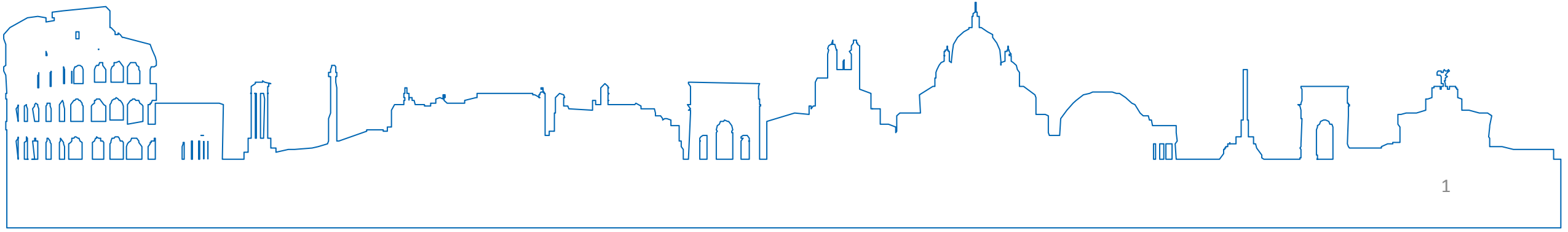
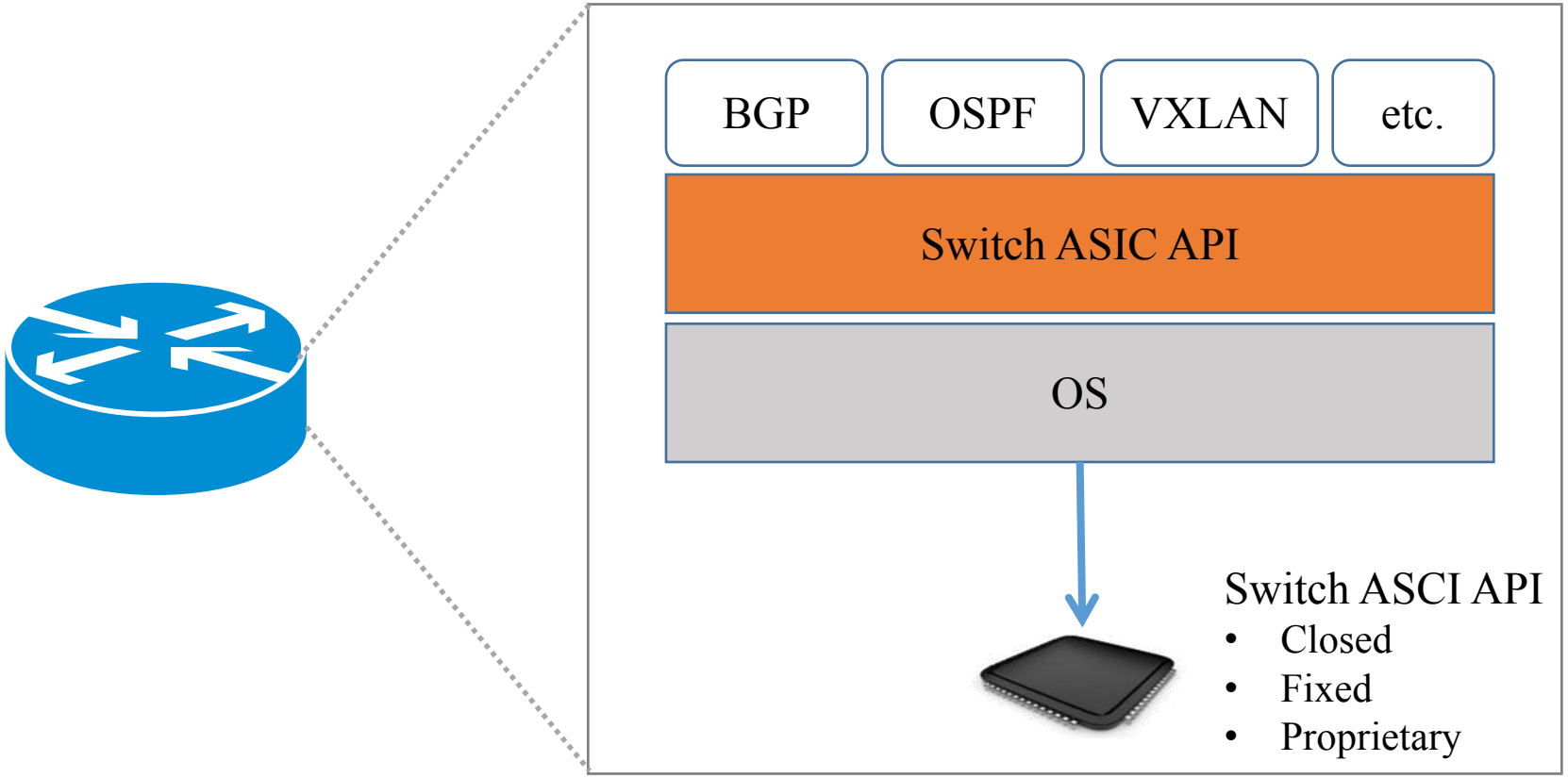
OPEN DAYLIGHT

