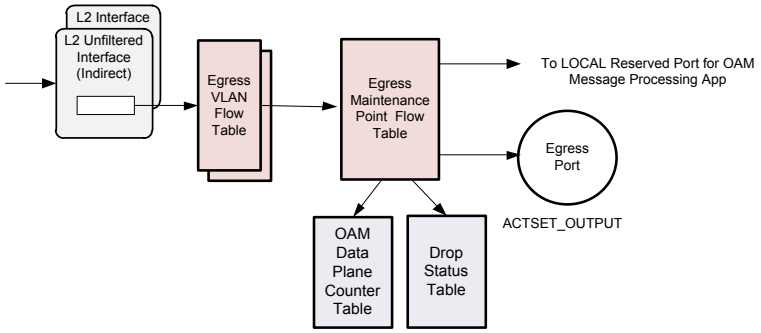

Figure 31 MPLS-TP Termination - Ethernet over MPLS-TP OAM PDU

While egress tables are not allowed to change the output port in the action set, note that they can still include output actions in an action list invoked with the Apply-Actions instruction. The Egress Maintenance Point Flow Table decides where the OAM frame will be processed by forwarding it either to CONTROLLER or LOCAL. It can output the frame or drop it by clearing the action set (which clears ACTSET_OUTPUT) and not providing a Goto-Table instruction.

3.9.3.2 G.8113.1 OAM for MPLS-TP

MPLS-TP OAM methods are described in ITU-T G.8113.1/Y.1372.1 [20]. This standard describes a method for leveraging Ethernet OAM for MPLS-TP. In particular, it describes methods and procedures for applying ITU-T Y.1731 [21] Protocol Data Units (PDUs) for MPLS-TP OAM.

Packet formats for G.8113.1 are shown in Figure 32.

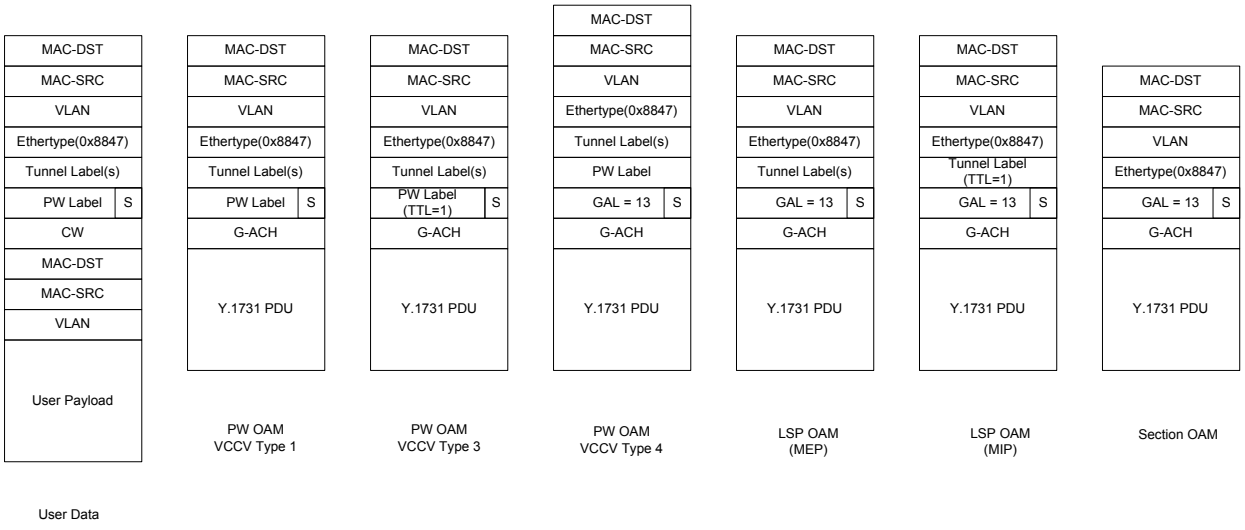

Figure 32 OAM MPLS-TP G.8113.1 Packet Formats

Note that G.8113.1 OAM only requires Down MEPs.

The OF-DPA objects added for OAM data frame processing at pseudo-wire initiation are shown in Figure 33. The differences from the pipeline without OAM mainly consist of adding the Check Drop Status table action, and adding the LM counter action support to the MPLS L2 VPN and MPLS Tunnel Label group entries.

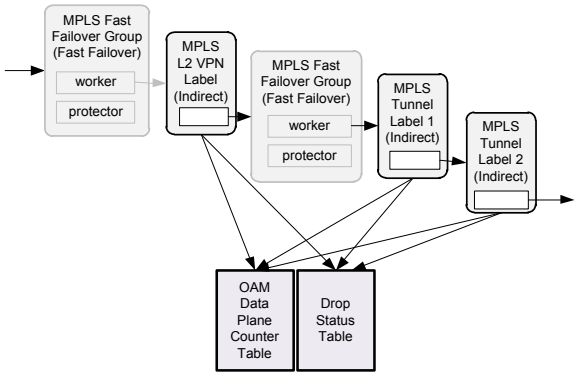

Figure 33 MPLS-TP Initiation - G.8113.1 OAM Data Frame

The OF-DPA objects added for OAM data frame processing at pseudo-wire termination are shown in Figure 34. The difference from the pipeline without OAM is the addition of LM count action to the MPLS table rules for PW and LSP labels.

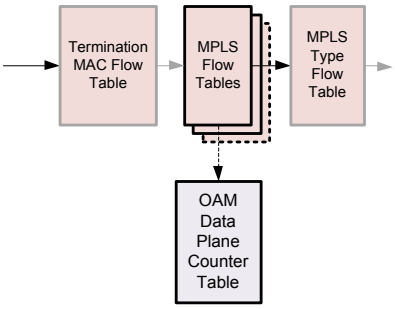

Figure 34 MPLS-TP Termination - G.8113.1 OAM Data Frame

For the LSR data frame Down MEP or MIP processing cases, receive counters are associated with MPLS Flow Table label matching as in Figure 34, and the transmit counters are associated with the MPLS Tunnel Label groups as in Figure 33.

Figure 35 shows the OF-DPA objects involved with injected OAM PDUs. The OAM processing engine injects Down MEP PDU frames through the LOCAL reserved port along with LMEP Id pipeline metadata. These are Y.1731 frames with Ethernet headers. Rules in the Ingress Port Flow Table match these frames by EtherType and non-zero LMEP Id, with Goto-Table instructions specifying the Injected OAM Flow Table. Rules in the Injected OAM Flow Table then match the LMEP Id to determine how to encapsulate the frame. For MPLS-TP initiation an Injected OAM Flow Table rule pushes a GAL (VCCV type 4), PW label, and ACH, and then sets a group for LSP processing, with no next table. Similarly, an Injected OAM Flow Table rule can push labels for LSP and Section OAM and set an MPLS interface group.

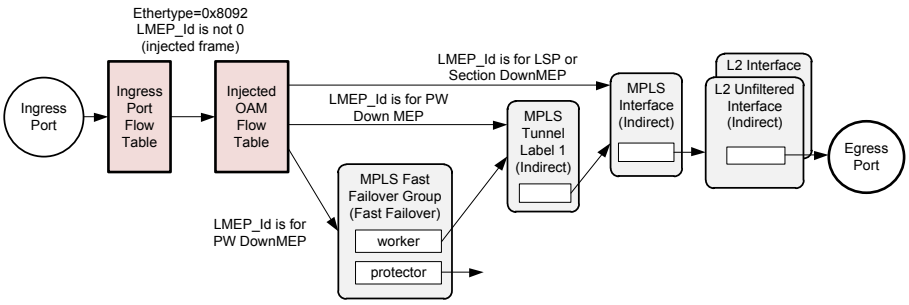

Figure 35 MPLS-TP - G.8113.1 OAM PDU

The pipeline fragment in Figure 36 shows the OF-DPA objects for processing OAM data frames for PW termination, LSP OAM, and Section OAM. For PW termination, OAM PDUs parsed according to the VCCV Type 1 or 3 packet format shown in Figure 32 are matched in the MPLS Flow Tables by PW label and ACH channel type. Similarly, VCCV 4 formats are recognized by the underlying bottom of stack GAL and similarly matched.

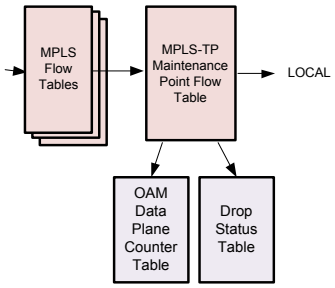

Figure 36 MPLS-TP - G.8113.1 OAM PDU Termination

The OAM message packet formats for LSP and Section OAM are processed in almost the same way, the difference being in the match fields used in MPLS Flow Table rules. For example, LSP MIP detection requires matching on the MPLS-TTL.

Once the MPLS Flow Table rules match an OAM frame, the actions assign the LMEP Id and pop the labels and associated headers, resulting in a Y.1731 frame with its Ethernet header. MPLS-TP Maintenance Point Flow table rules are then used to match the LMEP Id and opcode and determine whether to drop the frame, or LM count and output it to the LOCAL reserved port for processing by the OAM engine.

3.10 Protection Switching

OF-DPA supports 1:1 linear protection using Fast Failover group entry types.

The OpenFlow 1.3.4 specification requires Fast Failover group types to support liveness monitoring to determine which bucket to use for forwarding. Fast Failover groups can be configured with watch_port and watch_group parameters, only one of which is used to determine bucket liveness. Liveness monitoring works as follows:

- A port is considered live if it has the OFPPS_LIVE flag set in its port state. Port liveness may be managed by code outside of the OpenFlow portion of a switch or defined outside of the OpenFlow specification, such as Spanning Tree or a KeepAlive mechanism. The port must not be considered live (and the OFPPS_LIVE flag must be unset) if one of the port liveness mechanisms enabled on the switch consider the port not live, or if the port config bit OFPPC_PORT_DOWN indicates the port is down, or if the port state bit OFPPS_LINK_DOWN indicates the link is down.
- A bucket is considered live if either watch_port is not OFPP_ANY and the port watched is live, or if watch_group is not OFPG_ANY and the group watched is live.
- A group is considered live if at least one of its buckets is live.

OF-DPA introduces OAM Protection Liveness Logical Ports solely for the purpose of controlling the liveness property for OF-DPA MPLS Fast Failover group entry buckets. The local Protection Switching Process controls the liveness property for OAM Protection Liveness Logical Ports and uses this mechanism to cause one or more OF-DPA MPLS Protection Fast Failover groups to switch buckets.

Note: OAM Protection Liveness Logical Port configuration is read only and always reported as all zeroes.

The protection switching process and its linkage to the pipeline for switchover is illustrated in Figure 37.

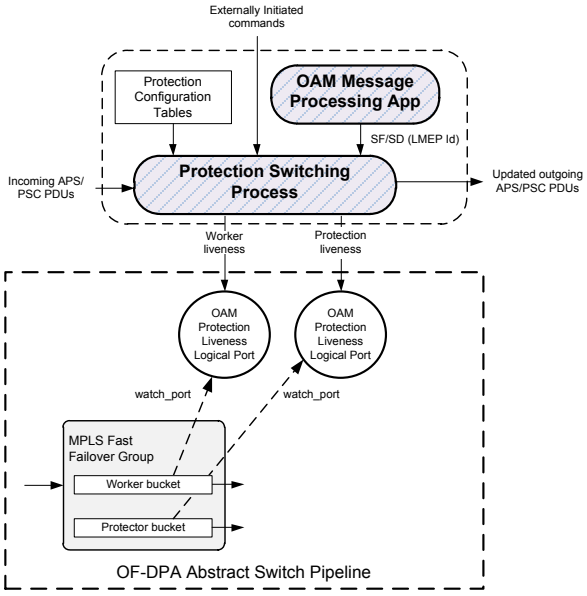

Figure 37 Protection Switching Process

OAM Protection Liveness Logical Ports have the assigned number range shown in Table 216. OAM Protection Liveness Logical Ports come predefined and require no configuration before being referenced. By default they are administratively up.

OF-DPA MPLS Fast Failover group entry buckets should be configured with watch_group OFPG_ANY and watch_port, an OAM Protection Liveness Logical Port.

The Protection Process can control the switchover from worker to protection path by changing the administrative state of an OAM Protection Liveness logical port. The Protection Process is configured with the OAM Protection Liveness logical port for use with a particular LMEP Id.

Note: Both the controller and a local OAM Engine can change the administrative status of an OAM Protection Liveness Logical Port. The local mechanism is implementation dependent.

Note: An OAM Protection Liveness Logical Port cannot be used as an IN_PORT or in an output action.

4 OF-DPA OBJECT DETAILS

OF-DPA 2.0 presents the application writer with a set of objects that can be programmed using OpenFlow 1.3.4. The programmable objects include flow tables, group table entries, and meter table entries. It also presents a set of objects that can be configured using an SDN configuration protocol such as OF-Config. The configurable objects include ports, queues, and OAM functions.

This section provides programming descriptions for these objects. For details consult the OF-DPA 2.0 TTP.

Flow tables have specific attributes, including entry types (rules) that have specific match fields, actions, and instructions. Flow entries can have "Goto-Table" instructions that determine the next table to process the packet. In other words, the flow entry programming determines the order in which packets traverse tables and accumulate actions in an action set. Actions in the action set are applied prior to the packet being forwarded when there is no next table specified. Specific forwarding actions, including egress packet edits, are for the most part included within the action sets of the group entries. OF-DPA 2.0 uses specific types of group entries to support different packet flow scenarios. Apply-actions instructions and action lists are also used for some VLAN tag and MPLS label packet editing, and to send packets to the controller.

In the general OpenFlow case, packets pass from flow table to flow table and can be arbitrarily modified between tables. To take advantage of this generality each table stage would need to include a packet parser. 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 fields that are passed between tables and used for matching or modification by "apply actions" instructions. It is not expected that this distinction will matter to applications.[24]

Section 3 showed the tables, group table entries, and ancillary objects used in the OF-DPA abstract pipeline for different packet processing flows.

Section 4.1 describes the OF-DPA flow tables in terms of their supported match fields, flow entry rule types, instructions, actions, expiration provisions, and statistics counters. In most tables built-in default

[24] In general, for a Packet Out the original ingress packet is sent to the Controller. The Controller would not always see conceptual modifications such as VLAN tag changes.

rules define table miss actions. However if the built-in default rule does not apply the packet is dropped. Section 4.2 describes the OF-DPA group table entry types and action set constraints. Section 4.4 describes the meter table entry types.

Ingress packets always have an associated Tunnel Id metadata value and may have an associated LMEP_Id. Packets that enter the data plane from physical ports always have a zero Tunnel. Packets from tunnel logical ports require a positive, non-zero Tunnel Id value to be assigned in order to identify the tenant forwarding domain. The Tunnel Id logically accompanies the packet through the pipeline so that when tenant packets are output, the Tunnel Id is supplied to the egress logical port.

4.1 Flow Tables

This section provides detailed descriptions of the OF-DPA flow tables in terms of match fields and actions.

4.1.1 Ingress Port Flow Table

The Ingress Port Flow Table is the first table in the pipeline and, by OpenFlow convention, is numbered zero. The Ingress Port Flow Table decides whether to forward the packet using the main pipeline or in an overlay forwarding domain.

The Ingress Port Flow Table presents what is essentially a de-multiplexing logic function as an OpenFlow table that can be programmed from the controller. Three classes of ingress packets can be dispatched:

- Normal packets from physical ports.
- Overlay packets from logical ports with a Tunnel Id.
- Injected OAM packets from the OAM engine.

Note: OF-DPA may prevent certain types of rules from being added to other tables unless there is appropriate flow entry in the Ingress Port Flow Table.

This table has a built-in default rule that matches normal VLAN packets (with zero Tunnel Id and zero LMEP Id) and forwards them to the VLAN Flow Table. The built-in default miss rule effectively prevents further processing on the frame and has a Goto-Table instruction specifying the Policy ACL Flow Table.

4.1.1.1 Flow Entry Types and Match Fields

The Ingress Port Flow Table supports the flow entry types listed in Table 5.

Table 5 Ingress Port Flow Table Entry Types

Type	Description
Normal Ethernet	Matches packets from local physical ports, identified by zero Tunnel Id. Normal Ethernet rules have Goto-Table instructions that can specify either the VLAN Flow Table or one of the Physical Port Trust Flow tables from

Type	Description
	Table 8.
Overlay Tunnel	Matches packets from a data center overlay tunnel logical port by Tunnel Id type. Overlay Tunnel rules have a Goto-Table instruction that specifies either the Bridging Flow Table or one of the Data Center Overlay Trust Flow tables from Table 8. The controller must add a rule or rules of this type before adding rules in other tables that need to match Data Center Overlay Tunnel packets.
OAM PDU Injection	Matches frames with non-zero LMEP Id pipeline metadata. This requires its match priority to be set lower than other rules that exact match on zero LMEP Id values. Goto-Table instruction specifies the Injected OAM Flow Table.

Note: Future versions of OF-DPA may support flow entry types for packets from other types of logical ports.

The Ingress Port Flow Table uses the match fields listed in Table 6. Note that QoS rules must have higher relative priority than overlapping non-QoS rules. DSCP rules must have higher relative priority than PCP rules that match the same flows.

Table 6 Ingress Port Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	Yes	Ingress port. Depending on rule may be omitted to match any IN_PORT. Exact match required in rules with a Goto-Table instruction specifying a QoS Trust Flow Table.
TUNNEL ID	32	No	No	Required non-zero in Data Center Overlay rules. Must be exact match in order to distinguish frames for different tenants. Must be zero in Normal Ethernet rules.
ETH_TYPE	16	No	Yes	Satisfies required prerequisite in rules with a Goto-Table instruction that specifies a Port DSCP Trust Flow Table. Only allowed values are 0x0800 and 0x86dd. Must be omitted in other rule types.
LMEP Id	32	No	No	Must be supplied as zero for other than injected OAM frames.

Note: If separate IPv4 and IPv6 classification rules are entered for the same IN_PORT, both actions sets must specify the same QoS Index value.

Note: The range of QoS Index values may vary by table and platform.

4.1.1.2 Instruction Types

The Ingress Port Flow Table supports the instructions listed in Table 7.

Table 7 Ingress Port Flow Table Instructions

Name	Argument	Description
Goto-Table	Table	Depending on rule type, one of: VLAN Flow Table; Bridging Flow Table; Injected OAM Flow Table; or a QoS Trust Flow Table (see Table 8)
Apply-Actions	Action list	Can contain at most one instance of each of the actions listed in Table 9.

Table 8 Ingress Port Flow Table QoS Next Table by Entry Type

Rule Type	Next Table	Description
Normal Ethernet IP DSCP	Physical Port DSCP Trust Flow Table	Rule must include the IP EtherType prerequisite and specify a QoS Trust Index
Normal Ethernet PCP	Physical Port PCP Trust Flow Table	Rule must also specify QoS Trust Index
Overlay Tunnel IP DSCP	Data Center Overlay DSCP Trust Flow Table	Rule must include the IP EtherType prerequisite and specify a QoS Trust Index
Overlay Tunnel PCP	Data Center Overlay PCP Trust Flow Table	Rule must also specify QoS Trust Index

4.1.1.3 Actions

The Ingress Port Flow Table action list can optionally assign a non-zero VRF value to the packet. OF-DPA defines VRF as a new pipeline metadata field that is used by L3 VPN routing. The VRF defaults to zero if not set. The Ingress Port Flow Table action list can also optionally set the packet QoS marking trust properties.

Table 9 Ingress Port Flow Table Action List

Name	Argument	Description
Set-Field	QoS Index	Sets 8-bit index value for PCP or DSCP trust table lookup. Optional, but required if next table is a QoS Trust table.
Set-Field	VRF	VRF for L3 lookups. Only applicable to Normal Ethernet Frame rules. Optional.

4.1.1.4 Counters and Flow Expiry

The Ingress Port Flow Table supports the basic table and flow entry counters listed in Table 10.

Table 10 Ingress Port 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 timeout ageing per entry is supported, as indicated in Table 11.

Table 11 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.

4.1.2 Injected OAM Flow Table

The Injected OAM Flow Table is used to process injected OAM frames matched in the Ingress Port Flow Table. These OAM frames are Ethernet frames containing Y.1731 PDUs and ingress with LMEP Id metadata. The Injected OAM Flow Table matches LMEP Id and forwards the frame to the appropriate point in the pipeline.

This table has a built-in default drop miss rule with a Clear-Actions instruction.

4.1.2.1 Flow Entry Types and Match Fields

The Injected OAM Flow Table supports the flow entry types listed in Table 12.

Table 12 Injected OAM Flow Table Entry Types

Type	Description
Ethernet Service Up MEP or MHF	Sets MPLS L2 Port and Tunnel Id for the frame with a Goto-Table instruction specifying the MPLS L2 Port Flow Table. Traffic Class and Color are as accompanies the frame at input. Provider OAM frame to peer PE or MHF reply to remote CE partner.
Section	Pushes G-ACH and GAL. Group is MPLS Interface with no Goto-Table instruction. Frame is forwarded to group.
LSP Down MEP or MHF	Pushes G-ACH, GAL, and tunnel label with no Goto-Table instruction. Group is MPLS Interface. Frame is forwarded to group. LSP OAM frame or MHF reply to peer.
PW VCCV Type 1 or 3	Pushes G-ACH and bottom-of-stack pseudo-wire label with no Goto-Table instruction. Group is either Tunnel 1 or MPLS Fast Failover. Frame is forwarded to group. PW OAM frame or MHF reply to peer.
VCCV Type 4	Pushes G-ACH, GAL, and pseudo-wire label with no Goto-Table instruction. Group is either Tunnel 1 or MPLS Fast Failover. Frame is forwarded to group. PW OAM frame or MHF reply to peer.

The Injected OAM Flow Table uses the match fields listed in Table 13.

Table 13 Injected OAM Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
LMEP Id	32	No	No	LMEP Id assigned by OAM Processing Engine. Required.
ETH_TYPE	16	No	No	Must be 0x8902.

4.1.2.2 Instruction Types

The Injected OAM Flow Table supports the instructions listed in Table 14.

Table 14 Injected OAM Flow Table Instructions

Name	Argument	Description
Goto-Table	Table	Only option is the MPLS L2 Port Flow Table. This is used for Up MEP and MHF frames.
Apply-Actions	Action list	Can contain at most one instance of each of the actions listed in Table 15.

4.1.2.3 Actions

The actions used in Injected OAM Flow Table flow entries are listed in Table 15.

Table 15 Injected OAM Flow Table Action List

Name	Argument	Description
Set-Field	MPLS L2 Port	Required in Ethernet Service Up MEP or MHF rule, not used otherwise.
Set-Field	Tunnel Id	Required in Ethernet Service Up MEP or MHF rule, not used otherwise.
Push-MPLS	EtherType	Must be 0x8847.
Set-Field	MPLS Label	
Set-Field	MPLS-BOS	
Set-Field	MPLS_TC	Optionally set MPLS TC field value to a constant value, can depend on Traffic Class and Color pipeline match fields supplied with packet.
Set MPLS TC From Tunnel Table	QoS Index	Optionally set MPLS TC on a Section or LSP label using the MPLS Tunnel Label Remark Action Table.
Set MPLS TC From VPN Table	QoS Index	Optionally set MPLS TC on a Pseudo-wire label using the MPLS VPN Label Remark Action Table.
Set-Field	MPLS-TTL	

Name	Argument	Description
Push-CW	-	Pushes zero word between bottom of stack label and payload. Will be converted into a G-ACH.
Set-Field	MPLS Data First Nibble	Set first nibble to 1 to make into an ACH
Set-Field	MPLS ACH Channel	Sets channel to 0x8902 for G.8113.1 OAM.

4.1.2.4 Counters and Flow Expiry

The Injected OAM Flow Table supports the basic table and flow entry counters listed in Table 16.

Table 16 Injected OAM 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 timeout ageing per entry is supported, as indicated in Table 17.

Table 17 Injected OAM 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.

4.1.3 VLAN Flow Table

The VLAN Flow Table is used for IEEE 801.Q VLAN assignment and filtering to specify how VLANs are to be handled on a particular port.[25] All packets must have an associated VLAN Id in order to be processed by subsequent tables.

[25] The VLAN Flow Table presents the hardware port and VLAN configuration bitmaps to the OpenFlow controller as a flow table.

Packets that do not match any entry in the VLAN table are dropped by a default miss rule with Clear Actions and Goto Table - Policy ACL instructions.

Note that IEEE defined BPDUs are always received untagged.[26]

The VLAN Flow Table can optionally set a VRF pipeline metadata value.

4.1.3.1 Flow Entry Types and Match Fields

The VLAN Flow Table supports the Flow Entry Types listed in Table 18. Flow entries are differentiated based on IN_PORT, whether or not the packet was tagged, and the VLAN Id in the tag. In addition OAM processing rules match packets by packet EtherType and MAC-DST.

Since version 1.0 OpenFlow has used a 16-bit field for VLAN Id. Since only the low order 12 bits are needed, OpenFlow defined special match field values to represent tagged and untagged packets. In particular, the VLAN Id 0x0000 (OFPVID_NONE, defined in the OpenFlow specification) is used for 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[27] (OFPVID_PRESENT | VLAN Id value). This convention must be followed in programming rules from the controller, both for match fields and for setting the VLAN Id field in an action. For further explanation consult the OpenFlow 1.3.4 specification.

The VLAN Flow Table participates in the implementation of Ethernet OAM maintenance points. For this the Goto-Table instruction must be programmed to specify the Ingress Maintenance Point Flow Table.

Note: OpenFlow does not intrinsically provide the ability to distinguish between S-Tags and C-Tags for IEEE 802.1Q. In OF-DPA, the convention followed is that an S-Tag always has EtherType 0x88a8 and a C-Tag always has EtherType 0x8100. A single tagged packet might be either, but for consistent results a double tagged packet outermost tag should always be an S-Tag and the inner tag a C-Tag. OF-DPA supports at most two tags.

Table 18 VLAN Flow Table Flow Entry Types

Type	Description
VLAN Filtering	Exact match on IN_PORT and VLAN_VID parsed from the packet. Matches packets with a VLAN_VID greater than zero. Must be programmed with VLAN_VID | OFPVID_PRESENT and no mask. The only instruction is Goto-Table and must specify the Termination MAC Flow Table. Tagged packets that do not match any rule are treated as VLAN_VIDs that are not allowed on the port and are dropped. Can optionally assign a VRF for routed packets.
Port VLAN	Exact match on IN_PORT and OFPVID_NONE. Action set must push a VLAN

[26] There are vendor specific BPDUs that are VLAN tagged.
[27] Although accepted by OF-DPA, IEEE 802.1Q indicates VLAN Id 4095 is a reserved value.

Type	Description
Assignment (Untagged Packet)	tag and assign a valid VLAN_VID. Optionally it can push another tag and set a new outermost VLAN Id. Note that before installing the Port VLAN Assignment rule there must already be a VLAN Filtering rule for the IN_PORT and outermost VLAN_VID. A Port VLAN Assignment Rule must have a Goto-Table instruction specifying the Termination MAC Flow Table. Untagged packets may be dropped if there is no VLAN assignment rule for the ingress port. Can optionally assign a VRF for routed packets.
Port VLAN Assignment (Priority Tagged Packet)	Exact match on IN_PORT and OFPVID_PRESENT. Action set must assign a valid VLAN_VID. Optionally it can push another tag and set a new outermost VLAN Id. Note that before installing the Port VLAN Assignment rule there must already be a VLAN Filtering rule for the IN_PORT and outermost VLAN_VID. A Port VLAN Assignment Rule must have a Goto-Table instruction specifying the Termination MAC Flow Table. Priority tagged packets may be dropped if there is no VLAN assignment rule for the ingress port. Can optionally assign a VRF for routed packets.
VLAN Allow All	Wildcard VLAN match for a specific IN_PORT. Essentially turns off VLAN filtering and/or assignment for a physical port, but must be programmed with both VLAN_VID and mask equal to OFPVID_PRESENT. Must have lower priority than any overlapping translation, filtering, MPLS, or VLAN assignment rule. A corresponding L2 Unfiltered Interface group entry should also be programmed for the port.
VLAN Translate Single Tag or Single Tag to Double Tag	Used to either modify the VLAN Id on a single tagged packet, or to optionally modify the VLAN Id and then push another tag onto a single tagged packet. Can also optionally assign a VRF for routed packets.
VLAN Translate Double Tag to Single Tag	Used in conjunction with the VLAN 1 Flow Table to pop the outer tag and optionally modify the remaining tag. Can also optionally assign a VRF for routed packets.
MPLS L2 Single Tag	Uses IN_PORT and VLAN Id to classify a packet for MPLS-TP pseudo-wire initiation and assign a Tunnel_Id and MPLS L2 Port. If the flow is to be enabled for Ethernet OAM then the Goto-Table instruction should specify the Ingress Maintenance Point Flow Table, otherwise the MPLS L2 Flow Table. The frame can be unchanged, the VLAN Id changed, or the tag removed. The rule also sets the value of the MPLS Type pipeline field, which in this version of OF-DPA must be VPWS.
MPLS L2 Double Tag	Used in conjunction with the VLAN 1 Flow Table to classify a double tagged packet to a pseudo-wire. Sets the OVID pipeline field to the outer VLAN Id before popping it in order to expose the inner tag for match by VLAN 1 Flow Table.

