Programmable Switch Hardware

ECE/CS598HPN

Radhika Mittal

Conventional SDN

- Programmable *control plane*.

- Data plane can support high bandwidth.
  - But has limited flexibility.

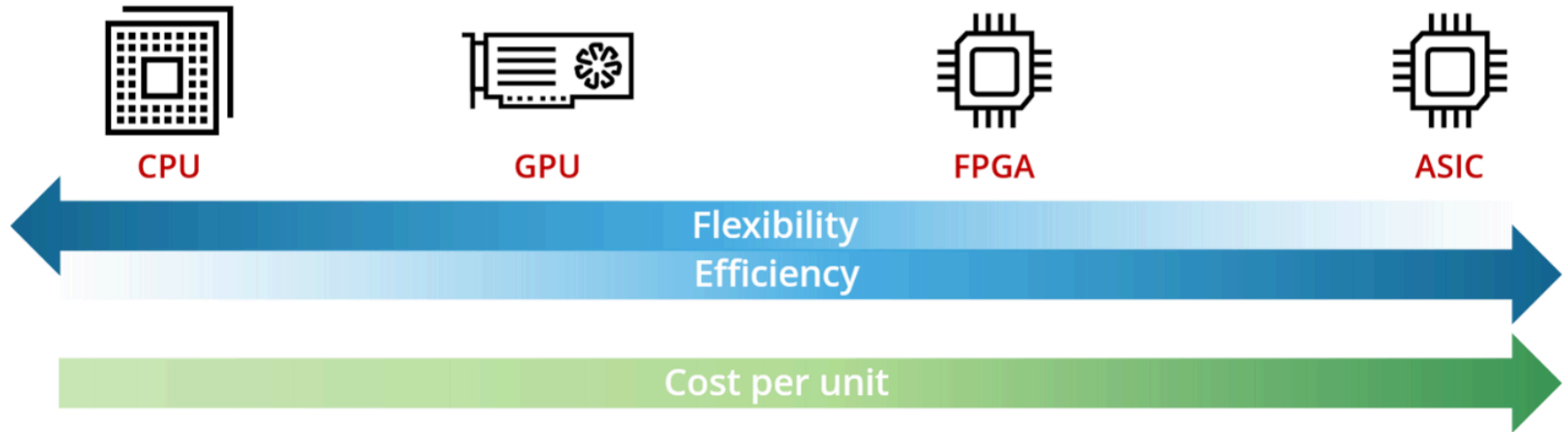- Restricted to conventional packet protocols.

Software Dataplane

- Very extensible and flexible.

- Extensive parallelization to meet performance requirements.
    - Might still be difficult to achieve 100's of Gbps.

- Significant cost and power overhead.

Programmable Hardware

- More flexible than conventional switch hardware.
- Less flexible than software switches.

- Slightly higher power and cost requirements than conventional switch hardware.
- Significantly lower than software switches.

Other alternatives?

| CPU | GPU | FPGA | ASIC |

← Flexibility

Efficiency →

Cost per unit →

Image copied from somewhere on the web.

Forwarding Metamorphosis: Fast Programmable Match-Action Processing in Hardware for SDN

Pat Bosshart, Glen Gibb, Hun-Seok Kim,
George Varghese, Nick McKeown, Martin Izzard,
Fernando Mujica, Mark Horowitz

What are the limitations of a fixed function switch?

Need for flexibility….

- Flexibility to:
    - Trade one memory size for another
    - Add a new table
    - Add a new header field
    - Add a different action

- SDN accentuates the need for flexibility
    - Gives programmatic control to control plane, expects to be able to use flexibility
    - OpenFlow designed to exploit flexibility.

What the Authors Set Out To Learn

- How to design a flexible switch chip?

- What does the flexibility cost?

RMT Switch Model

Enables flexibility through….

- Programmable parsing: support arbitrary header fields

- Ability to configure number, topology, width, and depths of match-tables.

- Programmable actions: allow a flexible set of actions (including arbitrary packet modifications).
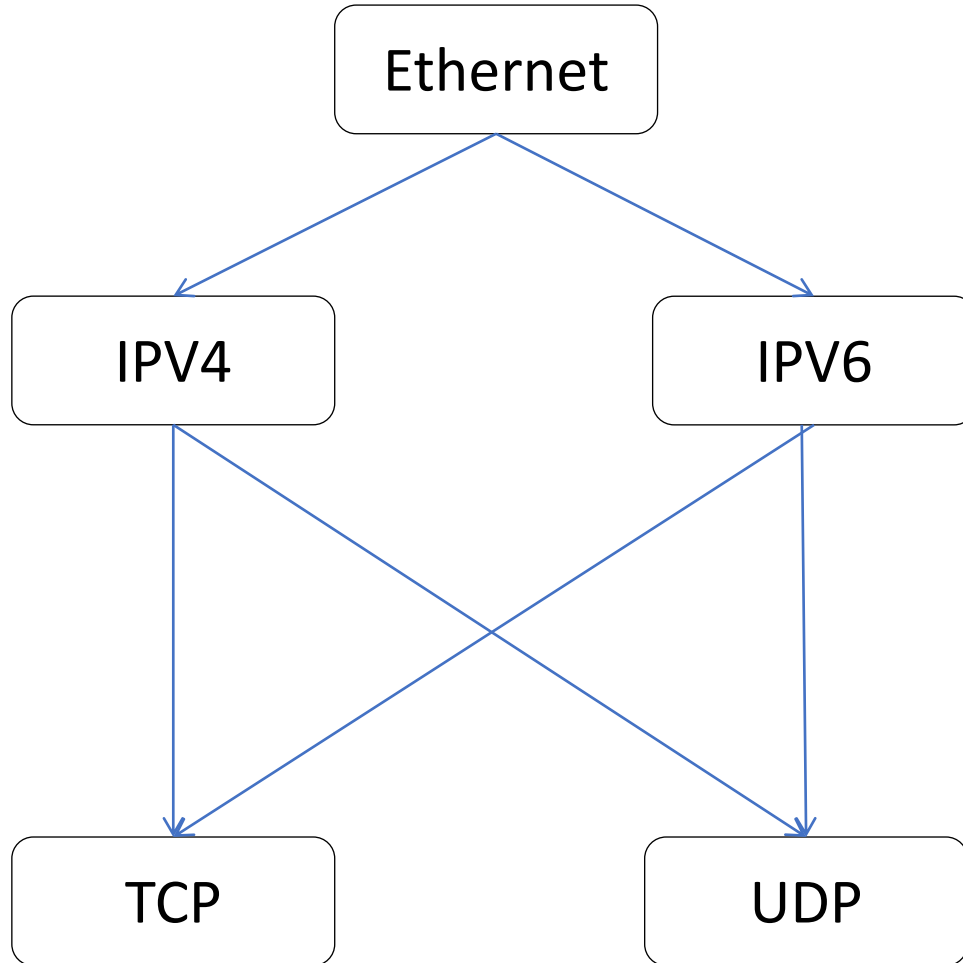
Design Considerations

