ref

https://github.com/anonymity12/p4-guide/blob/master/demo2/README.md

NOTE

DEMO2 要做什么

和demo1 相同，不过就是加上了 一个 prefix 的match count

```
83 control ingress(inout headers hdr,
84                 inout metadata meta,
85                 inout standard_metadata_t standard_metadata) {
86
87     direct_counter(CounterType.packets) ipv4_da_lpm_stats;
88
```

注意观察 demo 2 的 代码，我们认为，table 有一个 counters 的字段

```
 97     table ipv4_da_lpm {
 98         key = {
 99             hdr.ipv4.dstAddr: lpm;
100         }
101         actions = {
102             set_l2ptr;
103             drop_with_count;
104         }
105         default_action = drop_with_count;
106         counters = ipv4_da_lpm_stats;
107     }
```

我们可以通过cli 读出这个字段，命令是： `counter_read ipv4_da_lpm_stats 0`

DEMO2 的实验过程

按照readme 指导

1. 编译

命令是 :

```
p4c --target bmv2 --arch v1model demo2.p4_16.p4
```

```
myp4@myp4-VM:~/p4-guide/demo2$ p4c --target bmv2 --arch v1model demo2.p4_16.p4
myp4@myp4-VM:~/p4-guide/demo2$ l
demo2.p4_14.p4  demo2.p4_16.json  demo2.p4_16.p4  demo2.p4_16.p4i  demo2.p4_16.p4rt
```

2. 开启simple switch 并加载编译好的数据层面的json 文件（你可以理解json为配置，回调，合同，约定）

命令是:

```
sudo simple_switch --log-console -i 0@veth2 -i 1@veth4 -i 2@veth6 -i 3@veth8 -i 4@veth10 -i 5@veth12 -i 6@veth14 -i 7@veth16 demo2.p4_16.json
```

```
myp4@myp4-VM:~/p4-guide/demo2$ sudo simple_switch --log-console -i 0@veth2 -i 1@veth4 -i 2@veth6 -i 3@veth8 -i 4@ve
th10 -i 5@veth12 -i 6@veth14 -i 7@veth16 demo2.p4_16.json
[sudo] password for myp4:
Thrift port was not specified, will use 9090
Calling target program-options parser
[11:52:57.836] [bmv2] [D] [thread 23015] Set default default entry for table 'ingress.ipv4_da_lpm': ingress.drop_wi
th_count -
[11:52:57.837] [bmv2] [D] [thread 23015] Set default default entry for table 'ingress.mac_da': my_drop -
[11:52:57.837] [bmv2] [D] [thread 23015] Set default default entry for table 'egress.send_frame': my_drop -
Adding interface veth2 as port 0
[11:52:57.837] [bmv2] [D] [thread 23015] Adding interface veth2 as port 0
Adding interface veth4 as port 1
[11:52:57.843] [bmv2] [D] [thread 23015] Adding interface veth4 as port 1
Adding interface veth6 as port 2
[11:52:57.844] [bmv2] [D] [thread 23015] Adding interface veth6 as port 2
Adding interface veth8 as port 3
[11:52:57.846] [bmv2] [D] [thread 23015] Adding interface veth8 as port 3
Adding interface veth10 as port 4
[11:52:57.851] [bmv2] [D] [thread 23015] Adding interface veth10 as port 4
Adding interface veth12 as port 5
[11:52:57.853] [bmv2] [D] [thread 23015] Adding interface veth12 as port 5
Adding interface veth14 as port 6
[11:52:57.862] [bmv2] [D] [thread 23015] Adding interface veth14 as port 6
Adding interface veth16 as port 7
[11:52:57.863] [bmv2] [D] [thread 23015] Adding interface veth16 as port 7
Thrift server was started
```