Type	Description
MPLS L2 Untagged	Used to classify an untagged packet to for MPLS-TP pseudo-wire initiation and assign a Tunnel_Id and MPLS L2 Port. If enabled for Ethernet OAM the Goto-Table instruction should specify the Ingress Maintenance Point Flow Table, otherwise the MPLS L2 Flow Table. The frame can be unchanged or a VLAN tag pushed and a VLAN Id assigned. The rule also sets the value of the MPLS Type pipeline field, which in this version of OF-DPA must be VPWS.
MPLS L2 Untagged (Priority Tagged)	Used to classify a priority tagged packet for MPLS-TP pseudo-wire initiation and assign a Tunnel_Id and MPLS L2 Port. If enabled for Ethernet OAM the Goto-Table instruction should specify the Ingress Maintenance Point Flow Table, otherwise the MPLS L2 Flow Table. The frame can be unchanged, the VLAN Id can be set, or the tag removed. The rule also sets the value of the MPLS Type pipeline field, which in this version of OF-DPA must be VPWS.
MPLS L2 All Traffic on Port	Used to classify all packets on a port to an MPLS-TP pseudo-wire and assign a Tunnel_Id and MPLS L2 Port. Must be higher relative priority than any overlapping rule. If enabled for Ethernet OAM the Goto-Table instruction should specify the Ingress Maintenance Point Flow Table, otherwise the MPLS L2 Flow Table. The rule does not support any actions to modify the frame. It does set the value of the MPLS Type pipeline field, which in this version of OF-DPA must be VPWS.

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

Note: Untagged and priority tagged rules for the same IN_PORT must assign the same VLAN Id values in the same relative positions. That is, a priority tagged rule and an untagged rule for the same port must set the same (outermost and innermost) VLAN Id(s). Furthermore, an MPLS L2 untagged rule and a VLAN assignment rule for the same port must assign the same VLAN Id. In other words, there can only be one outermost and one innermost port VLAN Id.

Note: The VLAN Id assigned by a VLAN assignment rule must have already been allowed by a VLAN Filtering rule. The presence of the VLAN filtering rule must be treated as a prerequisite.

The VLAN Flow Table match fields are listed in Table 19.

Table 19 VLAN Flow Table Match Fields

Field	Bits	Maskable	Optional	Description

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	No	Ingress port. Must be a physical port (high order 16 bits zero).
VLAN_VID	16	Yes	No	Outer VLAN Id. The mask value can only be either 0x1fff for VLAN filtering rules and OFPVID_PRESENT for allow all VLAN packet rules. Must be exact for other rules.

4.1.3.2 Instruction Types

The VLAN table supports the instruction types listed in Table 20.

Table 20 VLAN Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	The VLAN Flow Table supports the actions specified in Table 21.
Goto-Table	Table	For VLAN filtering or Port VLAN assignment the next table should be the Termination MAC Flow Table, but if an Ethernet Service OAM MIP or MEP is programmed the next table would be the Maintenance Point Flow Table. For VLAN translation or MPLS double tag classification the next table can be the VLAN 1 Flow Table. A packet is dropped if it matches an entry that has no Goto-Table instruction.

4.1.3.3 Actions

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

Table 21 VLAN Flow Table Action List Actions

Name	Argument	Description
Set-Field	VLAN_VID, must be between 1 and 4094.	Sets the VLAN Id on the currently outermost VLAN tag.
Set-Field	VRF	Optionally sets the VRF pipeline field. VRF must be the same in all rules for the same VLAN.

Name	Argument	Description
Set-Field	OVID	Pipeline metadata field representing an outer tag VLAN Id that was popped, so that it can be used as a match field in the VLAN 1 Flow Table. Only required in flow entries with Goto-Table instructions specifying the VLAN 1 Flow Table. Must be programmed with OFPVID_PRESENT.
Set-Field	MPLS_TYPE	Pipeline metadata field used to distinguish MPLS flow type.
Push VLAN	EtherType	Used in translating single to double tag. EtherType (TPID) must be 0x8100[28].
Pop VLAN		Used in processing double tagged frames, where the Goto-Table instruction specifies the VLAN 1 Flow table.
Set-Field	MPLS L2 Port	For MPLS-TP pseudo-wire classification.
Set-Field	Tunnel_Id	For MPLS-TP pseudo-wire classification.

Note: The untagged packet action is similar to what was used in OpenFlow 1.0.

4.1.3.4 Counters and Flow Expiry

The VLAN Flow Table supports the table and flow entry counters listed in Table 22.

Table 22 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 timeout ageing per entry is supported, as indicated in Table 23.

[28] Outermost TPID can be set using the Egress TPID Flow Table, described in Section 4.2.5.

Table 23 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.

4.1.4 VLAN 1 Flow Table

The VLAN 1 Flow Table is used for double tag matching and actions. OpenFlow is defined such that it only can match an outermost VLAN tag. For OF-DPA, a VLAN Flow Table rule copies the VLAN Id to a pipeline metadata field (OVID), pops the tag, and does a Goto-Table specifying VLAN 1. Rules in the VLAN 1 Flow Table can now match both VLAN tags. Actions are similar to the single tag VLAN case.

The VLAN 1 Flow Table has a built-in default rule to restore the frame (i.e., push back the OVID tag) and a Goto-Table instruction specifying the Policy ACL Flow Table.

4.1.4.1 Flow Entry Types and Match Fields

The VLAN 1 Flow Table supports the Flow Entry Types listed in Table 24.

Table 24 VLAN 1 Flow Table Flow Entry Types

Type	Description
Stacked VLAN Assignment	Exact match on IN_PORT, VLAN_VID, and OVID. Can optionally change the outermost VLAN_VID, or change and push another outermost tag, which could be set to OVID. Must have a Goto-Table instruction that specifies the Termination MAC Flow Table. Can also assign a VRF.
MPLS L2 Stacked VLAN Assignment	Exact match on IN_PORT, VLAN_VID, and OVID. Can optionally change the VLAN_VID, or change and push another outermost tag. Must assign an MPLS L2 Port value. Must have a Goto-Table instruction that specifies the MPLS L2 Port table or the MPLS Maintenance Point table.

The VLAN 1 Flow Table match fields are listed in Table 25.

Table 25 VLAN 1 Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	No	Ingress port. Must be a physical port.

Field	Bits	Maskable	Optional	Description
VLAN_VID	16	No	No	Inner VLAN Id. Must be exact. Must be programmed with OFPVID_PRESENT bit set.
OVID	16	No	No	Outer VLAN_VID, previously set by a VLAN table flow entry. Must be exact and programmed with OFPVID_PRESENT bit set.

4.1.4.2 Instruction Types

The VLAN table supports the instruction types listed in Table 26.

Table 26 VLAN 1 Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	The VLAN 1 Flow Table supports the actions specified in Table 27.
Goto-Table	Table	For VLAN assignment must be the Termination MAC Flow Table. For MPLS-TP must be either the MPLS L2 Port table or Ingress Maintenance Point table.

4.1.4.3 Actions

The VLAN 1 Flow Table action list actions are as shown in Table 27.

Table 27 VLAN 1 Flow Table Action List Actions

Name	Argument	Description
Set-Field	VLAN_VID	
Set-Field	VRF	Optionally sets the VRF pipeline field. VRF must be the same in all rules for the same VLAN.
Push VLAN	EtherType	Used in translating single to double tag. EtherType (TPID) must be 0x8100
Pop VLAN		Used in processing double tagged frames.
Set-Field	MPLS L2 Port	For pseudo-wire classification.

Name	Argument	Description
Set-Field	Tunnel_Id	For pseudo-wire classification.
Set-Field	MPLS_TYPE	Marks flow type for MPLS-TP flows, currently only VPWS.

4.1.4.4 Counters and Flow Expiry

The VLAN 1 Flow Table supports the table and flow entry counters listed in Table 28.

Table 28 VLAN 1 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 timeout ageing per entry is supported, as indicated in Table 29.

Table 29 VLAN 1 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.

4.1.5 Ingress Maintenance Point Flow Table

The Ingress Maintenance Point Flow Table is used to instrument Ethernet MIPs and/or MEPs. All packets for an OAM enabled Ethernet Service flow should pass through this table, both control and data.

For Ethernet OAM control frames, rules in this table determine whether and how to process by matching on the EtherType, ingress port, and VLAN Id, along with the MDL and opcode parsed from the Y.1731 PDU. Options are to assign an LMEP Id and output the packet to the LOCAL reserved port, or to treat the same as a data frame. For data frames, rules in this table can apply loss measurement counters and then forward the packet for further pipeline processing.

A built in default rule has a Goto-Table instruction specifying the MPLS L2 Port Flow Table to pass through frames on miss. These should usually be data frames, assuming the table is programmed correctly.

4.1.5.1 Flow Entry Types and Match Fields

The Maintenance Point Flow Table implements the single flow entry type listed in Table 30.

Table 30 Ingress Maintenance Point Flow Table Entry Types

Name	Description
MIP LTM Process (our MDL)	Used to process Link Trace packets at the MIP level that are addressed to the local MIP MAC address.
MIP LBM Process (our MDL)	Used to process Loopback packets at the MIP level.
PDU Passthru higher MDL)	Treat higher level OAM frames as data.
PDU Drop (lower MDL)	Actively filter lower level OAM frames.

The Maintenance Point Flow Table match fields are listed in Table 31.

Table 31 Ingress Maintenance Point Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	Yes	Required to match OAM frames, and must be 0x8902. Must be omitted in rules that match data frames, and these rules must be lowest priority to avoid overlap.
VLAN_VID	16	No	Yes	Used in conjunction with the ingress port to match a particular user service flow. VLAN Id must be specified for port+VLAN service. When VLAN Id is specified it must be programmed with OFPVID_PRESENT.
IN_PORT	32	No	No	Used in conjunction with the VLAN Id to match a particular user service flow, or to match all packets on a port.
ETH DST	48	No	Yes	Destination MAC, used in "MIP LBM Process (our MDL)" rules to filter Loopback packets for a local MIP. Must be omitted otherwise.

Field	Bits	Maskable	Optional	Description
OAM_Y1731_OPCODE	8	No	Yes	Parsed from the IEEE 802.1ag/Y.1731 header.
OAM_Y1731_MDL	3	No	Yes	Parsed from the IEEE 802.1ag/Y.1731 header.

4.1.5.2 Instruction Types

The Maintenance Point Flow Table can have the instructions shown in Table 32. At least one must be specified. Since there is no next table the packet is dropped.

Table 32 Ingress Maintenance Point Flow Table Instructions

Name	Argument	Description
Clear Actions		Used to drop frames, usually in conjunction with an output to LOCAL action.
Apply Actions	Action List	Optional, Action List Actions are listed in Table 33.
Write Actions	Action Set	Optional, Action Set Actions are listed in Table 34
Goto Table	MPLS L2 Port	Optional, for data frames, or for higher MDL OAM frames that must be treated as data.

4.1.5.3 Actions

Table 33 Ingress Maintenance Point Flow Table Action List

Name	Argument	Description
Set-Field	LMEP Id	Identifies the particular maintenance point for processing the frame.

Table 34 Ingress Maintenance Point Flow Table Action Set

Name	Argument	Description
Output	Port	LOCAL, for processing by the local OAM engine with the indicated LMEP Id. Used for OAM frames needing to be processed at this Maintenance Point.

4.1.5.4 Counters and Flow Expiry

The Maintenance Point Flow Table counters are listed in Table 35.

Table 35 Ingress Maintenance Point 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

Maintenance Point Flow Table expiry provisions are as indicated in Table 36.

Table 36 Ingress Maintenance Point 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.

4.1.6 MPLS L2 Port Flow Table

The MPLS L2 Port Flow Table can be used for QoS classification in conjunction with MPLS tunnel origination. This table has a built-in default rule with a Goto-Table instruction specifying the L2 Policer Flow table.

Table 37 MPLS L2 Port Metadata Naming Convention

Type	Numbering	Description
Local	0x0000nnnn	Local client interface (UNI). Assigned by VLAN table.
Network	0x0002nnnn	Network interface (NNI). Assigned by terminating MPLS label.

It is recommended that this programming convention be followed even though it is not needed for VPWS flows. In the TTP file, this convention is incorporated in the variable declarations for <local_mpls_l2_port> and <network_mpls_l2_port>.

4.1.6.1 Flow Entry Types and Match Fields

The MPLS L2 Flow Table supports the Flow Entry Types listed in Table 38.

Table 38 MPLS L2 Port Flow Table Flow Entry Types

Type	Description
MPLS V4 DSCP and **MPLS V4 DSCP**	Classify IPv4 or IPv6 packets based on DSCP. ETH_TYPE prerequisite is required. Must set QoS Index and have a Goto-Table instruction that specifies the MPLS DSCP Trust Flow Table. Should have higher relative priority assignment than overlapping rules, or PCP rules that match the same packet.
MPLS PCP	Classify packets based on Ethernet header PCP. Must set QoS Index and have a Goto-Table instruction that specifies the MPLS PCP Trust Flow Table. Should have higher relative priority assignment than overlapping wildcard rules.

The MPLS L2 Port Flow Table match fields are listed in Table 39. Note that since matching both MPLS L2 Port and Tunnel Id will be required to program future VPLS flows, both fields are used for VPWS as well.

Table 39 MPLS L2 Port Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
MPLS L2 Port	32	No	No	Identifies a logical ingress point for the customer flow.
Tunnel Id	32	No	No	Identifies the customer flow across more than one logical ingress point.
ETH-TYPE	16	No	Yes	Prerequisite for IP header DSCP rules, where it is required. Can only be 0x0800 for IPv4 or 0x86dd for IPv6.

4.1.6.2 Instruction Types

The MPLS L2 Port table supports the instruction types listed in Table 40.

Table 40 MPLS L2 Port Flow Table Instructions

Name	Argument	Description

Name	Argument	Description
Apply-Actions	Action List	The supported actions are specified in Table 41.
Write-Actions	Action Set	Only used for VPWS. Only action is Group, which must indicate one of: MPLS L2 VPN Label or Fast Failover Protection Group.
Goto-Table	Table	Must specify one of MPLS L2 Port DSCP or MPLS L2 Port PCP. L2 Policer is only used in the built-in default rule.

4.1.6.3 Actions

The MPLS L2 Port table uses the actions as listed in Table 41.

Table 41 MPLS L2 Port Flow Table Action List

Name	Argument	Description
Set-Field	QoS Index	Required if next table is one of: MPLS L2 Port DSCP or MPLS L2 Port PCP.

4.1.6.4 Counters and Flow Expiry

The MPLS L2 Port Flow Table supports the table and flow entry counters listed in Table 42.

Table 42 MPLS L2 Port 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 timeout ageing per entry is supported, as indicated in Table 43.

Table 43 MPLS L2 Port Flow Table Expiry

Name	Bits	Description

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.

4.1.7 QoS DSCP and PCP Trust Flow Tables

The Trust Flow Tables are used for ingress packet QoS classification depending on position in the pipeline. They can flexibly map incoming packet header field values to QoS Traffic Class and Color metadata values according to different trust profiles.

There are two basic types of QoS Trust Flow Tables, depending on whether the Traffic Class and Color are to be determined from the IP header or from the 802.1p header. DSCP Trust Flow Tables map DSCP field values from IP headers. PCP Trust Flow Tables map 802.1p PCP and DEI values.

There are three sets of QoS Trust Flow Tables. One is used to classify normal Ethernet frames from local ports. Another is used for Data Center Overlay frames from Data Center Overlay logical ports. The third is used for MPLS pseudo-wire origination frames from an MPLS L2 Port.

4.1.7.1 Flow Entry Types and Match Fields

The QoS DSCP and PCP Trust Flow Table match fields are similar across table types and are listed in Table 44 (for DSCP) and Table 45 (for PCP).

Table 44 QoS DSCP Trust Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
QoS Index	8	No	No	Index value for selecting a DSCP Trust profile
ETH-TYPE	16	No	No	Prerequisite for IP header DSCP field. Only values can be 0x0800 for IPv4 or 0x86dd for IPv6.
IP DSCP	8	No	No	IP header DSCP value

Table 45 QoS PCP Trust Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
QoS Index	8	No	No	Index value for selecting a PCP Trust profile from the Ingress Port Flow Table.

Field	Bits	Maskable	Optional	Description
VLAN_VID	16	Yes	No	Only used as prerequisite - OFPVID_PRESENT must be set.
PCP	3	No	No	802.1p header PCP value
DEI	1	No	No	802.1p header DEI value

4.1.7.2 Instruction Types

The QoS DSCP and PCP Trust Flow Tables support the instruction types listed in Table 46.

Table 46 QoS DSCP and PCP Trust Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	Actions are from Table 47.
Goto-Table	Table	Specifies the VLAN Flow Table for physical Port Trust Tables. Specifies the Bridging Flow Table for Overlay Trust Tables. Specifies the MPLS Type Flow Table for MPLS L2 Port Trust Tables.

4.1.7.3 Actions

All QoS DSCP and PCP Trust Flow Tables have the same Action List actions for setting QoS Traffic Class and Color, as listed in Table 47. Both are optional but at least one must be specified.

By convention, larger values indicate higher priority Traffic Class assignments. Color values must be one of: green (0); yellow (1); red (2); reserved (3).

Table 47 QoS DSCP and PCP Trust Flow Table Action List

Name	Argument	Description
Set-Field	Traffic Class	Sets 4-bit traffic class value used in subsequent QoS processing. Overrides default port Traffic Class if specified. Optional.
Set-Field	Color	Sets 2-bit Color value used in subsequent QoS processing. Optional.

4.1.7.4 Counters and Flow Expiry

All QoS DSCP and PCP Trust Flow Tables support the table and flow entry counters in Table 48.

Table 48 QoS DSCP and PCP Trust 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

QoS DSCP and PCP Trust Flow Tables support only hard interval timeout ageing per entry as indicated in Table 49.

Table 49 QoS DSCP and PCP Trust 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.

4.1.8 L2 Policer Flow Table

The L2 Policer Flow Table provides a large bank of meters that can be applied to MPLS-TP UNI to NNI customer flows.

On a miss the built in default rule has a Goto-Table instruction that specifies the MPLS Type Flow Table. On miss no meter is applied.

Note: Platforms that do not provide L2 Policer meters only support the built-in default rule.

4.1.8.1 Flow Entry Types and Match Fields

Table 50 L2 Policer Flow Table Entry Types

Type	Description
MPLS L2 Port Meter	Apply meter based on MPLS customer flow identified by MPLS L2 Port and Tunnel Id.

Table 51 L2 Policer Flow Table Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite
Tunnel Id	32	No	No	Identifies particular MPLS L2 Port service flow.
MPLS_L2_PORT	32	No	No	Must only match local MPLS L2 Port types.

4.1.8.2 Instructions

Table 52 L2 Policer Flow Table Instructions

Name	Argument	Description
Apply Actions	Action Set	Actions are as in Table 53.
Goto-Table	Table	One of: the L2 Policer Actions Flow Table (if a meter is applied), otherwise the MPLS Type Flow Table.
Meter	Meter id	Meter to apply. As a result of meter application the packet Color may change.

4.1.8.3 Action List Actions

The following actions can be used in an L2 Policer Flow Table action list.

Table 53 L2 Policer Flow Table Actions

Name	Argument	Description
Set-Field	Color Actions Index	Index used to select a color based action in the L2 Policer Actions Flow Table. Required if meter is applied.

4.1.8.4 Counters and Flow Expiry

The L2 Policer Flow Table counters are listed in Table 54.

Table 54 L2 Policer Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Number of active flow entries in the table

Name	Bits	Type	Description
Duration (sec)	32	Per-entry	Seconds since this flow entry was installed

The L2 Policer Flow Table expiry provisions are as indicated in Table 55.

Table 55 L2 Policer 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.

4.1.9 L2 Policer Actions Flow Table

Rules in the L2 Policer Actions Flow Table match on the QoS Index assigned by the L2 Policer Table and the Color pipeline field from the meter application and then specify header QoS remark actions. This logically is part of the metering action. However in OpenFlow a separate table is required since the Color is assigned by the meter instruction and only available for matching after the L2 Policer Flow Table.

Note: Rules in this table must always set Color Based Actions index back to zero, as it this field will also be used as a profile index for Policy ACL Color Based Actions.

The L2 Policer Actions Flow Table has a built-in default Goto-Table rule specifying the MPLS Type Flow Table on a miss.

4.1.9.1 Flow Entry Types and Match Fields

Table 56 L2 Policer Actions Flow Table Entry Types

Type	Description
Red Actions	Action set to apply to a red packet for this Color Actions Index.
Yellow Actions	Action set to apply to a yellow packet for this Color Actions Index.
Green Actions	Action set to apply to a green packet for this Color Actions Index.

Table 57 L2 Policer Actions Flow Table Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite

Field	Bits	Maskable	Optional	Description or Prerequisite
Color	2	No	No	Packet color.
Color Actions Index	32	No	No	Index set by L2 Policer Flow Table actions.

4.1.9.2 Instructions

Table 58 L2 Policer Actions Flow Table Instructions

Name	Argument	Description
Write Actions	Action Set	Actions in Table 60 can be used in an Action Set.
Goto-Table	Table	One of: MPLS Type Flow Table, or (if dropping the packet) the Policy ACL Flow Table.
Apply-Actions	Action List	Actions in Table 59 can be used in an Action List.
Clear Actions		Optional, can be used when dropping a packet.

4.1.9.3 Actions

The actions in Table 59 can be used in an Action List.

Table 59 L2 Policer Actions Flow Table Action List Actions

Name	Argument	Optional	Description
Set-Field	Color Actions Index	No	Must set to zero.

The header remark actions in Table 60 can be used in an Action Set. All are optional.

Table 60 L2 Policer Actions Flow Table Action Set Actions

Name	Argument	Optional	Description
Set-Field	IP_DSCP	Yes	Ignored when applied if not an IP packet.
Set-Field	IP_ECN	Yes	Ignored when applied if not an IP packet.

Name	Argument	Optional	Description
Set-Field	VLAN_PCP	Yes	
Set-Field	Traffic Class	Yes	

4.1.9.4 Counters and Flow Expiry

The L2 Policer Actions Flow Table counters are listed in Table 61.

Table 61 L2 Policer Actions 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

The L2 Policer Actions Flow Table expiry provisions are as indicated in Table 62.

Table 62 L2 Policer Actions 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.

4.1.10 Termination MAC Flow Table

The Termination MAC Flow Table is used to decide whether to do IP routing or MPLS processing on a packet. It identifies IP routed and MPLS packets by their destination MAC, VLAN, and EtherType. Routed packet rule types use a Goto-Table instruction where the next table is one of the routing tables. MPLS packet rule types use a Goto-Table instruction to indicate that the next table is either the MPLS 0 (if supported on the platform) or the MPLS 1 Flow Table. The built-in default miss rule has a Goto-Table instruction specifying the Bridging Flow Table.

4.1.10.1 Flow Entry Types and Match Fields

The Termination MAC Flow Table implements the flow entry types listed in Table 63.

Table 63 Termination MAC Flow Table Entry Types

Name	Description
IPv4 Unicast MAC	Used to identify an IPv4 router MAC address. Relative priority must be assigned so as to be lower than any multicast MAC rule. Must have a Goto-Table instruction specifying the Unicast Routing Flow Table or the L3 Type Flow Table.
IPv6 Unicast MAC	Used to identify an IPv6 router MAC address. Relative priority must be assigned so as to be lower than any multicast MAC rule. Must have a Goto-Table instruction specifying the Unicast Routing Flow Table or the L3 Type Flow Table.
IPv4 Multicast MAC	Wildcard rule that recognizes all IPv4 multicast MAC addresses specified in RFC 1112 [39]. 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. Must have a Goto-Table instruction specifying the Multicast Routing Flow Table.
IPv6 Multicast MAC	Wildcard rule that recognizes all IPv6 MAC addresses specified in RFC 2464 [40]. 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. Must have a Goto-Table instruction specifying the Multicast Routing Flow Table.
MPLS	Used to identify an MPLS label switch router or edge device MAC address. Must have a Goto-Table instruction specifying either MPLS 0 or MPLS 1 (for platforms that do not support MPLS 0).

The Termination MAC Flow Table match fields are listed in Table 64. Strict rule priority must be observed to avoid potential conflicts.

Table 64 Termination MAC Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	Yes	Physical (local) input port. Limited support on some platforms.
ETH_TYPE	16	No	No	Prerequisite for IPv4 (0x0800), IPv6 (0x86dd), or MPLS (0x8847).

Field	Bits	Maskable	Optional	Description
ETH_DST	48	Yes	Yes	Ethernet destination MAC. Prefix maskable for only the specific multicast IP flow entries in Table 63.
VLAN_VID	16	No	Yes	Matches against the outermost VLAN Id. Must be either omitted or exact.

4.1.10.2 Instruction Types

The Termination MAC Flow Table can have the instructions shown in Table 65.

Table 65 Termination MAC Flow Table Instruction Set

Name	Argument	Description
Goto-Table	Table	Unicast MAC rules can specify the Unicast Routing Flow Table or the L3 Type Flow Table. Multicast MAC rules can only specify the Multicast Routing Flow Table. MPLS rules must specify MPLS Flow Table 0 on platforms that support it, otherwise must specify MPLS Flow Table 1. The packet may be dropped if a rule matches and there is no Goto-Table instruction.
Apply Actions	Action List	Optional. If supplied can only contain one action, output a copy to CONTROLLER.

4.1.10.3 Counters and Flow Expiry

The Termination MAC Flow Table counters are listed in Table 66.

Table 66 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 67.

Table 67 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.

4.1.11 L3 Type Flow Table

RFC 1112 [39] and RFC 2464 [40] dictate specific multicast destination MAC address formats for IPv4 and IPv6 multicast packets, respectively. Rules in the Termination MAC Flow Table can be used to match these formats and Goto the Multicast Routing Flow Table. Normally rules that match on unicast destination MAC addresses would Goto the Unicast Routing Flow Table.

However on certain platforms Termination MAC rules that match unicast destination MAC addresses can optionally Goto the L3 Type Flow table. In the L3 Type Flow table built-in rules match IP packets that have multicast IP destination addresses and go to the Multicast Routing Flow table. Packets that miss either of these match a default rule and go to the Unicast Routing Flow table.

This is a fixed table that contains only built-in rules that cannot be modified by the Controller.

Note: Support for the L3 Type Flow Table is platform dependent. Some platforms only provide the default miss rule.

4.1.11.1 Flow Entry Types and Match Fields

The L3 Type Flow Table rule types are listed in Table 68.

Table 68 L3 Type Flow Table Entry Types

Type	Prerequisite(s)	Actions
IPv4 Multicast IP	ETH_TYPE is 0x0800	Goto Multicast Routing.
IPv6 Multicast IP	ETH_TYPE is 0x86dd	Goto Multicast Routing.
Default		Goto Unicast Routing.

The L3 Type Flow Table match fields are listed in Table 69.

Table 69 L3 Type Flow Table Match Fields

Field	Bits	Maskable	Optional	Description

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	IPv4 or IPv6
IPv4_DST	32	Yes	No	Matches multicast IPv4 frames with a 224/8 destination address and 224.0.0.0 mask value.
IPv6_DST	128	Yes	No	Matches match multicast IPv6 frames with a FF00::/8 destination address and FF00:0:0:0:0:0:0:0 mask value.

4.1.11.2 Instruction Types

The L3 Type Flow Table built-in rules only have the Goto-Table instructions described above.

4.1.11.3 Actions

The L3 Type Flow Table has no actions.

4.1.11.4 Counters and Flow Expiry

The L3 Type Flow Table supports the table and flow entry counters in Table 70.

Table 70 L3 Type Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Fixed value of 3.
Duration (sec)	32	Per-entry	Seconds since this flow entry was installed. Always returns the number of seconds since the device was initialized.

The L3 Type Flow Table entries are fixed and do not age out.

4.1.12 MPLS Flow Tables

The MPLS pipeline can support three MPLS Flow Tables, MPLS Table 0, MPLS Table 1 and MPLS Table 2. An MPLS Flow Table lookup matches the label in the outermost MPLS shim header in the packets. MPLS Table 0 is only used to pop a protection label on platforms that support this table. MPLS Table 1 and MPLS Table 2 can be used for all label operations, including detection of an MPLS-TP Section OAM PDU. MPLS Table 1 and MPLS Table 2 are synchronized flow tables and updating one updates the other.

Note: Some implementations may only permit modification of MPLS 1, which also updates MPLS 2.

MPLS Table 0 only matches an outermost label that is not bottom of stack. MPLS Table 0 has a built-in default rule with a Goto-Table instruction specifying MPLS Table 1 on a miss. An MPLS Table 0 rule can pop the outermost label and have a Goto-Table instruction specifying the MPLS Table 1. If the pop exposes an OAM frame (GAL or RAL) and the label is enabled for OAM, the frame can be forwarded to an OAM engine for further processing.

MPLS Table 1 matches the outermost label. If the label is not bottom of stack, it can pop the outermost label with a Goto-Table instruction specifying MPLS Table 2, for matching the next label. Up to three labels can be explicitly matched in this way, depending on platform. An entry that matches bottom of stack set can only be added or modified to MPLS Table 1 or MPLS Table 2.

MPLS Table 1 can match a label with the reserved value GAL (13) that is bottom of stack, but only for Section OAM. GAL is not otherwise directly matched. The presence of a GAL in other OAM PDU use cases must be detected using the NEXT_LABEL_IS_GAL pipeline match field.

4.1.12.1 Flow Entry Types and Match Fields

MPLS Table 0 Flow supports the flow entry types in Table 71.

Table 71 MPLS Flow Table 0 Flow Entry Types

Type	Prerequisite(s)	Actions	Use Case
Pop Tunnel Label	MPLS_BOS = 0	Pop outermost tunnel label; Set EtherType to 0x8847; Decrement/check TTL; Optionally Copy TTL in; Optionally Copy EXP in; Goto-Table instruction specifies MPLS Table 1.	Pop Protection Label and forward based on an inner label
Pop Tunnel Label (LSP OAM MEP Data)	MPLS_BOS = 0	Pop outermost tunnel label; Set EtherType to 0x8847; Decrement/check TTL; Optionally Copy TTL in; Optionally Copy EXP in; Goto-Table instruction specifies MPLS Table 1. Increment OAM LM Counters for this LMEP.	MPLS-TP LSP OAM
Pop Tunnel Label (LSP OAM MEP Control)	MPLS_BOS = 0, Next Label is GAL	Set the LMEP Id, Goto-Table instruction specifies the MPLS-TP Maintenance Point Flow Table.	MPLS-TP LSP OAM

MPLS Tables 1 and 2 support the flow entry types in Table 72. These match the same fields but actions differ depending on the table and packet flow use case.