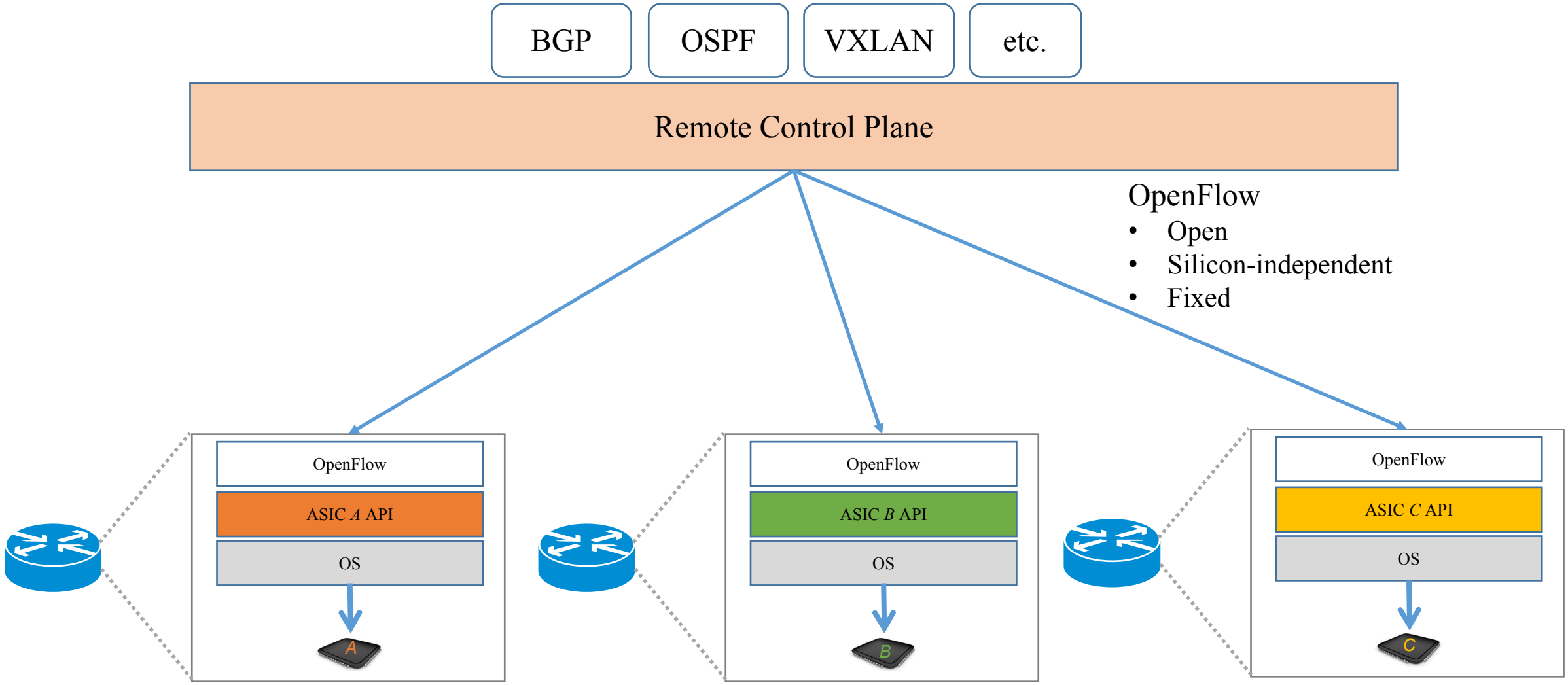
P4 Runtime integration with ODL



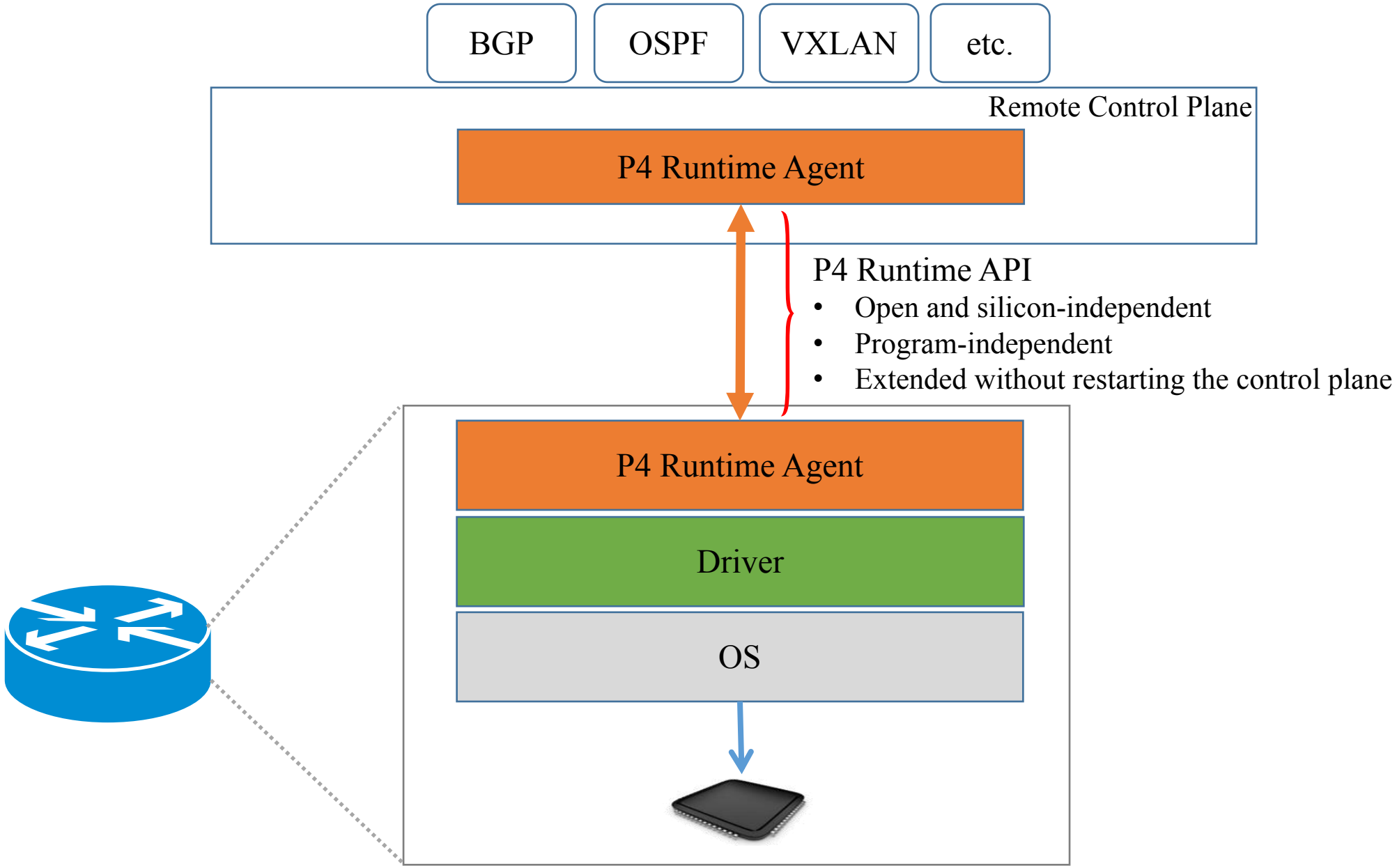
Traditional Network Equipment



In The SDN Era

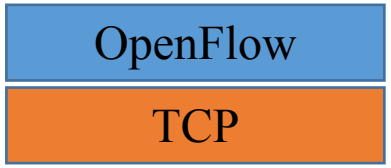


After P4 Generation

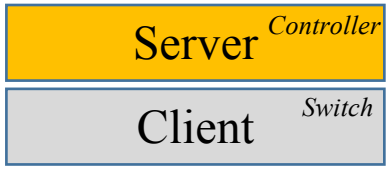


South Channel Compare and Contrast

- OpenFlow
  - OpenFlow protocol

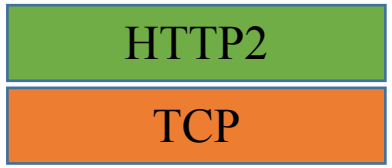


  - Client and Server



  - It only gives us a way to populate a set of well-known tables
  - Device discovery
  - 
  - etc.

