

Openflow Protocol

- Operates b/w controller, switch
- TCP used to exchange messages
(optional encryption)
- 3 classes of OF messages:
 - controller to switch
 - asynchronous (switch to controller)
 - symmetric (misc)
- Distinct from Open Flow API

Use Cases

Network Virtualizations

- Despite sharing the same physical links, we can divide the network into many virtual networks which feel isolated by themselves despite sharing physicals through SDN.

Network Telemetry

- Add meta data such as switch ID, arrival and departure times, queue delays etc. to a packet at each of the switches it passes through in order to generate reports at the end. for analysis.



Key usage enabled by SDN \rightarrow Generalized forwarding

\rightarrow Match plus action - matches bits in arriving packet, takes action.

Earlier - destination based forwarding : - based on IP.

Now - Generalized Forwarding :-

- ① Many header fields can determine action.
- ② Many actions possible : - drop / copy / modify / log packets

Simple packet handling rules :-

- 1) Match: pattern values in packet header fields
- 2) Actions: for matched packet, drop, forward, modify or send to controller.
- 3) Priority: Disambiguate overlapping patterns
- 4) Counters: # bytes and # packets.

OpenFlow Abstractions

① Router:

- match: longest destination IP prefix
- action: forward out a link

③ Firewall:

- match: IP and TCP/UDP port numbers
- action: permit or deny

② Switch:

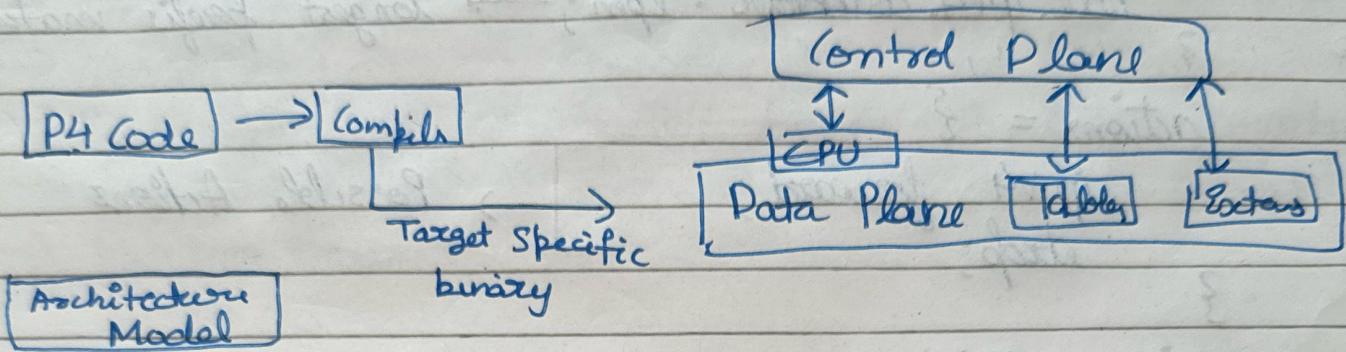
- match: destination MAC address
- action: forward or flood.

③ NAI:

- match: IP and port
- action: rewrite address and port.

→ Protocol Independent Switch Architecture (PISA)

P4 Programming



Lecture 20 (P4)



Eg table ipv4_lpm {

key = {

 } hdr.ipv4.dstAddr : lpm; → longest prefix match

actions = {

 } ip v4 forward;
 drop;

] → Possible Actions

size = 1024;

default action= drop();] → Default Action

Example : IP forwarding table

action ip v4_forward (macAddr_t dstAddr, egressSpec_t port) {

 } hdr.ethernet.srcAddr = hdr.ethernet.dstAddr;

 } hdr.ethernet.dstAddr = dstAddr;

 } hdr.ipv4.ttl = hdr.ipv4.ttl - 1;

 } standard_metadata.egress_spec = port;

}

Metadata

Standard

- Ingress port → Egress Spec → Egress Port
- Clone Spec → Instance Type → Drop.
- Recvict Port → Packet length → Enq. Timestamp
- Enq_qdepth → deq-timedelta → deq-q_depth
- Parser Error

Intrinsic

- ingress_global_timestamp → egress_global_timestamp
- if-field-list → most_grp → resubmit_flag
- egress_rid → checksum_error, → recursive_flag.

The intrinsic flags often extend some functionality and even though they can be called by P4, they are really actually enabled by the architecture.

Externs

Each architecture also defines a list of "externs"
i.e. black box functions whose interface is known
OR

Functions that are provided by the hardware

- Most targets contain specialized components which cannot be expressed in P4 (e.g. complex computations)
- At the same time, P4 should be target independent

Similar to Java interface where only signature is known but not implementation



Portable Switch Architecture

Language Constructs Basics

- Statically typed language.
- Base types and ~~derived~~ operators to derive compound ones.

- bool
- bit<w> bit string of width W
- int<w>



Data Plane



- Deals with forwarding
i.e. moving packets from
input port to the appropriate
output port
- Deals with protocols like IP

- Deals with FIB, optimized
for lookups etc.

- Northbound control interface
(historically private)

Control Plane

- Deals with routing
i.e. determining the path the
packet will take from src. to dest.
- Deals with protocols like
OSPF, BGP, RIP.

- Deals with RIB, collection
of link / path attributes

- Northbound configuration
interface via CLI



Lecture 21

P4

- The time taken by each packet to go through the pipeline is fixed.
- Does not support everything as memory is limited and time for processing is extremely less.

Header

header Ethernet_h {

 bit <48> dstAddr;

 bit <48> srcAddr;

 bit <16> etherType;

}

- Parsing a packet using extract() fills in the fields of the header from a network packet.
- A successful extract sets to true the validity bit of the extracted header.