Table 72 MPLS Flow Table 1 and 2 Flow Table Entry Types

Type	Prerequisite(s)	Actions	Use Case
Pop Tunnel Label and Forward Based on Next Label	MPLS_BOS = 0	Pop outermost tunnel label, EtherType stays 0x8847; Check TTL; Optionally Copy TTL in; Optionally Copy EXP in; Goto-Table instruction specifies MPLS Table 2.	Pop label and forward based on an inner label
Pop Tunnel Label and Forward Based on This Label (PHP)	MPLS_BOS = 0	Pop outermost tunnel label, EtherType stays 0x8847; Decrement/check TTL; Optionally Copy TTL in; Optionally Copy EXP in; Group: MPLS Interface; Goto-Table instruction specifies the Policy ACL Flow Table.	Pop and forward based on this label to a next hop router. (PHP)
Swap Tunnel Label and Forward Based on This Label	MPLS_BOS = 0	Decrement/check TTL; Group: MPLS Swap, MPLS ECMP, or MPLS Fast Failover; Goto-Table instruction optionally specifies the Policy ACL Flow Table.	Swap and forward based on this label (LSR)
Swap PW Label and Forward Based on This Label	MPLS_BOS = 1	Decrement/check TTL; Group: MPLS Swap MPLS ECMP, or MPLS Fast Failover; Goto-Table instruction optionally specifies the Policy ACL Flow Table.	Swap and forward based on this label (MS-PW, LSR)
Pop Tunnel Label and Forward Based on Next Label (MEP Data Frame)	MPLS_BOS = 0	Pop outermost tunnel label, EtherType stays 0x8847; Decrement/check TTL; Optionally Copy TTL in; Optionally Copy EXP in; Increment OAM LM Counters for this LMEP; Goto-Table instruction specifies MPLS Table 1.	MPLS-TP LSP OAM MEP
Pop Tunnel Label and Forward Based on Next Label	MPLS_BOS = 0, next label is GAL.	Set the LMEP Id; Make into a Y.1731 frame by popping the tunnel label, GAL, and ACH, setting the EtherType to 0x8902.	MPLS-TP LSP OAM MEP

Type	Prerequisite(s)	Actions	Use Case
(MEP Control Frame)		Goto-Table instruction specifies the MPLS-TP Maintenance Point Flow Table.	
Swap Tunnel Label and Forward Based on This Label (MIP Control)	MPLS_BOS = 0, TTL=1 and next label is GAL.	Set the LMEP Id; Make into a Y.1731 frame by popping the tunnel label, GAL, and ACH, and setting the EtherType to 0x8902. Goto-Table instruction specifies the MPLS-TP Maintenance Point Flow Table..	MPLS-TP LSP MIP
Swap PW Label and Forward Based on This Label (MIP Control)	MPLS_BOS=1, next label is GAL	Set the LMEP Id; Make into a Y.1731 frame by popping the tunnel label, GAL, and ACH, and setting the EtherType to 0x8902. Goto-Table instruction specifies the MPLS-TP Maintenance Point Flow Table.	MPLS-TP MIP (MS-PW, LSR)
L3 VPN Route	MPLS_BOS = 1	Decrement/check TTL and optionally copy in; optionally Copy EXP in; Set-Field VRF; Goto-Table instruction specifies the MPLS L3 Type Flow Table.	Pop and forward to MPLS L3 Type table.
L3 VPN Forward	MPLS_BOS = 1	Decrement/check TTL and optionally copy in; optionally Copy EXP in; Set-Field VRF; Goto-Table instruction specifies the MPLS L3 Type Flow Table.	Set output group.
L2 Switch (PW Terminate)	MPLS_BOS = 1	Pop outermost (pseudo-wire) label; pop outer L2 header; set MPLS Type to VPWS. Decrement/check TTL. Optionally pop CW, Set-Field MPLS L2 Port; Set-Field Tunnel_Id; Goto-Table instruction specifies the Bridging Flow Table. Optionally increment OAM LM Counters.	Pop, decap, and L2 forward
PW VCCV 1 (PW Terminate) (MEP Control)	MPLS_BOS = 1, MPLS_DATA_FIRST_NIBBLE is 1, MPLS_ACH_CHANNEL	Set the LMEP Id; Make into a Y.1731 frame by popping the VC label and ACH, and setting the EtherType to 0x8902. Set-Field	MPLS-TP PW termination. OAM PDU

Type	Prerequisite(s)	Actions	Use Case
	is 0x8902	MPLS L2 Port; Set-Field Tunnel_Id; Goto-Table Instruction specifies the MPLS-TP Maintenance Point Flow Table.	
PW VCCV 3 (PW Terminate) (MEP Control)	MPLS_BOS = 1, TTL=1, MPLS_DATA_FIRST_NIBBLE is 1, MPLS_ACH_CHANNEL is 0x8902	Set the LMEP Id; Make into a Y.1731 frame by popping the VC label and ACH, and setting the EtherType to 0x8902. Set-Field MPLS L2 Port; Set-Field Tunnel_Id; Goto-Table Instruction specifies the MPLS-TP Maintenance Point Flow Table.	MPLS-TP PW termination. OAM PDU
PW VCCV 4 (PW Terminate) (MEP Control)	MPLS_BOS = 0, next label is GAL, MPLS_DATA_FIRST_NIBBLE is 1, MPLS_ACH_CHANNEL is 0x8902	Set the LMEP Id; Make into a Y.1731 frame by popping the VC label, GAL, and ACH, and setting the EtherType to 0x8902. Set-Field MPLS L2 Port; Set-Field Tunnel_Id; Goto-Table Instruction specifies the MPLS-TP Maintenance Point Flow Table.	MPLS-TP PW termination. OAM PDU
Section OAM	MPLS_BOS is 1, Label is GAL (13), MPLS_DATA_FIRST_NIBBLE is 1, MPLS_ACH_CHANNEL is 0x8902	Make into a Y.1731 frame by popping the GAL and ACH, and setting the EtherType to 0x8902. Goto-Table instruction specifies the MPLS-TP Maintenance Point Flow Table.	MPLS-TP Section OAM PDU

The MPLS Flow Table match fields for all flow entry types are shown in Table 73.

Table 73 MPLS Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	Must be 0x8847. This is the OpenFlow required prerequisite for MPLS matching.
MPLS_BOS	1	No	No	Bottom of stack
MPLS_LABEL	20	No	No	Outermost label

Field	Bits	Maskable	Optional	Description
IN_PORT	32	Yes	Yes	Physical (local) input port. Field maskable only.
MPLS_TTL	8	No	Yes	TTL=1 is the only value matched, drop unless an OAM PDU (indicated by a match on NEXT_LABEL_IS_GAL).
MPLS_DATA_FIRST_NIBBLE	4	No	Yes	High order 4 bits of control word. 0x00 for CW, 0x01 for ACH.
MPLS_ACH_CHANNEL	16	No	Yes	Parsed from ACH Channel Type field if MPLS_DATA_FIRST_NIBBLE is 0x01. Only value is 0x8902, which is for G.8113.1.
NEXT_LABEL_IS_GAL	1	No	Yes	Parser peeks at next label, detects GAL and MPLS_BOS, and sets this pipeline match field.
VLAN_VID	16	No	Yes	Only used in Section OAM rules.

4.1.12.2 Instruction Types

The MPLS Flow Table can have the instructions shown in Table 74.

Table 74 MPLS Flow Table Instructions

Name	Argument	Description
Goto-Table	Table	Depends on rule type as detailed in Table 71 and Table 72.
Apply Actions	Action List	Allowed actions are listed in Table 75 and depend on rule type.
Write Actions	Action Set	Allowed actions are listed in Table 76 and depend on rule type.

4.1.12.3 Actions

The MPLS Flow Table action list supports the actions in Table 75. These are applied immediately to the frame.

Table 75 MPLS Flow Table Action List Actions

Name	Argument	Description
Pop Label	EtherType	Some BOS rules set the packet EtherType based on the value of MPLS_DATA_FIRST_NIBBLE. Otherwise argument should be 0x8847, although it may not always be relevant for setting the packet EtherType. Not used if Group specifies an MPLS Swap Label Group entry type.
Decrement TTL		This action must always be performed on frames that will be forwarded. If TTL is invalid after decrement packet is sent to Controller. This does not apply to VCCV Type 3 frames, which should be output for OAM processing.
Copy TTL in		Optional. Only applies if there is an inner label or IP packet and TTL is valid after decrement.
Copy TC in		Optional. Only applies if there is an inner label. Primarily used to set the inner label MPLS_TC or IP header TOS for PHP.
Set-Field	VRF	Required for L3 Route rules, otherwise not used.
Set-Field	MPLS L2 Port	Required for pseudo-wire termination.
Set-Field	MPLS Type	Required for pseudo-wire termination.
Set-Field	Tunnel Id	Required for pseudo-wire termination.
Set Field	QoS Index	The QoS Index used with the MPLS_TC field to access the MPLS Label Trust Flow Table to set packet Traffic Class and Color.
Set-Field	Traffic Class	Sets static Traffic Class value. Mutually exclusive with setting Traffic Class from lookup in the MPLS Label Trust Table.
Set-Field	L3_IN_PORT	Only used for L3 Multicast forwarding lookup, otherwise should be omitted.
Copy Field	MPLS_TC, PKT_REG(0)	Copies the MPLS_TC field from the shim header that will be used to access the MPLS Label Trust Flow Table to set packet Traffic Class and Color.

Name	Argument	Description
Pop CW or ACH		Control word expected, pop without checking sequence number. Only used in conjunction with popping a bottom of stack pseudo wire label for MPLS-TP termination for data frames.
Pop VLAN		Pop outermost VLAN tag, only used in conjunction with popping an outermost Ethernet header.
Pop L2 Header		Pop outermost Ethernet header. Header cannot have a VLAN tag (must have already been popped). Only used for MPLS-TP termination.
Set-Field	LMEP_Id	Indicates a layer MEP or MIP on this label, and that subsequent OAM PDU processing is required.
Set-Field	Protection_Index	Indicates whether this label is for the working or protection path. Optional. 0: Protection path 1: Working path

Notes:

- For MPLS_TP PW termination, actions must be listed in the following order: Pop Label, Pop CW or ACH, Pop VLAN (if any), Pop L2 Header.
- The hardware parses, recognizes, and validates various OAM frame formats based on the next underlying label (GAL or RAL) or TTL. It also parses, recognizes, and validates a G.8113.1 frame based on ACh, destination UDP port and TTL, etc. These are made available as match fields.
- Control word sequence number insertion, increment, or validation is not supported in this version of OF-DPA.
- Even though both tables contain the same rules, a rule in MPLS 2 cannot go to itself. A "Pop and Forward Based on Next Label" rule should not match in MPLS 2, but in the event one does, the results are undefined.

The MPLS Flow Table Write-Actions instruction can update the actions listed in Table 76 to the action set.

Table 76 MPLS Flow Table Action Set Actions

Name	Argument	Description
Group	Group id	Depends on rule type as detailed in Table 71 and Table 72.

4.1.12.4 Counters and Flow Expiry

The MPLS Flow Table counters are listed in Table 77.

Table 77 MPLS 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

MPLS Flow Table expiry provisions are as indicated in Table 78.

Table 78 MPLS Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Entry is removed in both MPLS Flow Tables. If zero or not specified, entry does not age out.

4.1.13 MPLS L3 Type Flow Table

The MPLS L3 Type Flow Table sets the EtherType and MPLS_TYPE for L3 VPN terminated packets. Built-in rules determine the IP version for the EtherType, and whether the packet is unicast or multicast. This is a fixed table and contains only built-in rules that cannot be modified by the Controller.

4.1.13.1 Flow Entry Types and Match Fields

The MPLS L3 Type Flow Table rule types are listed in Table 79.

Table 79 MPLS L3 Type Flow Table Entry Types

Type	Prerequisite(s)	Actions
L3 VPN Route (IPv4 Unicast)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x0800. Set MPLS_TYPE to L3 Unicast.
L3 VPN Route (IPv4 Multicast)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x0800. Set MPLS_TYPE to L3 Multicast.

Type	Prerequisite(s)	Actions
L3 VPN Route (IPv6 Unicast)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x86dd. Set MPLS_TYPE to L3 Unicast.
L3 VPN Route (IPv6 Multicast)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x86dd. Set MPLS_TYPE to L3 Multicast.
L3 VPN Forward (IPv4)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x0800. Rules in the MPLS tables have already set the MPLS_TYPE to L3 PHP.
L3 VPN Forward (IPv4)	ETH_TYPE is 0x8847	Pop label and set packet EtherType to 0x86dd. Rules in the MPLS tables have already set the MPLS_TYPE to L3 PHP.

The MPLS L3 Type Flow Table match fields are listed in Table 80

Table 80 MPLS L3 Type Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	MPLS (0x8847)
MPLS_TYPE	8	No	Yes	Differentiates routed from forwarded (PHP) packets.
MPLS_DATA_FIRST_NIBBLE	4	No	No	Used to determine IP version based on the IP header.
IPv4_DST	32	Yes	Yes	Used to determine if packet is multicast
IPv6_DST	128	Yes	Yes	Used to determine if packet is multicast

4.1.13.2 Instruction Types

The MPLS L3 Type Flow Table supports the instruction types listed in Table 81.

Table 81 MPLS L3 Type Flow Table Instructions

Name	Argument	Description

Name	Argument	Description
Apply Actions	Action List	Actions are from Table 47.
Goto-Table	Table	Must specify the MPLS Label Trust Flow Table.

4.1.13.3 Actions

The MPLS L3 Type Flow Table Action List actions pop the MPLS shim header and set packet MPLS_TYPE and EtherType, as listed in Table 82.

Table 82 MPLS L3 Type Flow Table Action List

Name	Argument	Description
POP Label	EtherType	Sets packet header EtherType to 0x0800 for IPv4 or to 0x86dd for IPv6.
Set-Field	MPLS Type	Sets the MPLS Type to L3 Unicast or L3 Multicast as needed.
Set-Field	MAC_DST	Sets the Destination MAC address to a multicast address for multicast frames.

4.1.13.4 Counters and Flow Expiry

The MPLS L3 Type Flow Table supports the table and flow entry counters in Table 83.

Table 83 MPLS L3 Type Flow Table Counters

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

The MPLS L3 Type Flow Table supports only hard interval timeout ageing per entry as indicated in Table 84

Table 84 MPLS L3 Type Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Always zero since built-in entries do not age out.

4.1.14 MPLS Label Trust Flow Table

The MPLS Label Trust Flow Table is used for QoS classification based on the MPLS label. It maps the MPLS_TC field from the MPLS shim header to QoS Traffic Class and Color metadata values.

4.1.14.1 Flow Entry Types and Match Fields

The MPLS Label Trust Flow Table match fields are listed in Table 85.

Table 85 MPLS Label Trust Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
QoS Index	8	No	No	Index value for selecting a trust profile
PKT_REG(0)	8	No	No	MPLS_TC value from the MPLS shim header used to determine QoS properties.

4.1.14.2 Instruction Types

The MPLS Label Trust Flow Table supports the instruction types listed in Table 86.

Table 86 MPLS Label Trust Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	Actions are from Table 47.
Goto-Table	Table	Must specify the MPLS Type table.

4.1.14.3 Actions

The MPLS Label Trust Flow Table Action List actions set packet Traffic Class and Color, as listed in Table 87.

Table 87 MPLS Label Trust Flow Table Action List

Name	Argument	Description
Set-Field	Traffic Class	Sets 4-bit traffic class value used in subsequent QoS processing. Overrides default port Traffic Class if specified. Optional.
Set-Field	Color	Sets 2-bit Color value used in subsequent QoS processing. Optional.

4.1.14.4 Counters and Flow Expiry

The MPLS Label Trust Flow Table supports the table and flow entry counters in Table 88.

Table 88 MPLS Label Trust 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

MPLS Trust Flow Table entries do not age out.

4.1.15 MPLS Type Flow Table

The MPLS Type Flow Table directs MPLS packets to forwarding tables by matching on the MPLS_TYPE pipeline field set in a previous table in the pipeline. Forwarding can include Bridging, Routing, or OAM processing. This is a fixed logic table where all rules are built-in; the Controller cannot add, delete, or modify rules in this table. The default on a miss (unrecognized MPLS_TYPE) is to drop the packet.

4.1.15.1 Flow Entry Types and Match Fields

Table 89 MPLS Type Flow Table Entry Types

Type	Description
VPWS	Goto Policy ACL
MPLS L3 Unicast	Goto Unicast Routing

Type	Description
MPLS L3 Multicast	Goto Multicast Routing
L3 PHP	Goto Policy ACL

Table 90 MPLS Type Flow Table Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite
MPLS_TYPE	16	No	Yes	Match on flow type

4.1.15.2 Instructions

Table 91 MPLS Type Flow Table Instructions

Name	Argument	Description
Goto-Table	Table	Depends on MPLS_TYPE.
Clear-Actions	-	Drop the packet.

4.1.15.3 Counters and Flow Expiry

The MPLS Type Flow Table does not have counters. MPLS Type Flow Table entries do not expire.

4.1.16 Bridging Flow Table

The Bridging Flow Table supports Ethernet packet switching for potentially large numbers of flow entries using the hardware L2 tables.

Note: The Policy ACL Flow Table is recommended for matching and forwarding BPDUs.

The Bridging Flow Table has the following types of rules: VLAN rules forward based on VLAN and MAC_DST for normal switched packets; and Overlay rules forward based on Tunnel Id and MAC_DST for tunneled packets. The naming convention for Tunnel IdId metadata is as in Table 92.

The built-in default rule also has a Goto Table instruction specifying the Policy ACL Flow Table.

Table 92 Tunnel Id Metadata Naming Convention

Type	Numbering	Description

Type	Numbering	Description
Data Center Overlay	0x0000nnnn	Identifies a data center overlay tenant forwarding domain.
MPLS-TP	0x0001nnnn	Identifies an MPLS-TP pseudo-wire forwarding domain. For future use.

4.1.16.1 Flow Entry Types and Match Fields

The Bridging Flow Table flow entry types are listed in Table 93. VLAN and Tunnel Id are mutually exclusive.

Table 93 Bridging Flow Table Flow Entry Types

Type	Description
Unicast VLAN	Matches switched unicast Ethernet frames by VLAN Id and MAC_DST. MAC_DST must be unicast and cannot be masked. VLAN Id must be present and non-zero. Tunnel Id must be masked or omitted.
Multicast VLAN	Matches switched multicast Ethernet frames by VLAN Id and MAC_DST. MAC_DST must be multicast and cannot be masked. VLAN Id must be present and non-zero. Tunnel Id must be masked or omitted.
DLF VLAN	Matches switched Ethernet frames by VLAN Id only. MAC_DST must be field masked and match any destination. Must have lower relative priority than any unicast or multicast flow entries that specify this VLAN. VLAN Id must be present and non-zero. Tunnel Id must be masked or omitted.
Unicast Overlay	Matches switched unicast Ethernet frames by Tunnel Id and MAC_DST. MAC_DST must be unicast and cannot be masked. Tunnel Id must be non-zero, type overlay tunnel (0x0000nnnn), and cannot be masked. VLAN Id must be masked or omitted.
Multicast Overlay	Matches switched multicast Ethernet frames by Tunnel Id and MAC_DST. MAC_DST must be multicast and cannot be masked. Tunnel Id must be non-zero, type overlay tunnel (0x0000nnnn), and cannot be masked. VLAN Id must be masked or omitted.
DLF Overlay	Matches switched Ethernet frames by Tunnel Id only. MAC_DST is must be field masked and match any destination. Must have lower relative priority than any unicast or multicast flow entries

Type	Description
	that specify this Tunnel Id. Tunnel Id must be non-zero, type overlay tunnel (0x0000nnnn), and cannot be masked. VLAN Id must be masked or omitted.

Note: Overlay rules must be given higher relative priority assignments than any potentially overlapping VLAN rules.

Match fields for flow entry types are described in Table 94.

Table 94 Bridging Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_DST	48	Yes	Yes	Ethernet destination MAC, allowed values depend on flow entry type. Exact match only (mask must be all 1's if supplied).
VLAN_VID	16	Yes	Yes	VLAN Id, allowed values depend on flow entry type.
TUNNEL ID	32	Yes	Yes	Identifies forwarding domain for data center overlay traffic. Allowed values depend on flow entry type.

4.1.16.2 Instruction Types

Default next table if no match is the ACL Policy Flow Table for VLAN and Data Center Overlay rules.

Table 95 Bridging Flow Table Instructions

Name	Argument	Description
Write-Actions	Action set	Only the actions in Section 4.1.16.3 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	If specified (e.g., not a drop), must be the Policy ACL Flow table for VLAN and Data Center Overlay rules.

4.1.16.3 Action Set

The Bridging Flow Table supports the actions in Table 96 by flow entry type. The OF-DPA API validates consistency of flow entry type and OF-DPA group entry type references.

Table 96 Bridging Flow Table Action Set

Flow Entry Type	Argument	Description
Unicast VLAN	Group id	Must be an OF-DPA L2 Interface group entry for the forwarding VLAN.
Multicast VLAN	Group id	Must be an OF-DPA L2 Multicast group entry for the forwarding VLAN.
DLF VLAN	Group id	Must be an OF-DPA L2 Flood group entry for the forwarding VLAN.
Unicast Overlay	Output	Must be an overlay tunnel logical port for the tenant overlay forwarding domain Tunnel Id.
Multicast Overlay	Group id	Must be an OF-DPA L2 Overlay Multicast sub-type group entry for the tenant overlay forwarding domain Tunnel Id.
DLF Overlay	Group id	Must be an OF-DPA L2 Overlay Flood sub-type group entry for the tenant overlay forwarding domain Tunnel Id.

4.1.16.4 Counters and Flow Expiration

The Bridging Flow Table counters are listed in Table 97.

Table 97 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 98.

Table 98 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.

4.1.17 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 sub-tables by IP protocol, and supports the flow entry types listed in Table 99. A single table number is used for both logical tables.

Table 99 Unicast Routing Flow Table Entry Types

Type	Table	Prerequisite(s)	Description
IPv4 Unicast Host	Table 100	EtherType=0x0800	Matches routed unicast IPv4 packets. Exact match on IP destination. The Goto-Table instruction specifies the Policy ACL Table.
IPv6 Unicast Host	Table 101	EtherType=0x86dd	Matches routed unicast IPv6 packets. Exact match on IP destination. The Goto-Table instruction specifies the Policy ACL Table.
IPv4 Unicast LPM	Table 100	EtherType=0x0800	Matches routed unicast IPv4 packets. Longest Prefix match on IP destination. The Goto-Table instruction specifies the Policy ACL Table.
IPv6 Unicast LPM	Table 101	EtherType=0x86dd	Matches routed unicast IPv6 packets. Longest Prefix match on IP destination. The Goto-Table instruction specifies the Policy ACL Table.

The Unicast Routing Flow Table can be partitioned into multiple virtual routing tables. The VRF pipeline metadata match field value identifies the virtual routing table to use for a particular packet lookup.

4.1.17.1 Flow Entry Types and Match Fields

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

Table 100 Unicast Routing Flow Table IPv4 Header Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	Must be 0x0800
VRF	16	No	Yes	If omitted or zero indicates the default routing table.
IPv4 DST	32	Yes	No	Must be a unicast IPv4 address. Prefix maskable only, mask used for LPM forwarding.

Table 101 Unicast Routing Flow Table IPv6 Header Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	Must be 0x86dd
VRF	16	No	Yes	If omitted or zero indicates the default routing table.
IPV6_DST	128	Yes	No	Must be a unicast IPv6 address. Prefix maskable only, used for LPM forwarding.

Note: Exact match rules must be given higher relative priority assignments than any LPM prefix match rules to avoid conflicts.

Note: Rules that specify a non-zero VRF must be assigned a higher relative priority than rules with a zero or omitted VRF value.

4.1.17.2 Instruction Types

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

Table 102 Unicast Routing Flow Table Instructions

Name	Argument	Description
Write-Actions	Action Set	Optional. Only the actions in Table 103 can be specified.
Apply-Actions	Action List	Optionally output to CONTROLLER.

Name	Argument	Description
Clear-Actions	-	Used to delete any earlier forwarding decision. The packet will be dropped unless the Policy ACL Table supplies a group or output action.
Goto-Table	Table	Must specify the Policy ACL Flow Table.

4.1.17.3 Action Set

The actions in Table 103 are supported.

Table 103 Unicast Routing Flow Table Action Set

Name	Argument	Description
Group	Group id	Must be an OF-DPA L3 Unicast, L3 ECMP, or MPLS L3 VPN Label Group Entry. Optional.
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 sent to the CPU for forwarding to the CONTROLLER. Similarly, a packet that exceeds the MTU is dropped and a copy sent to the CONTROLLER. Must be included with and only with a group action.

4.1.17.4 Counters and Flow Expiration

The Routing Flow Table counters are listed in Table 104.

Table 104 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 105.

Table 105 Unicast Routing 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.

4.1.18 Multicast Routing Flow Table

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

The Multicast Routing Flow Table can also support multiple virtual routing tables, matching on the packet VRF field value.

The Multicast Routing Flow Table is also organized as two mutually exclusive logical sub-tables by IP protocol, and supports the flow entry types listed in Table 106.

Table 106 Multicast Routing Flow Table Entry Types

Type	Table	Prerequisite(s)	Description
IPv4 Multicast	Table 107	EtherType=0x0800	Matches routed multicast IPv4 packets.
IPv6 Multicast	Table 108	EtherType=0x86dd	Matches routed multicast IPv6 packets.
IPv4 Multicast MPLS L3 VPN	Table 107	EtherType=0x0800	Matches routed multicast IPv4 packets.
IPv6 Multicast MPLS L3 VPN	Table 108	EtherType=0x86dd	Matches routed multicast IPv6 packets.

4.1.18.1 Flow Entry Types and Match Fields

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

Table 107 Multicast Routing Flow Table IPv4 Match Fields

Field	Bits	Maskable	Optional	Description

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	Must be 0x0800. Required prerequisite.
VLAN_VID	16	No	No	Used for source removal in for normal routed packets.
L3_IN_PORT	32	No	Yes	Used for source removal for MPLS L3 VPN terminated packets.
VRF	16	No	Yes	VRF. Identifies a virtual routing table partition.
IPV4_SRC	32	No	Yes	Cannot be bit masked, but can be omitted.
IPV4_DST	32	Yes	No	Must be an IPv4 multicast group address.

Table 108 Multicast Routing Flow Table IPv6 Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	No	Must be 0x86dd. Required prerequisite.
VLAN_VID	16	No	No	Used for source removal in for normal routed packets.
L3_IN_PORT	32	No	Yes	Used for source removal for MPLS L3 VPN terminated packets.
VRF	16	No	Yes	VRF.
IPV6_SRC	128	No	Yes	Cannot be bit masked, but can be omitted.
IPV6_DST	128	Yes	No	Must be an IPv6 multicast group address.

Note: Rules that specify a non-zero VRF must be assigned a higher relative priority than rules with a zero or omitted VRF field.

4.1.18.2 Instruction Types

A built-in default miss rule Goto-Table instruction specifies the Policy ACL Flow Table.

Table 109 Multicast Routing Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	Can only optionally specify output to CONTROLLER.
Write-Actions	Action Set	Optional. Only the actions in Table 110 can be specified.
Goto-Table	Table	Can only specify the Policy ACL Flow Table. In the event that there is no group entry referenced and no next table specified, the packet may be dropped unless the Policy ACL table supplies a forwarding decision.

4.1.18.3 Action Set

The Multicast Routing Table supports the actions in Table 110.

Table 110 Multicast Routing Flow Table Action Set

Name	Argument	Description
Group	Group id	Optional, if supplied can only be an OF-DPA L3 Multicast group entry.
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 sent to the CPU for forwarding to the CONTROLLER. Similarly, a packet that exceeds the MTU is dropped and a copy sent to the CONTROLLER. Required with and only with a multicast group.

4.1.18.4 Counters and Flow Expiration

The Multicast Routing Flow Table counters are as shown in Table 111.

Table 111 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 112.

Table 112 Multicast Routing 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.

4.1.19 MPLS-TP Maintenance Point Flow Table

The MPLS-TP Maintenance Point Flow Table determines where to forward a received Y.1732 OAM control frame for processing based on the LMEP Id and opcode. It matches the Y.1731 OAM PDUs that the MPLS Flow tables have rendered into Ethernet OAM frames. Options are to send the packet to the CONTROLLER or LOCAL reserved port, or to drop the packet.

4.1.19.1 Flow Entry Types and Match Fields

The Maintenance Point Flow Table implements the single flow entry type listed in Table 113.

The built in default miss rule voids any forwarding decision with a Clear-Actions instruction and hence drops frames.

Table 113 MPLS-TP Maintenance Point Flow Table Entry Types

Name	Description
MEP PDU Process G.8113.1	The only rule type.

The Maintenance Point Flow Table match fields are listed in Table 114.

Table 114 MPLS-TP Maintenance Point Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
LMEP_Id	32	No	No	Local identifier for the MEP or MIP

Field	Bits	Maskable	Optional	Description
OAM_Y1731_OPCODE	8	No	No	Parsed from the IEEE 802.1ag/Y.1731 header.
ETH_TYPE	16	No	No	OpenFlow prerequisite. MPLS label popping transforms packet into an Ethernet OAM frame.

A built-in default rule drops the packet if no rules match.

4.1.19.2 Instruction Types

The MPLS-TP Maintenance Point Flow Table can have the instructions shown in Table 115. At least one must be specified. Since there is no next table the packet is dropped.