- Chip size

- High frequency

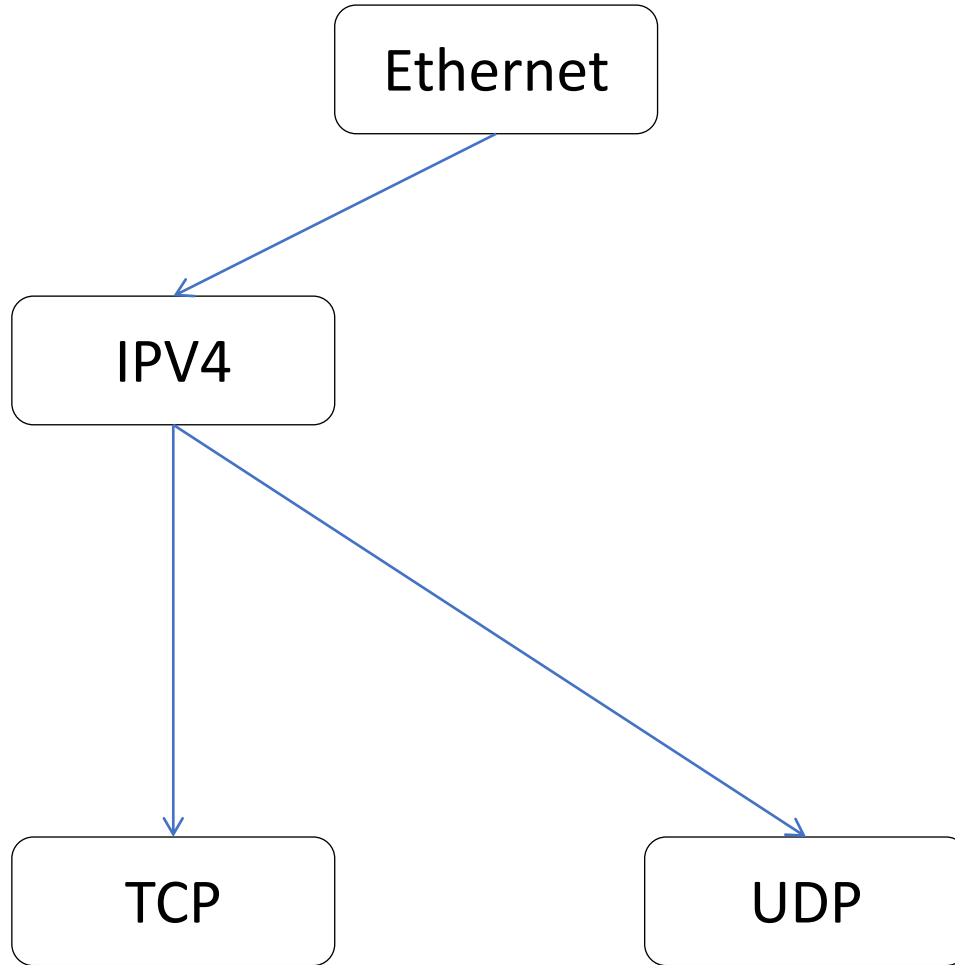- Wiring and crossbars

- Amount of memory

The RMT Abstract Model

- Parse graph
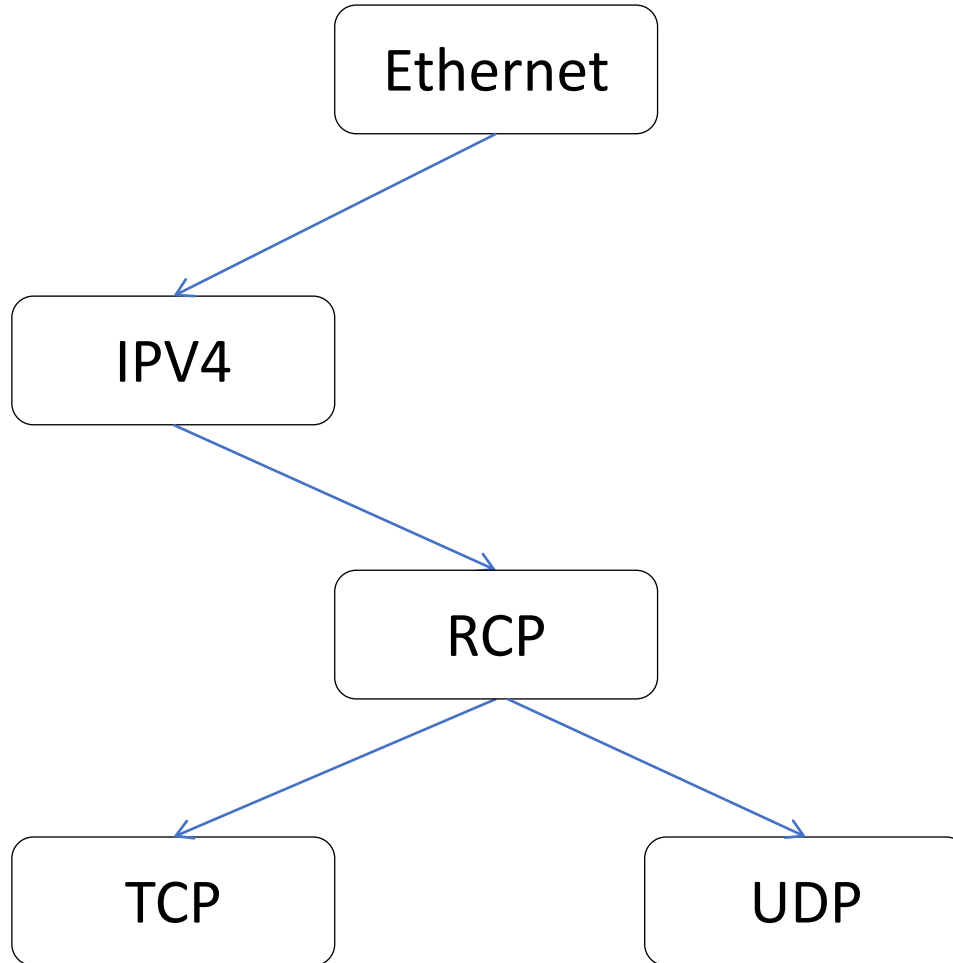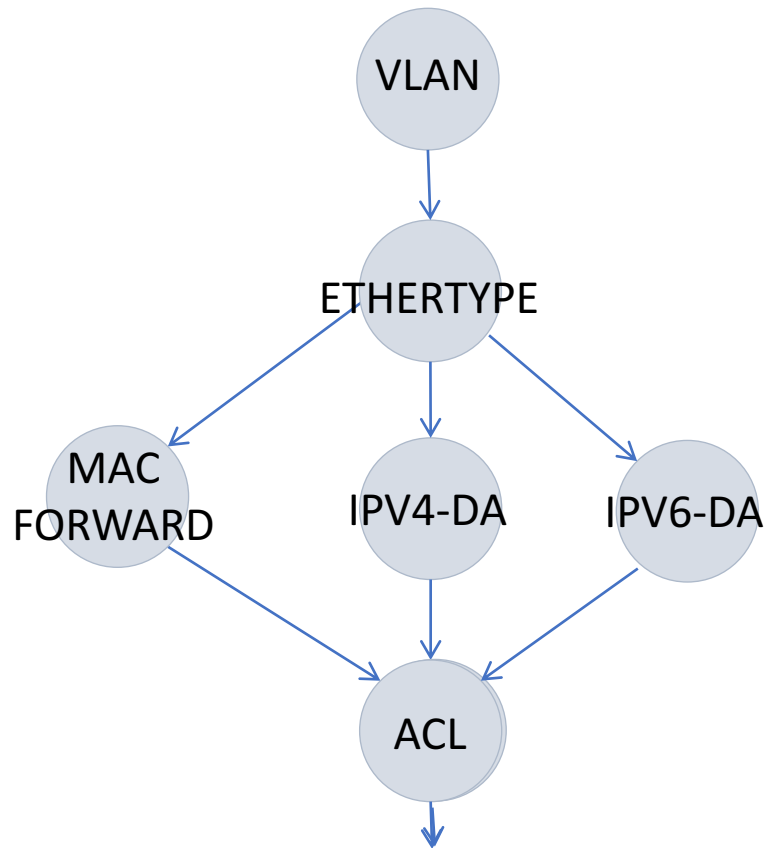- Table graph

Arbitrary Fields: The Parse Graph

Packet:	Ethernet	IPV4	TCP

Arbitrary Fields: The Parse Graph

Packet:

Ethernet	IPV4	TCP

Ethernet

IPV4

TCP

UDP

Arbitrary Fields: The Parse Graph

