

daPIPE

Data Plane

Incremental Programming

Environment

Mario Baldi


Politecnico di Torino (Technical University of Turin)

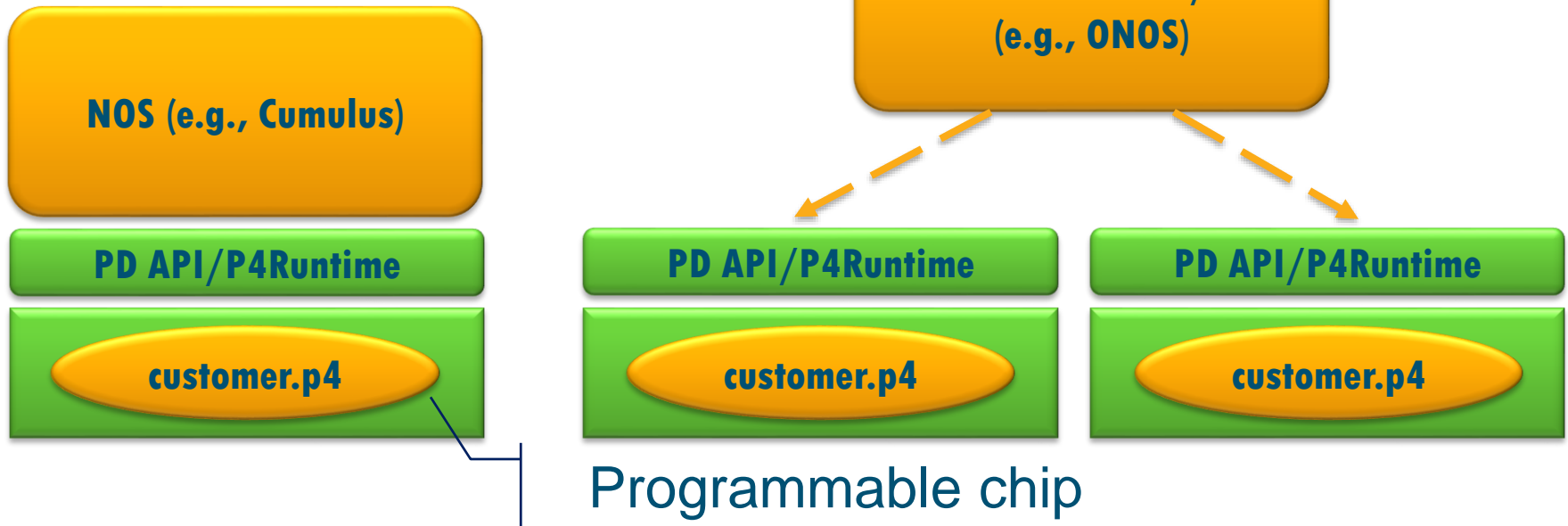
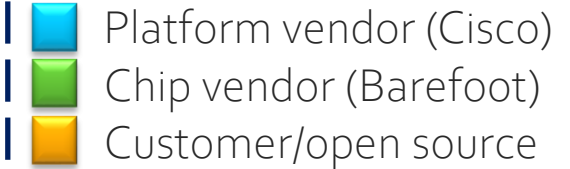
Department of Control and Computer Engineering

To set the context

Let's look into deployment options for programmable switches

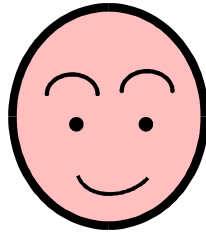
Whitebox Deployment

- Maximum flexibility 
- Maximum disruption/risk/work 

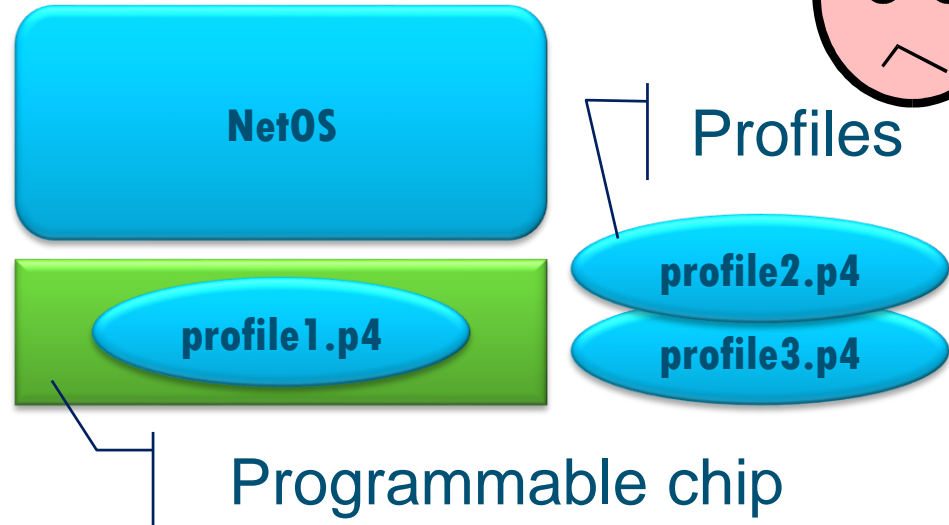
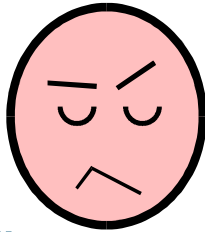


Turn-key Deployment

- Deployment as usual
 - Familiar features and interfaces
- Resource optimization
- Future proof
- Feature agility
- Streaming telemetry