Table 115 MPLS-TP Maintenance Point Flow Table Instructions

Name	Argument	Description
Clear Actions		Used to drop and stop processing the OAM frame. Required.
Apply Actions	Action List	Optional, Actions are listed in Table 116.

4.1.19.3 Actions

Table 116 MPLS-TP Maintenance Point Flow Table Actions

Name	Argument	Description
Output	Port	Can only be LOCAL for processing by the local OAM engine.

4.1.19.4 Counters and Flow Expiry

The Maintenance Point Flow Table counters are listed in Table 117.

Table 117 MPLS-TP Maintenance Point 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

Maintenance Point Flow Table expiry provisions are as indicated in Table 118.

Table 118 MPLS-TP Maintenance Point 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.

4.1.20 Policy ACL Flow Table

The Policy ACL Flow Table supports wide, multi-field matching. Most fields can be wildcard matched, and relative priority must be specified in all flow entries. The Policy ACL Flow Table provides actions to redirect packets to different destination groups. It can be used to output copies to the Controller, for example, ARP packets or BPDU frames. It includes a Meter instruction that can indicate application of a QoS policer.

The Policy ACL Flow Table is organized two mutually exclusive logical sub-tables. Flow entries in the IPv6 logical tables match only packets that require matching on IPv6 header fields. The non-IPv6 logical table matches any packet that does not require matching on IPv6 header fields. Entries can optionally supply either a VLAN Id or a Tunnel Id match field, but not both. Following the 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.

Flow entries must conform to match field prerequisite requirements defined in the OpenFlow specification or in this document. In particular, if a prerequisite field is identified for a particular match field then it must be explicitly provided. For example, to match a TCP source port the IP protocol must be 4 (TCP) and the EtherType must be 0x0800 (IPv4) or 0x86dd (IPv6).

The Policy ACL Flow Table can optionally have a Goto-Table instruction specifying the Color Based Actions Flow Table. In this case the flow entry must supply a Color Actions Index pipeline field. If no next table is specified then it is the last table in the pipeline before applying the action set.

The default on table miss is to do nothing. The packet will be forwarded using the output or group in the action set, if any. If the action set does not have a group or output action the packet is dropped.

4.1.20.1 Flow Entry Types and Match Fields

The Policy ACL Flow Table supports the flow entry types listed in Table 119.

Table 119 Policy ACL Flow Table Flow Entry Types

Type	Table	Prerequisite	Description
IPv4 VLAN	Table 120	EtherType cannot be 0x86dd. IN_PORT	Matches packets by VLAN Id except for IPv6. VLAN Id is optional but must be non-zero if supplied.

Type	Table	Prerequisite	Description
		must be a physical port.	Tunnel Id must not be supplied.
IPv6 VLAN	Table 121	EtherType must be 0x86dd, IN_PORT must be a physical port.	Matches only IPv6 packets by VLAN Id. VLAN Id is optional but must be non-zero if supplied. Tunnel Id must not be supplied.
IPv4 Overlay	Table 120	EtherType cannot be 0x86dd. IN_PORT must be a tunnel logical port.	Matches packets by Tunnel Id except for IPv6. A non-zero Tunnel Id is required and is not maskable. VLAN Id must not be supplied.
IPv6 Overlay	Table 121	EtherType must be 0x86dd. IN_PORT must be a tunnel logical port	Matches only IPv6 packets by Tunnel Id. A non-zero Tunnel Id is required and is not maskable. VLAN Id must not be supplied.
IPv4 MPLS-TP	Table 120	EtherType cannot be 0x86dd.	Matches packets by Tunnel Id and MPLS L2 port except for IPv6. A non-zero Tunnel Id is required and is not maskable. VLAN Id must not be supplied.
IPv6 MPLS-TP	Table 121	EtherType must be 0x86dd.	Matches only IPv6 packets by Tunnel Id and MPLS L2 port. A non-zero Tunnel Id is required and is not maskable. VLAN Id must not be supplied.

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

Table 120 Policy ACL Flow Table IPv4 Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite
IN_PORT	32	No	Yes	Physical or logical ingress port.
ETH_SRC	48	Yes	Yes	Ethernet source MAC
ETH_DST	48	Yes	Yes	Ethernet destination MAC
ETH_TYPE	16	No	Yes	Any value except 0x86dd. Explicit prerequisite must be 0x800 if IP fields are to be matched.

Field	Bits	Maskable	Optional	Description or Prerequisite
VLAN_VID	16	Yes	Yes	VLAN Id. Cannot be masked for a VLAN bridging rule that redirects to a different L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	No	Yes	802.1p priority field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
VLAN_DEI	1	No	Yes	802.1p drop eligibility indicator field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
TUNNEL ID	32	No	Yes	Tunnel forwarding domain. Applicable to data center overlay and MPLS-TP bridged flow entry types. IN_PORT must be a data center overlay tunnel or MPLS-TP logical port consistent with the tunnel-id range.
VRF	16	No	Yes	VRF.
IPV4_SRC	32	Yes	Yes	Matches SIP if EtherType = 0x0800
ARP_SPA	32	Yes	Yes	Matches ARP source protocol address if EtherType = 0x0806
IPV4_DST	32	Yes	Yes	Matches DIP if EtherType = 0x0800
IP_PROTO	8	No	Yes	IP protocol field from IP header if EtherType = 0x0800
IP_DSCP	6	No	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474 if EtherType = 0x0800
IP_ECN	2	No	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168 if EtherType = 0x0800
TCP_SRC	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 6
UDP_SRC	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 17
SCTP_SRC	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 132
ICMPV4_TYPE	8	No	Yes	If EtherType = 0x0800 and IP_PROTO = 1

Field	Bits	Maskable	Optional	Description or Prerequisite
TCP_DST	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 6
UDP_DST	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 17
SCTP_DST	16	No	Yes	If EtherType = 0x0800 and IP_PROTO = 132
ICMPv4_CODE	8	No	Yes	If EtherType = 0x0800 and IP_PROTO = 1
MPLS L2 PORT	16	No	Yes	MPLS L2 Port

Table 121 Policy ACL Flow Table IPv6 Match Fields

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	Yes	Physical or logical ingress port.
ETH_SRC	48	Yes	Yes	Ethernet source MAC
ETH_DST	48	Yes	Yes	Ethernet destination MAC
ETH_TYPE	16	No	Yes	Must be 0x86dd
VLAN_VID	16	Yes	Yes	VLAN Id. Cannot be masked for a VLAN bridging rule that redirects to a different L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	No	Yes	802.1p priority field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
VLAN_DEI	1	No	Yes	802.1p drop eligibility indicator field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
TUNNEL ID	32	No	Yes	Tunnel forwarding domain. Applicable to data center overlay and MPLS-TP bridged flow entry types. IN_PORT must be a data center overlay tunnel or MPLS-TP logical port consistent with the tunnel-id range. VLAN_VID must not be supplied.
VRF	16	No	Yes	VRF

Field	Bits	Maskable	Optional	Description
IPV6_SRC	128	Yes	Yes	Matches IPv6 SIP
IPV6_DST	128	Yes	Yes	Matches IPv6 DIP
IP_PROTO	8	No	Yes	Matches IPv6 Next header
IPV6_FLABEL	20	No	Yes	Matches IPv6 flow label
IP_DSCP	6	No	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474 if EtherType = 0x86dd
IP_ECN	2	No	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168 if EtherType = 0x86dd
TCP_SRC	16	No	Yes	If EtherType = 0x86dd and IP_PROTO = 6
UDP_SRC	16	No	Yes	If EtherType = 0x86dd and IP_PROTO = 17
SCTP_SRC	16	No	Yes	If EtherType = 0x86dd and IP_PROTO = 132
ICMPV6_TYPE	8	No	Yes	If EtherType = 0x86dd and IP_PROTO = 58
TCP_DST	16	No	Yes	If EtherType = 0x86dd 00 and IP_PROTO = 6
UDP_DST	16	No	Yes	If EtherType = 0x86dd and IP_PROTO = 17
SCTP_DST	16	No	Yes	If EtherType = 0x86dd and IP_PROTO = 132
ICMPv6_CODE	8	No	Yes	If EtherType = 0x86dd and IP_PROTO = 58
MPLS L2 PORT	16	No	Yes	MPLS L2 Port

Notes:

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

4.1.20.2 Instruction Types

The Policy ACL Flow Table instructions are shown in Table 122.

Table 122 Policy ACL Flow Table Instructions

Name	Argument	Description
Meter	Meter identifier	Optional. Apply the meter indicated. Meter entry must exist prior to installing the flow.
Goto-Table	Color Based Actions	Optional. If not supplied pipeline processing is terminated and the action set is applied.
Apply Actions	Action list	Optional. Only the actions in Table 123 can be specified.
Clear Actions		Used to clear the action set.
Write Actions	Action set	Only the actions in Table 124 or Table 125 can be specified, depending on rule type.

The packet is dropped if there is no group or output action in the action set.

4.1.20.3 Action List Actions

The Policy ACL Flow Table action lists support the actions listed in Table 123.

Table 123 Policy ACL Flow Table Action List Actions

Name	Argument	Description
Set-Field	Color	New packet drop precedence. Optional. Overrides the color set by the meter.
Set-Field	Color Actions Index	Index into Color Based Actions Flow Table. Required for Color Based Actions.
Set-Field	Traffic Class	New packet priority. Optional.

4.1.20.4 Action Set Actions

The Policy ACL Flow Table action set supports the actions listed in Table 124 for VLAN match rule types, and the actions in Table 125 for tunnel match rule types.

Table 124 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: L2 Interface, L2 Rewrite, L2 Multicast, L3 Unicast, L3 Multicast, or L3 ECMP; must respect 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-Field	VLAN PCP	Set outer VLAN priority marking independent of Color. Optional.
Set-Field	IP_ECN	Set ECN field marking. Applicable to IP packets independent of Color. Optional.
Set-Field	IP_DSCP	Set IP DSCP marking. Applicable to IP packets independent of Color. Optional.

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 that if the group entry type is L2 Interface. L2 Rewrite, or L2 Multicast then these checks will not be done.

Table 125 Policy ACL Flow Table Tunnel Flow Entry Action Set

Name	Argument	Description
Group	Group	Sets output group entry for multicast forwarding or flooding. Group entry must exist, and must be one of OF-DPA L2 Overlay Multicast or L2 Overlay Flood sub-type with a Tunnel Id for the tenant forwarding domain.
Output	ifNum	Sets output port for unicast forwarding. Must be a tunnel logical port consistent with the rule forwarding domain.
Set-Field	VLAN PCP	Set outer VLAN priority marking independent of Color. Optional.
Set-Field	IP_ECN	Set ECN field marking. Applicable to IP packets independent of Color. Optional.

Name	Argument	Description
Set-Field	IP_DSCP	Set IP DSCP marking. Applicable to IP packets independent of Color. Optional.

4.1.20.5 Counters and Flow Expiration

The Policy ACL Flow Table counters are listed in Table 126. These are applicable to both VLAN and Tenant flow entries.

Table 126 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 that hit this flow entry.
Received Bytes	64	Per-entry	Number of bytes that hit this flow entry.

Policy ACL Flow Table expiry provisions are shown in Table 127. Each flow entry can have its own timeout values.

Table 127 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.

4.1.21 Color Based Actions Flow Table

The Color Based Actions Flow Table provides packet editing actions based on the packet color. It is mainly used to set packet QoS fields such as DSCP or PCP.

Note: Actions in this table will override interface remark actions of the same type applied from group entries.

4.1.21.1 Flow Entry Types and Match Fields

The Maintenance Point Flow Table implements the single flow entry type listed in Table 128.

Table 128 Color Based Actions Flow Table Entry Types

Name	Description
Color Actions	The only rule type.

The Color Based Actions Flow Table match fields are listed in Table 129.

Table 129 Color Based Actions Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
Color	2	No	No	Packet color determined from prior stages
Color Actions Index	32	No	No	Set by Policy ACL Flow Table.

The default on a miss is to do nothing.

4.1.21.2 Instruction Types

The Color Based Actions Flow Table can have the instructions shown in Table 130. Since there is no next table, there are no Goto-Table or Write Metadata instructions.

Table 130 Color Based Actions Flow Table Instructions

Name	Argument	Description
Clear-Actions	none	Used to drop the packet.
Apply-Actions	Action List	Allowed actions are listed in Table 131
Write Actions	Action Set	Allowed actions are listed in Table 132.

4.1.21.3 Actions

The Color Based Actions Flow Table has the actions listed in Table 131 and Table 132.

Table 131 Color Based Actions Flow Table Action List Actions

Name	Argument	Description
Output	CONTROLLER	Send a copy to the controller.

Table 132 Color Based Actions Flow Table Action Set Actions

Name	Argument	Description
Set-Field	Traffic Class	New packet traffic class. Optional
Set-Field	VLAN PCP	New outer VLAN priority marking. Optional.
Set-Field	IP_ECN	New ECN field marking. Applicable to IP packets. Optional.
Set-Field	IP_DSCP	New IP DSCP marking. Applicable to IP packets. Optional.

4.1.21.4 Counters and Flow Expiry

The Color Based Actions Flow Table counters are listed in Table 133.

Table 133 Color Based Actions 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

Color Based Actions Flow Table expiry provisions are as indicated in Table 134.

Table 134 Color Based Actions Flow Table Expiry

Name	Bits	Description

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.

4.2 Egress Flow Tables

Egress flow tables permit matching in the context of the egress port after group entry processing and can only go to higher numbered egress flow tables. The first egress table action set contains just the output action, and no group actions can be added to it. Egress tables are restricted from changing the output port, however they can use the action list to send a copy to another port. They can remove the output action using a Clear Actions instruction, which will effectively and cause the packet to be dropped.[29]

OF-DPA 2.0 has four egress flow tables. For all of them the default behavior on a miss is to go to the "next" egress table or output the packet.

4.2.1 Egress VLAN Flow Table

The Egress VLAN Flow Table is used for VLAN translation and for recognizing frames that require OAM Maintenance Point processing, analogous to the VLAN Flow Table described earlier.

4.2.1.1 Flow Entry Types and Match Fields

The Egress VLAN Flow Table supports the Flow Entry Types listed in Table 135.

Table 135 Egress VLAN Flow Table Flow Entry Types

Type	Description
VLAN Translate Single Tag	Used to modify a single tagged packet. Can be used to remove the tag, change the VLAN Id, or to modify the VLAN Id and push another tag. L2 Unfiltered Interface must be last group.
VLAN Translate, Double Tag	Set the OVID metadata value to the outer VLAN Id and pop the outer tag. A Goto-Table instruction specifies the Egress VLAN 1 Flow table for further processing.

The Egress VLAN Flow Table match fields are listed in Table 19.

[29] Egress Flow Tables are a planned feature of OpenFlow 1.5.

Table 136 Egress VLAN Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ACTSET_OUTPUT	32	No	No	Egress physical port.
VLAN_VID	16	No	No	Outer VLAN Id. Exact match. Must include OFPVID_PRESENT.
ALLOW_VLAN_TRANSLATION	1	No	No	Must be 1 to apply rule.

4.2.1.2 Instruction Types

The Egress VLAN Flow table supports the instruction types listed in Table 137.

Table 137 Egress VLAN Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	The Egress VLAN Flow Table supports the actions specified in Table 138.
Clear-Actions		Can optionally be used to remove the output action and drop the packet
Goto-Table	Table	One of: Egress VLAN 1 Flow Table, Egress Maintenance Point Flow Table, or the Egress DSCP PCP Remark Table. The next table should be the Maintenance Point Flow Table if an Ethernet Service MIP or MEP configured.

4.2.1.3 Actions

The Egress VLAN Flow table uses Apply Actions for port VLAN tagging and assignment, as shown in Table 138.

Table 138 Egress VLAN Flow Table Action List

Name	Argument	Description
Set-Field	VLAN_VID	
Set-Field	OVID	Pipeline metadata field representing an outer tag VLAN Id that was popped, so that it can be used as a match field in

Name	Argument	Description
		the VLAN 1 Flow Table for double tag processing.
Push VLAN	TPID	Used in translating single to double tag. TPID must be 0x8100.
Pop VLAN		Used in processing double tagged frames, where the Goto-Table instruction specifies the Egress VLAN 1 Flow table.

4.2.1.4 Counters and Flow Expiry

The Egress VLAN Flow Table supports the table and flow entry counters listed in Table 139.

Table 139 Egress 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 timeout ageing per entry is supported, as indicated in Table 23 Table 140.

Table 140 Egress 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.

4.2.2 Egress VLAN 1 Flow Table

The Egress VLAN 1 Flow Table is used for double tag VLAN translation and matching. As with the VLAN table, the OVID pipeline metadata field is used so that the Egress VLAN 1 Flow Table can match on two VLAN tags.

4.2.2.1 Flow Entry Types and Match Fields

The VLAN 1 Flow Table supports the Flow Entry Types listed in Table 141.

Table 141 Egress VLAN 1 Flow Table Flow Entry Types

Type	Description
VLAN Assignment	Exact match on IN_PORT, VLAN_VID, and OVID. Can optionally: pop the tag (packet becomes untagged); set the VLAN Id (single tag); or set the VLAN Id, push a tag, and set the VLAN Id for the pushed tag (double tagged). Next table is Maintenance Point Flow Table if a MIP or MEP is configured, otherwise the packet is output. A Clear Actions instruction can be used to drop the packet.

The Egress VLAN 1 Flow Table match fields are listed in Table 142.

Table 142 Egress VLAN 1 Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ACTSET_OUTPUT	32	No	No	Egress port. Must be a physical port (high order 16 bits zero).
VLAN_VID	16	No	No	Inner VLAN Id.
OVID	16	No	No	Outer VLAN Id, set by a VLAN table flow entry.

4.2.2.2 Instruction Types

The VLAN table supports the instruction types listed in Table 143.

Table 143 Egress VLAN 1 Flow Table Instructions

Name	Argument	Description
Apply-Actions	Action List	The Egress VLAN 1 Flow Table supports the actions specified in Table 144.
Clear-Actions		Can optionally be used to remove the output action and drop the packet
Goto-Table	Table	Maintenance Point Flow Table if an Ethernet Service MIP or MEP configured. Unless there is a Clear Actions instruction, the packet is forwarded to its output port.

4.2.2.3 Actions

The Egress VLAN 1 Flow Table action list is as shown in Table 144.

Table 144 Egress VLAN 1 Flow Table Action List

Name	Argument	Description
Set-Field	VLAN_VID	Used to set the VLAN Id in the innermost tag or after pushing a new outermost tag.
Pop VLAN		Used to remove innermost tag from a double tagged frame, possibly leaving it untagged.
Push VLAN	TPID	Used push an outermost tag so the final frame is double tagged. TPID must be 0x8100.

4.2.2.4 Counters and Flow Expiry

The Egress VLAN 1 Flow Table supports the table and flow entry counters listed in Table 145.

Table 145 Egress VLAN 1 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 timeout ageing per entry is supported, as indicated in Table 146.

Table 146 Egress VLAN 1 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.

4.2.3 Egress Maintenance Point Flow Table

The Egress Maintenance Point Flow Table determines where to forward an OAM control frame for processing based on the LMEP Id and the opcode parsed from the Y.1731 PDU. Essentially it duplicates the earlier Maintenance Point Flow Table but as in egress.

The Egress Maintenance Point Flow Table is only used for Ethernet OAM.

4.2.3.1 Flow Entry Types and Match Fields

The Egress Maintenance Point Flow Table implements the flow entry types listed in Table 147.

Table 147 Egress Maintenance Point Flow Table Entry Types

Name	Description
Up MHF PDU Process (LBM)	Loopback frame addressed to MIP.
Up MHF PDU Process (Our LTM)	Link Trace frame at MIP's MDL.
PDU Passthru (higher MDL)	Client Service OAM PDU (higher level MDL)
PDU Process (Provider Up MEP)	Provider OAM PDU.
PDU Drop (lower MDL)	Active filter by MDL.
Data Passthru (Provider Up MEP)	Data frame at MEP.

The Egress Maintenance Point Flow Table match fields are listed in Table 148.

Table 148 Egress Maintenance Point Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	Yes	Prerequisite for OAM PDU.
ACTSET_OUTPUT		No	No	Output port
VLAN_VID	16	No	No	VLAN Id
OAM_Y1731_OPCODE	16	No	Yes	Parsed from the IEEE 802.1ag/Y.1731 header.
OAM_Y1731_MDL	3	No	Yes	Parsed from the IEEE 802.1ag/Y.1731 header.

Field	Bits	Maskable	Optional	Description
ETH_DST	48	No	Yes	Identifies unicast frames addressed to the MIP.

The built-in default on a miss is to pass.

4.2.3.2 Instruction Types

The Egress Maintenance Point Flow Table can have the instructions shown in Table 149. At least one must be specified. Since there is no next table the packet is dropped.

Table 149 Egress Maintenance Point Flow Table Instructions

Name	Argument	Description
Write Actions	Action Set	Actions are listed in Table 150.
Clear Actions		Can optionally be used to remove the output action in preparation for dropping the packet.
Apply Actions	Action List	Actions are listed in Table 151.
Goto-Table	Table	Must be the Egress DSCP PCP Remark Flow Table.

4.2.3.3 Actions

Table 150 Egress Maintenance Point Flow Table Action Set Actions

Name	Argument	Description
Set-Field	LMEP_Id	Sets the LMEP Id. Will be used by the OAM engine to process the PDU.

Table 151 Egress Maintenance Point Flow Table Action List Actions

Name	Argument	Description
Output		Can be LOCAL for processing by the local OAM engine.
Set-Field	LMEP_Id	Sets the LMEP Id. Will be used by the OAM engine to process the PDU.

4.2.3.4 Counters and Flow Expiry

Egress Maintenance Point Flow Table supports the table and flow entry counters listed in Table 152.

Table 152 Egress Maintenance Point 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 timeout ageing per entry is supported, as indicated in Table 153.

Table 153 Egress Maintenance Point 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.

4.2.4 Egress DSCP PCP Remark Flow Table

The Egress DSCP PCP Remark Flow Table is used to remark the IP TOS/DSCP and VLAN PCP/DEI fields based on the current Color and Traffic Class values.

4.2.4.1 Flow Entry Types and Match Fields

The Egress DSCP PCP Remark Flow Table implements the flow entry types listed in Table 154.

Table 154 Egress DSCP PCP Remark Flow Table Entry Types

Name	Description
Remark DSCP PCP (IPv4)	Apply current value of Color and Traffic Class for the specified port and QoS index. EtherType required as prerequisite. Can set both DSCP and PCP/DEI values.
Remark DSCP PCP (IPv6)	Apply current value of Color and Traffic Class for the specified port and QoS index. EtherType required as prerequisite. Can set both DSCP and PCP/DEI values.
Remark PCP	Apply current value of Color and Traffic Class for the specified port and QoS index. Can only set VLAN PCP/DEI values.

The Egress DSCP PCP Remark Flow Table match fields are listed in Table 155.

Table 155 Egress DSCP PCP Remark Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ETH_TYPE	16	No	Yes	Required as prerequisite for setting DSCP field.
Traffic Class	4	No	No	Current packet traffic class
Color	2	No	No	Current packet color
ACTSET Output		No	No	Egress port to which this rule applies

The built-in default on a miss is to pass.

4.2.4.2 Instruction Types

The Egress DSCP PCP Remark Flow Table can have the instructions shown in Table 156. At least one must be specified. Since there is no next table the packet is dropped.

Table 156 Egress DSCP PCP Remark Flow Table Instructions

Name	Argument	Description
Apply Actions	Action List	Actions are listed in Table 157.
Goto-Table	Table	Can only be the Egress TPID Flow Table.

4.2.4.3 Actions

Table 157 Egress DSCP PCP Remark Flow Table Action List Actions

Name	Argument	Description
Set-Field	VLAN_PCP	Optional
Set-Field	VLAN_DEI	Optional

Name	Argument	Description
Set-Field	DSCP	Optional, only applies to IP packets.

4.2.4.4 Counters and Flow Expiry

The Egress DSCP PCP Remark Flow Table supports the table and flow entry counters listed in Table 158.

Table 158 Egress DSCP PCP Remark 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 timeout ageing per entry is supported, as indicated in Table 159.

Table 159 Egress DSCP PCP Remark 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.

4.2.5 Egress TPID Flow Table

The Egress TPID Flow Table is used to set the TPID (EtherType) of the outermost VLAN tag. In most cases this table does not need to be programmed, unless it is desired set the outermost tag TPID to an S-Tag (0x88a8). This and 0x8100 are the only two TPID values recognized by OpenFlow.

This table has a built-in table miss entry that sets the outermost VLAN tag TPID to 0x8100, regardless what value might be in the frame. If the packet is untagged the TPID cannot be changed.

4.2.5.1 Flow Entry Types and Match Fields

The Egress TPID Flow Table implements the flow entry types listed in Table 160.

Table 160 Egress TPID Flow Table Entry Types

Name	Description

Name	Description
S-Tag	If the packet has at least one VLAN Tag, set the TPID of the outermost tag to 0x88a8.

The Egress TPID Remark Flow Table match fields are listed in Table 155.

Table 161 Egress TPID Flow Table Match Fields

Field	Bits	Maskable	Optional	Description
ACTSET_OUTPUT	32	No	No	Egress Port.
VLAN_VID	16	Yes	No	Prerequisite verifying that the packet is tagged. Can only match value OFPVID_PRESENT with mask OFPVID_PRESENT.

4.2.5.2 Instruction Types

The Egress TPID Flow Table can have the instructions shown in Table 162. Only one can be programmed per port. Since there is no next table the packet is output.

Table 162 Egress TPID Flow Table Instructions

Name	Argument	Description
Apply Actions	Action List	Actions are listed in Table 157.

4.2.5.3 Actions

The actions in the Action List must be entered in the order given. The only way OpenFlow has for setting the TPID is when a VLAN tag is pushed. This Action List copies the VLAN Id, then pops the tag, pushes a new tag with the intended TPID, then sets the VLAN Id to the copied value.

Table 163 Egress TPIPD Flow Table Action List Actions

Name	Argument	Description
Copy Field	PACKET_REG(1)	Copy the VLAN Id to a temporary register.
POP VLAN	-	After copying the VLAN Id, pops the tag.

Name	Argument	Description
PUSH VLAN	ETH_TYPE	Must be 0x88a8.
Set-Field	PACKET_REG(1)	Sets the VLAN Id to the copied value.

4.2.5.4 Counters and Flow Expiry

The Egress TPID Flow Table supports the table and flow entry counters listed in Table 164

Table 164 Egress TPID 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 timeout ageing per entry is supported, as indicated in Table 159.

Table 165 Egress TPID 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.

4.3 Group Table

Most forwarding actions are embodied in group table entries. OF-DPA supports a defined set of group table entry types and enforces type checking consistency.

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, effectively partitioning the group table identity name space to create logical sub-tables. The naming convention is shown in Table 166.

Table 166 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 Data Center Overlay 9: OF-DPA MPLS Label 10: OF-DPA MPLS Forwarding 11: OF-DPA L2 Unfiltered Interface

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

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

4.3.1 OF-DPA L2 Interface Group Entries

OF-DPA L2 Interface Group entries are of OpenFlow indirect type, with a single action bucket. They are used for egress VLAN filtering and tagging. The identifier scheme is described in Table 166. If a specific set of VLANs is allowed on a port, appropriate group entries must be defined for the VLAN and port combinations.

Note: OF-DPA uses the L2 Interface group declaration to configure the port VLAN filtering behavior. This approach was taken since OpenFlow does not support configuring VLANs on physical ports.

Note: The VLAN status on the port cannot be changed or translated to a different VLAN by the Egress VLAN Flow Table.

4.3.1.1 Naming Convention

Table 167 details the OF-DPA L2 Interface group entry identifier sub-fields that encode combinations of egress port and VLAN Id.

Table 167 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 Interface)

4.3.1.2 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 168 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. Optional, for untagged ports.
Set Field	ALLOW VLAN TRANSLATION	Set to zero to prevent egress pipeline VLAN translation.

4.3.1.3 Counters

OF-DPA L2 Interface group entry counters are as shown in Table 169.

Table 169 OF-DPA L2 Interface Group Entry Counters

Name	Bits	Type	Description

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

4.3.2 OF-DPA L2 Unfiltered Interface Group Entries

L2 Unfiltered Interface Group entries are of OpenFlow indirect type, with a single action bucket. OF-DPA L2 Unfiltered Interface group entries are similar to L2 Interface group entries except that they are used for forwarding to ports where VLAN filtering and tagging is not required on the egress port.

As with L2 Interface group entries, OF-DPA uses the L2 Unfiltered Interface groups to configure the port to not do VLAN filtering. Thus an output port cannot have both L2 Interface and L2 Unfiltered Interface groups defined for it.

4.3.2.1 Naming Convention

Table 170 details the OF-DPA L2 Unfiltered Interface group entry identifier sub-field encodings.

Table 170 OF-DPA L2 Unfiltered Interface Group Naming Convention

Field	Bits	Description
Port identifier	[15:0]	Identifies a physical port (ifNum)
Reserved	[27:16]	Must be zero
Type	[31:28]	11 (L2 Unfiltered Interface)

4.3.2.2 Action Buckets

The single action bucket, detailed in Table 171, specifies the output port. This is essentially the same as for the L2 Interface group action bucket but without the VLAN tag Pop action.

Table 171 OF-DPA L2 Unfiltered Interface Group Bucket Actions

Field	Argument	Description
Output	Port	Physical output port.
Set Field	ALLOW VLAN	Set to 1 to permit egress pipeline VLAN translation.

Field	Argument	Description
	TRANSLATION	

4.3.2.3 Counters

OF-DPA L2 Interface group entry counters are as shown in Table 172.

Table 172 OF-DPA L2 Unfiltered Interface Group Entry Counters

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

4.3.3 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 only the Policy ACL Flow Table has the ability to use it in its Write-Actions instruction.