Packet:	Ethernet	IPV4	RCP	TCP

```
                    Ethernet
                        |
                        v
                      IPV4
                        |
                        v
                      RCP
                     /    \
                    v      v
                  TCP      UDP
```
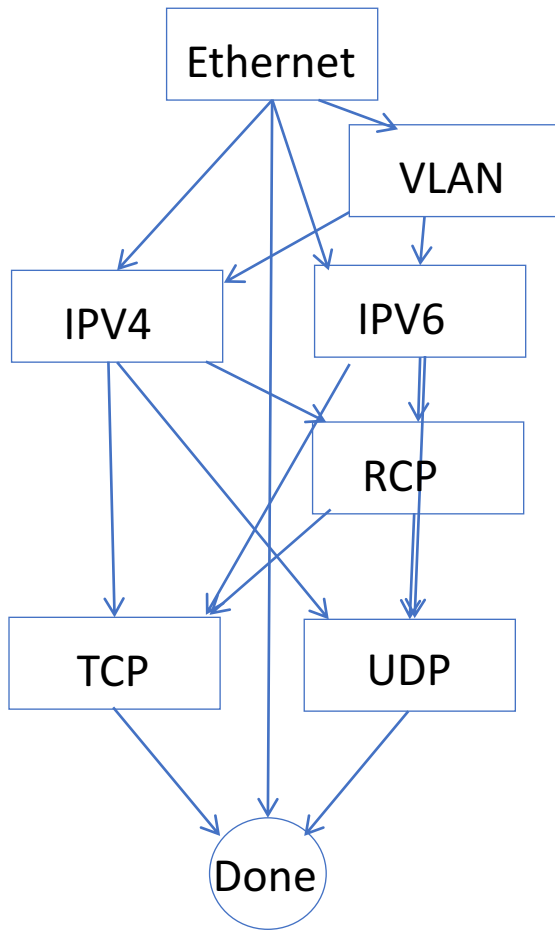
Reconfigurable Match Tables: The Table Graph