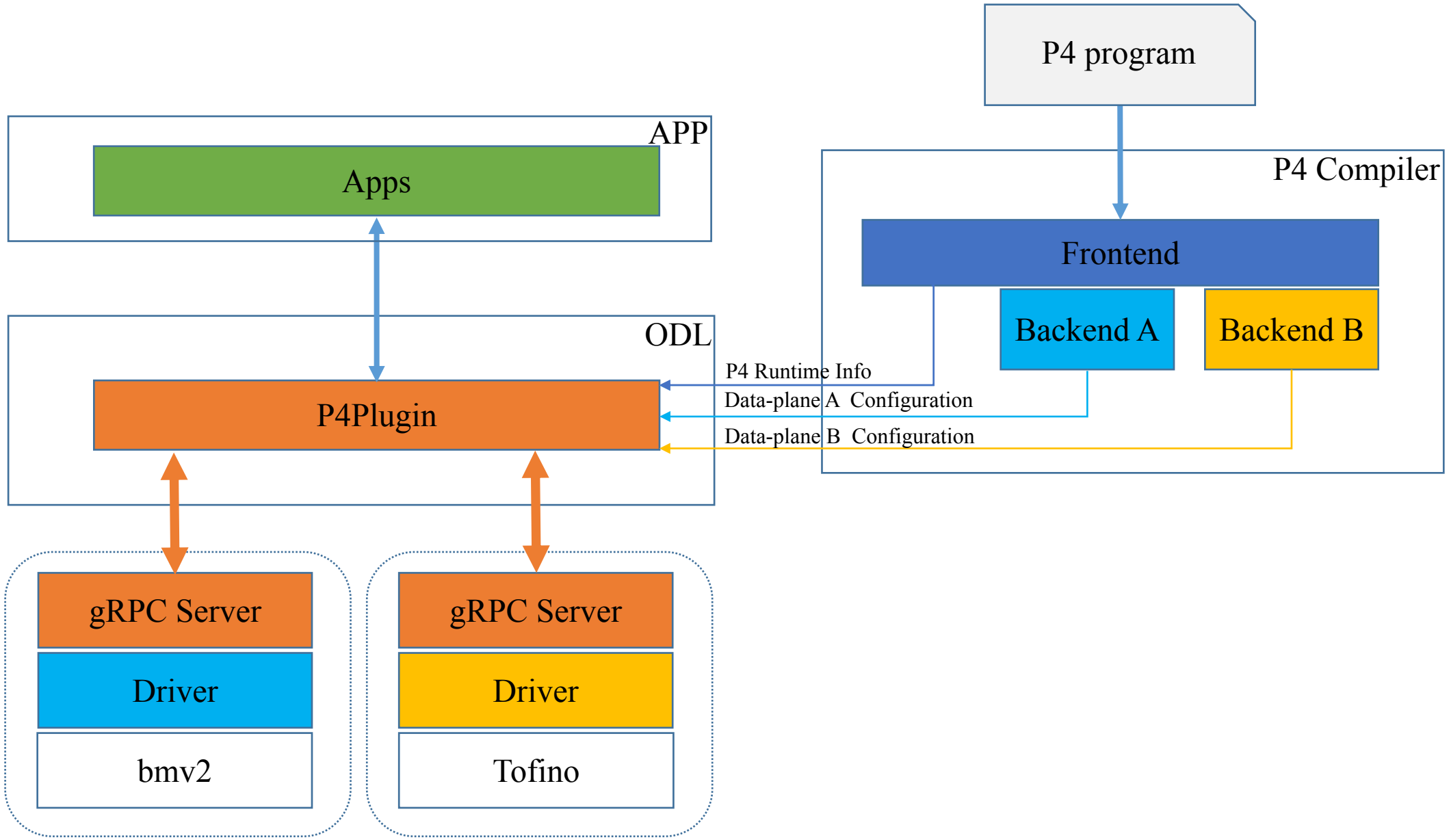
- P4 Runtime
  - gRPC/protoBuf



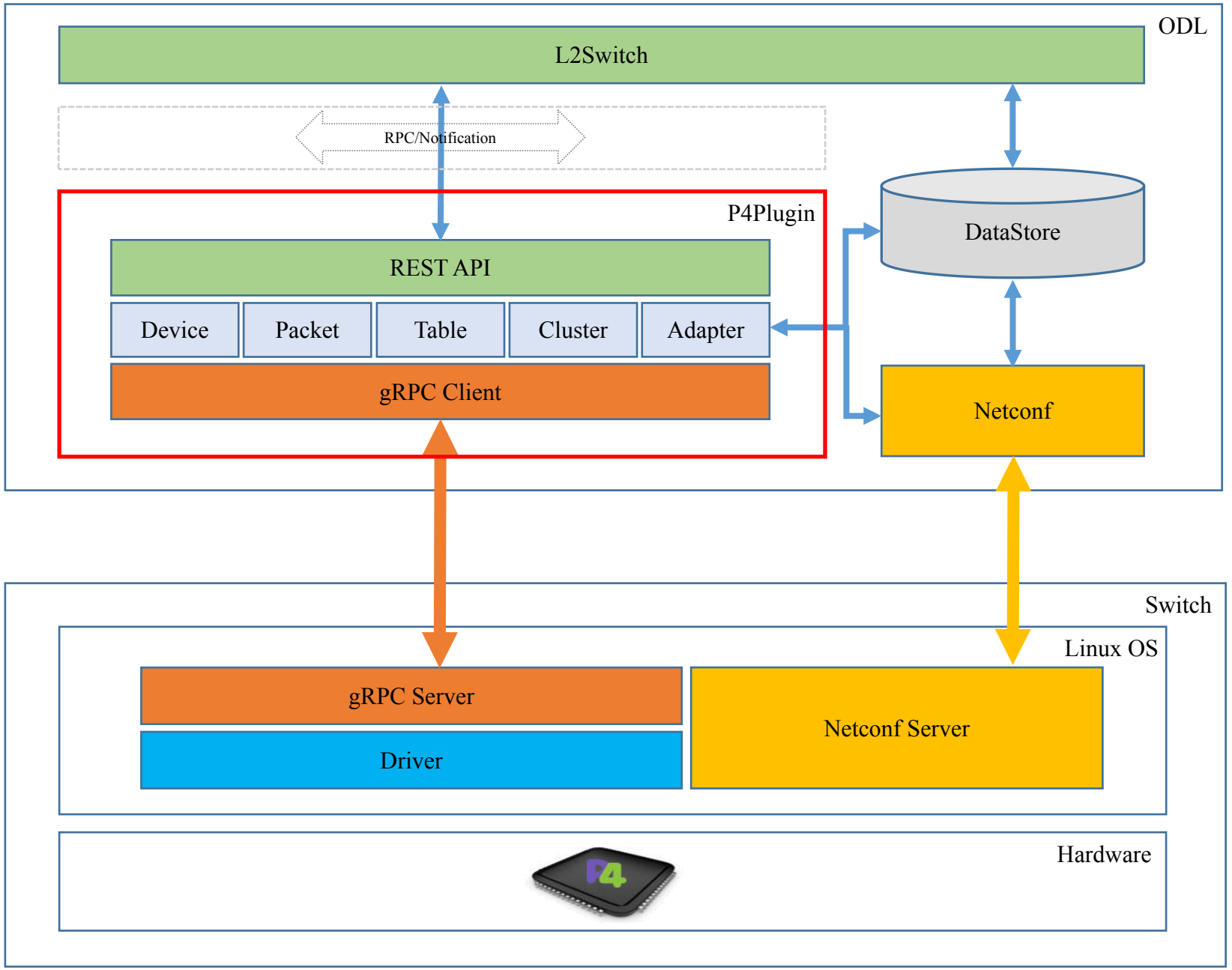
  - Client and Server

  - It is a protocol-independent API, it doesn't be tied to any specific network protocols
  - 
  - Set pipeline config
  - etc.