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 a chained L2 Interface group entry. Note that 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.

4.3.3.1 Naming Convention

Table 173 details the OF-DPA L2 Rewrite group entry identifier sub-fields that encode the type and VLAN Id.

Table 173 OF-DPA L2 Rewrite Group Entry Type Naming Convention

Field	Bits	Description
Id	[27:0]	Index to differentiate group entries of this type

Field	Bits	Description
Type	[31:28]	1 (OF-DPA L2 Rewrite)

4.3.3.2 Action Buckets

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

Table 174 OF-DPA L2 Rewrite Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Must chain to a L2 Interface group entry. Required.
Set-Field	MAC_SRC	Re-write the source MAC. Optional.
Set-Field	MAC_DST	Re-write the destination MAC. Optional.
Set-Field	VLAN-id	Re-write the VLAN Id. Optional.

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

4.3.3.3 Counters

OF-DPA L2 Rewrite group entry counters are as shown in Table 175 for completeness.

Table 175 OF-DPA L2 Rewrite Group Entry Counters

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

4.3.4 OF-DPA L3 Unicast Group Entries

OF-DPA L3 Unicast group entries are used to supply the routing next hop and output interface for packet forwarding. To properly route a packet from either the Routing Flow Table or the Policy ACL Flow Table, the forwarding flow entry must reference an OF-DPA L3 Unicast Group entry.

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, and is inherited by chained group entries. It is not visible to the controller and hence cannot be modified or read.

All packets must have a VLAN tag. A chained L2 Interface group entry must be in the same VLAN as assigned by the OF-DPA L3 Unicast Group entry.

4.3.4.1 Naming Convention

The naming convention for OF-DPA L3 Unicast Group entries is shown in Table 176.

Table 176 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)

4.3.4.2 Action Buckets

The single action bucket is as shown in Table 177.

Table 177 OF-DPA L3 Unicast Bucket Actions

Field	Argument	Description
Group	Group-id	Must chain to a L2 Interface group entry. ALLOW-IN_PORT permits the chained group entry output action to include the packet IN_PORT. Required.
Set-Field	MAC_DST	Write the next hop destination MAC. Required.
Set-Field	MAC_SRC	Write the source MAC corresponding to the L3 output interface. Required.
Set-Field	VLAN-id	Write the VLAN Id corresponding to the L3 output interface. Required.

4.3.4.3 Counters

The OF-DPA L3 Unicast group entry counters are as shown in Table 178.

Table 178 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

4.3.5 OF-DPA L2 Multicast Group Entries

OF-DPA L2 multicast group entries are of OpenFlow ALL type. There can be multiple action buckets, each referencing an output port by chaining to an OF-DPA L2 Interface Group entry.

Note: By OpenFlow default, 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.

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.

4.3.5.1 Naming Convention

OF-DPA L2 Multicast group entries use the naming convention in Table 179.

Table 179 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)

4.3.5.2 Action Buckets

The contents of OF-DPA L2 Multicast Group entry buckets can contain only the value shown in Table 180.

Table 180 OF-DPA L2 Multicast Bucket Actions

Field	Argument	Description

Field	Argument	Description
Group	Group-id	Must chain to a L2 Interface group entry whose VLAN Id name component matches the VLAN Id component of this group entry's name.

4.3.5.3 Counters

The VL2 Multicast group entry counters are as shown in Table 181.

Table 181 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

4.3.6 OF-DPA L2 Flood Group Entries

The OF-DPA L2 Flood Group entries are used by VLAN Flow Table wildcard (destination location forwarding, or DLF) rules. Like OF-DPA L2 Multicast group entry types they are of OpenFlow ALL type. The action buckets each encode an output port. Each OF-DPA L2 Flood Group entry bucket forwards a replica to an output port, except for packet IN_PORT.

The main difference from OF-DPA L2 Multicast Group entries is how they are processed in the hardware. Note that there can be only one OF-DPA L2 Flood Group Entry per VLAN.

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.

4.3.6.1 Naming Convention

OF-DPA L2 Flood group entries follow the naming convention shown in Table 182.

Table 182 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

Field	Bits	Description
Type	[31:28]	4 (OF-DPA L2 Flood)

4.3.6.2 Action Buckets

The contents of the OF-DPA L2 Flood Group Entry action buckets can contain only the values shown in Table 183.

Table 183 OF-DPA L2 Flood Bucket Actions

Field	Argument	Description
Group	Group-id	Must chain to L2 Interface group entry whose VLAN Id name component is the same as the VLAN Id in this entry's name.

4.3.6.3 Counters

The OF-DPA L2 Multicast group entry counters are as shown in Table 184.

Table 184 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

4.3.7 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.

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 and the VLAN Id in the name of the chained OF-DPA L2 Interface group.

4.3.7.1 Naming Convention

Table 185 details the OF-DPA L3 Interface group entry identifier sub-fields.

Table 185 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)

4.3.7.2 Action Buckets

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

Table 186 OF-DPA L3 Interface Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Must chain to one if the L2 Interface group entry types, which can output the packet to IN_PORT. The VLAN Id component of the chained group entry's name 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	MAC_DST	Write the destination MAC corresponding to the IP multicast group according to RFC 1112 or RFC 2464. Optional, only required for multicast packets after L3 VPN termination.
Set-Field	VLAN-id	Write the VLAN Id corresponding to the L3 output interface. 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.

4.3.7.3 Counters

OF-DPA L3 Interface group entry counters are as shown in Table 175 for completeness.

Table 187 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

4.3.8 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.

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 or MPLS L3 VPNs 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 that any chained OF-DPA L2 Interface Group entries must be in the same VLAN as the OF-DPA L3 Multicast group entry. However chained OF-DPA L3 Interface Group entries must be in different VLANs from the OF-DPA L3 Multicast Group entry, and from each other.

4.3.8.1 Naming Convention

The naming convention for OF-DPA L3 Multicast Group entries is shown in Table 188.

Table 188 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)

4.3.8.2 Action Buckets

The action buckets contain the values shown in Table 189.

Table 189 OF-DPA L3 Multicast Bucket Actions

Field	Argument	Description
Group	Group-id	Can chain to one of: L3 Interface; L2 Interface; L3 Fast Failover; MPLS Fast Failover; or MPLS L3 VPN Label group entry types. Chained group entry names must conform to the VLAN Id requirements above.

4.3.8.3 Counters

The OF-DPA L3 Multicast group entry counters are as shown in Table 190.

Table 190 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

4.3.9 OF-DPA L3 ECMP Group Entries

OF-DPA L3 ECMP group entries are of OpenFlow type SELECT. For IP routing the action buckets reference the OF-DPA L3 Unicast group entries that are members of the multipath group for ECMP forwarding. Figure 38 illustrates this OF-DPA L3 ECMP Group entry usage.

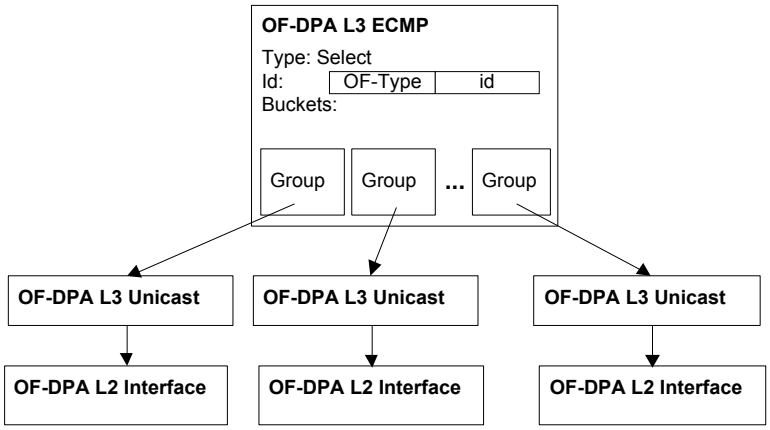

Figure 38 OF-DPA L3 ECMP Group Entry Usage

An OF-DPA L3 ECMP Group entry can also be used in a Provider Edge Router. In this packet flow it can chain to either an MPLS L3 Label group entry or to an MPLS Fast Failover group entry.

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.

4.3.9.1 Naming Convention

The naming convention for OF-DPA L3 ECMP Group entries is as shown in Table 191.

Table 191 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)

4.3.9.2 Action Buckets

The action buckets contain the single value listed in Table 192.

Table 192 OF-DPA L3 ECMP Group Entry Bucket Actions

Field	Argument	Description
Group	Group-id	May chain to an OF-DPA L3 Unicast, L3 Fast Failover, MPLS Fast Failover, or MPLS L3 VPN Label group entry.

4.3.9.3 Counters

The OF-DPA L3 ECMP group entry counters are as shown in Table 193.

Table 193 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 (sec)	32	Per-entry	Seconds since this group entry was installed

4.3.10 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 that 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 5.1.3.

4.3.10.1 OF-DPA L2 Overlay Group Sub-Types

There are four OF-DPA L2 Overlay Group sub-types. 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 that 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 sub-type name component.

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.

OF-DPA L2 Overlay Flood Over Multicast Tunnels group entries can have 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.

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

An OF-DPA L2 Overlay Multicast Over Unicast Tunnels group entry can have multiple tunnel logical port buckets. 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.

An OF-DPA L2 Overlay Multicast Over Multicast Tunnels group entry can have at most one bucket, which specifies 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.

4.3.10.2 Naming Convention

The naming convention for OF-DPA L2 Overlay group entries is shown in Table 194.

Table 194 OF-DPA L2 Overlay Group Entry Naming Convention

Field	Bits	Description

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

4.3.10.3 Action Buckets

The action buckets for all OF-DPA L2 Overlay Group Entry sub-types contain the values shown in Table 195.

Table 195 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.

4.3.10.4 Counters

The OF-DPA L2 Overlay Flood group entry counters are as shown in Table 196. These counters are individually maintained by sub-type.

Table 196 OF-DPA L2 Overlay 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 (sec)	32	Per-entry	Seconds since this group entry was installed

4.3.11 OF-DPA MPLS Interface Group Entry

An OF-DPA Interface Group Entry is of OpenFlow type INDIRECT. It is used to set the outgoing L2 header to reach the next hop label switch router or provider edge router.

4.3.11.1 Naming Convention

The naming convention for OF-DPA MPLS group entries is shown in Table 197.

Table 197 OF-DPA MPLS Interface Group Entry Naming Convention

Field	Bits	Description
Index	[23:0]	Used to differentiate MPLS group entries of the same sub-type
Sub-Type	[27:24]	Identifies the type of the entry: 0: OF-DPA MPLS Interface
Type	[31:28]	9 (OF-DPA MPLS Label)

4.3.11.2 Action Buckets

The action bucket for the OF-DPA MPLS Interface Group entry type contains the actions listed in Table 198.

Table 198 OF-DPA MPLS Interface Group Entry Bucket Actions

Field	Argument	Description
Set-Field	MAC-DST	Destination MAC address. Required.
Set-Field	MAC-SRC	Source MAC address. Required.
Set-Field	VLAN_VID	VLAN. Required.
OAM_LM_TX_Count	LMEP_Id [, Traffic Class]	Used for Section OAM Loss Measurement. Optional.
Group	Group entry	May chain to either an OF-DPA L2 Interface or an L2 Unfiltered Interface group entry type. Required.

4.3.11.3 Counters

The OF-DPA MPLS Interface Group entry type counters are as shown in Table 199.

Table 199 OF-DPA MPLS Interface Group Type 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

4.3.12 OF-DPA MPLS Label Group Entries

OF-DPA Label Group entries are of OpenFlow INDIRECT type. There are four MPLS label Group entry sub-types, all with similar structure. These can be used in different configurations to push up to three labels for tunnel initiation or LSR swap.

4.3.12.1 Naming Convention

The naming convention for OF-DPA MPLS Label group entries is shown in Table 194. Different sub-types are used for different action sets.

Table 200 OF-DPA MPLS Label Group Entry Naming Convention

Field	Bits	Description
Index	[23:0]	Used to differentiate MPLS group entries of the same sub-type
Sub-Type	[27:24]	Identifies the type of the entry: 1: OF-DPA MPLS L2 VPN Label 2: OPF-DPA MPLS L3 VPN Label 3: OF-DPA MPLS Tunnel Label 1 4: OF-DPA MPLS Tunnel Label 2 5: OF-DPA MPLS Swap Label
Type	[31:28]	9 (OF-DPA MPLS Label)

The contents of the action buckets vary according to the type of label.

4.3.12.2 MPLS VPN Label Action Buckets

OF-DPA MPLS VPN Label groups are used for VPWS and L3 VPN initiation.

The action bucket for OF-DPA MPLS VPN Label sub-type can contain the actions shown in Table 201.

Table 201 OF-DPA MPLS VPN Label Group Bucket Actions

Field	Argument	Description
Push L2 Header		Pushes an outermost Ethernet header on the packet. All fields are initialized to zero. If the outermost header has a VLAN tag it must be pushed using a Push VLAN action. Used only in MPLS L2 VPN group entries for MPLS-TP pseudo-wire initiation.
Push VLAN	EtherType = 0x8100	Used only in MPLS L2 VPN group entries to push a single VLAN tag on a newly pushed Ethernet header. Optional.
Push MPLS Header	EtherType = 0x8847	Used in all MPLS Label group entry types to push a new MPLS shim header on the frame immediately after the Ethernet header, including any VLAN tags. The new header becomes the outermost MPLS label.
Push CW		Used only in MPLS L2 VPN group entries to push a 32-bit Control Word after the PW label. The Control Word is initialized to all zeroes and is pushed between the MPLS shim label and the packet payload. This action can only be executed once and only after a Push MPLS Header for the bottom of stack (PW) label.
Set-Field	MPLS Label	20-bit value, set on the outermost MPLS header label field.
Set-Field	BOS	1-bit outermost MPLS header bottom of stack field.
Set-Field	TC	3-bit value for outermost MPLS header EXP field. Cannot be used in the same action set as Set TC from table action. Overrides any value set by an MPLS Flow Table action.
Set MPLS_TC From VPN Table	QoS Index	Index into MPLS VPN label remark table to lookup EXP value based on packet Traffic Class and Color pipeline match field values.
Set-Field	TTL	Eight-bit value for outermost MPLS TTL field. Cannot be used in the same action set as a Copy TTL outwards action. Overrides any value set by an MPLS Flow Table action.
Copy TTL outwards		Used only in MPLS L3 VPN Label group entries to copy

Field	Argument	Description
		the TTL from the IP header.
Set VLAN PRI From VPN Table	QoS Index	Sets the 802.1p priority field in the L2 header. Index into MPLS VPN Label Remark Action table to lookup VLAN_PRI value based on packet Traffic Class and Color values. Not used if final packet does not have a VLAN tag.
Group	Group entry	Possible values: OF-DPA MPLS Interface; OF-DPA MPLS Fast Failover; OF-DPA MPLS Tunnel Label 1 group entry

4.3.12.3 MPLS Tunnel Label 1 Action Buckets

The action bucket for OF-DPA MPLS Tunnel Label 1 sub-type can contain the actions shown in Table 202.

Table 202 OF-DPA MPLS Tunnel Label 1 Group Bucket Actions

Field	Argument	Description
Push MPLS Header	EtherType = 0x8847	Push a new MPLS shim header on the frame immediately after the Ethernet header, including any VLAN tags. The new header becomes the outermost MPLS label.
Set-Field	MPLS Label	20-bit value for outermost MPLS label field.
Set-Field	TC	3-bit value for outermost MPLS header EXP field. Cannot be used in the same action set as Set TC from table action. Overrides any value set by an MPLS Flow Table action.
Copy TC outwards		Copy EXP field from the previously outermost shim header which is now immediately under this tunnel label.
Set MPLS_TC From Tunnel Table	QoS Index	Index into MPLS Tunnel Label Remark Action to lookup EXP value based on packet traffic class and color values.
Set VLAN PRI From Tunnel Table	QoS Index	Sets the 802.1p priority field if outermost label. Index into the MPLS Tunnel Label Remark Action to lookup VLAN_PRI value based on packet Traffic Class and Color values. Not used if final packet does not have a VLAN tag.
Set-Field	TTL	Eight-bit value for outermost MPLS TTL field. Overrides a copy out action if both are supplied.

Field	Argument	Description
Copy TTL outwards		Copy TTL from next MPLS shim or IP header.
Group	Group entry	Possible values: OF-DPA MPLS Interface; OF-DPA MPLS Fast Failover; OF-DPA MPLS Tunnel Label 2 group entry

4.3.12.4 MPLS Tunnel Label 2 Action Buckets

The action bucket for OF-DPA MPLS Tunnel Label 2 sub-type can contain the values shown in Table 203.

Table 203 OF-DPA MPLS Tunnel Label 2 Actions

Field	Argument	Description
Push MPLS Header	EtherType = 0x8847	Push a new MPLS shim header on the frame immediately after the Ethernet header, including any VLAN tags. The new header becomes the outermost MPLS label.
Set-Field	MPLS Label	20-bit value for outermost MPLS label field.
Set-Field	TC	3-bit value for outermost MPLS header EXP field. Cannot be used in the same action set as Set TC from table action. Overrides any value set by an MPLS Flow Table action.
Copy TC outwards		Copy EXP field from the previously outermost shim header which is now immediately under this tunnel label.
Set MPLS_TC from Tunnel Table	QoS Index	Index into the MPLS Tunnel Label Remark Action table to lookup EXP value based on packet traffic class and color values.
Set VLAN PRI From Tunnel Table	QoS Index	Sets the 802.1p priority field if outermost label. Index into the MPLS Tunnel Label Remark Action to lookup VLAN_PRI value based on packet Traffic Class and Color values. Not used if final packet does not have a VLAN tag.
Set-Field	TTL	Eight-bit value for outermost MPLS TTL field.
Copy TTL outwards		Copy TTL from next MPLS shim.
Group	Group entry	Only possible value is an OF-DPA MPLS Interface group

Field	Argument	Description
		entry.

4.3.12.5 MPLS Swap Label Action Buckets

The action bucket for the OF-DPA MPLS Swap Label sub-type can contain the values shown in Table 204.

Table 204 OF-DPA MPLS Swap Label Actions

Field	Argument	Description
Set-Field	MPLS Label	20-bit value to replace that in the outermost MPLS label field.
Set-Field	TC	3-bit value for outermost MPLS header EXP field. Cannot be used in the same action set as Set TC from table action. Overrides any value set by an MPLS Flow Table action.
Copy TC outwards		Copy EXP field from next shim. If next header is not an MPLS shim, operation is a NOP and TC is unchanged.
Set MPLS_TC From VPN Table	QoS Index	Index into MPLS VPN Label Remark Action table to lookup EXP value based on packet traffic class and color values.
Set-Field	TTL	Eight-bit value for outermost MPLS TTL field.
Copy TTL outwards		Copy TTL from next MPLS shim or IP header. If next header is neither MPLS nor IP, operation is a NOP.
Set VLAN PRI From VPN Table	QoS Index	Sets the 802.1p priority field if outermost label. Index into the MPLS VPN Label Remark Action to lookup VLAN_PRI value based on packet Traffic Class and Color values. Not used if final packet does not have a VLAN tag.
Group	Group entry	Possible values: OF-DPA MPLS Interface; OF-DPA MPLS Fast Failover; OF-DPA MPLS Tunnel Label 1 group entry.
Check Drop Status	LMEP Id, 0	For LCK

4.3.12.6 Counters

The counters supported for all OF-DPA MPLS Label sub-type group entries are as shown in Table 205.

Table 205 OF-DPA MPLS Label 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 (sec)	32	Per-entry	Seconds since this group entry was installed

4.3.13 OF-DPA MPLS Fast Failover Group Entry

OF-DPA MPLS Fast Failover Group Entries are of OpenFlow FAST FAILOVER type. The action buckets chain to an MPLS label group. Fast Failover group entry types have two buckets, for working and protection paths. By convention the bucket selection is ordered, with the first bucket listed being the working bucket and the second bucket the protection bucket. Working and protection must be programmed in this order. The working bucket is always selected if its liveness condition is satisfied (i.e., the path working is functioning). If the liveness condition is not satisfied, then the group entry fails over to other bucket, assuming it is "live."

Note: The bucket liveness "watch_port" should be programmed with an OAM Protection Liveness Logical Port, and the "watch_group" should be programmed to ANY.

Note: The first (working) bucket is assigned protection index 1 to conform to standard nomenclature [37]. The second (protection) bucket is assigned protection index 0. The protection index is independent of bucket order.

4.3.13.1 Naming Convention

The naming convention for OF-DPA MPLS group entries is shown in Table 206.

Table 206 OF-DPA MPLS Fast Failover Group Entry Naming Convention

Field	Bits	Description
Index	[23:0]	Used to differentiate MPLS group entries of the same sub-type
Sub-Type	[27:24]	Identifies the type of the entry: 6: OF-DPA MPLS Fast Failover
Type	[31:28]	10: OF-DPA MPLS Forwarding

4.3.13.2 Action Buckets

The actions for the working and protection path buckets are shown in Table 207. Both buckets should be programmed with the same types of groups. If this convention is not followed, results may be unpredictable.

MPLS-TP L2 VPN flows should use MPLS L2 VPN Label groups. Likewise, MPLS L3 VPN flows should use MPLS L3 VPN Label groups.

Table 207 OF-DPA MPLS Fast Failover Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Depending on the packet flow can be one of: MPLS L2 VPN Label; MPLS L3 Label; MPLS Swap Label; MPLS Tunnel Label 1; or MPLS Tunnel Label 2.
Watch_Port	OAM Protection Liveness Logical Port	Determines liveness of bucket. The bucket is live if the logical port is operational. The first live bucket is used.

4.3.13.3 Counters

The counters for the OF-DPA MPLS Fast Failover Group entry are as shown in Table 208.

Table 208 OF-DPA MPLS Fast Failover 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

4.3.14 OF-DPA MPLS ECMP Group Entry

OF-DPA MPLS ECMP Group Entries are of OpenFlow SELECT type and used for MPLS ECMP multipath forwarding.

All buckets must reference the same type of OF-DPA MPLS Group. If this convention is not followed results may be unpredictable.

4.3.14.1 Naming Convention

The naming convention for OF-DPA MPLS ECMP group entries is shown in Table 209.

Table 209 OF-DPA MPLS ECMP Group Entry Naming Convention

Field	Bits	Description
Index	[23:0]	Used to differentiate MPLS group entries of the same sub-type
Sub-Type	[23:24]	Identifies the type of the entry: 8: OF-DPA MPLS ECMP
Type	[31:28]	10: OF-DPA MPLS Forwarding

4.3.14.2 Action Buckets

OF-DPA MPLS ECMP Group type entry buckets each have a single group as shown in Table 210.

Table 210 OF-DPA MPLS ECMP Group Entry Bucket Actions

Field	Argument	Description
Group	Group entry	Depending on the packet flow can be one of: MPLS Fast Failover or MPLS Swap Label.

Note: The maximum number of buckets available in any OF-DPA MPLS ECMP Group entry is platform dependent.

4.3.14.3 Counters

The OF-DPA MPLS ECMP Group entry type counters are as shown in Table 211.

Table 211 OF-DPA MPLS 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 (sec)	32	Per-entry	Seconds since this group entry was installed

4.4 Meters

This section describes the OF-DPA Meter Table. OF-DPA metering contains extensions to support RFC 2697, 2698, and 4115 token bucket meters. They are based on the Meter Table definitions in OpenFlow 1.3, but extended to allow specifying the policer mode and for changing the packet Color field.

This version of OF-DPA supports metering MPLS L2 flows using the L2 Policer Flow Table as well as metering all flows using the Policy ACL Flow Table. OpenFlow 1.3.4 specifies that a flow entry can have a single meter instruction that is evaluated before other instructions, in particular, before an Apply-Actions instruction. OF-DPA supports "color set" meter bands whose only action is to set the packet Color pipeline field. A Policy ACL Flow Table rule that has a meter instruction should not specify a Set-Field Color action using an Apply-Actions or Write-Actions instruction.

Color set meter bands can have mode and color-awareness properties. Color-blind operation effectively treats all incoming packets as Green, regardless of color. The default is color-blind. OF-DPA Meter Bands are programmed using OpenFlow Meter Band Experimenter fields.

Note: Color-based metering has been proposed as a future OpenFlow feature [36]. Although an effort has been made to align OF-DPA meters with this proposal as much as is feasible, there are a number of differences, including how the configured parameters are used to measure rate and burst using the token bucket algorithm.

4.4.1 Meter Table Entries

A Meter Table entry contains the parameters shown in Table 212. OF-DPA supports Meter Table entries with two meter bands. A default "Green" band with rate and burst of zero is assumed but not explicitly configured. Furthermore the Red band rates must be greater than the Yellow band rates. In OpenFlow terms the range determines which band applies for a particular packet.

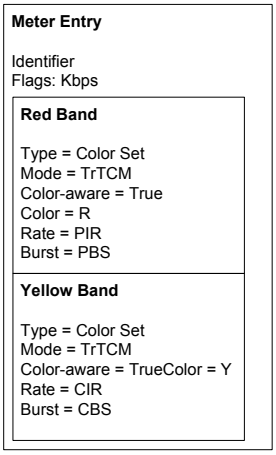

Figure 39 Meter Entry Example (TrTCM)

Figure 39 shows an example two rate three color meter (TrTCM) entry. In this example, both Yellow and Red bands are defined as Color Set, the only option. The TrTCM mode and color-aware property is specified in the bands, both of which must be configured the same. The interpretation of rate and burst properties is as specified in RFC2698. The port queuing function would typically use the packet color to determine the packet drop precedence[30]. Packets that fall in the Red band are colored Red and are usually dropped.

Table 212 lists the configuration parameters for OF-DPA Meters. All OF-DPA meters require a burst parameter.

Table 212 OF-DPA Meter Table Entry Parameters

Field	Description
Meter Id	Meter instance.
Flags	Bit position: 0: Kbps (only one of Kbps or Packets set) 1: Packets 2: Burst (must always be set) 3: Stats (per-band counters, platform dependent)
Meter Bands	Two meter bands, Yellow (1), and Red (2). Configuration is as described in Table 214.
Counters	Per-meter entry counters are specified in Table 213

Table 213 Meter Entry Counters

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

4.4.2 Meter Bands

Meter bands are configured using the entries in Table 214. The OpenFlow Drop type can only be used for Red bands.

[30] In this version of OF-DPA default queue drop precedence parameters are used and are not configurable.

OF-DPA meter operation is described in detail in Table 2and Table 3 in Section 3.8.2.

Table 214 Meter Band Configuration Parameters

Field	Description
Type	Meter bands must be Color Set. A Yellow band sets the Color to 1. A Red band sets the Color to 2.
Rate	Minimum rate for applying this meter. This value programs the token bucket CIR or PIR in the units specified by the meter entry.
Burst	Burst size for packets to be in profile. This value programs the token bucket CBS, PBS, or EBS depending on mode.
Color-Aware	0: Color-Blind 1: Color-Aware
Mode	1: TrTCM 2: SrTCM 3: Modified TrTCM (RFC4115) All other values reserved.
Counters	OF-DPA meter band counters are as in Table 215.

Note: All meters must support both rate and burst, and all bands must be configured with a burst size greater than zero. Both meter bands in a meter specification must have the same mode and color-awareness.

Note: For single rate meters, both bands must be programmed with the same rate value. The programmed rate and burst values are interpreted as in Table 3.

Table 215 Meter Band Counters

Name	Bits	Type	Description
Packet Count	64	Per-band	In-band packet count.
Byte Count	64	Per-band	In-band byte count

Note: Per-band counters are not supported on all platforms.

5 CONFIGURATION

This section describes the properties of OF-DPA features that must be configured by some mechanism outside of the OpenFlow protocol. These are requirements independent of configuration approach. Configuration information is modeled using a combination of UML diagrams and attribute tables. Possible options for control include OF-Config [14], OVSDB [16], or some local CLI.

Section 5.1 describes the OF-DPA physical and logical port configuration, and Section 5.2 describes queue configuration. OAM message processing configuration diagrams are included in Section 5.3, and protection configuration is in Section 5.4

5.1 Ports

This section lists the OF-DPA supported properties for physical, reserved, and logical ports. Ports are identified using a 32-bit ifNum value. The most significant two bytes indicate the type of port. Port types with their numbering conventions are listed in Table 216.

Table 216 Port Type Numbering Conventions

Type	Numbering	Description
Physical	0x0000 xxxx	Physical (front panel) port
Reserved	0xFFFF xxxx	Reserved ports as defined in the OpenFlow specification.
Overlay Tunnel Logical Port	0x0001 xxxx	Logical port used for VXLAN overlay tunnels. Can be a VXLAN Tunnel Endpoint Network Port, or a VXLAN Access Port. Port number assigned by configuration.
OAM Protection Liveness Logical Port	0xF000 xxxx	Logical port used for OAM protection liveness testing. These ports are predefined and do not need to be configured in order to be used. The default operational state is up (LIVE). Can be administratively configured

Type	Numbering	Description
		down, in which case the operational state will change to 1 (down).

5.1.1 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 40, from OF-Config 1.2, shows the UML port configuration data model.