3. 开启cli 进行 表项的下发

命令是 :

```
table_set_default ipv4_da_lpm my_drop
table_set_default mac_da my_drop
table_set_default send_frame my_drop

table_add ipv4_da_lpm set_l2ptr 10.1.0.1/32 => 58
table_add mac_da set_bd_dmac_intf 58 => 9 02:13:57:ab:cd:ef 2
table_add send_frame rewrite_mac 9 => 00:11:22:33:44:55

table_add ipv4_da_lpm set_l2ptr 10.1.0.200/32 => 81
table_add mac_da set_bd_dmac_intf 81 => 15 08:de:ad:be:ef:00 4
table_add send_frame rewrite_mac 15 => ca:fe:ba:be:d0:0d
```

```
[13:48:41.533] [bmv2] [T] [thread 23445] bm_table_add_entry
[13:48:41.533] [bmv2] [D] [thread 23445] Entry 1 added to table 'ingress.ipv4_da_lpm'
[13:48:41.533] [bmv2] [D] [thread 23445] Dumping entry 1
Match key:
* hdr.ipv4.dstAddr    : LPM       0a0100c8/32
Action entry: ingress.set_l2ptr - 51,

[13:48:46.045] [bmv2] [T] [thread 23445] bm_table_add_entry
[13:48:46.045] [bmv2] [D] [thread 23445] Entry 0 added to table 'ingress.mac_da'
[13:48:46.045] [bmv2] [D] [thread 23445] Dumping entry 0
Match key:
* meta.fwd_metadata.l2ptr: EXACT     00000051
Action entry: ingress.set_bd_dmac_intf - f,8deadbeef00,4,

[13:48:52.446] [bmv2] [T] [thread 23445] bm_table_add_entry
[13:48:52.446] [bmv2] [D] [thread 23445] Entry 1 added to table 'egress.send_frame'
[13:48:52.446] [bmv2] [D] [thread 23445] Dumping entry 1
Match key:
* meta.fwd_metadata.out_bd: EXACT     00000f
Action entry: egress.rewrite_mac - cafebabed00d,

[13:49:44.427] [bmv2] [T] [thread 23445] bm_table_add_entry
[13:49:44.427] [bmv2] [D] [thread 23445] Entry 1 added to table 'ingress.mac_da'
[13:49:44.427] [bmv2] [D] [thread 23445] Dumping entry 1
Match key:
* meta.fwd_metadata.l2ptr: EXACT     0000003a
Action entry: ingress.set_bd_dmac_intf - 9,21357abcdef,2,
```

```
action:              rewrite_mac
runtime data:        00:11:22:33:44:55
Entry has been added with handle 0
RuntimeCmd: table_add ipv4_da_lpm set_l2ptr 10.1.0.200/32 => 81
Adding entry to lpm match table ipv4_da_lpm
match key:           LPM-0a:01:00:c8/32
action:              set_l2ptr
runtime data:        00:00:00:51
Entry has been added with handle 1
RuntimeCmd: table_add mac_da set_bd_dmac_intf 81 => 15 08:de:ad:be:ef:00 4
Adding entry to exact match table mac_da
match key:           EXACT-00:00:00:51
action:              set_bd_dmac_intf
runtime data:        00:00:0f   08:de:ad:be:ef:00       00:04
Entry has been added with handle 0
RuntimeCmd: table_add send_frame rewrite_mac 15 => ca:fe:ba:be:d0:0d
Adding entry to exact match table send_frame
match key:           EXACT-00:00:0f
action:              rewrite_mac
runtime data:        ca:fe:ba:be:d0:0d
Entry has been added with handle 1
RuntimeCmd: table_add mac_da set_bd_dmac_intf 58 => 9 02:13:57:ab:cd:ef 2
Adding entry to exact match table mac_da
match key:           EXACT-00:00:00:3a
action:              set_bd_dmac_intf
runtime data:        00:00:09   02:13:57:ab:cd:ef       00:02
Entry has been added with handle 1
RuntimeCmd: table_set_default ipv4_da_lpm my_drop
Error: Table ipv4_da_lpm has no action my_drop
RuntimeCmd:
```

4. CLI 下读取counter

命令是： `counter_read ipv4_da_lpm_stats 0`

```
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=0, bytes=0)
```

5. 发送一个pkt 以期待计数器的变化。

PS: scapy 回话是在以太网接口用来发送pkt 的(我们此时使用的虚拟以太网接口: `veth`)，值得注意的是，任何在以太网接口上进行报文收发的行为，都需要root权限，所以请记得使用 `sudo scapy`