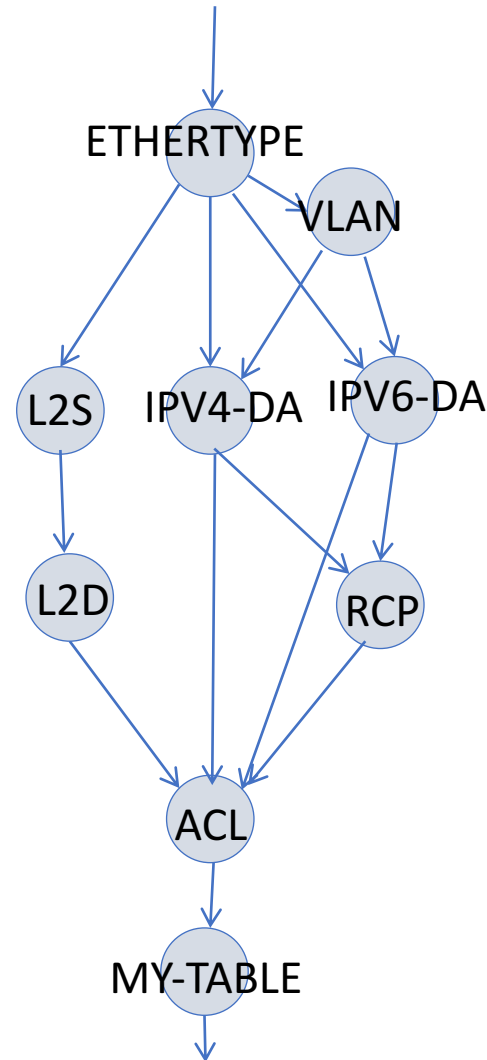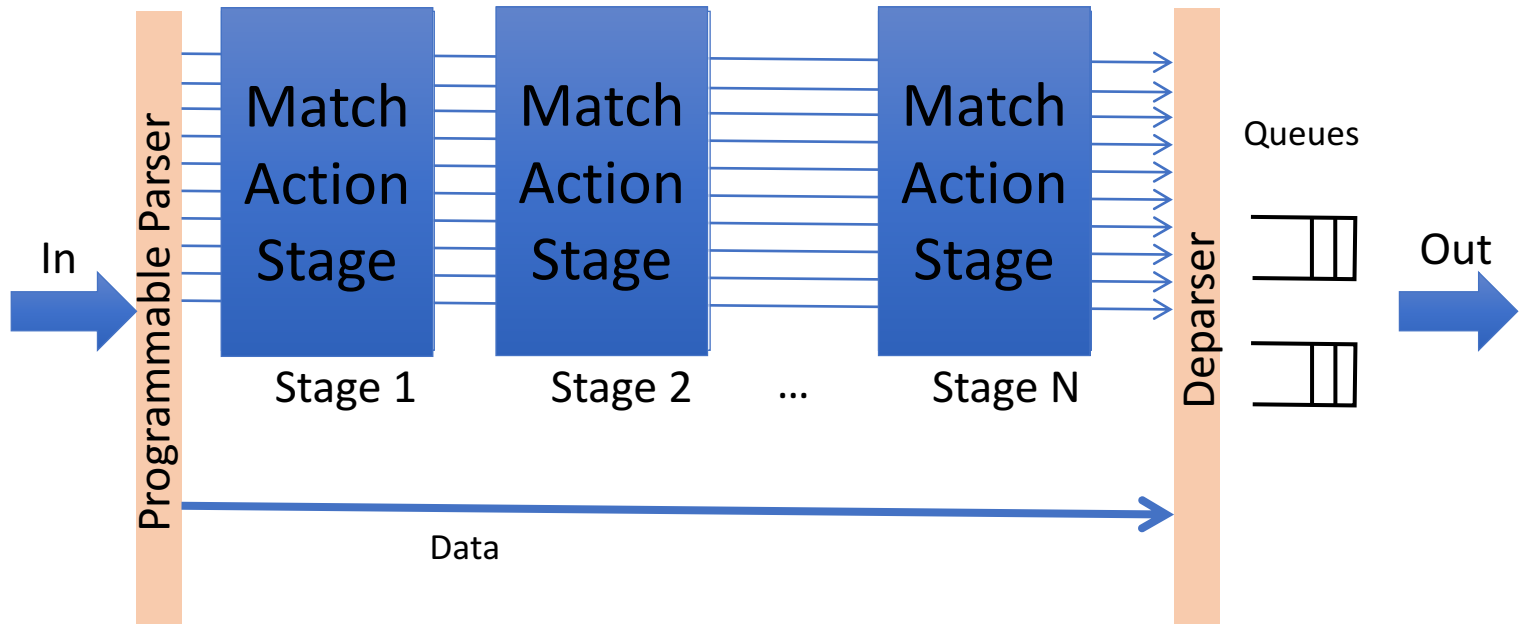# Changes to Parse Graph and Table Graph