- No flexibility
 - No custom feature and protocol support

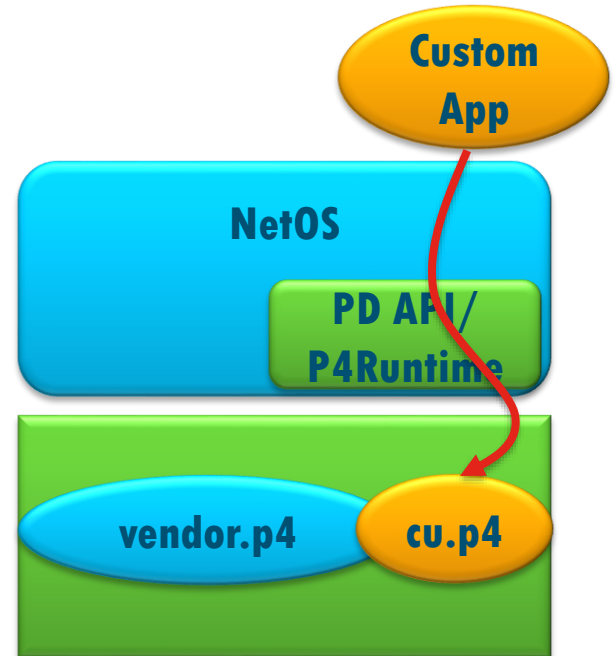


Hybrid Deployment

- Best of breed
- Deployment as usual
 - Familiar features and interfaces
- Minimum development effort
 - Leverage existing functions in building new features



Minimize disruption
and risk!



Challenges

Do not break what works

- Vendor data plane code is well tested
- ... and we don't want to need regression testing

Don't want to show, don't want to see

- Vendor code and custom code may be confidential
- Not practical to familiarize with a lot of vendor code to just write a few lines

Resource availability

- Still “limited” on current chips

Data/control plane dependence

- Net OS should keep working
- Net OS should not be aware of custom data plane functions

In a nutshell

P4 and its ecosystem were not
designed for
incremental programming



Single
programmer



Single source
code



Single control
plane

We need to explicitly support

Incremental Programming

How can we address these challenges?

Identify constraints
on new code

Enforce those
constraints on
custom code

Challenges

Do not break what works

- Vendor data plane code is well tested
- ...and we don't want to need regression testing

Don't want to show, don't want to see

- Vendor code and custom code may be confidential
- Not practical to familiarize with a lot of vendor code to just write a few lines