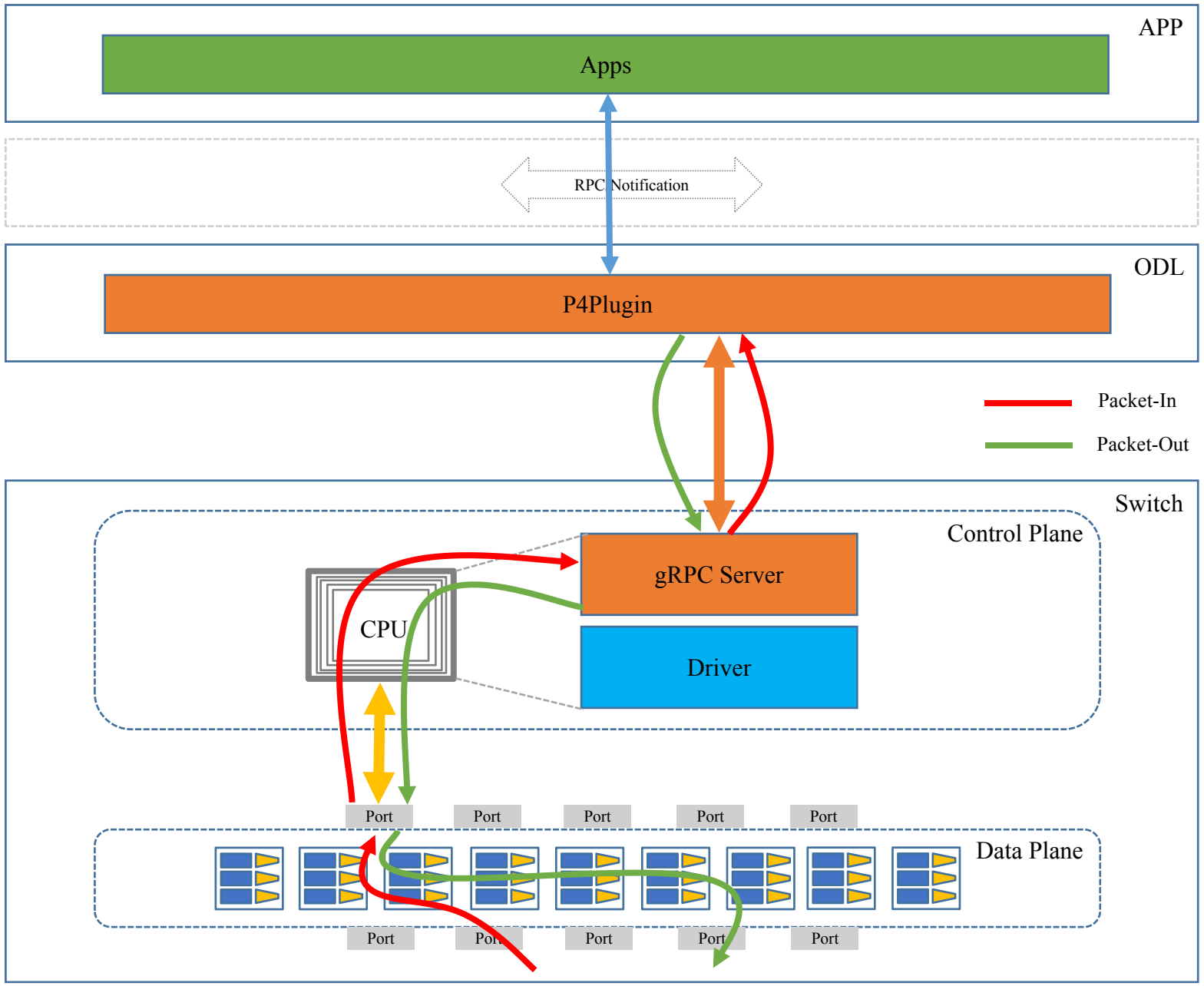
How to Use P4 Runtime



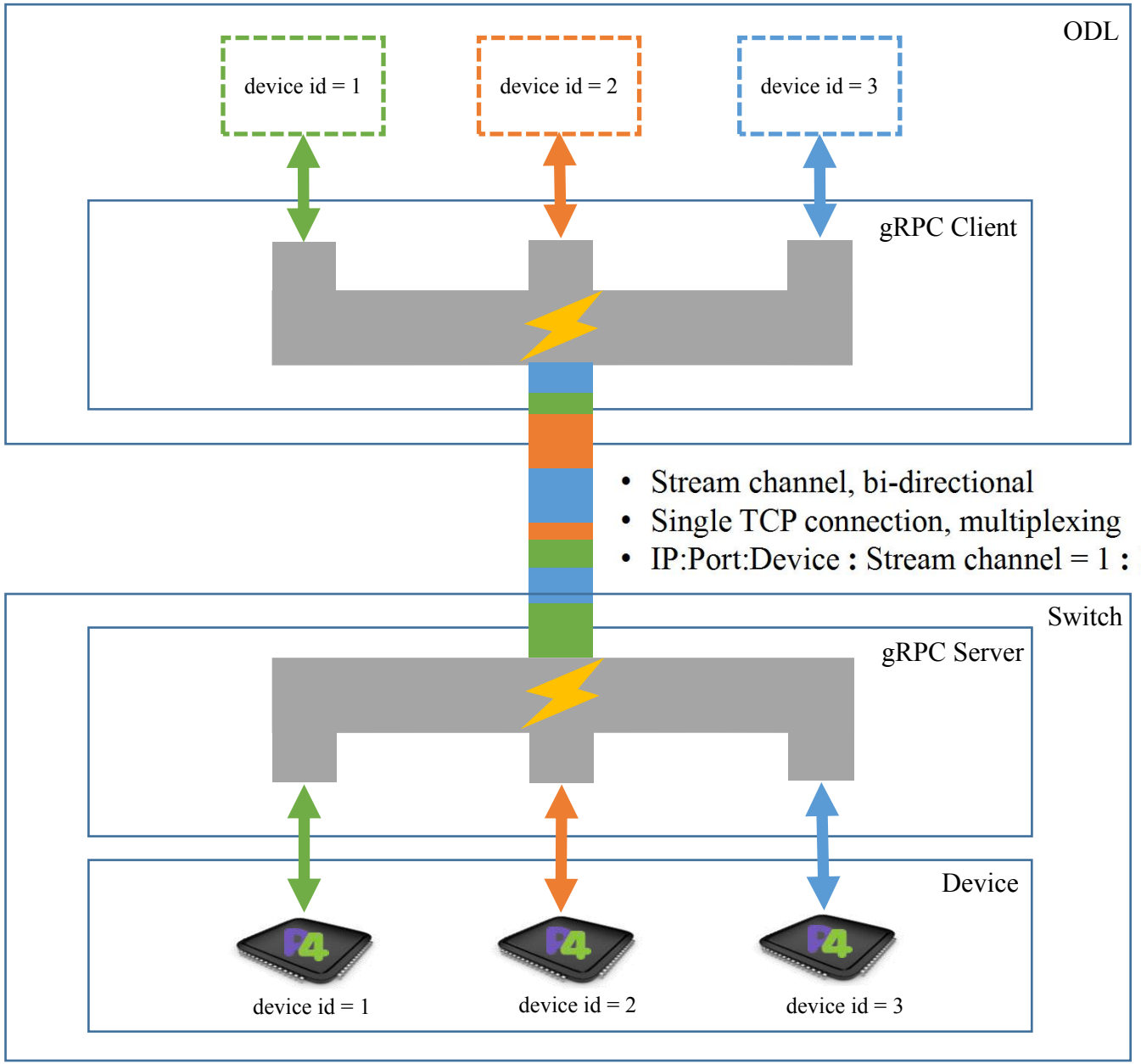
Introduce to P4Plugin in ODL



Packet-In/Out



Stream Channel



How to Implement Program Independent

```
service P4Runtime {
  rpc Write(WriteRequest) returns  ...
  rpc Read(ReadRequest)  returns ...
  ...
}

message WriteRequest {
  uint64 device_id = 1;
  Uint128 election_id = 2;
  repeated Update updates = 3;
}

message Update {
  enum Type {
    UNSPECIFIED = 0;
    INSERT = 1;
    MODIFY = 2;
    DELETE = 3;
  }
  Type type = 1;
  Entity entity = 2;
}
```

```
message Entity {
  oneof entity {
    TableEntry table_entry = 2;
     ...
  }
}

message TableEntry {
  uint32 table_id = 1;
  repeated FieldMatch match = 2;
  TableAction action = 3;
}

message TableAction {
  oneof type {
    Action action = 1;
    ...
  }
}

message Action {
  uint32 action_id = 1;
  message Param {
    uint32 param_id = 2;
    bytes value = 3;
  }
  repeated Param params = 4;
}
```

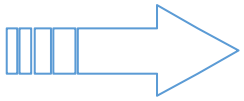
```
message FieldMatch {
  uint32 field_id = 1;
   ...
  message Exact {
    bytes value = 1;
  }
  message Ternary {
    bytes value = 1;
    bytes mask = 2;
  }
  message LPM {
    bytes value = 1;
    int32 prefix_len = 2;
  }
   oneof field_match_type {
   Exact exact = 2;
   Ternary ternary = 3;
   LPM lpm = 4;
    ...
  }
}
```

- Map-like message sequences, each message contains an unique ID and a value.
- Due to the use of **bytes**, the value will not be subject to any restrictions.
- What about API exposed to SDN Apps?

P4 Runtime Infomation

```
header_type ethernet_t { ... }
header_type ipv4_t { ... }
header ethernet_t ethernet;
header ipv4_t ipv4;
parser start { ... }
parser parse_ethernet { .. }
action set_nhop(nhop_ipv4, port)  { ... }

table ipv4_lpm {
    reads {
        ipv4.dstAddr : lpm;
    }
    actions {
        set_nhop;
        _drop;
    }
    size: 1024;
}

control ingress {
    if(valid(ipv4) and ipv4.ttl > 0) {
        apply(ipv4_lpm);
        apply(forward);
    }
}
```

simple_router.p4



Compile

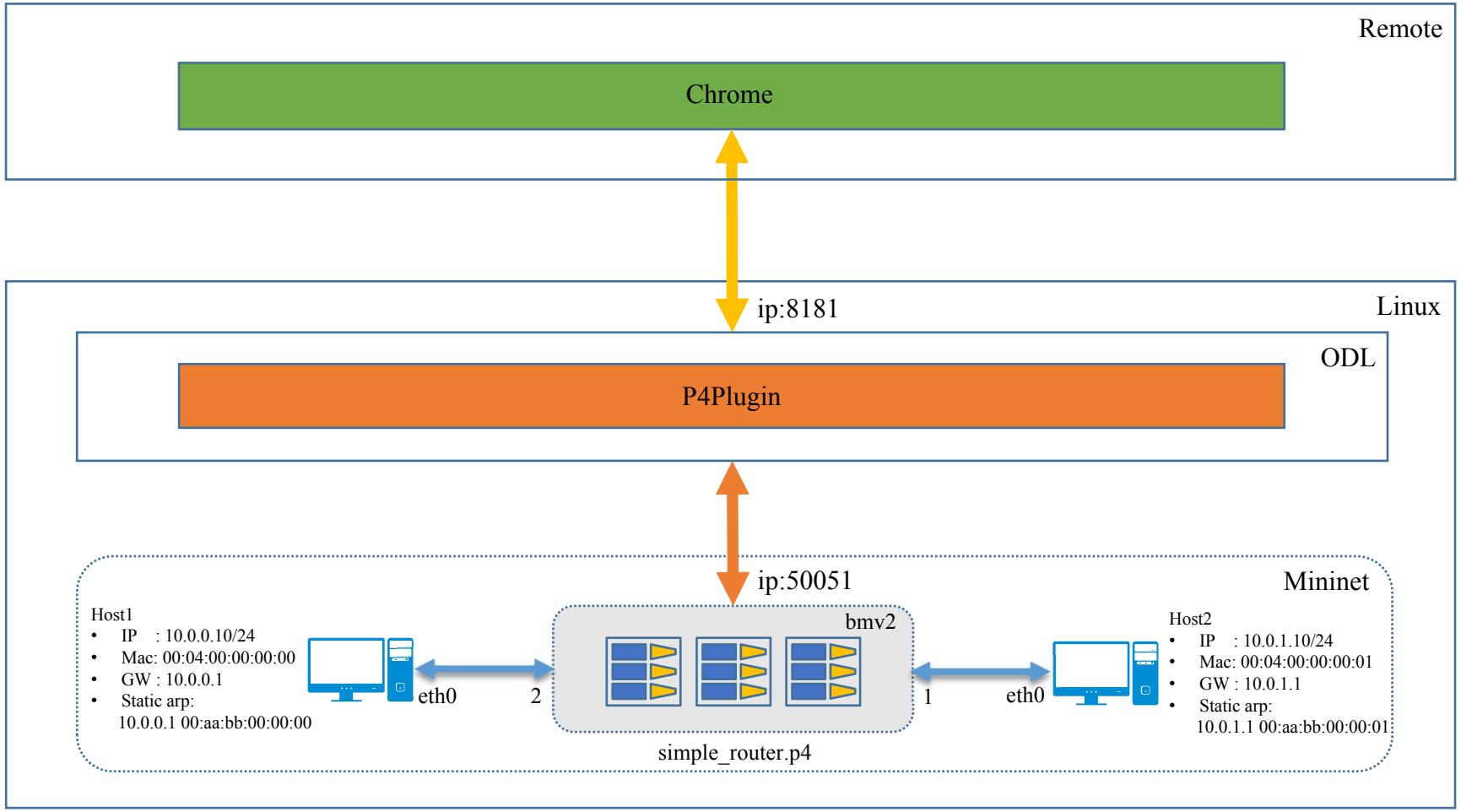
```
tables {
  preamble {
    id: 33581985
    name: "ipv4_lpm"
    alias: "ipv4_lpm"
  }
  match_fields {
    id: 1
    name: "ipv4.dstAddr"
    bitwidth: 32
    match_type: LPM
  }
  action_refs {
    id: 16812204
  }
  action_refs {
    id: 16784184
  }
  action_refs {
    id: 16800567
    annotations:
"@defaultonly()"
  }
  size: 1024
}
```

```
actions {
  preamble {
    id: 16812204
    name: "set_nhop"
    alias: "set_nhop"
  }
  params {
    id: 1
    name:
"nhop_ipv4"
    bitwidth: 32
  }
  params {
    id: 2
    name: "port"
    bitwidth: 9
  }
}
actions {
  preamble {
    id: 16784184
    name: "_drop"
    alias: "_drop"
  }
}
```

simple_router.proto.txt

ODL & P4 Demo

- Simple router based on bmv2



Detailed Steps I

- Compile

```
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ p4c-bm2-ss
simple_router.p4 --p4v 14 -o simple_router.json
simple_router.p4(107): warning: -1: negative value with unsigned type
  add_to_field(ipv4.ttl, -1);
               ^
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ p4c-bm2-ss
simple_router.p4 --p4v 14 --p4runtime-file simple_router.proto.txt --p4runtime-format text
simple_router.p4(107): warning: -1: negative value with unsigned type
  add_to_field(ipv4.ttl, -1);
               ^
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$
```