Parse Graph

Table Graph

17

Other alternatives to RMT
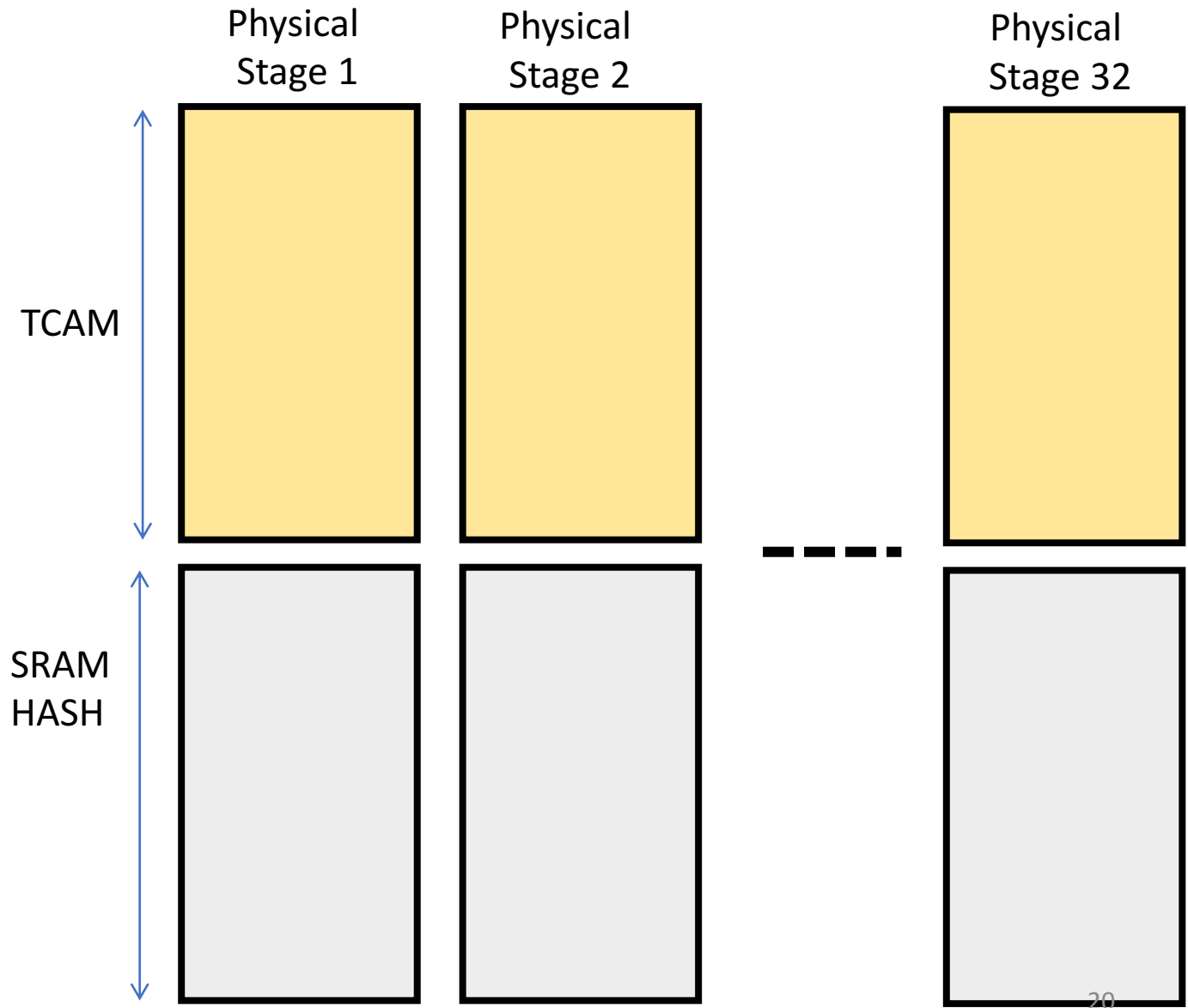
- SMT (Single Match-Action)

- MMT (Multiple Match-Action)

What are the limitations?

Match/Action Forwarding Model

In

Programmable Parser

Match Action Stage

Match Action Stage

...

Match Action Stage

Stage 1   Stage 2   ...   Stage N

Data

Deparser

Queues

Out

Logical to physical mapping

Physical Stage 1 Physical Stage 2 Physical Stage 32
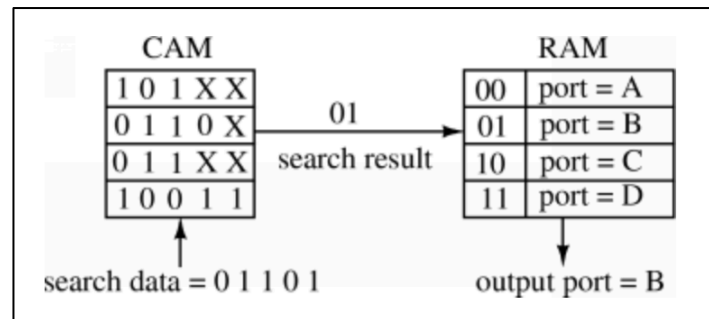
TCAM