Resource availability

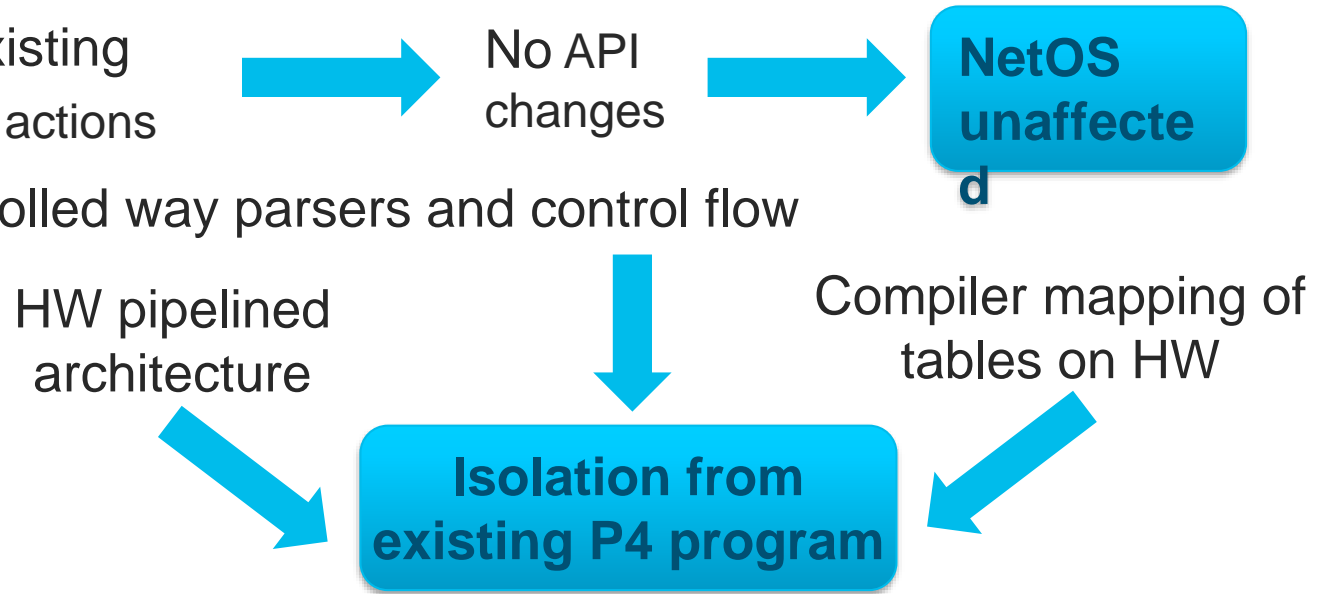
- Still "limited" on current chips

Data/control plane dependence

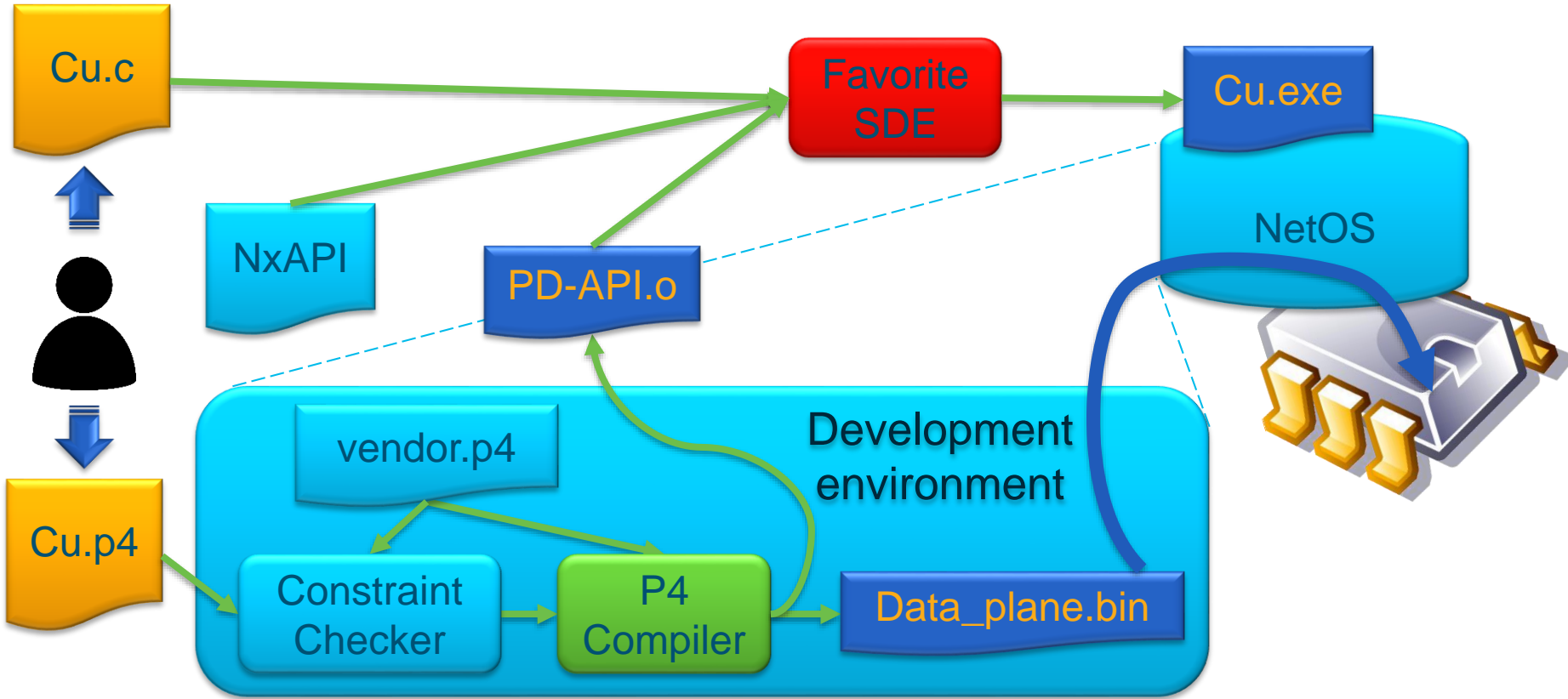
- NXOS should keep working
- NXOS should not be aware of custom data plane functions

Do's and Don't's

- Do add new
 - Headers, parsers, tables, actions
- Do not modify existing
 - Headers, tables, actions
- Modify in a controlled way parsers and control flow