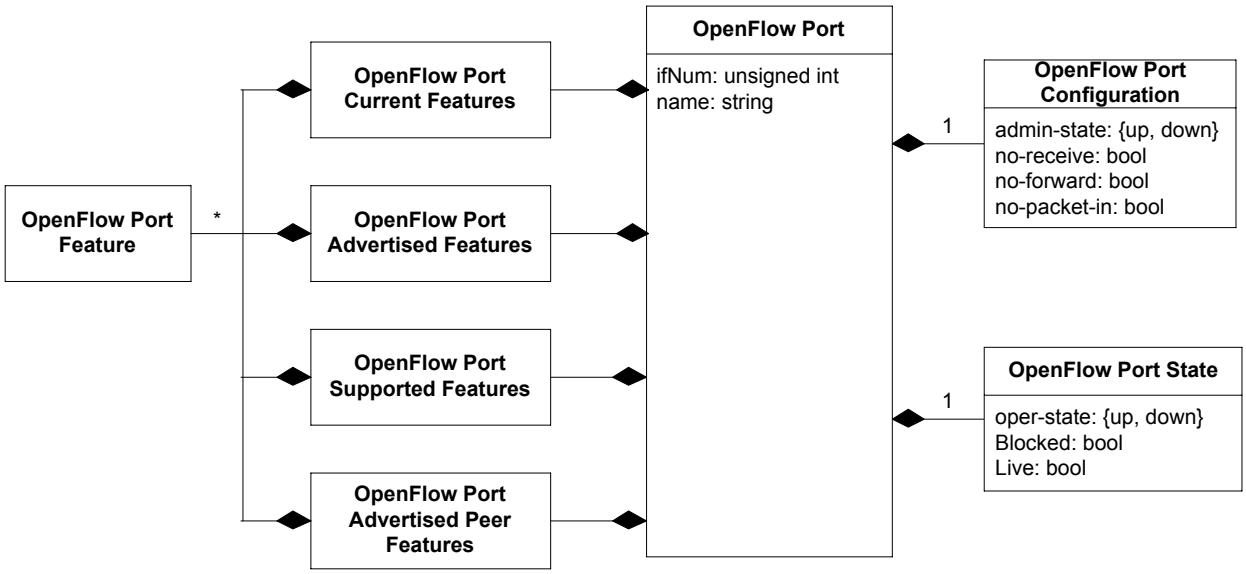

Figure 40 Port Properties Configuration

5.1.1.1 Features

OpenFlow Port Features are further modeled in terms of the sub-classes shown in Figure 41, also following OF-Config 1.2.

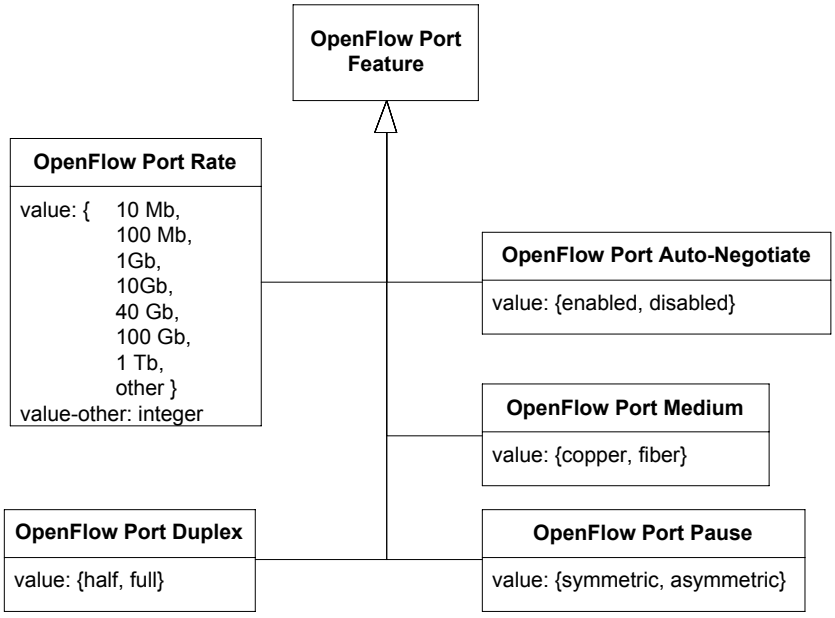

Figure 41 OpenFlow Feature Sub-Classes

OF-DPA supports the physical port features listed in Table 217.

Table 217 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 (operational) state is up (0), live (4), or down (1). Generally a port is live if operationally up.
Current Features	32	No	OF-DPA supports the feature bitmap in Table 218. A one indicates the feature is currently active.
Advertised Features	32	No	OF-DPA supports the feature bitmap in Table 218. A zero bit indicates the feature is not available.

Name	Bits	Configurable?	Description
Supported Features	32	No	OF-DPA supports the features in Table 218. A zero bit indicates the feature is not supported.
Peer Features	32	No	Bitmap indicating capabilities advertised by the peer from Table 218.
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 or CONTROLLER reserved port.

Table 218 shows the port features bitmap referenced from the table above and the OpenFlow Port Features sub-classes in Figure 41.

Table 218 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

Feature	Bit	Description
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

5.1.1.2 Counters

OF-DPA supports the port counters listed in Table 219.

Table 219 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

Name	Bits	Description
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

5.1.2 Reserved Ports

OF-DPA supports the reserved ports listed in Table 220. These do not require configuration and are listed for general information.

Table 220 OF-DPA Reserved Ports

Name	Required	Description	Use	Supported?
ALL	Yes	All interfaces, regardless of any other forwarding considerations.	Output	No
IN_PORT	Yes	Used to send packets to the ingress port to override OpenFlow default behavior. OF-DPA uses group ALLOW-IN_PORT property instead. Not to be confused with the IN_PORT match field.	Output	No
CONTROLLER	Yes	The OpenFlow controller. Output destination for sending packets to the Agent which, in turn, sends to the OpenFlow Controller in a Packet_In message. Also can optionally be used to indicate the source of packets received by the Agent in a Packet_Out message.	Input or output	Yes
TABLE	Yes	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

Name	Required	Description	Use	Supported?
ANY	Yes	Special value used in some requests.	Neither	Yes
LOCAL	No	Used to send and receive packets with the local Network Protection App. Analogous to Controller but the destination is a local OAM engine rather than the Agent. The exact mechanism is implementation dependent.	Input or output	Yes, for OAM
NORMAL	No	OpenFlow hybrid mode use.	Output	No
FLOOD	No	Legacy port for VLAN accommodation. Use an L2 Flood Group entry instead.	Output	No

5.1.3 Logical Ports

Logical Ports are used to model functionality as external of the OpenFlow pipeline, such as adding or removing tunnel encapsulation headers. 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. Thus, packets received from logical ports may not be the same as were transferred on the wire.

Logical port external function parameters are configured on the Logical Port. OF-DPA uses Logical ports for overlay tunnels and as objects for MPLS Protection Group entry buckets to watch.

5.1.3.1 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 42, which shows Tunnel Logical Ports as abstract classes.

Note that there are two Tunnel Logical Port sub-types 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. OF-DPA supports VXLAN [27] overlays.

Note: VXLAN support is hardware platform and version dependent.

The Tunnel Endpoint abstract class provides necessary configuration parameters common to different protocol sub-classes. 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 [25] port extension tag (ETAG) on that port.

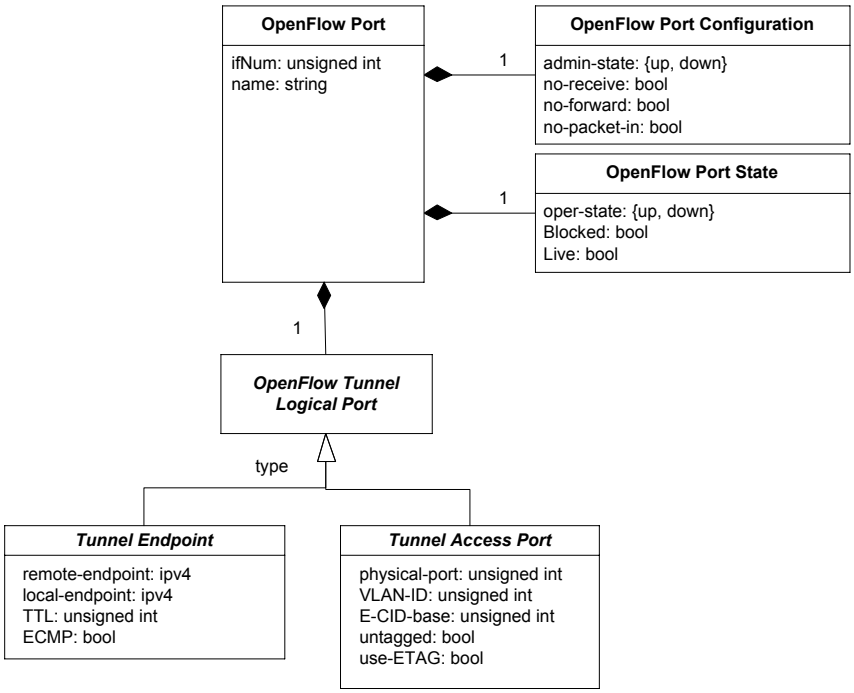

Figure 42 Tunnel Logical Port Configuration

5.1.3.2 VXLAN Tunnel Logical Port Configuration

VXLAN Logical Port configuration is shown in the UML diagram in Figure 43. 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 that 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 [25].

Note that 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.

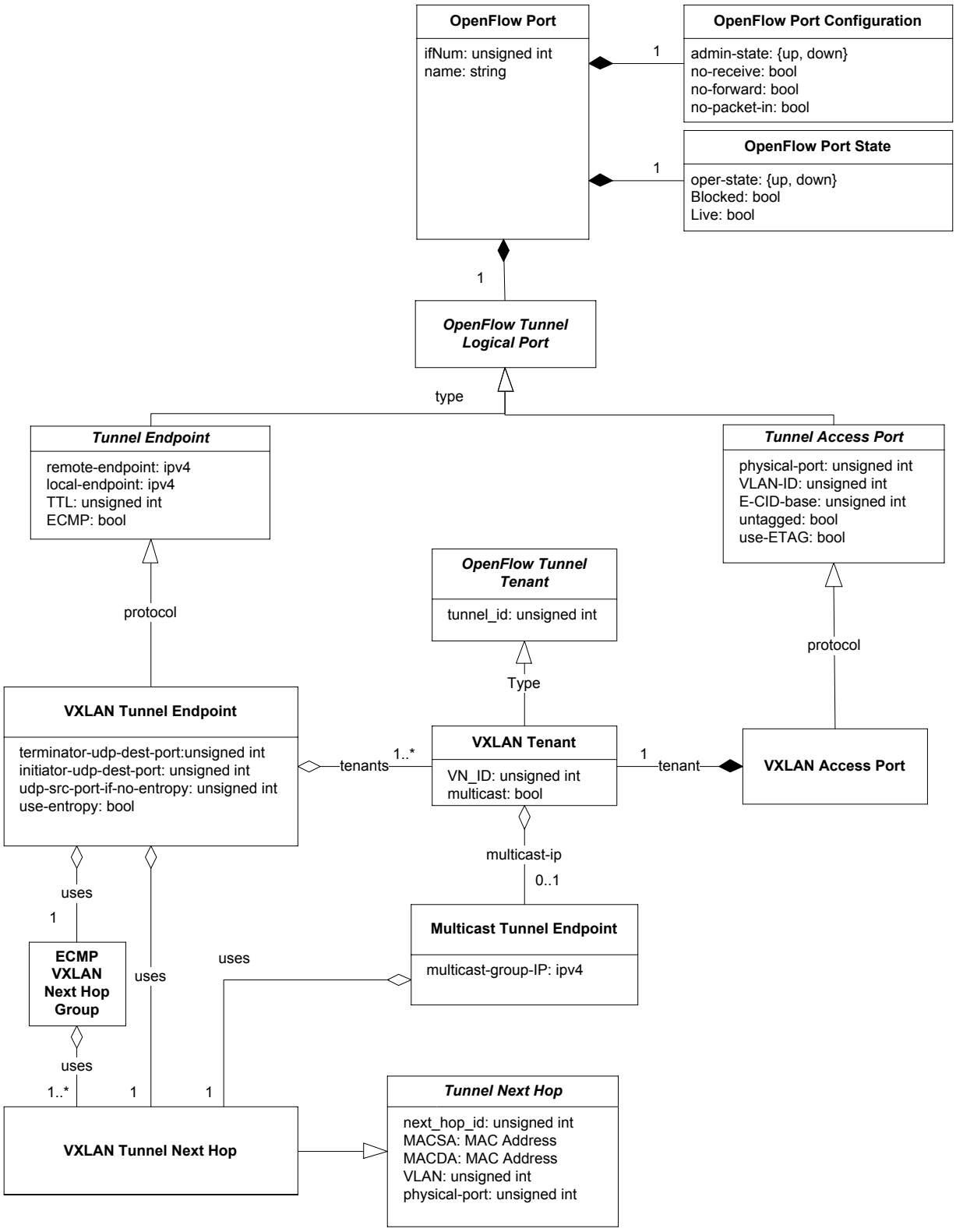

Figure 43 VXLAN Tunnel Configuration

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 221 thru Table 224.

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 221 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 222 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 223 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 224 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

5.1.3.3 OAM Protection Liveness Logical Ports

OAM Protection Liveness Logical Ports have no configuration. They are predefined in a specific range and default to administratively up. The operational state is always the same as the administrative state.

OAM Protection Liveness Logical Ports are used as a data object to control Fast Failover bucket liveness state.

5.2 Queues

OF-DPA supports eight queues per standard port.

5.2.1 Configuration

OF-DPA queue configuration parameters are listed in Table 225. 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 225 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.

5.2.2 Counters

OF-DPA queues counters are as shown in Table 226.

Table 226 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

5.3 OAM Message Processing

This section describes configuration for a local Network Protection App. Similar configuration could be used for Controller hosted Network Protection App. The Configuration model is intended to closely follow the parameters defined in OAM standards [20][21][35].

The classes described in this section can be realized as OF-Config resources. The companion YANG file describes an augment for OF-Config 1.2.

5.3.1 MPLS-TP Ethernet OAM Configuration

OAM Engine configuration for Ethernet OAM follows IEEE 802.1ag [35]. The configuration information model for Fault Management is diagrammed in Figure 44 and for Performance Monitoring in Figure 45.

5.3.1.1 Fault Management

Configuration is relative to the network device being configured. For each MEG that the device has at least one maintenance point for, it can be:

- A MEP, for which it needs Local_MEP and one or more Remote_MEP configured. If there are intermediate MIPs between the MEPs then one or more Remote_MIP objects could be configured for them.
- A MIP, for which it needs Local_MIP configuration

A client level MEG can specify that MIPs are automatically created on nodes where its server level MEG has a MEP configured.

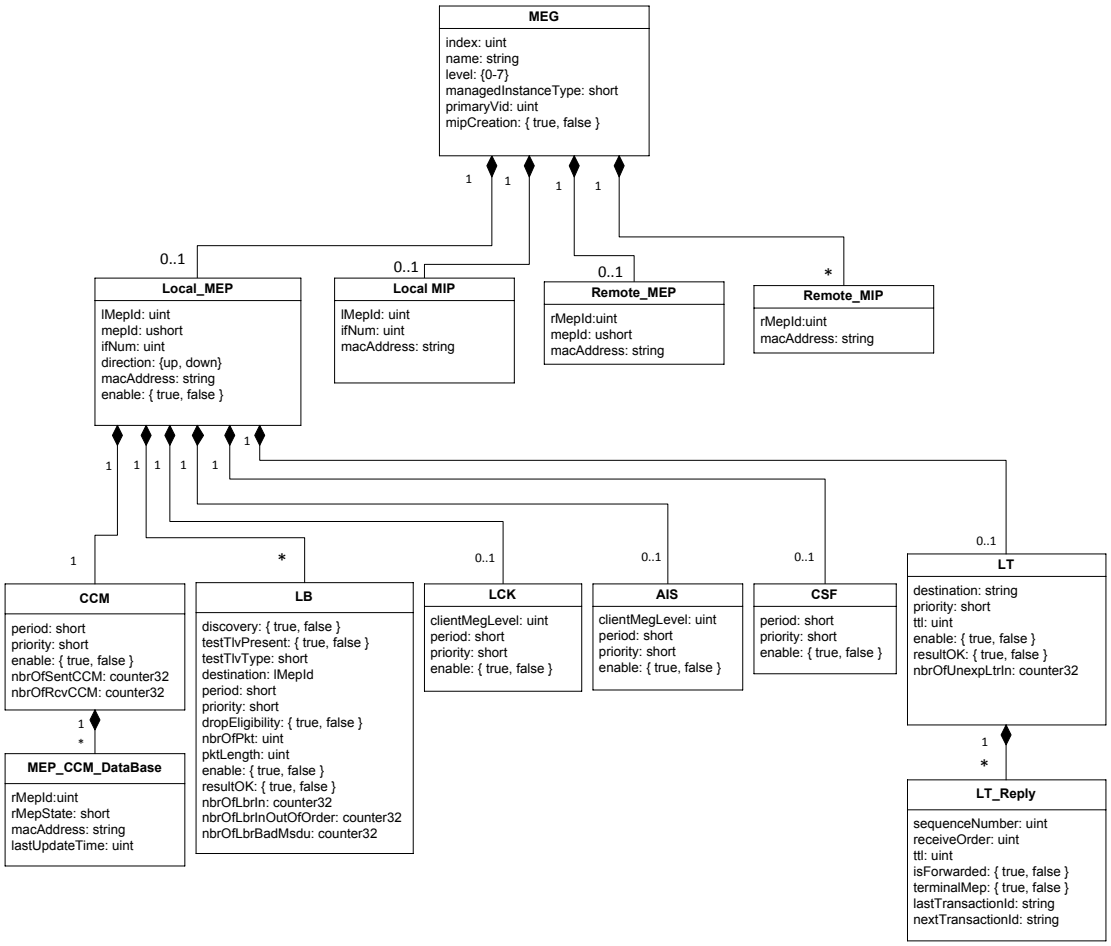

Figure 44 OAM Ethernet Fault Management Configuration

The following tables give details of the configurable objects described in Figure 44.

Table 227 Ethernet OAM MEG

Field	Bits	Access	Description
index	32	Read Create	Local Id for MEG. Required.
name	384	Read Write	48 byte MEG Id, including format and length bytes.
level	3	Read Write	0-7
managedInstanceType	4	Read Write	0: Ethernet 1: VPWS 2: Reserved
primaryVid	12	Read Write	Primary VLAN Id
mipCreation	1	Read Write	0: False (manual) 1: True (automatic)

Table 228 Ethernet OAM Local MIP

Field	Bits	Access	Description
lMepId	32	Read Write	LMEP_ID metadata that uniquely identifies this maintenance point when receiving or transmitting OAM frames
ifNum	32	Read Create	Interface number
macAddress	48	Read Only	Local (source) MAC address

Table 229 Ethernet OAM Local MEP

Field	Bits	Access	Description

Field	Bits	Access	Description
lMepId	32	Read Write	LMEP_ID metadata uniquely that identifies this maintenance point when receiving or transmitting OAM frames
mepId	16	Read Create	0-8191. Value used in PDU.
ifNum	32	Read Create	Interface number
direction	1	Read Write	2: up 1: down
macAddress	48	Read Write	Local (source) MAC address
enable	1	Read Write	0: False 1: True

Table 230 Ethernet OAM Remote MEP

Field	Bits	Access	Description
rMepId	32	Read Write	LMEP_ID metadata uniquely that identifies this maintenance point when receiving or transmitting OAM frames
mepId	16	Read Write	0-8191. Value used in PDU. Default value 1
macAddress	48	Read Write	Remote (destination) MAC address

Table 231 Ethernet OAM Remote MIP

Field	Bits	Access	Description
rMepId	32	Read Write	LMEP_ID metadata uniquely that identifies this maintenance point when receiving or transmitting OAM frames

Field	Bits	Access	Description
macAddress	48	Read Write	Remote (destination) MAC address

Table 232 Ethernet OAM CCM

Field	Bits	Access	Description
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
priority	3	Read Write	Traffic Class
enable	1	Read Write	0: Disabled 1: Enabled
nbrOfSentCCM	32	Read Only	Count of CCM frames sent
nbrOfRcvCCM	32	Read Only	Count of CCM frames received

Table 233 Ethernet OAM MEP CCM Database

Field	Bits	Access	Description
rMepId	16	Read Only	0-8191. From received PDU.

Field	Bits	Access	Description
rMepState	16	Read Only	1: Idle 2: Start (The timer has not expired since the state machine was reset, and no valid CCM has yet been received). 3: Failed (The timer has expired, both since the state machine was reset, and since a valid CCM was received). 4: Ok (The timer has not expired since a valid CCM was received).
macAddress	48	Read Only	Remote (destination) MAC address
lastUpdateTime	32	Read Only	

Table 234 Ethernet OAM Loopback

Field	Bits	Access	Description
discovery	1	Read Write	0: False 1: True
testTlvPresent	1	Read Write	0: False 1: True
testTlvType	16	Read Write	0: Null without CRC-32 1: Null with CRC-32 2: PRBS 2^31-1 without CRC-32 3: PRBS 2^32-1 with CRC=32
destination	32	Read Write	rMepId
period	16	Read Write	0-60 seconds, in milliseconds Default: 1000 (1 second)
priority	3	Read Write	Traffic Class
dropEligibility	1	Read Write	DEI for outer VLAN tag

Field	Bits	Access	Description
nbrOfPkt	32	Read Write	Number of packets to send
pktLength	32	Read Write	64-9600
enable	1	Read Write	0: disabled 1: enabled
resultOK	1	Read Write	0: False 1: True
nbrOfLBRIn	32	Read Only	Received replies counter
nbrOfLBRInOutOfOrder	32	Read Only	Received replies that were out of order counter
nbrOfLBRBadMsdu	32	Read Only	Received replies with payload errors counter.

Table 235 Ethernet OAM LCK

Field	Bits	Access	Description
clientMegLevel	3	Read Write	0-7
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
priority	16	Read Write	Traffic Class
enable	1	Read Write	0: Disabled 1: Enabled

Table 236 Ethernet OAM AIS

Field	Bits	Access	Description
clientMegLevel	3	Read Write	0-7
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
priority	16	Read Write	Traffic Class
enable	1	Read Write	0: Disabled 1: Enabled

Table 237 Ethernet OAM CSF

Field	Bits	Access	Description
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
priority	16	Read Write	Traffic Class
enable	1	Read Write	0: Disabled 1: Enabled

Table 238 Ethernet OAM Link Trace

Field	Bits	Access	Description
destination	48	Read Write	Target MAC Address (TA)
priority	3	Read Write	Traffic Class

Field	Bits	Access	Description
ttl	8	Read Write	Time to live
enable	1	Read Write	0: Disabled 1: Enabled
resultOK	1	Read Only	0: False 1: True
nbrOfUnexpectedLtrIn	32	Read Only	

Table 239 Ethernet OAM Link Trace Reply

Field	Bits	Access	Description
sequenceNumber	32	Read Only	
receiveOrder	32	Read Only	
ttl	32	Read Only	
isForwarded	1	Read Only	0: False 1: True
terminalMep	1	Read Only	0: False 1: True
lastTransactionId	64	Read Only	
nextTransactionId	64	Read Only	
ltrRelayAction	32	Read Only	

Field	Bits	Access	Description
ltrIngressAction	32	Read Only	
ltrIngressAddress	48	Read Only	
ltrIngressPortIdSubtype	32	Read Only	
ltrIngressPortId	32	Read Only	
ltrEgressAction	32	Read Only	
ltrEgressAddress	48	Read Only	
ltrEgressPortIdSubtype	32	Read Only	
ltrEgressPortId	32	Read Only	
ltrOrgSpecificTlv	-	Read Only	String up to 1500 bytes
ltrSenderIdTlv	-	Read Only	String up to 1500 bytes

5.3.1.2 Performance Monitoring

Note that OF-DPA 2.0 only partially supports performance monitoring. It does support sending and receiving performance monitoring OAM PDUs and associated configuration. However the OpenFlow pipeline does not currently have instrumentation support for Loss Measurement counters and Delay Measurement timestamps.

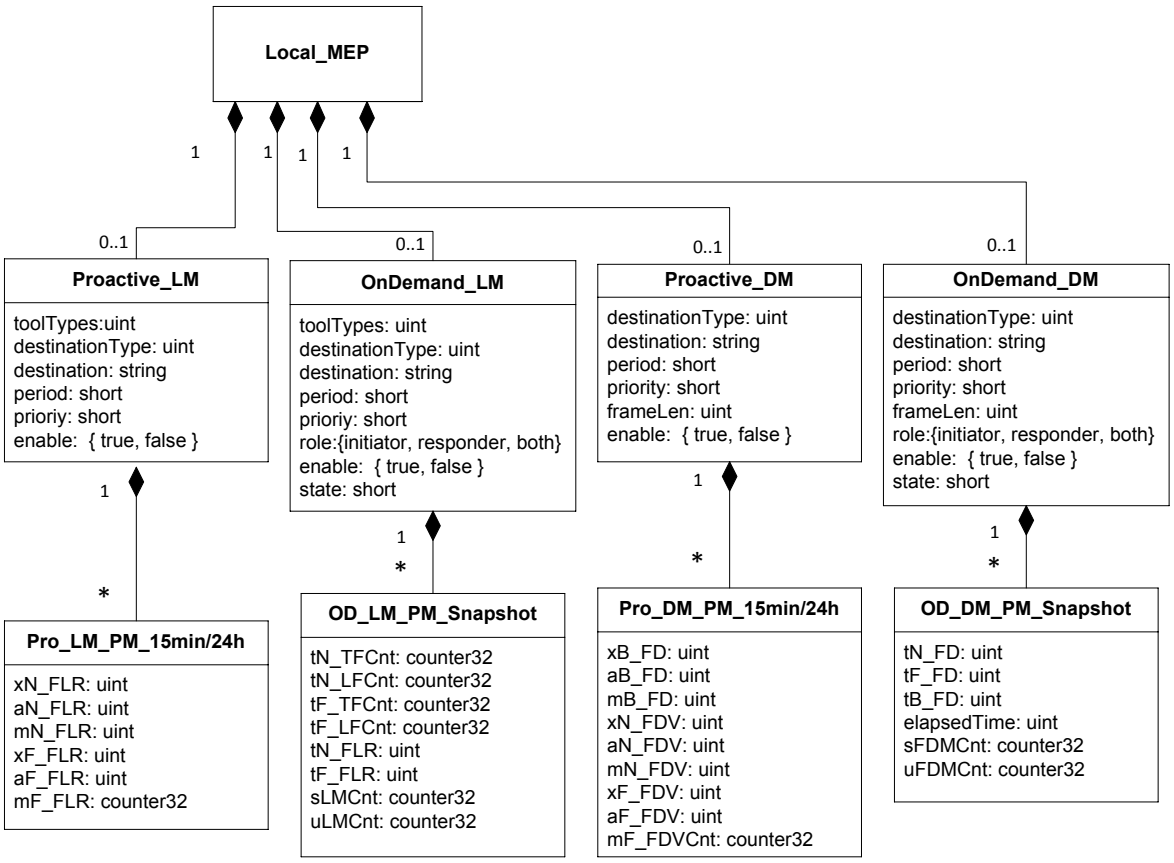

Figure 45 OAM Ethernet Performance Monitoring Configuration

The following tables give details of the configurable objects.

Table 240 Ethernet OAM Proactive LM

Field	Bits	Access	Description
toolTypes	3	Read Write	0: CCM 1: LM 2: Reserved
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id

Field	Bits	Access	Description
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
priority	3	Read Write	Traffic Class
enable	1	Read Write	0: Disabled 1: Enabled

Table 241 Ethernet OAM Proactive LM PM 15min/24h Counters

Field	Bits	Access	Description
xN_FLR	32	Read Only	Maximum Near-end Frame Loss Ratio, in milli-percent
aN_FLR	32	Read Only	Average Near-end Frame Loss Ratio, in milli-percent
mN_FLR	32	Read Only	Minimum Near-end Frame Loss Ratio, in milli-percent
xF_FLR	32	Read Only	Maximum Far-end Frame Loss Ratio, in milli-percent
aF_FLR	32	Read Only	Average Far-end Frame Loss Ratio, in milli-percent
mF_FLR	32	Read Only	Minimum Far-end Frame Loss Ratio, in milli-percent

Table 242 Ethernet OAM On-Demand LM

Field	Bits	Access	Description

Field	Bits	Access	Description
toolTypes	3	Read Write	0: CCM 1: LM 2: Reserved
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
priority	3	Read Write	PCP value
role	2	Read Write	0: initiator 1: responder 2: both
enable	1	Read Write	0: Disabled 1: Enabled
state	1	Read Write	0: stopped 1: running

Table 243 Ethernet OAM On Demand LM Snapshot

Field	Bits	Access	Description
tN_TFCnt	32	Read Only	Total Near-end transmitted frames since the single-ended LM initiated
tN_LFCnt	32	Read Only	Total Near-end lost frames since the single-ended LM initiated

Field	Bits	Access	Description
tF_TFCnt	32	Read Only	Total Far-end transmitted frames since the single-ended LM initiated
tF_LFCnt	32	Read Only	Total Far-end lost frames since the single-ended LM initiated
tN_FLR	32	Read Only	Total Near-end Frame Loss Ratio since the single-ended LM initiated, in milli-percent
tF_FLR	32	Read Only	Total Far-end Frame Loss Ratio since the single-ended LM initiated, in milli-percent
sLMCnt	32	Read Only	Number of successful loss measurements
uLMCnt	32	Read Only	Number of unsuccessful loss measurements

Table 244 Ethernet OAM Proactive DM

Field	Bits	Access	Description
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
priority	3	Read Write	PCP value

Field	Bits	Access	Description
frameLength	32	Read Write	64 - 9600
enable	1	Read Write	0: disabled 1: enabled

Table 245 Ethernet OAM Proactive DM PM 15min/24h Counters

Field	Bits	Access	Description
xB_FD	32	Read Only	Maximum Bi-directional Frame Delay, in microseconds
aB_FD	32	Read Only	Average Bi-directional Frame Delay, in microseconds
mB_FD	32	Read Only	Minimum Bi-directional Frame Delay, in microseconds
xN_FDV	32	Read Only	Maximum Near-end Frame Delay Variation, in microseconds
aN_FDV	32	Read Only	Average Near-end Frame Delay Variation, in microseconds
mN_FDV	32	Read Only	Minimum Near-end Frame Delay Variation, in microseconds
xF_FDV	32	Read Only	Maximum Far-end Frame Delay Variation, in microseconds
aF_FDV	32	Read Only	Average Far-end Frame Delay Variation, in microseconds
mF_FDV	32	Read Only	Minimum Far-end Frame Delay Variation, in microseconds

Table 246 Ethernet OAM On-Demand DM

Field	Bits	Access	Description
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
priority	16	Read Write	TC value to put in MPLS label
role	2	Read Write	0: initiator 1: responder 2: both
enable	1	Read Write	0: disabled 1: enabled
state	1	Read Write	0: stopped 1: running
frameLength	32	Read Write	64 - 9600

Table 247 Ethernet OAM On Demand DM Snapshot

Field	Bits	Access	Description
tN_FD	32	Read Only	i-th on-demand frame delay measurement result in near-end direction, in microseconds
tF_FD	32	Read Only	i-th on-demand frame delay measurement result in far-end direction, in microseconds

Field	Bits	Access	Description
tB_FD	32	Read Only	i-th on-demand frame delay measurement bi-directional result, in microseconds
elapsed time	32	Read Only	Elapsed time since beginning of delay measurement, in 10 ms.
sFDMCnt	32	Read Only	Number of successful delay measurements until now
uFDMCnt	32	Read Only	Number of unsuccessful delay measurements until now

5.3.2 MPLS-TP G.8113.1 OAM Configuration

OAM Engine configuration for G.8113.1 OAM follows [20]. The configuration information model is shown in Figure 46 and Figure 47.