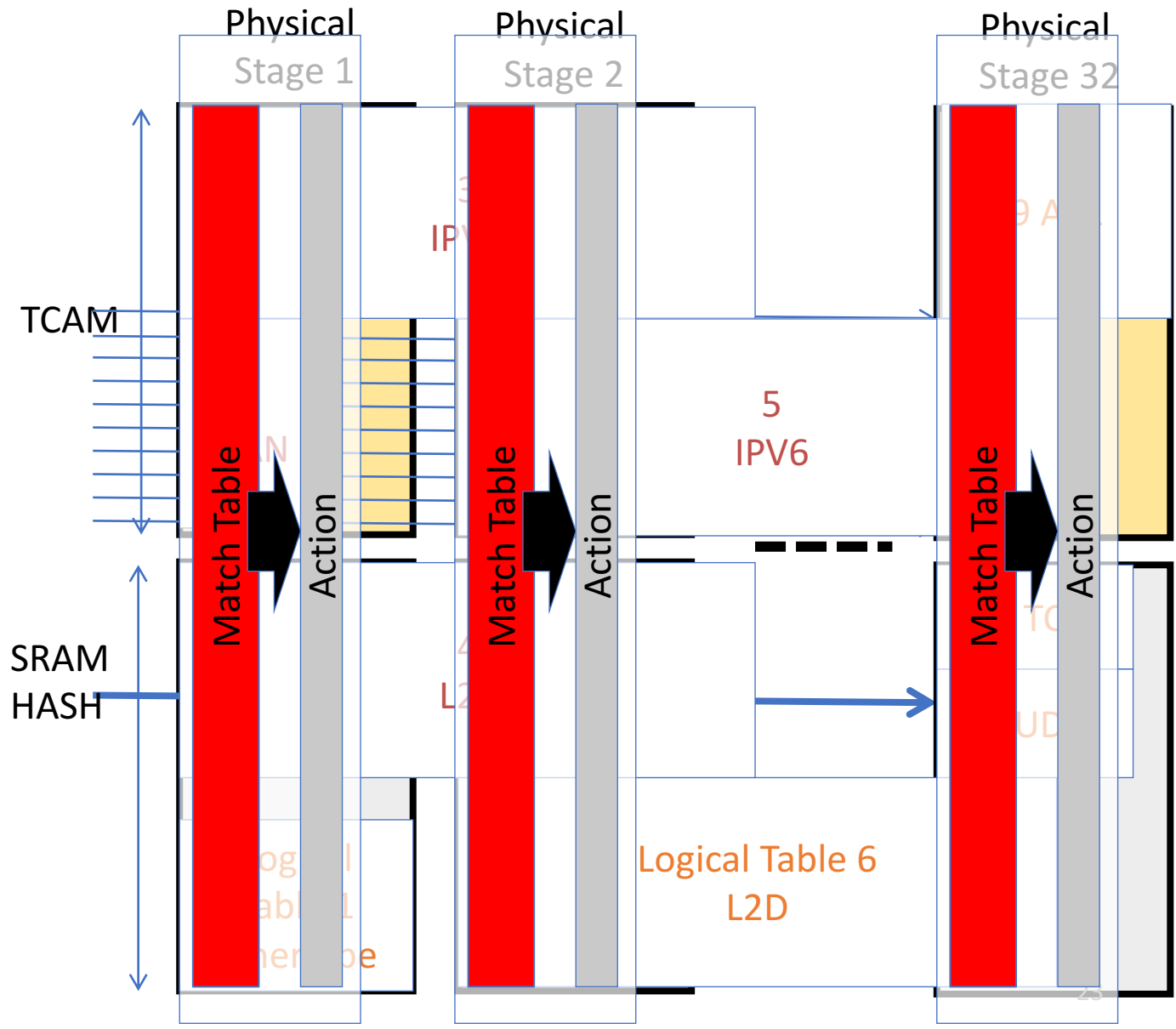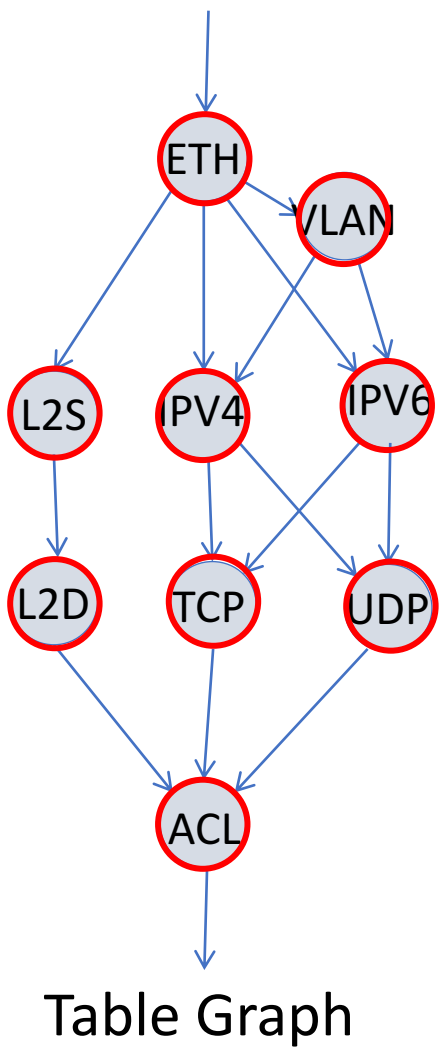SRAM HASH

Tiny Detour: CAMs and RAMs

- RAM:
  - Looks up the value associated with a memory address.

- CAM
  - Looks up memory address of a given value.
  - Two types:
    - Binary CAM: Exact match (matches on 0 or 1)
      - Can be implemented using SRAM.
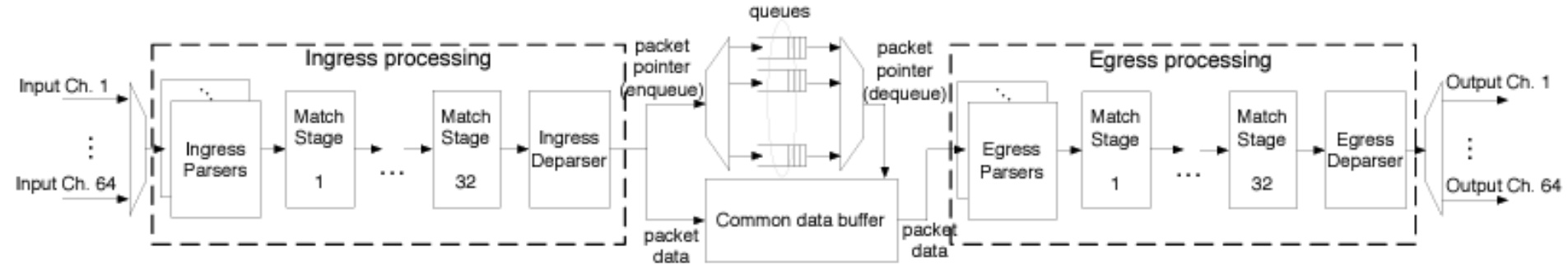   - Ternary CAM (TCAM): Allows wildcard (matches on 0, 1, or X).