Customer Programming Workflow



daPIPE

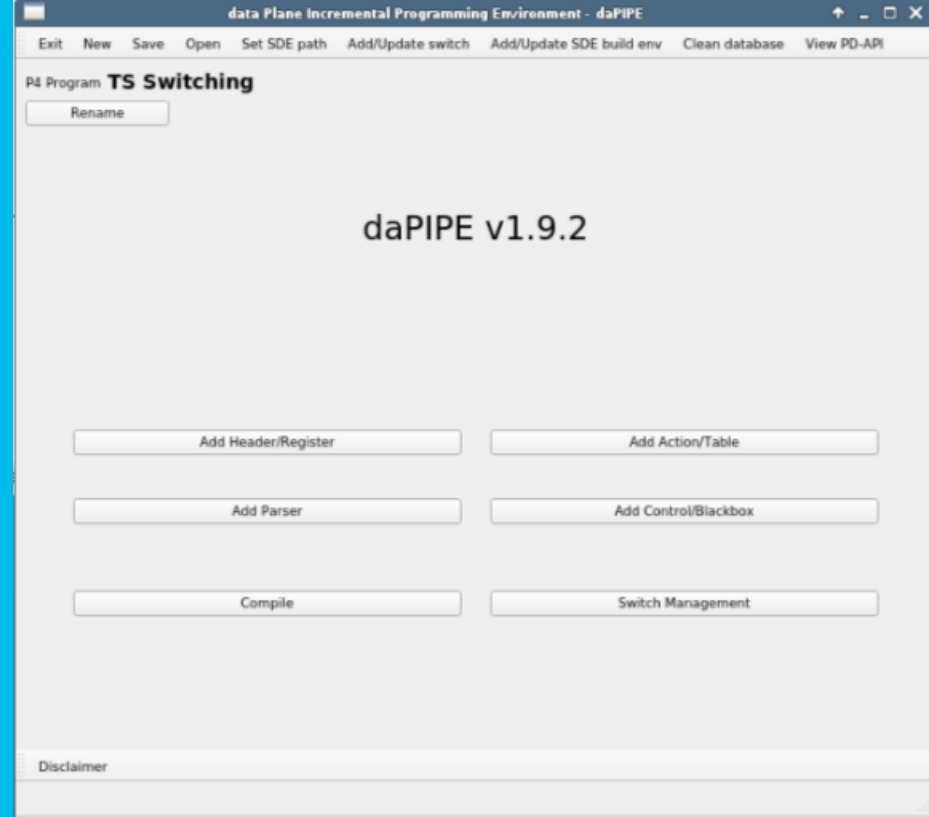
Data

Plane

Incremental

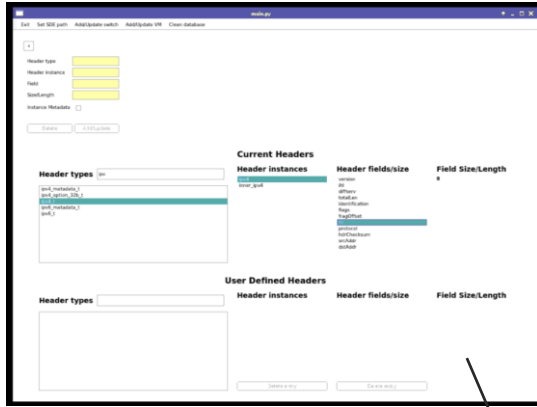
Programming

Environment

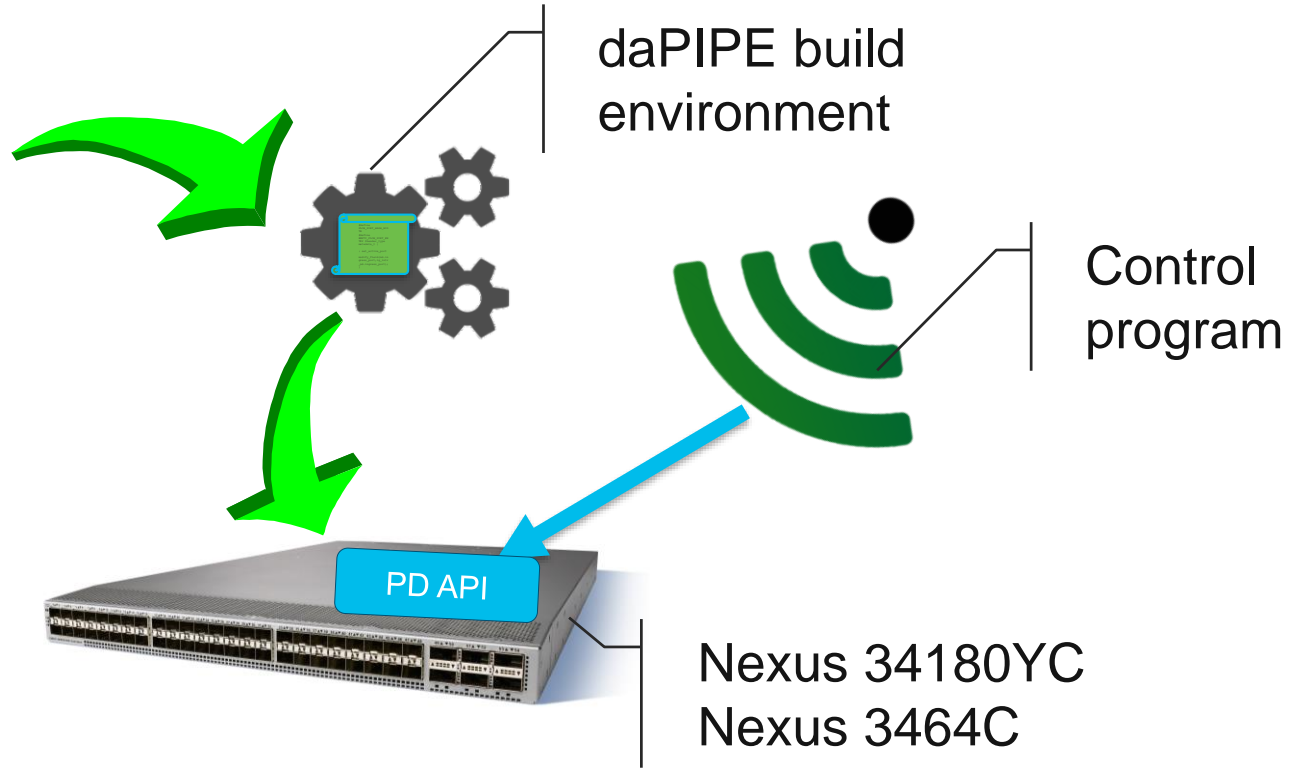


*Support developers
and streamline their task
(while enforcing constraints)*

Components of the Solution

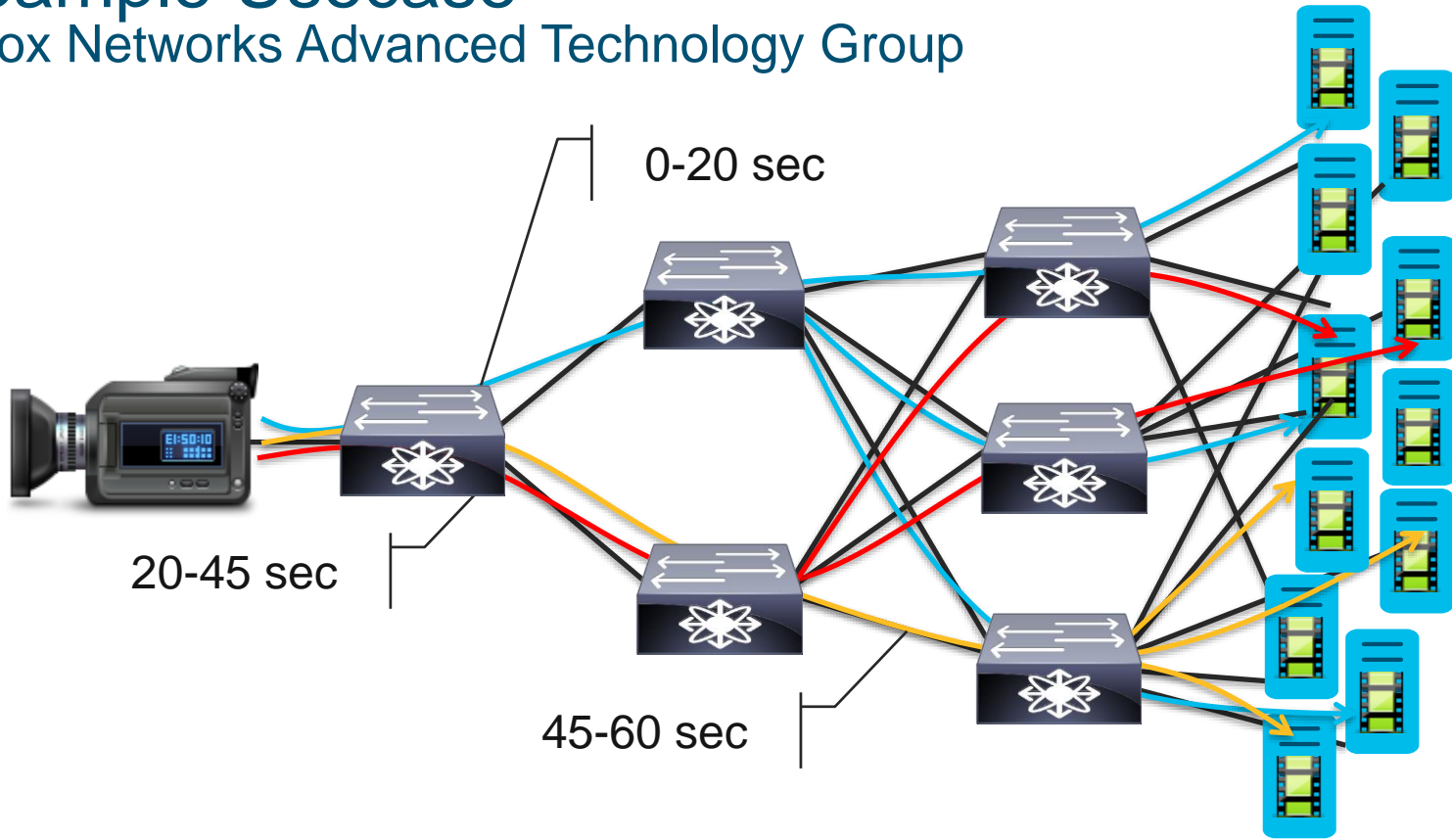


daPIPE Graphical User Interface



Sample Usecase

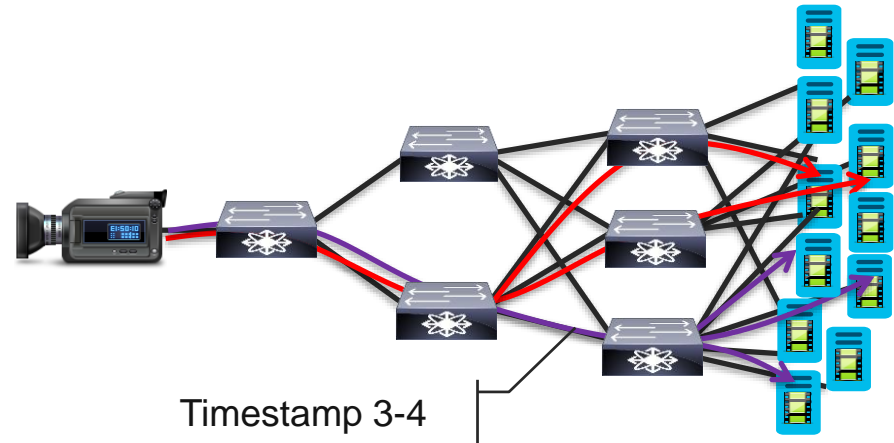