- Start ODL

```
hll@hll:/home/opendaylight/odl/new/p4plugin$ ./karaf/target/assembly/bin/karaf
karaf: JAVA_HOME not set; results may vary
Apache Karaf starting up. Press Enter to open the shell now...
100%
[========================================================================]
Karaf started in 0s. Bundle stats: 10 active, 10 total
  ________                ________          .__ .__     .__   __
  \_____  \ ______  ____  ____ \_____ \ _____ ___.__.|  | |__| ____ |  |___/  |_
   /   |   \\____ \_/ __ \/    \ |    |  \\__  \<   |  ||  | |  |/ ___\|  | \  __\
  /    |    \  |_> >  ___/|   |  \|    `   \/ __ \\___  ||  |_|  / /_/  >  Y \  |
  \_______  /   __/ \___  >___|  /_______  (____  / ____||____/__\___  /|___|
/__|
        \/|__|     \/    \/     \/    \/\/        /_____/     \/
```

Hit '<tab>' for a list of available commands
and '[cmd] --help' for help on a specific command.
Hit '<ctrl-d>' or type 'system:shutdown' or 'logout' to shutdown OpenDaylight.

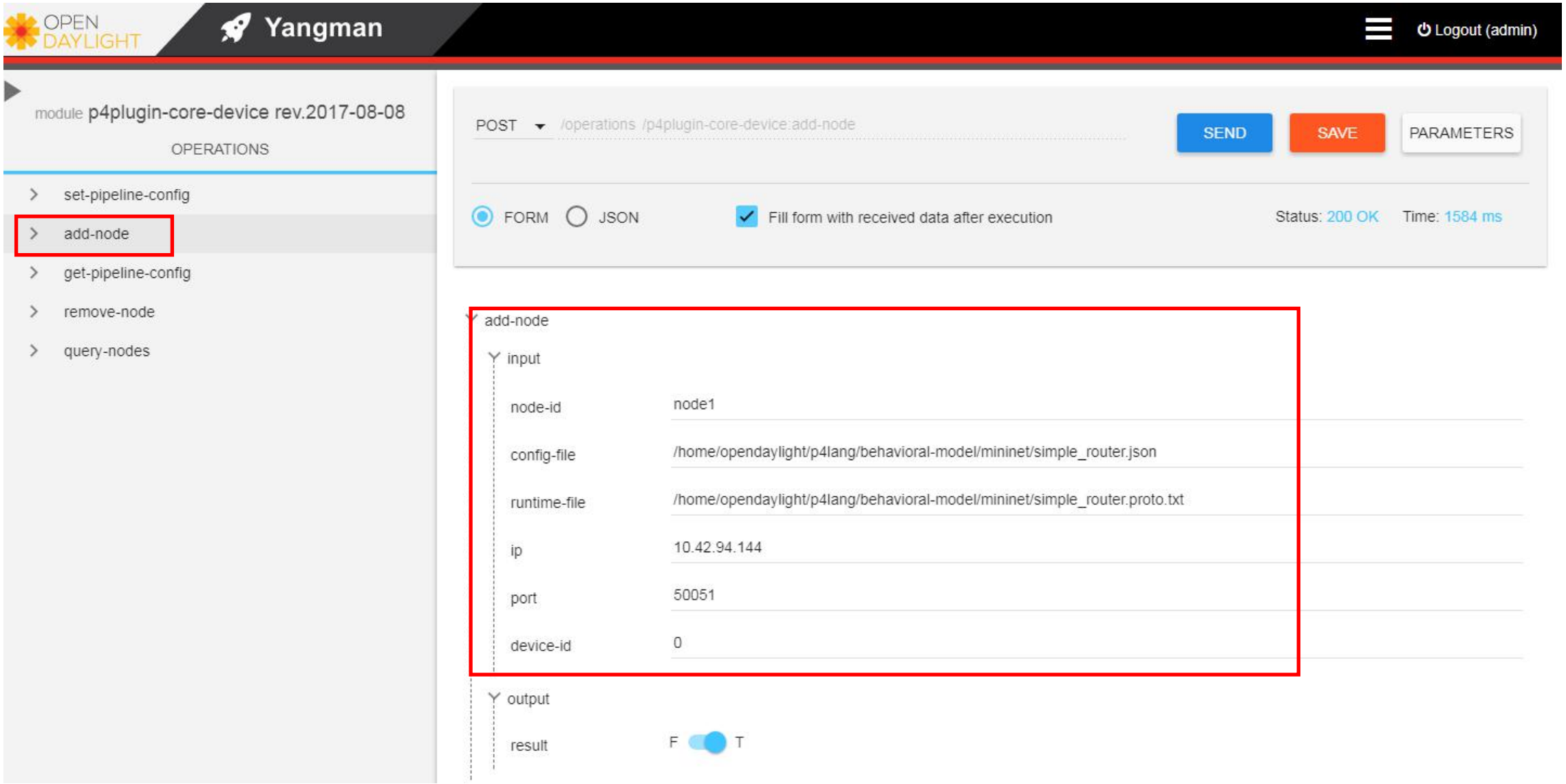
```
opendaylight-user@root>feature:install odl-p4plugin-a
odl-p4plugin-adapter   odl-p4plugin-all
opendaylight-user@root>feature:install odl-p4plugin-all
```

- Start Mininet

```
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ sudo
python ./1sw_grpc_demo.py --behavioral-
exe ../targets/simple_switch_grpc/simple_switch_grpc --cpu-port 64
[sudo] password for hll:
*** Creating network
*** Adding hosts:
h1 h2
*** Adding switches:
s1
*** Adding links:
(h1, s1) (h2, s1)
*** Configuring hosts
h1 h2
*** Starting controller

*** Starting 1 switches
s1 Starting P4 switch s1.
../targets/simple_switch_grpc/simple_switch_grpc --no-p4 -i 1@s1-eth1 -i 2@s1-eth2 --
nanolog ipc:///tmp/bm-0-log.ipc --device-id 0 --log-console -- --cpu-port 64
P4 switch s1 has been started.

**********
h1
default interface: eth0              10.0.0.10          00:04:00:00:00:00
**********
**********
h2
default interface: eth0              10.0.1.10          00:04:00:00:00:01
**********
Ready !
*** Starting CLI:
mininet>
```