5.3.2.1 Fault Management

OAM configuration is relative to the network device being configured. For each MEG that the device has at least one maintenance point for, that maintenance point can be:

- A MEP, for which it needs Local_MEP and at least one Remote_MEP configured. If there are intermediate MIPs between the MEPs then one or more Remote_MIP objects could be configured for them.
- A MIP, for which it needs Local_MIP configuration

A client layer MEG can specify that MIPs are automatically created on nodes where its server layer MEG has a MEP configured. This facilitates generating AIS frames, for example. Figure 26 illustrates examples of where MIPs would be configured over server layer MEPs.

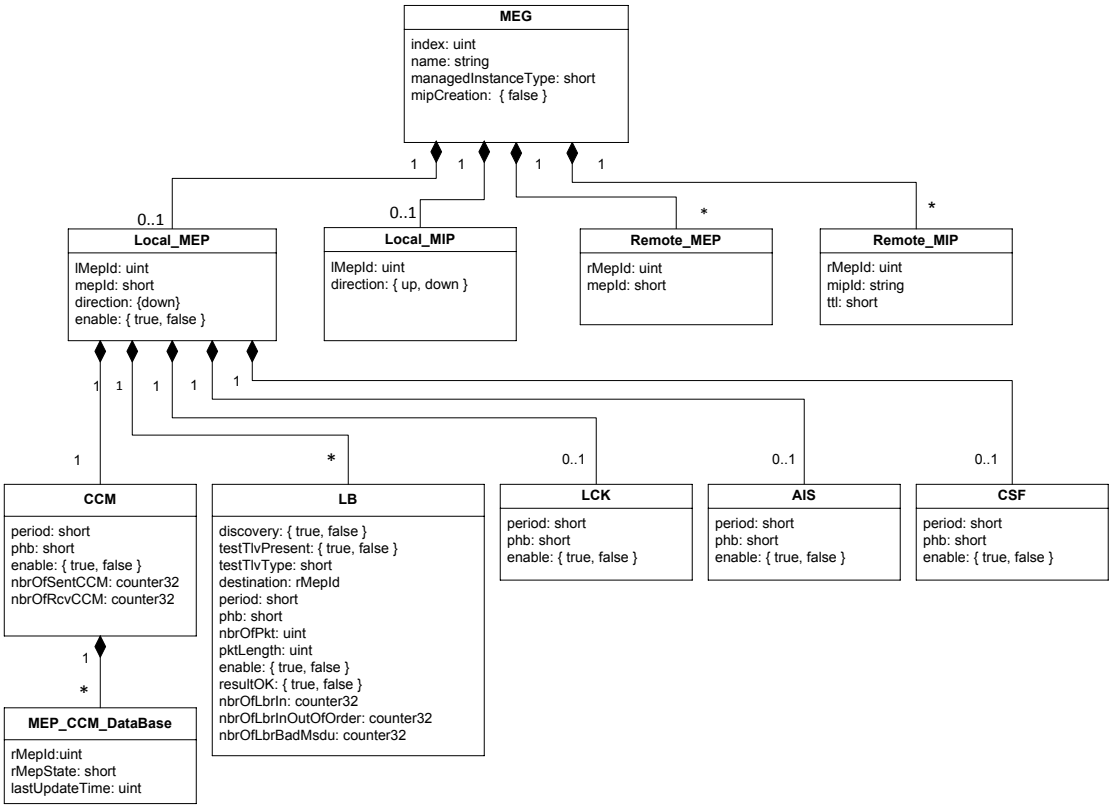

Figure 46 OAM G.8113.1 Fault Management Configuration

The following tables give details of the configurable objects in Figure 46.

Table 248 G.8113.1 OAM MEG

Field	Bits	Access	Description
index	32	Read Create	Local Id for MEG.
name	112	Read Write	14 byte MEG Id
managedInstanceType	4	Read Write	0: Section 1: LSP 2: PW
mipCreation	1	Read Write	0: False (manual) 1: True (automatic)

Table 249 G.8113.1 OAM Local MEP

Field	Bits	Access	Description
lMepId	32	Read Write	LMEP Id metadata used to identify the maintenance point in flow tables.
mepId	32	Read Create	Value used in PDU.
direction	1	Read Write	1: down
enable	1	Read Write	0: False 1: True

Table 250 G.8113.1 OAM Remote MEP

Field	Bits	Access	Description
rMepId	32	Read Write	Local identifier
mepId	16	Read Write	0-8191. Value used in PDU.

Table 251 G.8113.1 OAM Remote MIP

Field	Bits	Access	Description
rMepId	32	Read Write	Local identifier
mepId	16	Read Write	0-8191. Value used in PDU.
ttl	8	Read Write	Hop count

Table 252 G.8113.1 OAM Local MIP

Field	Bits	Access	Description
lMepId	32	Read Create	LMEP Id metadata used to identify the maintenance point in flow tables.
direction	1	Read Write	0: up 1: down

Table 253 G.8113.1 OAM CCM

Field	Bits	Access	Description
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
phb	3	Read Write	Traffic Class
enable	1	Read Write	0: False 1: True
nbrOfSentCCM	32	Read Only	Count of CCM messages sent
nbrOfRcvCCM	32	Read Only	Count of CCM messages received

Table 254 G.8113.1 OAM MEP CCM Database

Field	Bits	Access	Description
rMepId	16	Read Only	0-8191

Field	Bits	Access	Description
rMepState	16	Read Only	1: Idle 2: Start (The timer has not expired since the state machine was reset, and no valid CCM has yet been received). 3: Failed (The timer has expired, both since the state machine was reset, and since a valid CCM was received). 4: Ok (The timer has not expired since a valid CCM was received).
lastUpdateTime	32	Read Only	

Table 255 G.8113.1 OAM Loopback

Field	Bits	Access	Description
discovery	1	Read Write	0: False 1: True
testTlvPresent	1	Read Write	0: False 1: True
testTlvType	2	Read Write	0: Null signal (all zero) w/o CRC-32 1: Null signal (all zero) with CRC-32 2: PRBS 2^31 w/o CRC-32 3: PRBS 2^31 with CRC-32
destination	32	Read Write	rMepId
period	16	Read Write	0-60 seconds, in milliseconds Default: 1000 (1 second)
phb	3	Read Write	Traffic Class
nbrOfPkt	32	Read Write	Number of packets to send
pktLength	32	Read Write	64-9600

Field	Bits	Access	Description
enable	1	Read Write	0: disabled 1: enabled
resultOK	1	Read Write	0: false 1: true
nbrOfLbrIn	32	Read Only	Received replies counter
nbrOfLbrInOutOfOrder	32	Read Only	Received replies that were out of order counter
nbrOfLbrInWithBadMsdu	32	Read Only	Received replies with payload errors counter.

Table 256 G.8113.1 OAM LCK

Field	Bits	Access	Description
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
phb	3	Read Write	Traffic Class
enable	1	Read Write	0: disabled 1: enabled

Table 257 G.8113.1 OAM AIS

Field	Bits	Access	Description
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
phb	3	Read Write	Traffic Class

Field	Bits	Access	Description
enable	1	Read Write	0: disabled 1: enabled

Table 258 G.8113.1 OAM CSF

Field	Bits	Access	Description
period	3	Read Write	4: 1 sec 6: 1 min Others invalid.
phb	16	Read Write	Traffic Class
enable	1	Read Write	0: disabled 1: enabled

5.3.2.2 Performance Monitoring

Note that OF-DPA 2.0 only partially supports performance monitoring. It does support sending and receiving performance monitoring OAM PDUs and associated configuration. However the OpenFlow pipeline does not currently have instrumentation support for Loss Measurement counters and Delay Measurement timestamps.

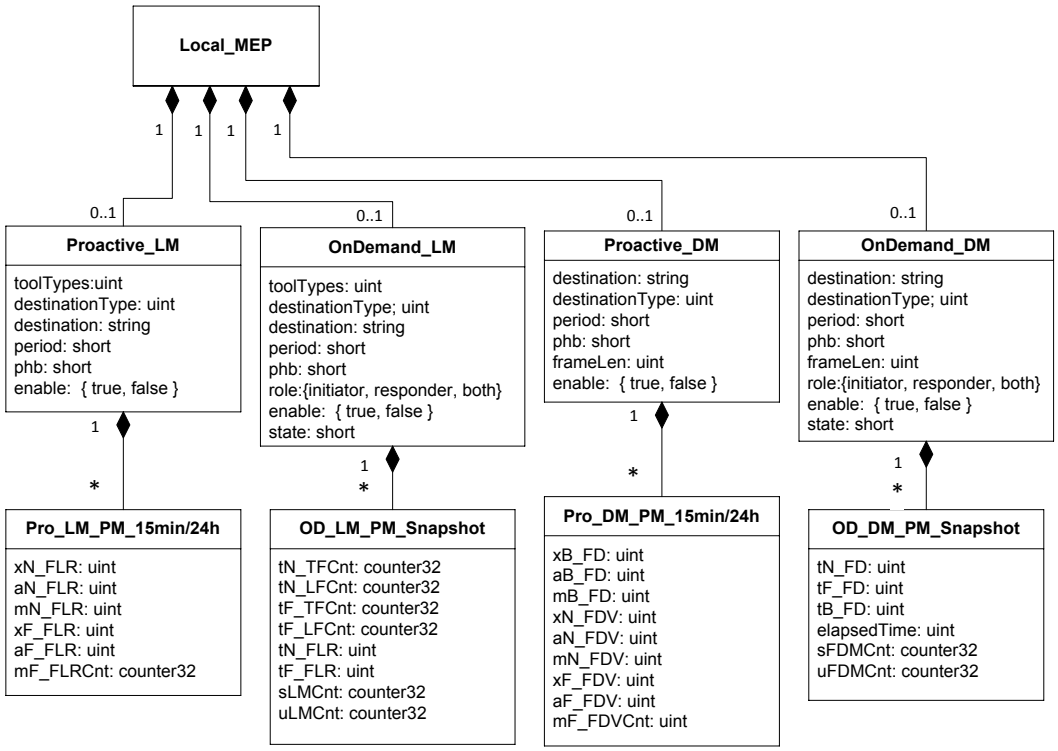

Figure 47 OAM G.8113.1 Performance Monitoring Configuration

The following tables give details of the configurable objects.

Table 259 G.8113.1 OAM Proactive LM

Field	Bits	Access	Description
toolTypes	3	Read Write	0: CCM 1: LM 2: Reserved
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id

Field	Bits	Access	Description
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
phb	16	Read Write	TC value to put in MPLS label
enable	1	Read Write	0: disabled 1: enabled

Table 260 G.8113.1 OAM Proactive LM PM 15min/24h Counters

Field	Bits	Access	Description
xN_FLR	32	Read Only	Maximum Near-end Frame Loss Ratio, in milli-percent
aN_FLR	32	Read Only	Average Near-end Frame Loss Ratio, in milli-percent
mN_FLR	32	Read Only	Minimum Near-end Frame Loss Ratio, in milli-percent
xF_FLR	32	Read Only	Maximum Far-end Frame Loss Ratio, in milli-percent
aF_FLR	32	Read Only	Average Far-end Frame Loss Ratio, in milli-percent
mF_FLR	32	Read Only	Minimum Far-end Frame Loss Ratio, in milli-percent

Table 261 G.8113.1 OAM On-Demand LM

Field	Bits	Access	Description

Field	Bits	Access	Description
toolTypes	3	Read Write	0: CCM 1: LM 2: Reserved
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
phb	16	Read Write	Traffic Class
role	2	Read Write	0: initiator 1: responder 2: both
enable	1	Read Write	0: disabled 1: enabled
state	1	Read Write	0: stopped 1: running

Table 262 G.8113.1 OAM On Demand LM Snapshot

Field	Bits	Access	Description
tN_TFCnt	32	Read Only	Total Near-end transmitted frames since the single-ended LM initiated (units of 1e-6)

Field	Bits	Access	Description
tN_LFCnt	32	Read Only	Total Near-end lost frames since the single-ended LM initiated (units of 1e-6)
tF_TFCnt	32	Read Only	Total Far-end transmitted frames since the single-ended LM (units of 1e-6)
tF_LFCnt	32	Read Only	Total Far-end lost frames since the single-ended LM initiated (units of 1e-6)
tN_FLR	32	Read Only	Total Near-end Frame Loss Ratio since the single-ended LM initiated (units of 1e-6)
tF_FLR	32	Read Only	Total Far-end Frame Loss Ratio since the single-ended LM initiated (units of 1e-6)
sLMCnt	32	Read Only	Number of successful loss measurements
uLMCnt	32	Read Only	Number of unsuccessful loss measurements

Table 263 G.8113.1 OAM Proactive DM

Field	Bits	Access	Description
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min

Field	Bits	Access	Description
phb	16	Read Write	Traffic Class
frameLength	32	Read Write	64 - 9600
enable	1	Read Write	0: disabled 1: enabled

Table 264 G.8113.1 OAM Proactive DM PM 15min/24h Counters

Field	Bits	Access	Description
xB_FD	32	Read Only	Maximum Bi-directional Frame Delay, in microseconds
aB_FD	32	Read Only	Average Bi-directional Frame Delay, in microseconds
mB_FD	32	Read Only	Minimum Bi-directional Frame Delay, in microseconds
xN_FDV	32	Read Only	Maximum Near-end Frame Delay Variation, in microseconds
aN_FDV	32	Read Only	Average Near-end Frame Delay Variation, in microseconds
mN_FDV	32	Read Only	Minimum Near-end Frame Delay Variation, in microseconds
xF_FDV	32	Read Only	Maximum Far-end Frame Delay Variation, in microseconds
aF_FDV	32	Read Only	Average Far-end Frame Delay Variation, in microseconds
mF_FDV	32	Read Only	Minimum Far-end Frame Delay Variation, in microseconds

Table 265 G.8113.1 OAM On-Demand DM

Field	Bits	Access	Description
destinationType	1	Read Write	0: ICC-based MEP Id 1: ICC-based MIP Id
destination	112	Read Write	2 bytes for ICC-based MEP Id, 14 bytes for ICC-based MIP Id
period	3	Read Write	0: Invalid 1: 3.33 ms 2: 10 ms 3: 100 ms 4: 1 sec 5: 10 sec 6: 1 min 7: 10 min
phb	16	Read Write	TC value to put in MPLS label
role	2	Read Write	0: initiator 1: responder 2: both
frameLength	32	Read Write	64 - 9600
enable	1	Read Write	0: disabled 1: enabled
state	1	Read Write	0: stopped 1: running

Table 266 G.8113.1 OAM On Demand DM Snapshot

Field	Bits	Access	Description
tN_FD	32	Read Only	i-th on-demand frame delay measurement result in near-end direction, in microseconds

Field	Bits	Access	Description
tF_FD	32	Read Only	i-th on-demand frame delay measurement result in far-end direction, in microseconds
tB_FD	32	Read Only	i-th on-demand frame delay measurement bi-directional result, in microseconds
elapsedTime	32	Read Only	Elapsed time since beginning of delay measurement, in 10 ms.
sFDMCnt	32	Read Only	Number of successful delay measurements until now
uFDMCnt	32	Read Only	Number of unsuccessful delay measurements until now

5.4 Protection

The Protection Process can control the switchover from worker to protection path by changing the administrative state of an OAM Protection Liveness logical Port. The Protection Process uses the information in the Fast Failover Configuration Table to determine which port to update for a particular path.

5.4.1 MPLS-TP Linear Protection

The Protection Process can control the switchover from worker to protection path by changing the administrative state of an OAM Protection Liveness logical Port. The Protection Process uses the information in the MPLS-TP linear protection configuration tables to determine which port to update for a particular path.

The MPLS-TP Linear Protection data model is diagrammed below.

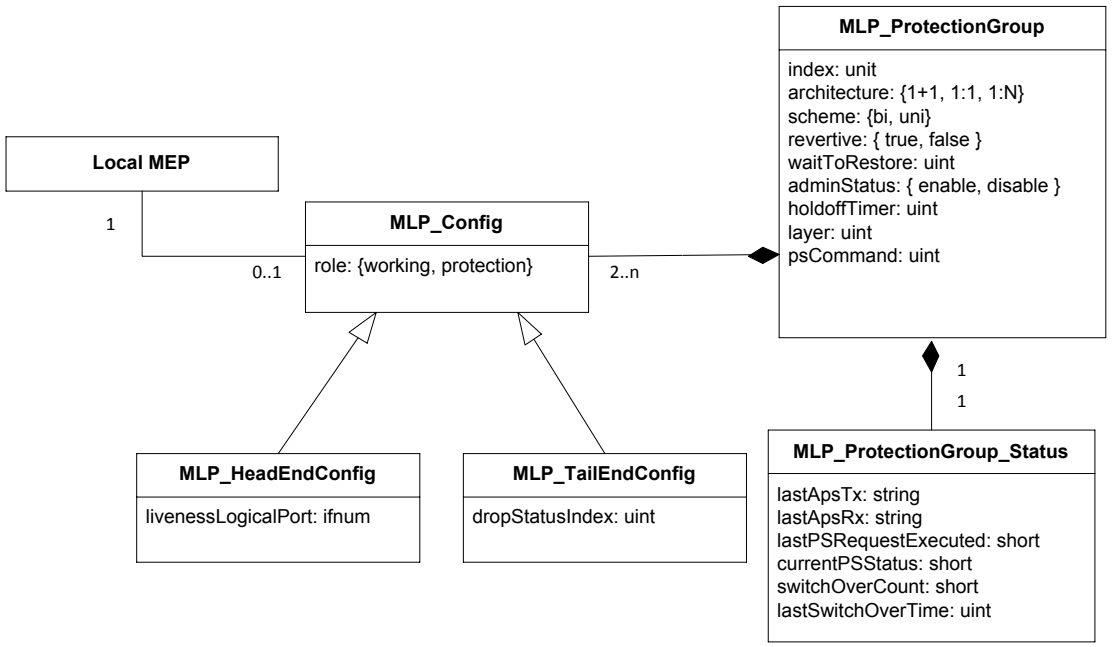

Figure 48 MPLS-TP Linear Protection Configuration

The following tables give details of the configurable objects described in Figure 48.

Table 267 MLP Protection Group

Field	Bits	Access	Description
index	32	Read Create	Local identifier
architecture	2	Read Write	0: Invalid 1: 1:1 2: 1+1 (for future use) 3: 1:N (for future use)
scheme	1	Read Write	0: unidirectional 1: bidirectional
layer	1	Read Write	0: tunnel 1: VC
revertive	1	Read Write	0: False 1: True

Field	Bits	Access	Description
waitToRestore	16	Read Write	In minutes
holdOffTimer	16	Read Write	In seconds
psCommand	16	Read Write	Last command 0: Invalid 1: noCmd 2: clear 3: lockoutOfProtection 4: forcedSwitchWorkToProtect 5: manualSwitchWorkToProtect 6: manualSwitchProtectToWork
adminStatus	1	Read Write	0: disabled 1: enabled

Table 268 MLP Protection Group Status

Field	Bits	Access	Description
lastApsTx	-	Read Only	String value, last APS PDU received
lastApsRx	-	Read Only	String value, last APS PDU sent
lastPsRequestExecuted	16	Read Only	Permitted values: 0: Clear 1: Lockout of protection 2: Signal fail for protection 3: forced switch 4: Signal fail for work 5: Signal degrade 6: Manual switch 7: Wait-to-restore 8: Exercise 9: Reverse request 10: Do not revert 11: No request

Field	Bits	Access	Description
currentPsStatus		Read Only	Permitted values: 0: Invalid 1: Lockout of protection 2: Signal fail for protection 3: forced switch 4: Signal fail for work 5: Signal degrade 6: Manual switch 7: Wait-to-restore 8: Exercise 9: Reverse request 10: Do not revert 11: No request
switchOverCount	16	Read Only	Number of switchovers since system start
lastSwitchOverTime	32	Read Only	Time tick since last switch protection was executed

Table 269 MLP Config

Field	Bits	Access	Description
role	1	Read Write	0: protection 1: working

The MPLS-TP Head End Protection objects configure head end failover objects. For 1:1 protection one primary and one backup would be configured. For 1:N multiple backup objects can be configured. The path is identified by the LMEP related to the object.

Table 270 MLP Head End Config

Field	Bits	Access	Description
livenessLogicalPort	32	Read Write	ifNum of the OAM Protection Liveness logical port being watched by the Fast Failover protection bucket for this path.

Table 271 MLP Tail End Config

Field	Bits	Access	Description
dropStatusIndex	32	Read Write	Index into the Drop Status table for the Protection Group entry that controls whether to drop or allow this path. Entry type is 1, head end protection.

6 VENDOR EXTENSION FEATURES

OF-DPA provides vendor extensions for source MAC learning, L3 forwarding IN_PORT control, MPLS and OAM actions and pipeline match fields, and new ancillary object types.

In many cases the vendor extension features only affect the OpenFlow abstract switch and can be accommodated by the existing OpenFlow 1.3.4 protocol. In others, an OpenFlow 1.3.4 agent and compatible controller can be extended using the OpenFlow Experimental facility to add new protocol elements as needed.

6.1 Source MAC Learning

OpenFlow 1.3.4 does not provide for flow tables that provide different views into the same database, e.g., using different lookup keys for different purposes.[31]

The Bridging Flow Table contains MAC forwarding entries and is looked up by MAC_DST and either VLAN or Tunnel_id. An exact match hit in the table sets the id of the group entry for forwarding the packet. However if there is no exact match, a flow entry that matches the VLAN or Tunnel_id but wildcards the ingress port can provide a flooding forwarding 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 or Tunnel_id. If there is a hit, the output interface is compared against the IN_PORT. If there is a mismatch, an entry for this MAC and VLAN or Tunnel_id is added to the table along with the interface. If there is a hit but the interface values are different, it means the end station has moved and the entry needs to be updated accordingly.[32]

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 or Tunnel_id miss, or if the source has moved, 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.

Note: The learning port depends on the packet flow. Bridging and routing learns physical ports, while overlay tunnels learn tunnel endpoint logical ports.

[31] This is addressed to some degree by the table synchronization extensions in OpenFlow 1.4.
[32] This is subject, of course, to security policy.

6.1.1 Controller Managed Learning

Learning is managed by the OpenFlow Controller as follows. For a MAC-SRC miss, OF-DPA will send a PACKET_IN message to the Controller with a "no match" reason code, a "virtual" table id that indicates the learning lookup, and the MAC-SRC, VLAN or Tunnel_Id, 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.

Note: If learning is enabled, the affected MAC entries must be programmed in the Bridging Flow Table even if only doing routing.

6.1.1.1 Configuration

The configuration options are listed in Table 272.

Table 272 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.

6.2 Additional Group Properties

OF-DPA adds the vendor extension property "ALLOW-IN_PORT" to OF-DPA L3 Interface group entries. This property applies to the group entry and to any referenced group entries. L3 Interface group entries automatically come with the property set, and it cannot be overridden. This obviates the need for special protocol support in OpenFlow 1.3.4.

6.3 MTU Check

OF-DPA adds an MTU check with the TTL check described in the OpenFlow 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 set the appropriate ICMP destination unreachable reason code.

This approach does not require special protocol support in OpenFlow 1.3.4.

6.4 Table Numbering

OF-DPA table number assignments are shown in Table 273.

Table 273 Flow Table Number Assignments

Flow Table Name	Number
Ingress Port Flow Table	0
Port DSCP Trust Flow Table	5
Port PCP Trust Flow Table	6
Tunnel DSCP Trust Flow Table	7
Tunnel PCP Trust Flow Table	8
Injected OAM Flow Table	9
VLAN Flow Table	10
VLAN 1 Flow Table	11
Ingress Maintenance Point Flow Table	12
MPLS L2 Port Flow Table	13
MPLS L2 Port DSCP Trust Flow Table	15
MPLS L2 Port PCP Trust Flow Table	16
Termination MAC Flow Table	20
L3 Type Flow Table	21
MPLS 0 Flow Table	23

Flow Table Name	Number
MPLS 1 Flow Table	24
MPLS 2 Flow Table	25
MPLS-TP Maintenance Point Flow Table	26
MPLS L3 Type Flow Table	27
MPLS Label Trust Flow Table	28
MPLS Type Flow Table	29
Unicast Routing Flow Table	30
Multicast Routing Flow Table	40
Bridging Flow Table	50
L2 Policer	55
L2 Policer Actions	56
Policy ACL Flow Table	60
Color Based Actions Flow Table	65
Egress VLAN Flow Table	210
Egress VLAN 1 Flow Table	211
Egress Maintenance Point Flow Table	226
Egress DSCP PCP Remark Flow Table	230
Egress TPID Flow Table	235
Source MAC Learning "Virtual" Flow Table	254

6.4.1 Egress Tables

The first egress table is the Egress VLAN Flow Table. The first egress table number is thus fixed at 210.

The tables higher numbered than 210, e.g. 211, 226, and 230, listed in Table 273 can only be used as egress tables and cannot be used as ingress tables.[33]

All packet flows that forward to an output port are processed by the egress tables. The default on an egress table miss is to terminate processing and output the packet.

6.5 Experimenter Features

This section describes extensions to OpenFlow required by OF-DPA 2.0 that require protocol support in order to be used with OpenFlow 1.3.4. These are encoded using the Experimenter facilities described in the standard.

Note that some of the extensions described in this section have already been proposed for or included in later versions of OpenFlow and also in the OpenFlow extensions packs. Although OF-DPA is intended for the widely deployed OpenFlow 1.3.4, accommodations have been made where possible to ease upgrading to later versions when these become widely deployed.

The experimenter id used for these features in OF-DPA is the Broadcom OUI: 00-10-18. The following details the type codes used in the various experimenter headers.

```
enum ofdpa_experimenter_type {
    OFDPA_EXP_MPLS_VPN_LABEL_REMARK_MOD_MSG =          6,
    OFDPA_EXP_MPLS_VPN_LABEL_REMARK_MULTIPART =        7,
    OFDPA_EXP_MPLS_TUNNEL_LABEL_REMARK_MOD_MSG =       8,
    OFDPA_EXP_MPLS_TUNNEL_LABEL_REMARK_MULTIPART =     9,
};

enum ofdpa_msg_mod_command {
    OFDPA_MSG_MOD_ADD =                            0,
    OFDPA_MSG_MOD_MODIFY =                             1,
    OFDPA_MSG_MOD_DELETE =                             2,
};

enum ofdpa_color_values {
    OFDPA_COLOR_GREEN =                0,
    OFDPA_COLOR_YELLOW =               1,
    OFDPA_COLOR_RED =                  2,
    OFDPA_COLOR_ALL =                  0xFF,
}

enum ofdpa_traffic_class_values {
    OFDPA_TRAFFIC_CLASS_BE =           0,
    OFDPA_TRAFFIC_CLASS_AF1 =          1,
    OFDPA_TRAFFIC_CLASS_AF2 =          2,
```

[33] This numbering convention does not apply to the “virtual” MAC learning table.


```
    OFDPA_TRAFFIC_CLASS_AF3 =           3,
    OFDPA_TRAFFIC_CLASS_AF4 =           4,
    OFDPA_TRAFFIC_CLASS_EF =            5,
    OFDPA_TRAFFIC_CLASS_CS6 =           6,
    OFDPA_TRAFFIC_CLASS_CS7 =           7,
    OFDPA_TRAFFIC_CLASS_ALL =           0xFF,
}

enum ofdpa_all_any_values {
    OFDPA_LMEP_ALL =                    0xFFFFFFFF,
    OFDPA_INDEX_ALL =                   0xFFFFFFFF,
    OFDPA_MPLS_TC_ALL =                 0xFF,
    OFDPA_VLAN_PCP_ALL =                0xFF,
    OFDPA_VLAN_DEI_ALL =                0xFF,
};
```

6.5.1 Action Tables

This section describes the encodings for the new OF-DPA action tables.

6.5.1.1 MPLS Label Remark Action Tables

The MPLS VPN Label Remark Action tables can be invoked from MPLS VPN and Swap Label Group entries using optional Set-MPLS_TC-From-VPN-Table and Set-PCP-DFI-From-VPN-Table actions. The Set-MPLS_TC-From-VPN_Table action uses the packet Traffic Class and Color and a supplied QoS Index argument to perform a Set-Field(MPLS_TC) action on the outermost label. The Set-PCP-DFI-From-VPN_Table action uses the packet Traffic Class and Color along with a supplied QoS Index argument to perform Set-Field(PCP) and Set-Field(DEI) actions on the outermost VLAN tag.