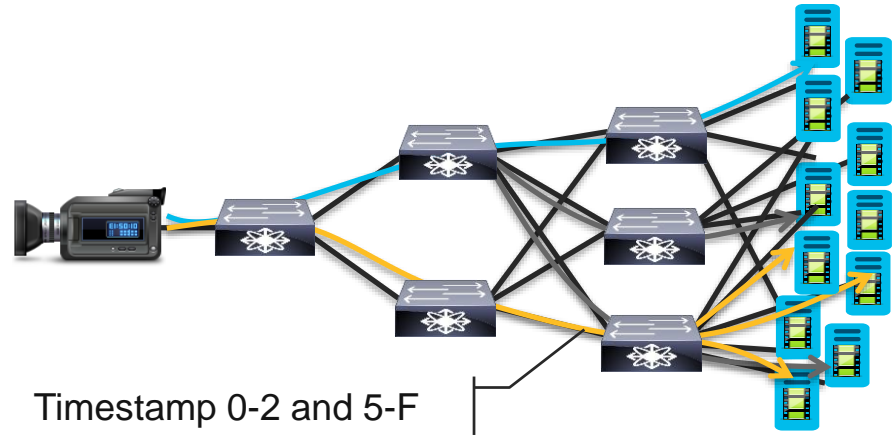
Fox Networks Advanced Technology Group



https://github.com/FOXNEOAdvancedTechnology/ts_switching_P4

Specification

- A switch shall forward packets based on the **RTP timestamp** they contain
- If sent to 239.1.1.1, change destination address to 239.3.3.3 when **RTP timestamp** is
 - Between 0 and 2
 - Between from 5 and F
- If sent to 239.2.2.2, change destination address to 239.3.3.3 when **RTP timestamp** is
 - Between 3 and 4