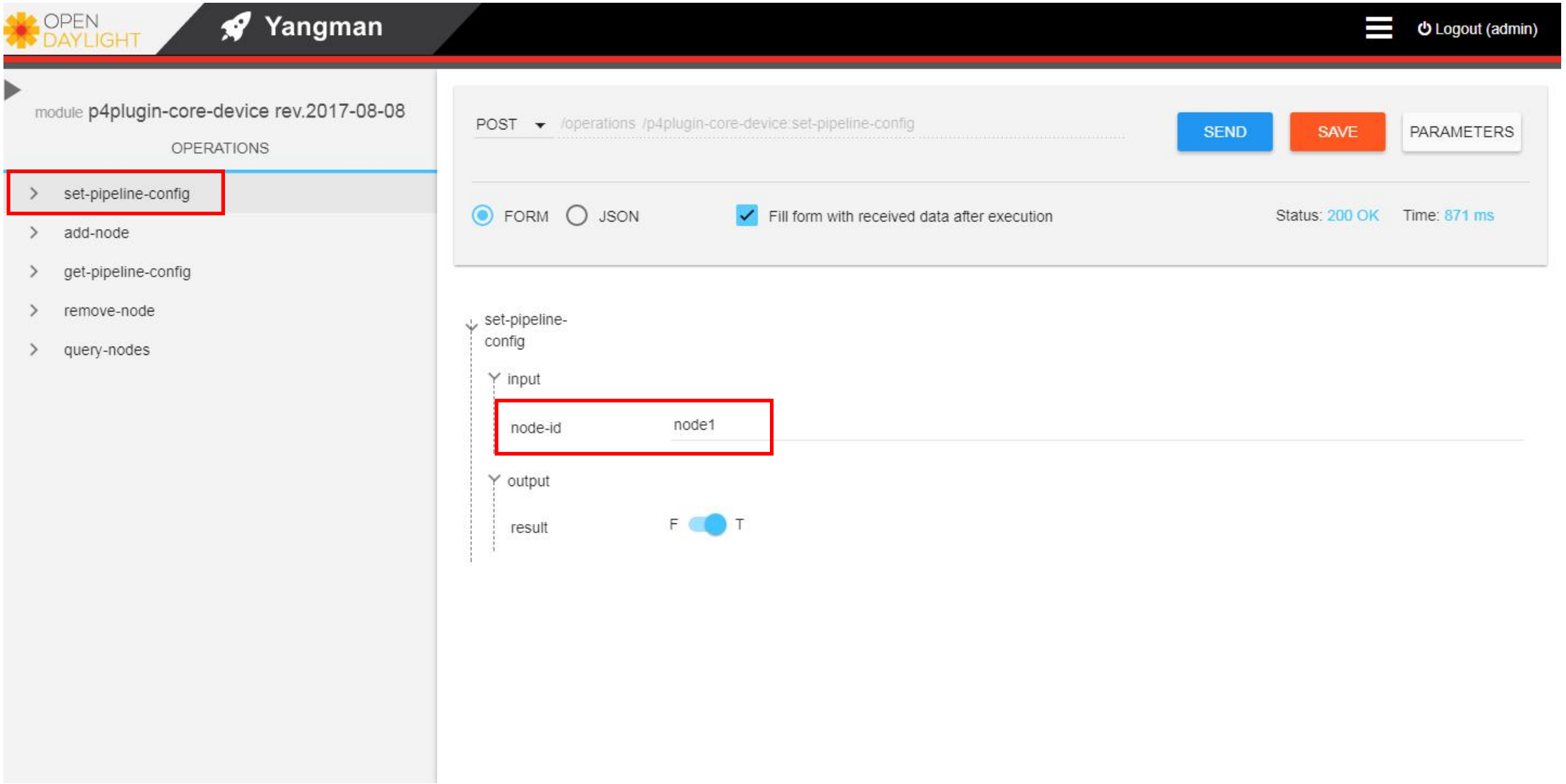
Detailed Steps II

- Add node



Detailed Steps III

- Set pipeline config



Detailed Steps IV

- TCP connection

```
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ netstat -tuan | grep 50051
tcp6       0      0 :::50051                :::*                    LISTEN

hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ netstat -tuan | grep 50051
tcp6       0      0 :::50051                :::*                    LISTEN
tcp6       0      0 10.42.94.144:38464      10.42.94.144:50051      ESTABLISHED
tcp6       0      0 10.42.94.144:50051      10.42.94.144:38464      ESTABLISHED
```

- Captured packet

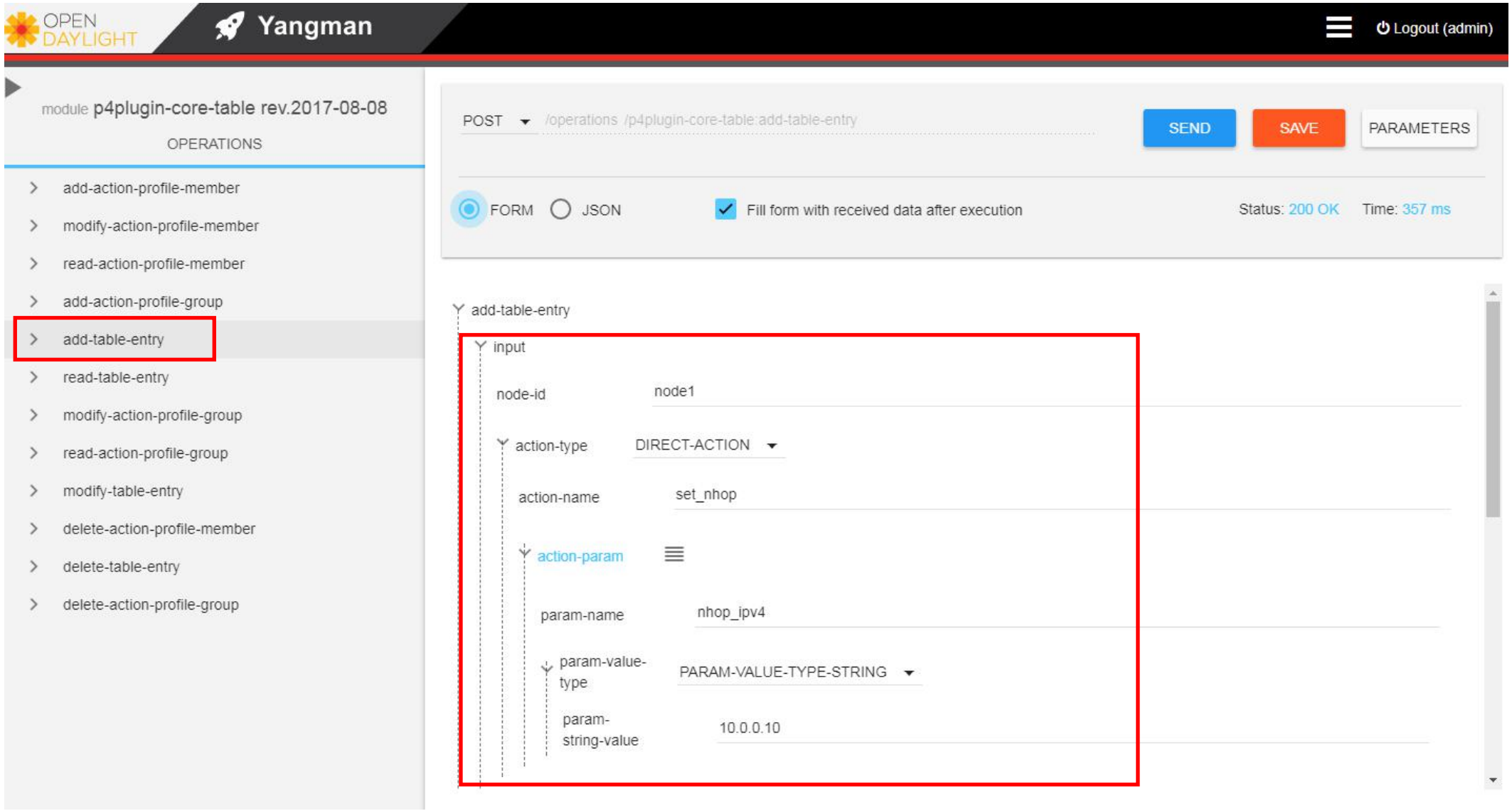
```
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ sudo tcpdump -i lo port 50051
tcpdump: listening on lo, link-type EN10MB (Ethernet), capture size 262144 bytes
11:22:33.475071 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [S], seq 2577324257 ...
11:22:33.475088 IP 10.42.94.144.50051 > 10.42.94.144.38464: Flags [S.], seq 3272263691 ...
11:22:33.475101 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [.], ack 1, win 86, ... length 0
11:22:33.475375 IP 10.42.94.144.50051 > 10.42.94.144.38464: Flags [P.], seq 1:28, ack 1 ... length 27
11:22:33.475387 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [.], ack 28... length 0
11:22:33.478076 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [P.], seq 1:71, ack 28 ... length 70
11:22:33.478089 IP 10.42.94.144.50051 > 10.42.94.144.38464: Flags [.], ack 71 ... length 0
11:22:33.478180 IP 10.42.94.144.50051 > 10.42.94.144.38464: Flags [P.], seq 28:37, ack 71 ... length 9
11:22:33.479914 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [P.], seq 71:80, ack 37 ... length 9
11:22:33.482040 IP 10.42.94.144.38464 > 10.42.94.144.50051: Flags [P.], seq 80:208, ack 37... length 128
11:22:33.482076 IP 10.42.94.144.50051 > 10.42.94.144.38464: Flags [.], ack 208 ... length 0
```