Similarly, the MPLS Tunnel Label Remark Action tables can be invoked from MPLS Tunnel Label Group entries using optional Set-MPLS_TC-From-Tunnel-Table and Set-PCP-DFI-From-Tunnel-Table actions. These also use the packet Traffic Class and Color and a supplied QoS Index argument to perform a Set-Field(MPLS_TC) action or a Set-Field(PCP) and/or Set-Field(DEI) action(s) on the outermost VLAN tag.

While similar to match-action tables in some respects, these tables are invoked in the context of evaluating an Action Set to perform Set-Field actions. These tables are modified and read using new Experimenter messages.

Entries in both VPN and Tunnel Remark Tables have the fields shown in Table 274.

Table 274 MPLS Label Remark Table Entry Fields

Field	Bits	Description
Index	4	Mapping profile index
Traffic Class	4	Traffic Class
Color	2	Color

Field	Bits	Description
MPLS_TC	3	Three-bit EXP value to use with the MPLS label.
VLAN_PCP	3	Three-bit PRI value to use in the outermost VLAN tag.
VLAN_DEI	1	DEI value to use in the outermost VLAN tag.

An experimenter message type is used to support modifications to the MPLS VPN Label Remark Action Table. This is only used as a Controller/Switch message.

```
struct ofdpa_mpls_vpn_label_remark_action_mod_msg {
    struct ofp_experimenter_header header; /*16 bytes, OFDPA_EXPERIMENTER_ID */
                           /* exp_type = OFDPA_EXP_MPLS_VPN_LABEL_REMARK_MOD_MSG */

    /* Experimenter-defined arbitrary additional data. */
    uint32_t        command;            /* one of OFDPA_MSG_MOD_* */
    uint32_t        index;              /* index */
    uint8_t         traffic_class;      /* traffic class */
    uint8_t         color;              /* color */
    uint8_t         mpls_tc;            /* MPLS TC value to set */
    uint8_t         vlan_pcp;           /* outer vlan PCP to set */
    uint8_t         vlan_dei;           /* outer vlan DEI to set */
    uint8_t         pad[3];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_mpls_vpn_label_remark_action_mod_msg) == 32);
```

An experimenter multipart message type is used to read status of the MPLS VPN Label Remark Action Table. It is used for both request and reply messages.

```
struct ofdpa_mpls_vpn_label_remark_action_multipart_request {
    struct ofp_multipart_request  header; /*16 bytes, OFPMP_EXPERIMENTER */

    /* Body for ofp_multipart_request/reply of type OFPMP_EXPERIMENTER. */
    struct ofp_experimenter_multipart_header exp_hdr; /* 8 bytes */
                           /* exp_type = OFDPA_EXP_MPLS_VPN_LABEL_REMARK_MULTIPART */

    /* Experimenter-defined arbitrary additional data. */
    uint32_t        index;              /* OFDPA_INDEX_ALL to get all */
    uint8_t         traffic_class;      /* OFDPA_TRAFFIC_CLASS_ALL to get all */
    uint8_t         mpls_tc             /* OFDPA_MPLS_TC_ALL to get all*/
    uint8_t         vlan_pcp            /* OFDPA_VLAN_PCP_ALL to get all */
    uint8_t         vlan_dei            /* OFDPA_VLAN_DEI_ALL to get all*/
};
OFP_ASSERT(sizeof(struct ofdpa_mple_vpn_label_remark_action_multipart_request)== 32);

struct ofdpa_mpls_vpn_label_remark_action_multipart_reply {
    struct ofp_multipart_reply   header; /*16 bytes, OFPMP_EXPERIMENTER */

    /* Body for ofp_multipart_request/reply of type OFPMP_EXPERIMENTER. */
    struct ofp_experimenter_multipart_header exp_hdr; /* 8 bytes */
                           /* exp_type = OFDPA_EXP_MPLS_VPN_LABEL_REMARK_MULTIPART */

    /* Experimenter-defined arbitrary additional data. */
```



```
    uint32_t        index;              /* entry index */
    uint8_t         traffic_class;      /* entry traffic class */
    uint8_t         mpls_tc             /* entry mpls_tc value */
    uint8_t         vlan_pcp            /* entry vlan_pcp value */
    uint8_t         vlan_dei            /* engry vlan_dei value */
};
OFP_ASSERT(sizeof(struct ofdpa_mple_vpn_label_remark_action_multipart_reply)==32);
```

An experimenter message type is used to support modifications to the MPLS Tunnel Label Remark Action Table. This is only used as a Controller/Switch message.

```
struct ofdpa_mpls_tunnel_label_remark_mod_msg {
    struct ofp_experimenter_header header; /*16 bytes, OFDPA_EXPERIMENTER_ID */
                          /* exp_type = OFDPA_EXP_MPLS_TUNNEL_LABEL_REMARK_MOD_MSG */

    /* Experimenter-defined arbitrary additional data. */
    uint32_t        command;            /* one of OFDPA_MSG_MOD_* */
    uint32_t        index;              /* index */
    uint8_t         traffic_class;      /* traffic class */
    uint8_t         color;              /* color */
    uint8_t         mpls_tc;            /* MPLS TC value to set */
    uint8_t         vlan_pcp;           /* outer vlan PCP to set */
    uint8_t         vlan_dei;           /* outer vlan DEI to set */
    uint8_t         pad[3];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_mpls_tunnel_label_remark_action_mod_msg)==32);
```

An experimenter multipart message type is used to read status of the MPLS Tunnel Label Remark Action Table. It is used for both request and reply messages.

```
struct ofdpa_mpls_tunnel_label_remark_action_multipart_request {
    struct ofp_multipart_request  header; /*16 bytes, OFPMP_EXPERIMENTER */

    /* Body for ofp_multipart_request/reply of type OFPMP_EXPERIMENTER. */
    struct ofp_experimenter_multipart_header exp_hdr;/* 8 bytes */
                    /* exp_type = OFDPA_EXP_MPLS_TUNNEL_LABEL_REMARK_MULTIPART */

    /* Experimenter-defined arbitrary additional data. */
    uint32_t        index;              /* OFDPA_INDEX_ALL to get all */
    uint8_t         traffic_class;      /* OFDPA_TRAFFIC_CLASS_ALL to get all */
    uint8_t         mpls_tc             /* OFDPA_MPLS_TC_ALL to get all*/
    uint8_t         vlan_pcp            /* OFDPA_VLAN_PCP_ALL to get all */
    uint8_t         vlan_dei            /* OFDPA_VLAN_DEI_ALL to get all*/
};
OFP_ASSERT(sizeof(struct
ofdpa_mple_tunnel_label_remark_action_multipart_request)==32);


struct ofdpa_mpls_tunnel_label_remark_action_multipart_reply {
    struct ofp_multipart_reply   header; /*16 bytes, OFPMP_EXPERIMENTER */

    /* Body for ofp_multipart_request/reply of type OFPMP_EXPERIMENTER. */
    struct ofp_experimenter_multipart_header exp_hdr;/* 8 bytes */
             /* exp_type = OFDPA_EXP_MPLS_TUNNEL_LABEL_REMARK_MULTIPART */

    /* Experimenter-defined arbitrary additional data. */
    uint32_t        index;              /* entry index */
    uint8_t         traffic_class;      /* entry traffic class */
```



```
    uint8_t         mpls_tc             /* entry mpls_tc value */
    uint8_t         vlan_pcp            /* entry vlan_pcp value */
    uint8_t         vlan_dei            /* engry vlan_dei value */
};
OFP_ASSERT(sizeof(struct ofdpa_mple_tunnel_label_remark_action_multipart_reply)==32);
```

6.5.1.2 Color Set Meter Band

The experimenter definitions to support OF-DPA 2.0 color set meter bands is given below. This is the only meter band type supported and replaces the corresponding fields in Meter modification messages. Both color set meter bands must be the same mode.

```
/* OF-DPA Experimenter Meter Band types */
enum ofdpa_meter_band_exp_type {
      OFDPA_OFPMBT_COLOR_SET =      1,
 };

/* OF-DPA Experimenter Color Set Meter Band Modes */
enum ofdpa_color_set_band_mode {
      OFDPA_COLOR_SET_BAND_MODE_TRTCM =         1,
      OFDPA_COLOR_SET_BAND_MODE_SRTCM =         2,
      OFDPA_COLOR_SET_BAND_MODE_RFC4115 =       3
};

/* OFPMT_EXPERIMENTER band for Color Set */

struct ofp_meter_band_experimenter_color_set {
    uint16_t      type;             /* OFPMBT_EXPERIMENTER (0xFFFF) */
    uint16_t      len;              /* Length in bytes of this band. */
    uint32_t      rate;             /* Committed rate for this band. */
    uint32_t      burst_size;       /* Excess (burst) for this band. */
    uint32_t      experimenter;     /* Experimenter Id: (00-00-10-18) */
    uint16_t      exp_type;         /* OFPMBT_COLOR_SET (1) */
    uint8_t       mode;             /* OFDPA_COLOR_SET_BAND_MODE_*  */
    uint8_t       color-aware;      /* Color-blind (0) or color-aware (1) */
    uint8_t       color;            /* New color, one of Yellow(1), Red(2) */
    unit8_t       pad[3];           /* Align on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofp_meter_band_experimenter_color_set) == 24);
```

6.5.2 Actions

OF-DPA 2.0 requires a number of new actions which are described in Table 275.

Table 275 OF-DPA Experimenter Actions

Action	Argument	Description
Push_L2_Header	None	Push a new outermost Ethernet header on the packet. All fields are initialized to zero.

Action	Argument	Description
Pop_L2_Header	None	Pop the outermost Ethernet header from the packet. The header cannot have a VLAN tag.
Push_CW	None	Push MPLS-TP PW Control Word. The Control Word is initialized to zero.
Pop_CW_or_ACH	None	Pop MPLS-TP PW Control Word or Associated Channel Header. Used in MPLS bottom of stack label match actions.
Copy_MPLS_TC_In	None	Copy EXP field value in. Used in conjunction with popping an MPLS label.
Copy_MPLS_TC_Out	None	Copy EXP value out, used when pushing a label.
Set_MPLS_TC_From_VPN_Table	QoS Index [, Traffic Class, Color]	MPLS label EXP field remark action based on packet Traffic Class, Color, and the provided QoS index argument to look up the new value in the MPLS VPN Label Remark Action Table. Traffic Class and Color are the current values of the pipeline match fields and are supplied by the switch.
Set_MPLS_TC_From_Tunnel_Table	QoS Index [, Traffic Class, Color]	Same as above but for Tunnel Groups. Uses the MPLS Tunnel Remark Action Table. Traffic Class and Color are the current values of the pipeline match fields and are supplied by the switch.
Set_MPLS_PCPDEI_From_VPN_Table	QoS Index [, Traffic Class, Color]	Outer VLAN tag remark action for PCP and DEI fields based on packet Traffic Class, Color, and the provided QoS index argument to look up the new value in an MPLS VPN Label Remark Action Table. Traffic Class and Color are the current values of the pipeline match fields and are supplied by the switch.
Set_MPLS_PCPDEI_From_Tunnel_Table	QoS Index [, Traffic Class, Color]	Same as above but for Tunnel Groups. Uses the MPLS Tunnel Remark Action Table.
Dec_TTL_MTU	None	Decrement TTL and do MTU check.

Action	Argument	Description
Copy Field	Source and destination fields.	Follows EXT-320. Source and destination can be Header or Pipeline fields. Only apply copy field is supported. See EXT-320 for encoding and other details.

The action types are programmed using the following assignments:

```
enum ofdpa_action_exp_types {
    OFDPA_ACT_PUSH_L2_Header =                          1,
    OFDPA_ACT_POP_L2_Header =                           2,
    OFDPA_ACT_PUSH_CW =                                 3,
    OFDPA_ACT_POP_CW_OR_ACH =                           4,
    OFDPA_ACT_COPY_TC_IN =                              5,
    OFDPA_ACT_COPY_TC_OUT =                             6,
    OFDPA_ACT_SET_MPLS_TC_FROM_VPN_TABLE =              7,
    OFDPA_ACT_DEC_TTL_MTU =                            13,
    OFDPA_ACT_SET_MPLS_PCPDEI_FROM_VPN_TABLE =         16,
    OFDPA_ACT_SET_MPLS_TC_FROM_TUNNEL_TABLE =          17,
    OFDPA_ACT_SET_MPLS_PCPDEI_FROM_TUNNEL_TABLE =      18,
};
```

Experimenter encodings are required for new actions. These follow the models in the OpenFlow specification and also include the experimenter header (struct_ofp_action_exprimenter_header).

Actions that have no arguments just consist of an experimenter action header and a type code value from `ofdpa_action_exp_type`. These are: Push L2 header; Pop L2 header; Push CW; Pop CW or ACH; Copy TC In; and Copy TC out.

```
enum ofdpa_class_and_color_values {
    OFDPA_CTR_CLASS_ANY =           0xFFFF,
    OFDPA_CTR_COLOR_ANY =           0xFFFF
}

struct ofdpa_action_experimenter  {
    struct ofp_action_experimenter_header header; /* 8 bytes, OFDPA_EXPERIMENTER_ID */
    uint16_t        exp_type;           /* One of above action types*/
    uint8_t         pad[6];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_action_experimenter) == 16);
```

The Set MPLS TC From VPN Table sets the MPLS TC field in the MPLS shim header from the value in the MPLS VPN Label Remark Table. Similarly, the Set MPLS PCP DEI From VPN Table action sets VLAN priority and drop indication header fields from a value in the MPLS VPN Label Remark Table.

```
struct ofdpa_action_set_mpls_tc_from_vpn_table  {
    struct ofp_action_experimenter_header header; /* 8 bytes, OFDPA_EXPERIMENTER_ID */
    uint16_t        exp_type;           /* OFDPA_ACT_SET_MPLS_TC_FROM_VPN_TABLE */
    uint16_t        qos_index;          /* qos index argument */
    uint8_t         traffic_class;      /* traffic class, supplied by switch */
    uint8_t         color;              /* color, supplied by switch */
```



```
    uint8_t         pad[2];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_action_set_mpls_tc_from_vpn_table) == 16);

struct ofdpa_action_set_mpls_pcpdei_from_vpn_table  {
    struct ofp_action_experimenter_header header; /* 8 bytes, OFDPA_EXPERIMENTER_ID */
    uint16_t        exp_type;           /* OFDPA_ACT_SET_MPLS_PCP_DEI_FROM_VPN_TABLE
*/
    uint16_t        qos_index;          /* qos index argument */
    uint8_t         traffic_class;      /* traffic class, supplied by switch */
    uint8_t         color;              /* color, supplied by switch */
    uint8_t         pad[2];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_action_set_mpls_pcpdei_from_vpn_table) == 16);
```

Similar to the above actions, the Set MPLS TC From Tunnel Table and Set MPLS PCP DEI From Tunnel Table actions set header fields by accessing a value in an MPLS Tunnel Label Remark Table. The same arguments are used.

```
struct ofdpa_action_set_mpls_tc_from_tunnel_table  {
    struct ofp_action_experimenter_header header; /* 8 bytes, OFDPA_EXPERIMENTER_ID */
    uint16_t        exp_type;           /* OFDPA_ACT_SET_MPLS_TC_FROM_TUNNEL_TABLE */
    uint16_t        qos_index;          /* qos index argument */
    uint8_t         traffic_class;      /* traffic class, supplied by switch */
    uint8_t         color;              /* color, supplied by switch */
    uint8_t         pad[2];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_action_set_mpls_tc_from_tunnel_table) == 16);

struct ofdpa_action_set_mpls_pcp_dei_from_tunnel_table  {
    struct ofp_action_experimenter_header header; /* 8 bytes, OFDPA_EXPERIMENTER_ID */
    uint16_t        exp_type;     /* OFDPA_ACT_SET_MPLS_PCP_DEI_FROM_TUNNEL_TABLE */
    uint16_t        qos_index;          /* qos index argument */
    uint8_t         traffic_class;      /* traffic class, supplied by switch */
    uint8_t         color;              /* color, supplied by switch */
    uint8_t         pad[2];             /* align message on 64-bit boundary */
};
OFP_ASSERT(sizeof(struct ofdpa_action_set_mpls_pcp_dei_from_tunnel_table) == 16);
```

6.5.3 Match Fields

OF-DPA introduced header and pipeline metadata match fields that require Experimenter encoding are described in Table 276.

Table 276 OF-DPA Experimenter Match Fields

Field	Bits	Masked	Prerequisite	Description
VRF	16	No		Pipeline metadata. Virtual Router, used to select virtual routing table.

Field	Bits	Masked	Prerequisite	Description
Traffic Class	4	No		Pipeline metadata. QoS traffic class. Eight values supported, encoded as 16-bit field.
Color	2	No		Pipeline metadata. Drop precedence. Values are: 00: Green 01: Yellow 10: Red 11: reserved Encoded as 16 bit field.
VLAN DEI	1	No	VLAN tag	Drop eligibility indication from 802.1Q.
QoS Index	8	No	None	Pipeline metadata. Used when looking up Traffic Class and Color based on packet contents in QoS Trust Flow tables.
LMEP Id	32	No	None	Pipeline metadata. Used to identify a local MEP or MIP instance.
MPLS_TTL	8	No	ETH-TYPE=0x8847	
MPLS L2 Port	32	Yes		Pipeline metadata. Used to identify an MPLS-TP pseudo wire endpoint.
L3_IN_PORT	32	No		Pipeline metadata. Used to identify an L3 VPN endpoint.
OVID	16	No	VLAN tag	Pipeline metadata. Outer VLAN Id, which has been popped in the VLAN Flow Table, to enable double tag matching in the VLAN 1 Flow Table.

Field	Bits	Masked	Prerequisite	Description
MPLS_DATA_FIRST_NIBBLE	4	No	ETH-TYPE=0x8847 and MPLS_BOS=1	Determine if data (0000b) or control (0001b)
MPLS_ACH_CHANNEL	16	No	MPLS control frame - MPLS_DATA_FIRST_NIBBLE is one	
MPLS_NEXT_LABEL_IS_GAL	1	No	ETH-TYPE=0x8847	Pipeline metadata derived from the packet parser "peeking" at the next label.
OAM_Y1731_MDL	3	No	ETH-TYPE=0x8902	OAM PDU Maintenance Domain Level
OAM_Y1731_OPCODE	8	No	ETH-TYPE=0x8902	OAM PDU opcode
COLOR_ACTIONS_INDEX	32	No	None	Pipeline metadata. Used to identify an entry in the Color Based Actions Flow Table.
Protection_Index	8	No		Indicates whether label is for protection path (0). Other values represent working paths.
ETH_SUB_TYPE	8	No	ETH_TYPE=0x8809	Identify EFM frame
ACTSET_OUTPUT	32	No		Output port assigned by match table or group output action. Only used in egress tables. Encoding defined in EXT-233. Read only from Controller.

Field	Bits	Masked	Prerequisite	Description
MPLS_TYPE	16	No	varies	Bit mask. Indicates flow type: 0: None 1: VPWS 2: Reserved 4: OAM 8: L3 Unicast 16: L3 Multicast 32: L3 PHP all others reserved
PACKET_REGS(N)	64	No	None	Packet Registers (encoding defined in EXT-244). Hold a temporary copy of a header field for later use. PACKET_REGS(0) is used for MPLS_TC. PACKET_REGS(1) is used for VLAN_VID. Both can be used as match fields in rules.
ALLOW_VLAN_TRANSLATION	1	No	None	Used in egress pipeline to enable VLAN translation. Set in L2 Unfiltered Interface group entry type.

The match fields are programmed using the following assignments:

```
/* OF-DPA Experimenter Match Field types */
enum ofdpa_match_exp_fields {
    OFDPA_OXM_VRF =                         1,
    OFDPA_OXM_TRAFFIC_CLASS =               2,
    OFDPA_OXM_COLOR =                       3,
    OFDPA_OXM_DEI =                         4,
    OFDPA_OXM_QOS_INDEX =                   5,
    OFDPA_OXM_LMEP_ID =                     6,
    OFDPA_OXM_MPLS_TTL =                    7,
    OFDPA_OXM_MPLS_L2_Port =                8,
    OFDPA_OXM_L3_IN_PORT =                  9,
    OFDPA_OXM_OVID =                        10,
    OFDPA_OXM_MPLS_DATA_FIRST_NIBBLE =      11,
    OFDPA_OXM_MPLS_ACH_CHANNEL =            12,
    OFDPA_OXM_MPLS_NEXT_LABEL_IS_GAL =      13,
    OFDPA_OXM_OAM_Y1731_MDL =               14,
    OFDPA_OXM_OAM_Y1731_OPCODE =            15,
    OFDPA_OXM_COLOR_ACTIONS_INDEX =         16,
```



```
    OFDPA_OXM_PROTECTION_INDEX =          21,
    OFDPA_OXM_ETH_SUB_TYPE =              22,
    OFDPA_OXM_MPLS_TYPE =                 23,
    OFDPA_OXM_ALLOW_VLAN_TRANSLATION =    24
}
```

These use the OpenFlow Experimenter OXM TLV definition as follows:

```
struct ofdpa_oxm_match_field  {
    uint32_t      oxm_header;       /* oxm_class = OFPXMC_ EXPERIMENTER */
    uint32_t      experimenter;     /* Experimenter ID: 00-00-10-18 */
};
OFP_ASSERT(sizeof(struct ofdpa_oxm_match_field) == 8);
```

The `oxm_class` field in the header is set to `OFPXMC_EXPERIMENTER`. Following the examples in the ONF Extensions documents, this is followed by a 2 byte subtype field containing a value from `ofdpa_match_exp_fields`. This is followed by an argument as specified in Table 276. The payload length is given in the `oxm_length` field, so that the total length of the field is 4+`oxm_length`. Note that the minimum argument field size is a byte. An OXM TLV does not need to be padded or aligned.

6.5.4 Extension Fields

OF-DPA uses fields defined in the OpenFlow extensions documents [11].

EXT-244 defines the encoding for Packet Registers. The Class ID of this extension is:

```
OFPXMC_PACKET_REGS = 0x8001

/* Structure for OXM field output match. */
struct onf_oxm_packet_regs  {
    uint32_t      oxm_header; /* oxm_class = OFPXMC_PACKET_REGS,
                                 oxm_field = <N>. */
    uint64_t      value;      /* Packet Register value. */
};
OFP_ASSERT(sizeof(struct onf_oxm_packet_regs) == 12);
```

EXT-320 defines the encoding for the Copy Field action, reproduced here. Table features encodings are reproduced here and described in detail in the "Copy Field action Extension" document.

```
ONF_EXPERIMENTER_ID = 0x4F4E4600

/* Action types */
enum onf_act_exp_type {
    ONFACT_ET_COPY_FIELD = 3200, /* Copy-Field action. */
};

/* Action structure for ONFACT_ET_COPY_FIELD. */
struct onf_act_copy_field {
    uint16_t type;            /* OFPAT_EXPERIMENTER. */
    uint16_t len;             /* Length is padded to 64 bits. */
    uint32_t experimenter;    /* ONF_EXPERIMENTER_ID. */
    uint16_t exp_type;        /* ONFACT_ET_COPY_FIELD. */
    uint8_t pad2[2];
    uint16_t n_bits;          /* Number of bits to copy. */
    uint16_t src_offset;      /* Starting bit offset in source. */
```



```
    uint16_t dst_offset;      /* Starting bit offset in destination. */
    uint8_t pad[2];           /* Align to 32 bits. */

/* Followed by:
    * - Exactly 8, 12 or 16 bytes containing the oxm_ids, then
    * - Enough all-zero bytes (either 0 or 4) to make the action a whole
    * multiple of 8 bytes in length */

    uint32_t oxm_ids[0];      /* Source and destination OXM headers */
};
OFP_ASSERT(sizeof(struct ofp_action_copy_field) == 20);


/* Action types */
enum onf_tfp_exp_type {
    ONFTFP_ET_WRITE_COPYFIELD = 3200, /* Write Copy-Field property. */
    ONFTFP_ET_APPLY_COPYFIELD = 3201, /* Apply Copy-Field property. */
};
```

The table feature properties `ONFTFP_ET_WRITE_COPYFIELD` and `ONFTFP_ET_APPLY_COPYFIELD` use the following structure:

```
/* Table Feature Property structure for ONFTFP_ET_WRITE_COPYFIELD
            * and ONFTFP_ET_APPLY_COPYFIELD. */
struct onf_tfp_copy_field {
    uint16_t type;            /* One of OFPTFPT_EXPERIMENTER,
                              OFPTFPT_EXPERIMENTER_MISS. */
    uint16_t length;          /* Length in bytes of this property. */
    uint32_t experimenter;    /* ONF_EXPERIMENTER_ID. */
    uint32_t exp_type;        /* ONFTFP_ET_WRITE_COPYFIELD or
                              ONFTFP_ET_APPLY_COPYFIELD. */
/* Followed by:
    * - Exactly (length - 4) bytes containing the oxm_ids, then
    * - Exactly (length + 7)/8*8 - (length) (between 0 and 7)
    * bytes of all-zero bytes */

    uint32_t oxm_ids[0];      /* Array of OXM headers */
};
OFP_ASSERT(sizeof(struct onf_tfp_copy_field) == 12);
```


APPENDIX A :REFERENCES

[1] "Software Defined Networking Definition," Open Networking Foundation, https://www.opennetworking.org/sdn-resources/sdn-definition

[2] OpenFlow 1.0 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.0.0.pdf, December, 2009

[3] OpenFlow 1.3.0 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.0.pdf, June, 2012

[4] OpenFlow 1.3.1 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.1.pdf, September, 2012

[5] OpenFlow 1.3.2 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.2.pdf, April, 2013

[6] OpenFlow 1.3.3 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.3.pdf, September, 2013

[7] OpenFlow 1.3.4 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.3.4.pdf, March, 2014

[8] OpenFlow 1.3.5 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.3.5.pdf, April, 2015

[9] OpenFlow 1.4.0 Specification, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.4.1.pdf

[10] OpenFlow Extensions 1.3.x Package, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-extensions-1.3.x-package.zip, October, 2013

[11] OpenFlow Extensions 1.3.x Package 2, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-extensions-1.3.x-pack2.zip , January, 2015

[12] OpenFlow Table Type Patterns 1.0, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/OpenFlow%20Table%20Type%20Patterns%20v1.0.pdf , August, 2014

[13] OpenFlow Data Plane Abstraction (OF-DPA™): Abstract Switch Specification, Version 1.0, https://github.com/Broadcom-Switch/of-dpa/blob/master/doc/OFDPA_OASS-ETP101-R.PDF, February, 2014

[14] OpenFlow Management and Configuration Protocol 1.2 (OF-Config 1.2), https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow-config/of-config-1.2.pdf, 2014

[15] OF-Config 1.2 Yang Model, https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow-config/of-config1.zip, 2014

[16] Pffaf, B., and Davie, B, "The Open vSwitch Database Management Protocol," RFC7047, December, 2013

[17] Katz, D., and Ward, D., "Bidirectional Forwarding Detection," RFC 5880, June, 2010

[18] Aggarwal, R., Kompella, K., Nadeau, T., Swallow, G., "Bidirectional Forwarding Detection (BFD)for MPLS Label Switched Paths (LSPs)," RFC 5884, June, 2010

[19] Allan, D.,Swallow, G., and Drake, J., "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile," RFC 6428, November, 2011

[20] International Telecommunications Union, "Operations, administration and maintenance mechanism for MPLS-TP in packet transport networks", Recommendation ITU-T G.8113.1/Y.1372.1, November, 2012

[21] International Telecommunications Union, "Operations, administration and maintenance mechanism for MPLS-TP networks using the tools defined for MPLS", Recommendation ITU-T G.8113.2/Y.1372.2, November, 2012

[22] International Telecommunications Union, "OAM functions and mechanisms for Ethernet based networks," Recommendation ITU-T Y.1731, November, 2013

[23] "Ryu, a Software Defined Network Framework," http://osrg.github.io/ryu/

[24] "OpenDaylight Home Page," http://www.opendaylight.org/

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

[26] IEEE Standard for Local and metropolitan area networks – Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications – Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks, IEEE Std 802.3ah™ - 2002

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

[28] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and Weiss, W., "An Architecture for Differentiated Services," RFC2475, December, 1998

[29] Heinanen, J. and Guerin, R., "A Single Rate Three Color Marker," RFC 2697, September, 1999

[30] Heinanen, J. and Guerin, R., "A Two Rate Three Color Marker," RFC 2698, September, 1999

[31] Aboul-Magd, O., Rabie, S., "Differentiated Service Two-Rate, Three-Color Marker with Efficient Handling of in-Profile Traffic," RFC 4115, July, 2005

[32] "Ethernet Service Attributes Phase 2," MEF Technical Specification 10.2, October, 2009.

[33] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation", RFC 1701, October 1994

[34] Vigoureux, M., Ward, D., and Betts, M., "Requirements for Operations, Administration, and Maintenance (OAM) in MPLS Transport Networks," RFC 5860, May, 2010

[35] IEEE Standard for Local and metropolitan area networks – Virtual Bridged Local Area Networks – Virtual Bridged Local Area Networks – Amendment 5: Connectivity Fault Management, IEEE Std 802.1ag™-2007

[36] Open Networking Foundation Extensibility Working Group JIRA EXT-407, available to ONF member companies at https://rs.opennetworking.org/bugs/browse/EXT-407.

[37] International Telecommunications Union, "Characteristics of Ethernet transport network equipment functional blocks", Recommendation ITU-T G.8021/Y.1341, May, 2012

[38] Martini, L., Rosen, E., El-Aawar, N., and Heron, G., "Encapsulation Methods for Transport of Ethernet over MPLS Networks," RFC 4448, April, 2006

[39] Deering, S., "Host Extensions for IP Multicasting," RFC 1112, August 1989

[40] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks," RFC 2464, December 1998