Incremental Programming Unique Advantage

- Leverage existing features
 - Protocol parsing up to UDP messages
 - Layer 2-3 forwarding, including multicast packet forwarding
 - Multicast routing (offered by the operating system)
- Focus on new feature
 - Write just a few lines of P4 code and control code

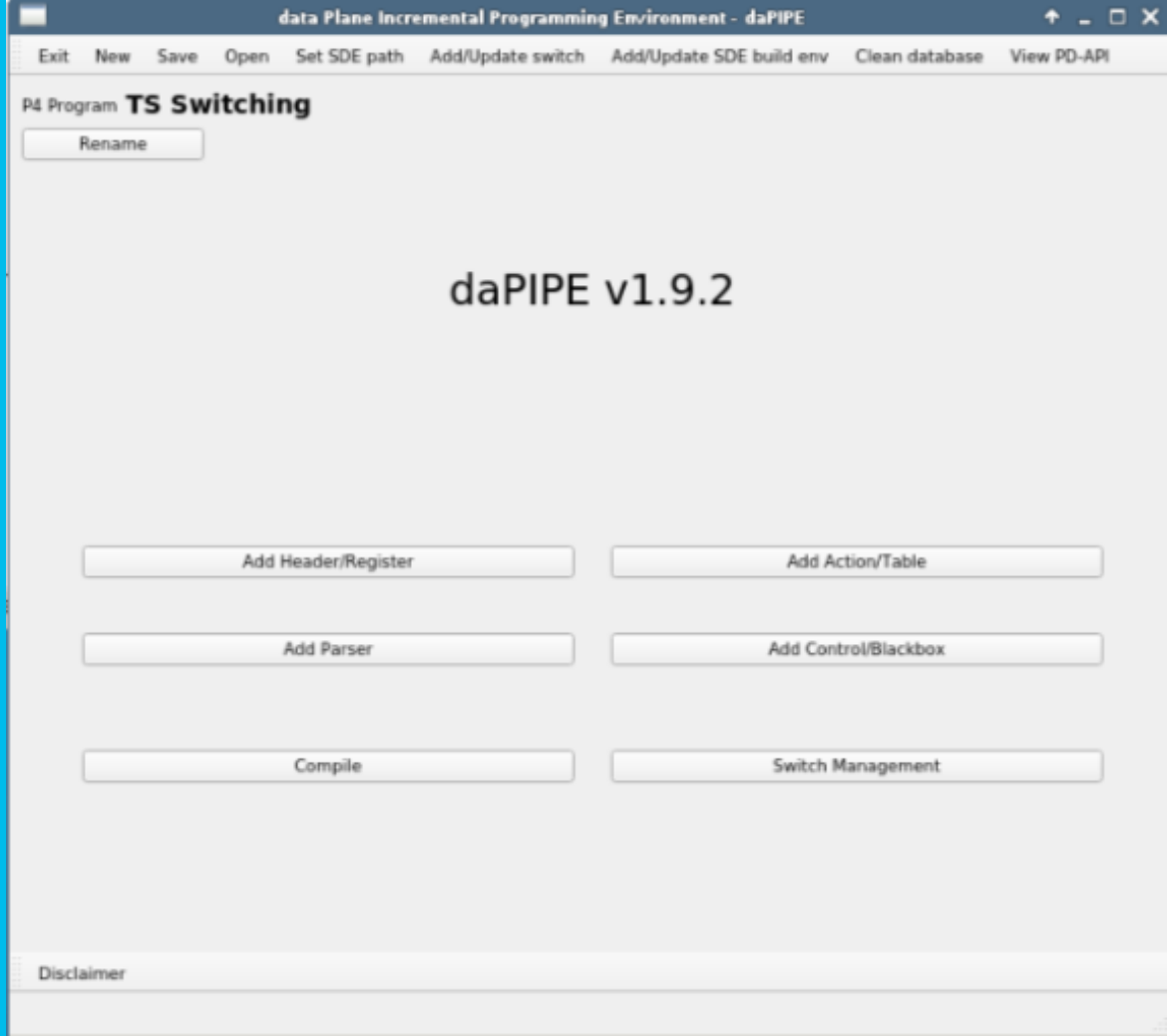
daPIPE bonus feature:

no need to deal with the complexity of pre-existing code

Development Workflow

- Browse available (stock) metadata
- Define custom **headers** and metadata
- Specify **parser(s)** and their hook(s) in existing (stock) parsers
- Define custom **tables** and **actions**
- Specify **control flow**
- Compile and load on chip
- Develop control plane functionalities