- Log

```
hll@hll:/home/opendaylight/p4lang/behavioral-model/mininet$ more /tmp/p4s.s1.log
Thrift port was not specified, will use 9090
Calling target program-options parser
Adding interface s1-eth1 as port 1
[09:50:25.359] [bmv2] [D] [thread 4008] Adding interface s1-eth1 as port 1
Adding interface s1-eth2 as port 2
[09:50:25.425] [bmv2] [D] [thread 4008] Adding interface s1-eth2 as port 2
Server listening on 0.0.0.0:50051
P4Runtime SetForwardingPipelineConfig
[10:11:56.557] [bmv2] [D] [thread 4025] Set default entry for table 'ipv4_lpm': NoAction -
[10:11:56.557] [bmv2] [D] [thread 4025] Set default entry for table 'forward': NoAction -
[10:11:56.557] [bmv2] [D] [thread 4025] Set default entry for table 'send_frame': NoAction -
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Processing packet received on port 1
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Parser 'parser': start
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Parser state 'start' has no switch, going to
default next state
[10:11:56.559] [bmv2] [T] [thread 4761] [0.0] [cxt 0] Bytes parsed: 0
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Extracting header 'ethernet'
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Parser state 'parse_ethernet': key is 86dd
[10:11:56.559] [bmv2] [T] [thread 4761] [0.0] [cxt 0] Bytes parsed: 14
[10:11:56.559] [bmv2] [T] [thread 4761] [0.0] [cxt 0] Skipping checksum 'cksum' verification
because target field invalid
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Parser 'parser': end
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Pipeline 'ingress': start
[10:11:56.559] [bmv2] [T] [thread 4761] [0.0] [cxt 0] simple_router.p4(153) Condition "and" is
false
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Pipeline 'ingress': end
[10:11:56.559] [bmv2] [D] [thread 4761] [0.0] [cxt 0] Egress port is 0
[10:11:56.559] [bmv2] [D] [thread 4762] [0.0] [cxt 0] Pipeline 'egress': start
[10:11:56.559] [bmv2] [T] [thread 4762] [0.0] [cxt 0] Applying table 'send_frame'
[10:11:56.559] [bmv2] [D] [thread 4762] [0.0] [cxt 0] Looking up key:
* standard_metadata.egress_port: 0000
```

Detailed Steps IV

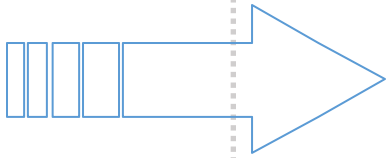
- Add table entry (Form format)



Detailed Steps IV (continued)

- Json format

```
{
  "add-table-entry": {
    "input": {
      "node-id": "node1",
      "action-name": "set_nhop",
      "action-param": [
        {
          "param-name": "port",
          "param-string-value": "1"
        },
        {
          "param-name": "nhop_ipv4",
          "param-string-value": "10.0.0.10"
        }
      ],
      "priority": "0",
      "controller-metadata": "0",
      "table": "ipv4_lpm",
      "field": [
        {
          "field-name": "ipv4.dstAddr",
          "lpm-string-value": "10.0.0.0",
          "lpm-prefixLen": "24"
        }
      ]
    }
  }
}
```



- Data

```
P4Runtime Write
updates {
 type: INSERT
 entity {
  table_entry {
   table_id: 33581985
   match {
    field_id: 1
    lpm {
     value: "\n\000\000\000"
     prefix_len: 24
    }
   }
   action {
    action {
     action_id: 16812204
     params {
      param_id: 2
      value: "\000\001"
     }
     params {
      param_id: 1
      value: "\n\000\000\n"
     }
    }
   }
   controller_metadata: 0
  }
 }
}
```

Detailed Steps V

- Ping test

```
*** Starting CLI:
mininet>
mininet> pingall
*** Ping: testing ping reachability
h1 -> h2
h2 -> h1
*** Results: 0% dropped (2/2 received)
mininet> h1 ping h2 -c 4
PING 10.0.1.10 (10.0.1.10) 56(84) bytes of data.
64 bytes from 10.0.1.10: icmp_seq=1 ttl=63 time=1.32 ms
64 bytes from 10.0.1.10: icmp_seq=2 ttl=63 time=1.42 ms
64 bytes from 10.0.1.10: icmp_seq=3 ttl=63 time=2.07 ms
64 bytes from 10.0.1.10: icmp_seq=4 ttl=63 time=1.14 ms

--- 10.0.1.10 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3004ms
rtt min/avg/max/mdev = 1.149/1.494/2.070/0.347 ms
```

- Packet processing log

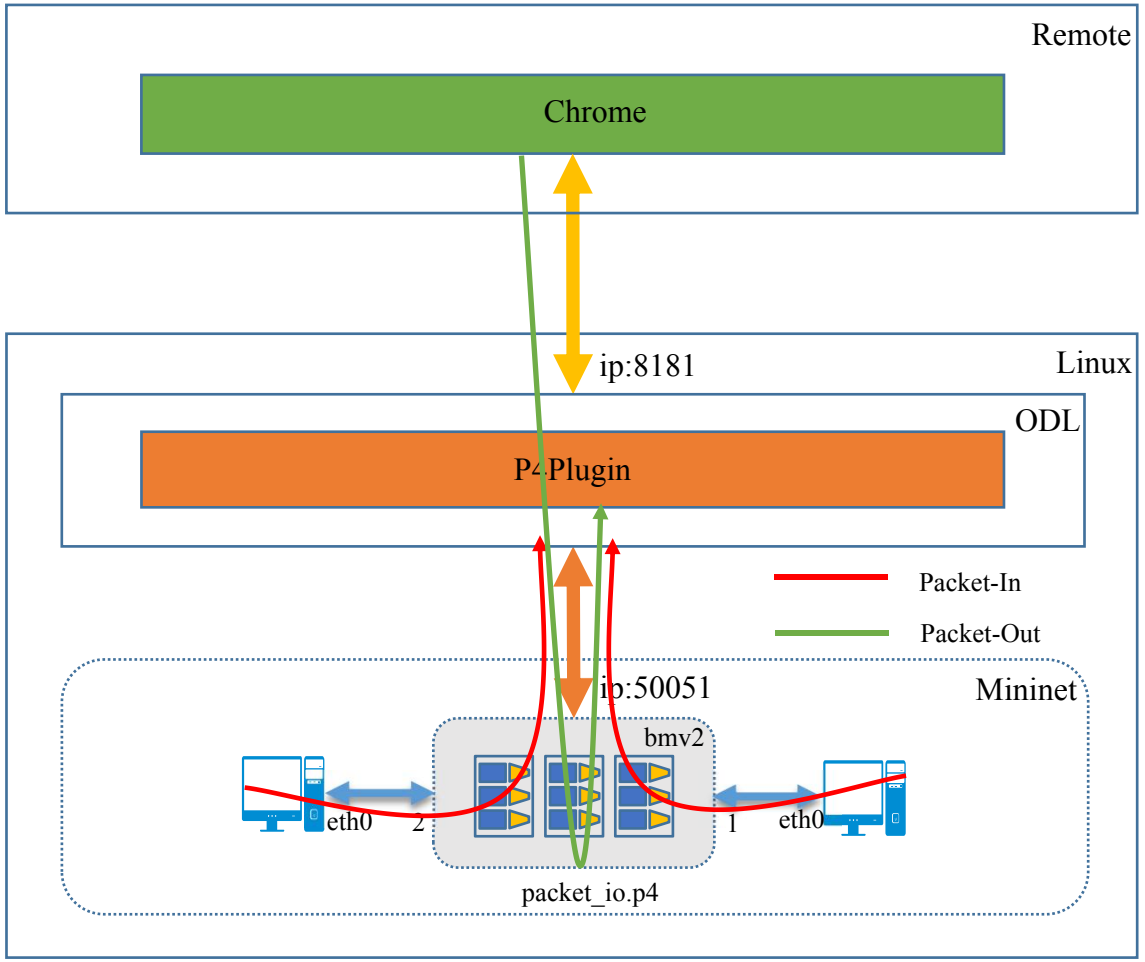
```
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Processing packet received on port 1
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Parser 'parser': start
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Parser state 'start' has no switch, going to default next state
[15:57:51.603] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Bytes parsed: 0
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Extracting header 'ethernet'
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Parser state 'parse_ethernet': key is 0800
[15:57:51.603] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Bytes parsed: 14
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Extracting header 'ipv4'
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Parser state 'parse_ipv4' has no switch, going to default next state
[15:57:51.603] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Bytes parsed: 34
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Verifying checksum 'cksum': true
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Parser 'parser': end
[15:57:51.603] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Pipeline 'ingress': start
[15:57:51.603] [bmv2] [T] [thread 27525] [25.0] [cxt 0] simple_router.p4(153) Condition "and" is true
[15:57:51.603] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Applying table 'ipv4_lpm'
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Looking up key:
* ipv4.dstAddr       : 0a00010a
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Table 'ipv4_lpm': hit with handle 1
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Dumping entry 1
Match key:
* ipv4.dstAddr       : LPM      0a000100/24
Action entry: set_nhop - a00010a,2,

[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Action entry is set_nhop - a00010a,2,
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Action set_nhop
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] simple_router.p4(104) Primitive nhop_ipv4, port) { ...
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] simple_router.p4(104) Primitive port) { ...
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] simple_router.p4(107) Primitive add_to_field(ipv4.ttl, -1)
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Applying table 'forward'
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Looking up key:
* routing_metadata.nhop_ipv4: 0a00010a

[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Table 'forward': hit with handle 1
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Dumping entry 1
Match key:
* routing_metadata.nhop_ipv4: EXACT     0a00010a
Action entry: set_dmac - 400000001,

[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Action entry is set_dmac - 400000001,
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] Action set_dmac
[15:57:51.604] [bmv2] [T] [thread 27525] [25.0] [cxt 0] simple_router.p4(121) Primitive dmac) { ...
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Pipeline 'ingress': end
[15:57:51.604] [bmv2] [D] [thread 27525] [25.0] [cxt 0] Egress port is 2
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Pipeline 'egress': start
[15:57:51.604] [bmv2] [T] [thread 27528] [25.0] [cxt 0] Applying table 'send_frame'
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Looking up key:
* standard_metadata.egress_port: 0002

[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Table 'send_frame': hit with handle 1
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Dumping entry 1
Match key:
* standard_metadata.egress_port: EXACT     0002
Action entry: rewrite_mac - 10203040506,

[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Action entry is rewrite_mac - 10203040506,
[15:57:51.604] [bmv2] [T] [thread 27528] [25.0] [cxt 0] Action rewrite_mac
[15:57:51.604] [bmv2] [T] [thread 27528] [25.0] [cxt 0] simple_router.p4(136) Primitive smac) { ...
[15:57:51.604] [bmv2] [T] [thread 27528] [25.0] [cxt 0] simple_router.p4(138) Primitive
modify_field(standard_metadata.egress_port, 2)
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Pipeline 'egress': end
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Deparser 'deparser': start
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Updating checksum 'cksum'
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Deparsing header 'ethernet'
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Deparsing header 'ipv4'
[15:57:51.604] [bmv2] [D] [thread 27528] [25.0] [cxt 0] Deparser 'deparser': end
[15:57:51.604] [bmv2] [D] [thread 27530] [25.0] [cxt 0] Transmitting packet of size 98 out of port 2
```