Tiny Detour: CAMs

Line No.	Address (Binary)	Output Port
1	101XX	A
2	0110X	B
3	011XX	C
4	10011	D

CAM

| 1 0 1 X X |
| 0 1 1 0 X |
| 0 1 1 X X |
| 1 0 0 1 1 |

01 search result

RAM

00	port = A
01	port = B
10	port = C
11	port = D

search data = 0 1 1 0 1

output port = B

Logical to physical mapping

Table Graph

Physical Stage 1

Physical Stage 2

Physical Stage 32

TCAM

SRAM HASH

Match Table

Action

5
IPV6

Logical Table 6
L2D

Overall architecture

Configurable Parser

Arbitrary Fields: The Parse Graph

Packet: | Ethernet | IPV4 | TCP |
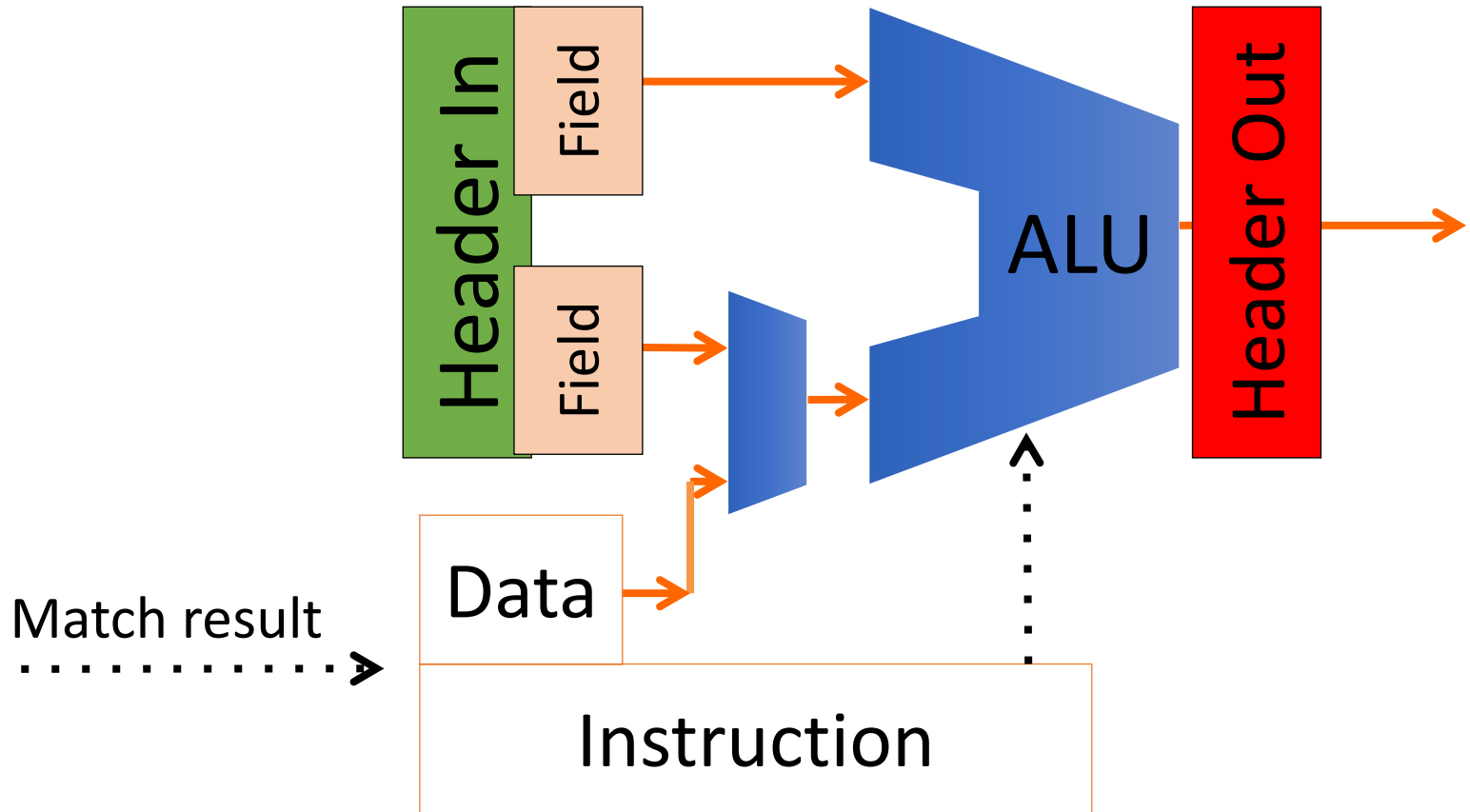
Ethernet