Main window



Existing header view

main.py

Exit Set SOE path Add/Update switch Add/Update VM Clean database

Header type

Header instance

Field

Size/Length

Instance Metadata

Delete Add/Update

Current Headers

Header types	Header instances	Header fields/size	Field Size/Length
<input type="text" value="ipv"/> ipv4_metadata_t ipv4_option_32b_t ipv4_t ipv6_metadata_t ipv6_t	ipv4 inner_ipv4	version ihl diffserv totallLen identification flags fragOffset ttl protocol hdrChecksum srcAddr dstAddr	8

User Defined Headers

Header types	Header instances	Header fields/size	Field Size/Length
<input type="text"/>			

Delete entry Delete entry

4

Header type

Header instance

Field

Size/Length

Instance Metadata

Delete

Add/Update

Current Headers

Header types

```

acl_metadata_t
egress_intrinsic_metadata_for_mirror_buffer_t
egress_intrinsic_metadata_for_output_port_t
egress_intrinsic_metadata_from_parser_aux_t
egress_intrinsic_metadata_t
egress_metadata_t
erspan_header_t3_t
ethernet_t
fabric_header_cpu_t
fabric_header_t
fabric_header_timestamp_t
fabric_metadata_t

```

Header instances

Header fields/size

Field Size/Length

User Defined Headers

Header types

rtp_t

Header instances

rtp

Header fields/size

padding
extension
version

Field Size/Length

2

Delete entry

Delete entry

Adding RTP header

Current Parsers

Selected parser: parse_udp Selected hook point:

Parsers

```

parse_set_prio_high
parse_set_prio_med
parse_sflow
parse_snap_header
parse_tcp
parse_udp
parse_vlan
parse_vxlan
start
USER PARSERS:

```

Hook points

udp.dstPort

Potential Hook points

Parser name

parse_rtp

Hook value

 Set as default

Delete

Add/Update

Clear

```

extract(rtp);
return ingress;

```

Current parser:

Current hook point:

Available Headers

Header types

```

rtp_t
acl_metadata_t
egress_intrinsic_metadata_for_mirror_buffer_t
egress_intrinsic_metadata_for_output_port_t
egress_intrinsic_metadata_from_parser_aux_t
egress_intrinsic_metadata_t
egress_metadata_t
erspan_header_t3_t
ethernet_t
fabric_header_cpu_t

```

Header instances

Header fields/size

Field Size/Length

Adding RTP parser

Resulting Parsing Code

Stock code
Custom code
Autom. code

```
...
header_type ethernet_t {
    fields {
        dstAddr : 48;
        srcAddr : 48;
        etherType : 16;
    }
}
header ethernet_t ethernet;
...
header_type rtp_t {
    fields {
        version : 2;
        padding : 1;
        ...
        sequence_number : 16;
        timestamp : 32;
        SSRC : 32;
    }
}
header rtp_t rtp;
...
```

```
...
parser parse_ethernet {
    extract(ethernet);
    return select(latest.etherType)
{
    ETHERTYPE_IPV4 : parse_ipv4;
    default: ingress;
}
}
parser parse_udp {
    extract(udp);
    return parse_rtp;
}
...
parser parse_rtp {
    extract(rtp);
    return ingress;
}
...
}
```

Add action

main.py

Exit Set SDE path Add/Update switch Add/Update VM Clean database Save Open

Actions

New action name: take_video(dstIP)

```
modify_field(ipv4.dstAddr,dstIP);
```

Delete action Add action

Metadata Headers

- gre_t
- hash_metadata_t
- i2e_metadata_t
- icmp_t
- igmp_t
- ingress_intrinsic_metadata_for_mirro
- ingress_intrinsic_metadata_for_tm_t
- ingress_intrinsic_metadata_from_par
- ingress_intrinsic_metadata_t
- ingress_metadata_t
- ingress_parser_control_signals
- intrinsic_metadata_t
- ipv4_metadata_t
- ipv4_option_32b_t
- ipv4_t**

Variables name

- inner_ipv4
- ipv4

Fields

- diffserv
- dstAddr
- flags
- fragOffset
- hdrChecksum
- identification
- ihl
- protocol
- srcAddr
- totalLen
- ttl
- version

Tables

New table name:

Delete table Add table

Available Actions

- acl_deny
- acl_mirror
- acl_permit
- acl_redirect_ecmp
- acl_redirect_nexthop
- acl_stats_update
- compute_lkp_ipv4_hash
- compute_lkp_ipv6_hash
- compute_lkp_non_ip_hash
- compute_other_hashes
- copy_to_cpu
- copy_to_cpu_with_reason
- decap_genv_inner_ipv4
- decap_genv_inner_ipv6
- decap_genv_inner_non_ip

User defined tables

Disclaimer

Adding a table

main.py

Exit Set SDE path Add/Update switch Add/Update VM Clean database Save Open

Actions

New action name:

Metadata Headers Variables name Fields

```
modify_field(ipv4.dstAddr,dstIP);
```

```
acl_metadata_t
egress_intrinsic_metadata_for_mirror
egress_intrinsic_metadata_for_output
egress_intrinsic_metadata_from_pars
egress_intrinsic_metadata_t
egress_metadata_t
erspan_header_t3_t
ethernet_t
fabric_header_cpu_t
fabric_header_t
fabric_header_timestamp_t
fabric_metadata_t
fabric_payload_header_t
genv_t
global_config_metadata_t
```

Delete action Add action

Tables

New table name:

Available Actions

User defined tables

```
reads {
  ipv4.dstAddr : exact;
  rtp.timestamp : range;
}
actions {
  take_video;
  drop;
}
size : 16384;
```

```
terminate_pw
terminate_tunnel_inner_ethernet_ipv4
terminate_tunnel_inner_ethernet_ipv6
terminate_tunnel_inner_ipv4
terminate_tunnel_inner_ipv6
terminate_tunnel_inner_non_ip
terminate_vppls
tunnel_check_pass
tunnel_lookup_miss
unicast_replica_from_rid
update_ingress_bd_stats
urpf_bd_miss
urpf_miss
USER ACTIONS:
take_video(dstIP)
```

Delete table Add table

Disclaimer

Define control flow

main.py

Exit Set SDE path Add/Update switch Add/Update VM Clean database Save Open

Ingress Pipeline beginning

```
apply(shedule_table);
```

Delete control Add to pipeline

Controls

- egress
- ingress
- USER Controls:

Tables

- acl_stats
- adjust_lkp_fields
- bd_flood
- capture_tstamp
- compute_ipv4_hashes
- compute_ipv6_hashes
- compute_non_ip_hashes
- compute_other_hashes
- cpu_packet_transform
- dmac
- drop_stats
- ecmp_group
- egress_bd_map
- egress_bd_stats
- egress_outer_bd_map
- egress_port_mapping
- egress_system_acl
- egress_vlan_xlate
- egress_vni
- fabric_ingress_det_lkp

Disclaimer

Compile
and upload to
switch

main.py

Exit Set SDE path Add/Update switch Add/Update VM Clean database Save Open

Switch address

Remote compilation

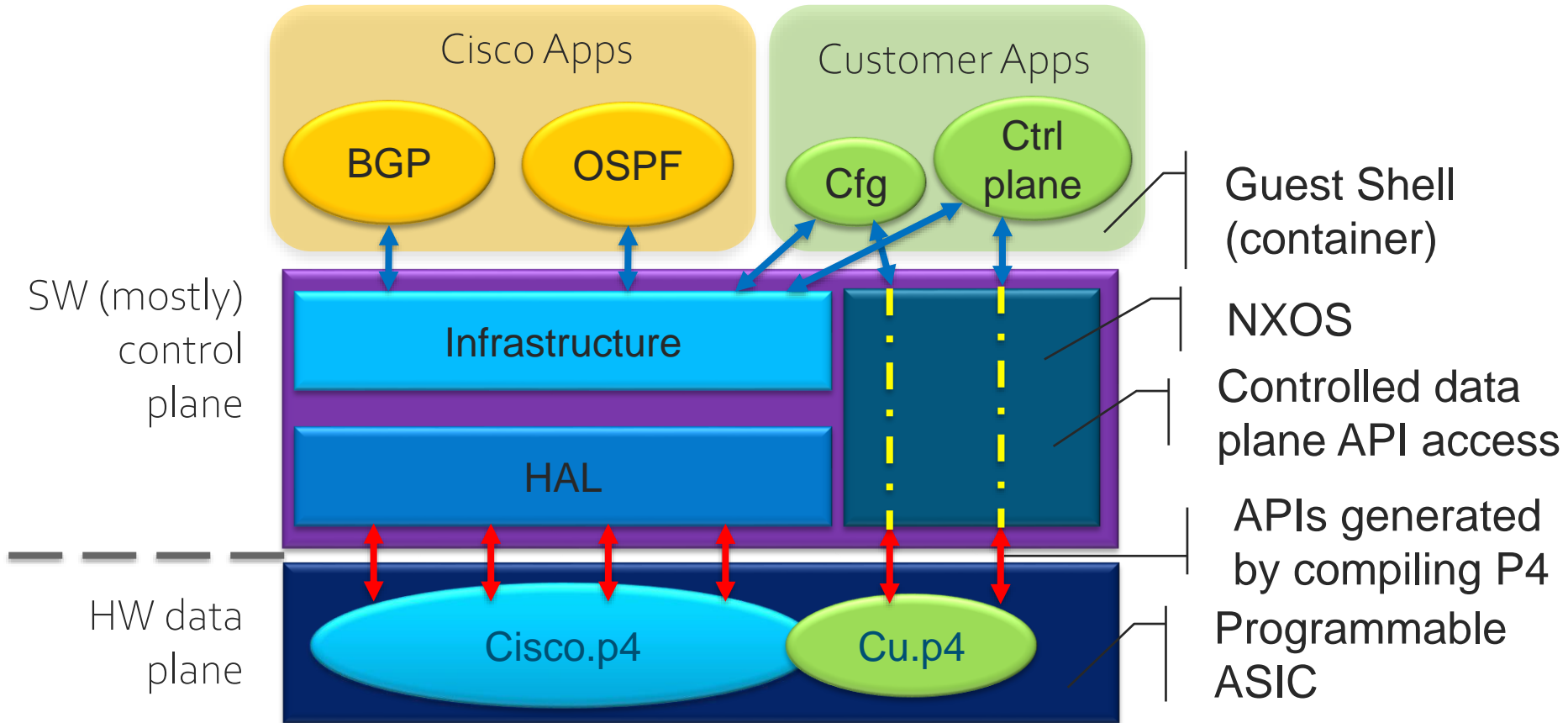
Remote IP address

Username

Password

Disclaimer

Control Plane and NetOS Support



Open Challenges

- On the customer side
 - Debugging
 - Access to the right level of knowledge on the stock P4 program
- On switching system vendor side
 - Support model
 - Troubleshooting issues
 - Identify whether related to stock code of customer code
- On programmable ASIC vendor side
 - Offer technical support directly to the end customer for chip/compiler related problems