命令是：

```
sudo scapy

fwd_pkt1=Ether() / IP(dst='10.1.0.1') / TCP(sport=5793, dport=80)
drop_pkt1=Ether() / IP(dst='10.1.0.34') / TCP(sport=5793, dport=80)

# Send packet at layer2, specifying interface
<p class="mume-header " id="send-packet-at-layer2-specifying-interface"></p>

sendp(fwd_pkt1, iface="veth2")
sendp(drop_pkt1, iface="veth2")

fwd_pkt2=Ether() / IP(dst='10.1.0.1') / TCP(sport=5793, dport=80) / Raw('The quick brown fox jumped over the lazy dog.')
sendp(fwd_pkt2, iface="veth2")
```

scapy 在`veth2`上发送一个pkt

```
myp4@myp4-VM:~/p4-guide/demo1$ sudo scapy
[sudo] password for myp4:
INFO: Can't import python gnuplot wrapper . Won't be able to plot.
INFO: Can't import PyX. Won't be able to use psdump() or pdfdump().
WARNING: No route found for IPv6 destination :: (no default route?)
INFO: Can't import python Crypto lib. Won't be able to decrypt WEP.
INFO: Can't import python Crypto lib. Disabled certificate manipulation tools
Welcome to Scapy (2.2.0)
>>> fwd_pkt1=Ether() / IP(dst='10.1.0.1') /TCP(sport=5793, dport=80)
>>> drop_pkt1=Ether() / IP(dst='10.1.0.34') / TCP(sport=5793, dport=80)
>>> sendp(fwd_pkt1, iface="veth2")
.
Sent 1 packets.
>>>
```

simple switch 这边 立刻就处理完成

```
[14:08:01.858] [bmv2] [T] [thread 23021] [0.0] [cxt 0] demo2.p4_16.p4(111) Primitive hdr.ethernet.dstAddr = dmac
[14:08:01.858] [bmv2] [T] [thread 23021] [0.0] [cxt 0] demo2.p4_16.p4(112) Primitive standard_metadata.egress_spe
= intf
[14:08:01.858] [bmv2] [T] [thread 23021] [0.0] [cxt 0] demo2.p4_16.p4(113) Primitive hdr.ipv4.ttl = hdr.ipv4.ttl
1
[14:08:01.858] [bmv2] [D] [thread 23021] [0.0] [cxt 0] Pipeline 'ingress': end
[14:08:01.858] [bmv2] [D] [thread 23021] [0.0] [cxt 0] Egress port is 2
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Pipeline 'egress': start
[14:08:01.858] [bmv2] [T] [thread 23024] [0.0] [cxt 0] Applying table 'egress.send_frame'
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Looking up key:
* meta.fwd_metadata.out_bd: 000009

[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Table 'egress.send_frame': hit with handle 0
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Dumping entry 0
Match key:
* meta.fwd_metadata.out_bd: EXACT     000009
Action entry: egress.rewrite_mac - 1122334455,

[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Action entry is egress.rewrite_mac - 1122334455,
[14:08:01.858] [bmv2] [T] [thread 23024] [0.0] [cxt 0] Action egress.rewrite_mac
[14:08:01.858] [bmv2] [T] [thread 23024] [0.0] [cxt 0] demo2.p4_16.p4(137) Primitive hdr.ethernet.srcAddr = smac
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Pipeline 'egress': end
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Deparser 'deparser': start
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Updating checksum 'cksum'
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Updating checksum 'cksum_0'
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Deparsing header 'ethernet'
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Deparsing header 'ipv4'
[14:08:01.858] [bmv2] [D] [thread 23024] [0.0] [cxt 0] Deparser 'deparser': end
[14:08:01.858] [bmv2] [D] [thread 23026] [0.0] [cxt 0] Transmitting packet of size 54 out of port 2
```

ps: 从p4-guide/demo2/the_log_when_demo2 可以看到完整的scapy 发送报文 导致simple switch（虚拟交换机/交换机模拟器）产生的 报文处理log

6. 确认计数值确实发生了变化

可以与4 中的0 相比，这里， 你看到了pkt=1，bytes = 54.

```
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=1, bytes=54)
```

发送一个因为目的ip 不存在而drop 的pkt，以及一个带有raw data 的pkt

完整的变化结果：

```
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=0, bytes=0)
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=1, bytes=54)
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=1, bytes=54)
RuntimeCmd: counter_read ipv4_da_lpm_stats 0
this is the direct counter for table ingress.ipv4_da_lpm
ipv4_da_lpm_stats[0]=  BmCounterValue(packets=2, bytes=153)
```

PS: pkt 的size 和raw data 有关，raw data:

```
The quick brown fox jumped over the lazy dog.
```

是45 个字符，也就是45B，所以看到第三次的pkt size 是 99（54 + 45），统计看到的bytes 是从54 （`54 + 99 = 153`）跃迁到 153, 那么我们可以认为：

> 一个不带任何data 的pkt， 其size 是 54。带了多少字节的data， 就会让这个pkt 增加多少字节的总长度。