IPV4   IPV6
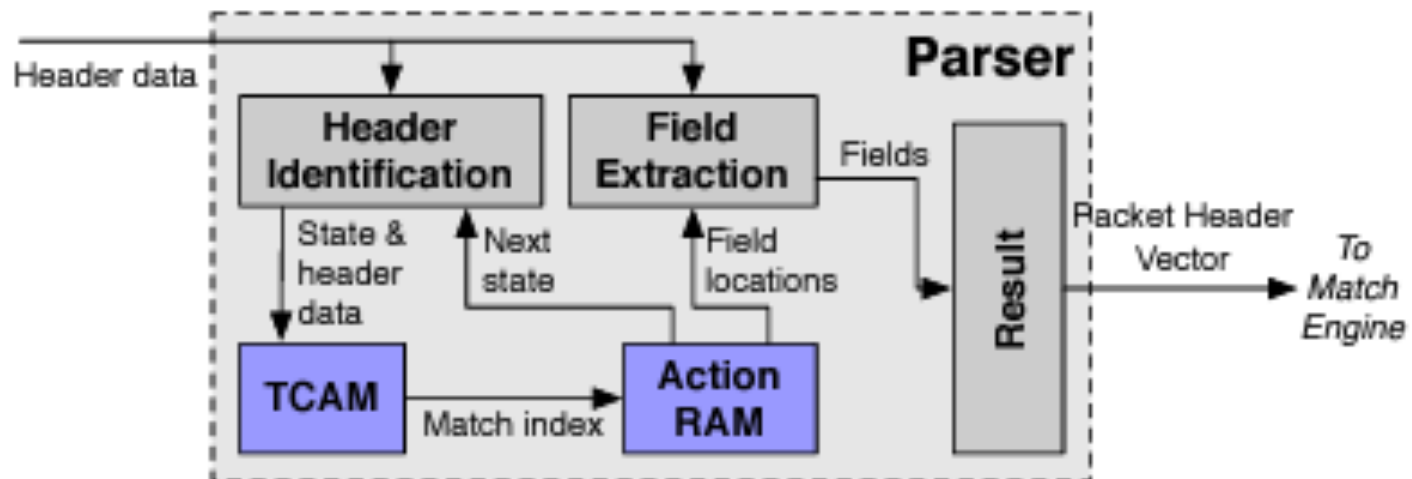
TCP   UDP
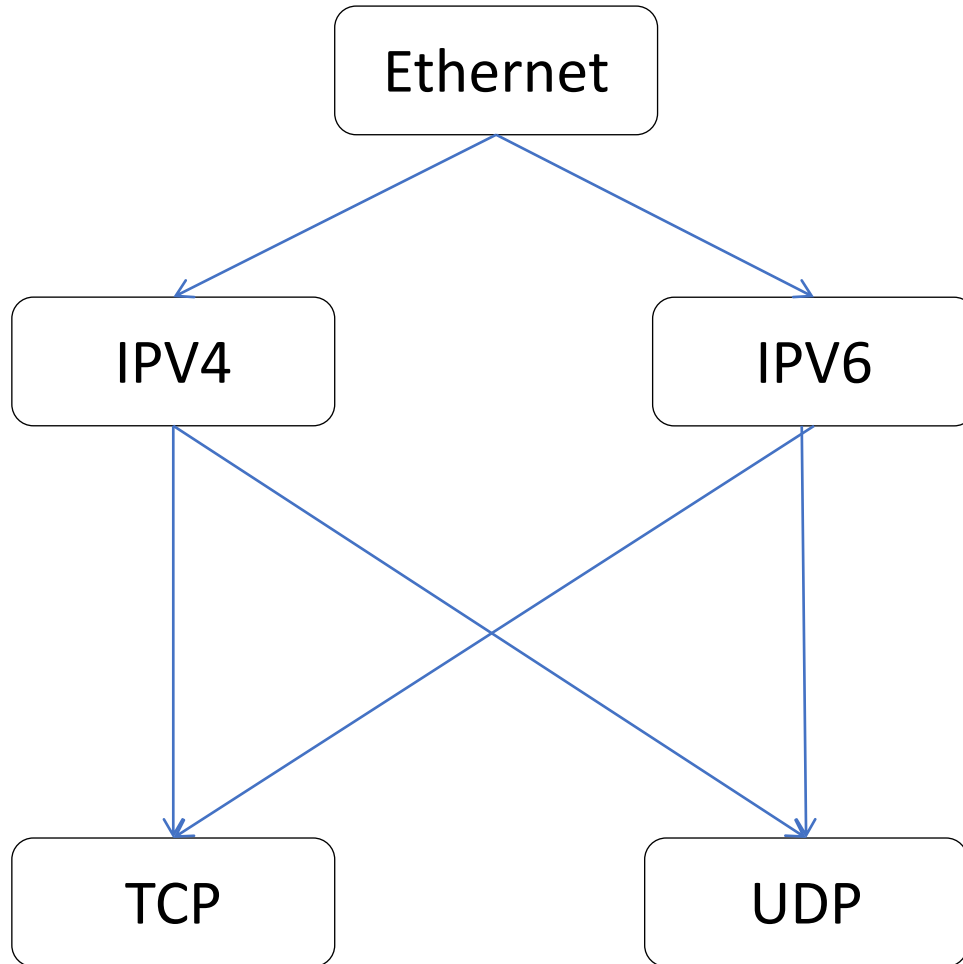
Action Processing Model

RMT Switch Design

- 64 x 10Gb ports
  - 960M packets/second
  - 1GHz pipeline
- Programmable parser
- 32 Match/action stages

- Huge TCAM: 10x current chips
  - 64K TCAM words x 640b
- SRAM hash tables for exact matches
  - 128K words x 640b
- 224 action processors per stage
- All OpenFlow statistics counters

Summary

- Conventional switch chip are inflexible
- SDN demands flexibility…sounds expensive…
- How do they do it: The RMT switch model
- *Flexibility costs less than 15%*

How is this paradigm different from active networking?

What are the limitations on flexibility?

Since 2013....

- 2013: RMT switch was commercialized (2013)
  - Barefoot Tofino
  - 6.5Tb/s

- 2017: Adoption at Alibaba, Baidu, Tencent, AT&T

- 2019: Barefoot acquired by Intel

- *Jan 2023: Intel stopped future investment in it;*
  - *(will continue to support existing products and customers).*

On research….

- Disaggregated RMT (SIGCOMM'17)
- Enabling stateful processing (HotNets'20)
- Trio (SIGCOMM'22) – Juniper's programmable chipset

- And many others….