ODL & P4 Demo

- Packet in/out based on bmv2

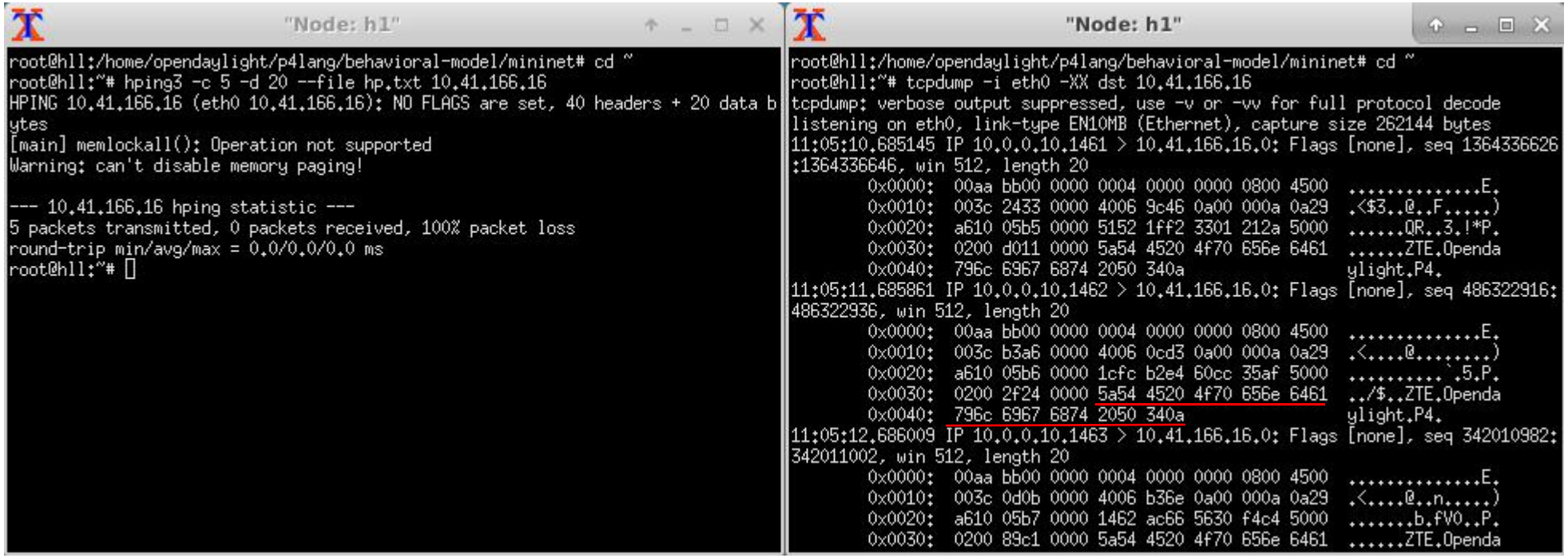
packet_io.p4

```
#define CPU_PORT 64

parser start {
    return ingress;
}

action redirect() {
    modify_field(standard_metadata.egress_spec, CPU_PORT);
}

table t_redirect {
    actions { redirect; }
    default_action: redirect();
}

control ingress {
    apply(t_redirect);
}

control egress { }
```



Detailed Steps I

- Packet in
  - Transmit and Capture



  - Receive log

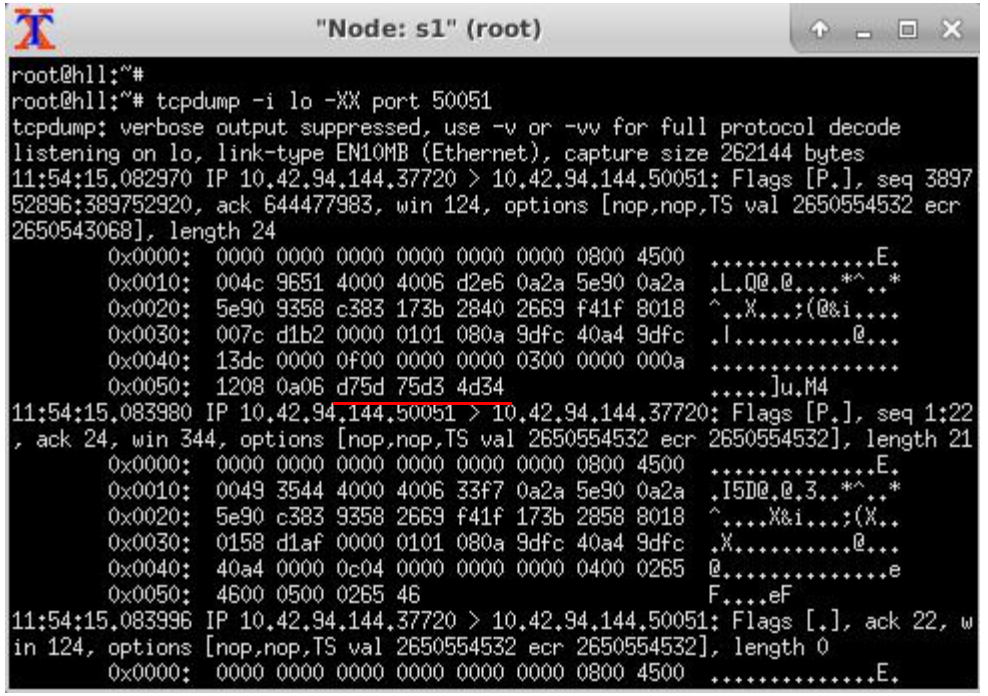
```
2017-11-17 11:05:13,687 | INFO  | fault-executor-6 | P4RuntimeStub                    | 286 - org.opendayli
ght.p4plugin.core-impl - 0.1.0.SNAPSHOT | Receive packet from node = node1, body = 00aabb000000000400000000
08004500003cd5f700004006ea810a00000a0a29a61005b800003e38cdca12bd3fc65000020036f800005a5445204f70656e6461796
c696768742050340a.
```

Detailed Steps II

- Packet out
  - ➢ Transmit and Capture

YangUI

* “11110000” base64 decode = d75d75d34d34



- ➢ Receive log

```
2017-11-17 11:21:52,218 | INFO  | 1480256041-14325 | P4RuntimeStub                    | 286 - org.ope
ght.p4plugin.core-impl - 0.1.0.SNAPSHOT | Transmit packet = d75d75d34d34 to node = node1.
2017-11-17 11:21:52,222 | INFO  | fault-executor-9 | NotificationServiceProvider      | 286 - org.ope
ght.p4plugin.core-impl - 0.1.0.SNAPSHOT | Notification publish!
2017-11-17 11:21:52,222 | INFO  | fault-executor-9 | P4RuntimeStub                    | 286 - org.ope
ght.p4plugin.core-impl - 0.1.0.SNAPSHOT | Receive packet from node = node1, body = d75d75d34d34.
```

Thank you.