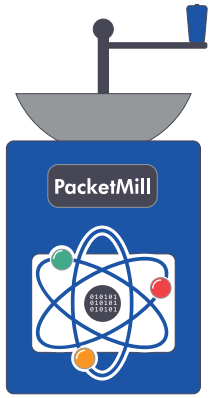
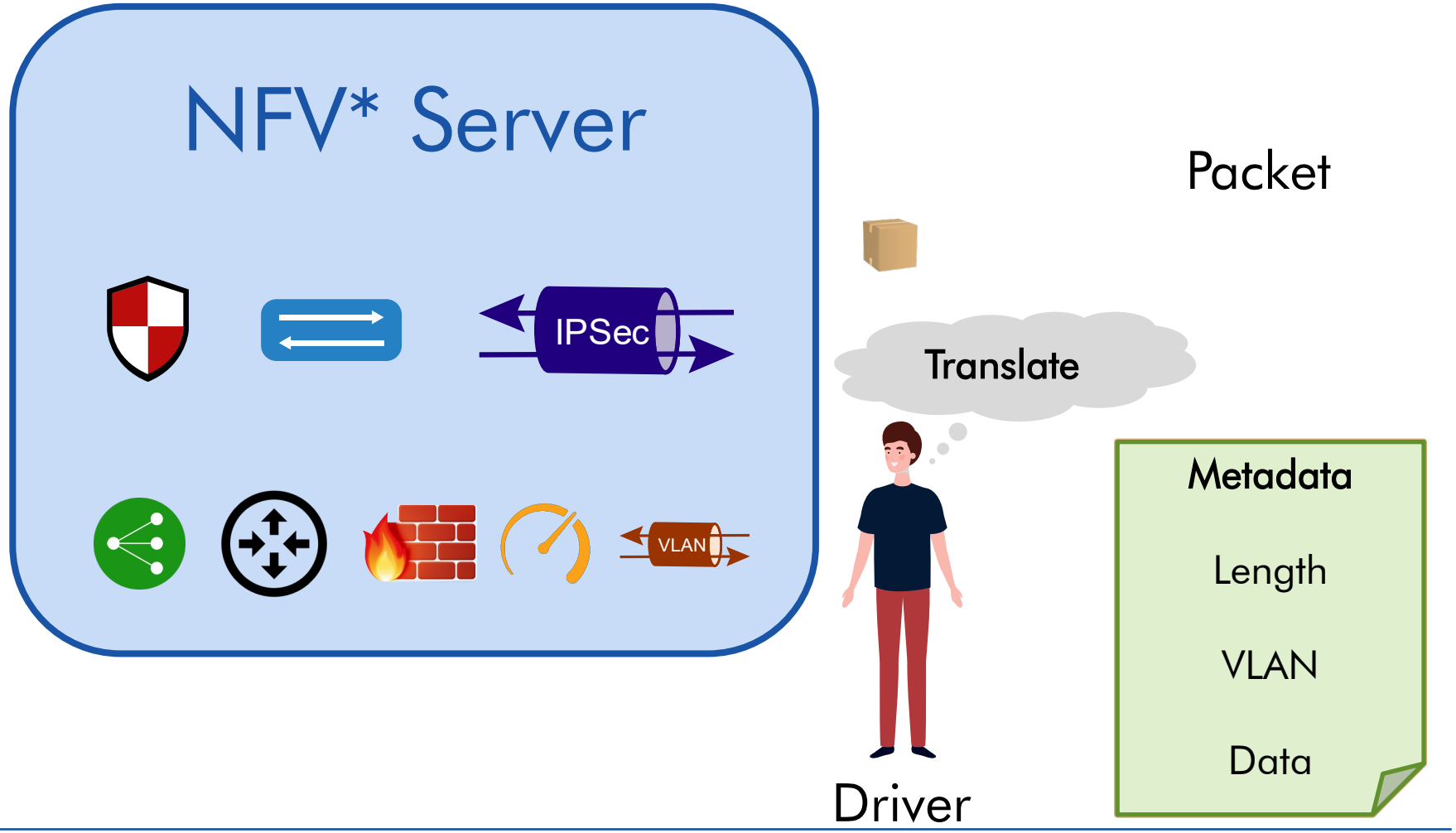


PacketMill: Toward Per-Core 100-Gbps Networking

Alireza Farshin*, Tom Barbette*, Amir Roozbeh*+, Gerald Q. Maguire Jr.*, Dejan Kostić*

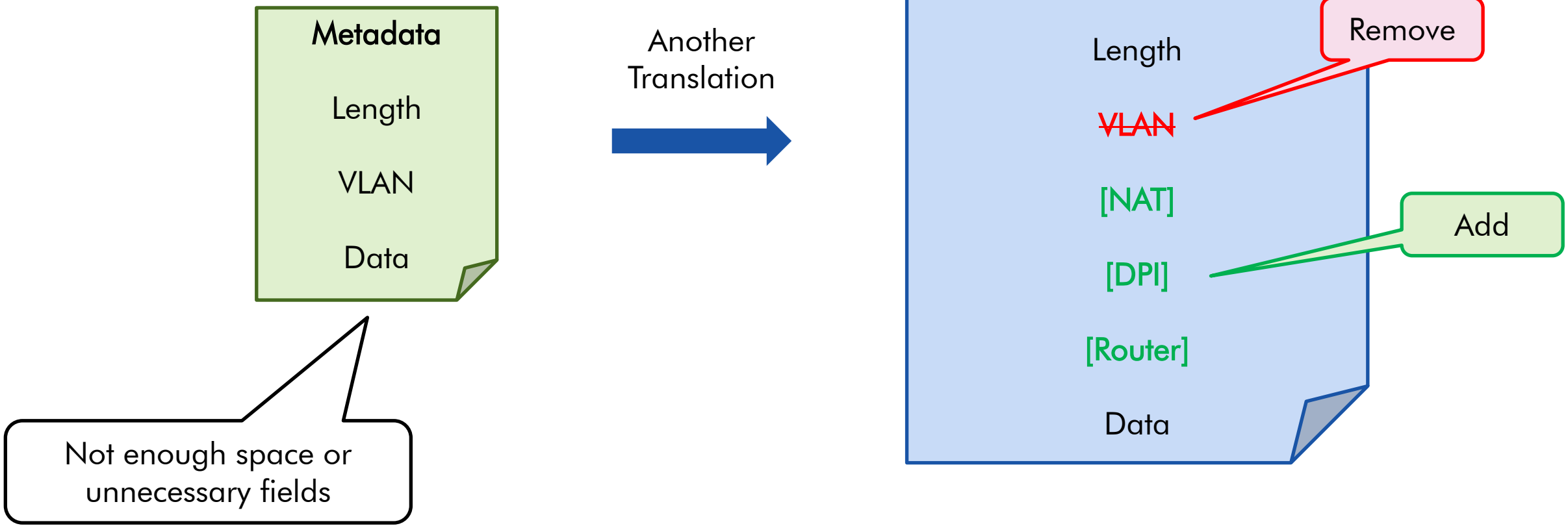
* KTH Royal Institute of Technology + Ericsson Research

A Story of Packet Delivery

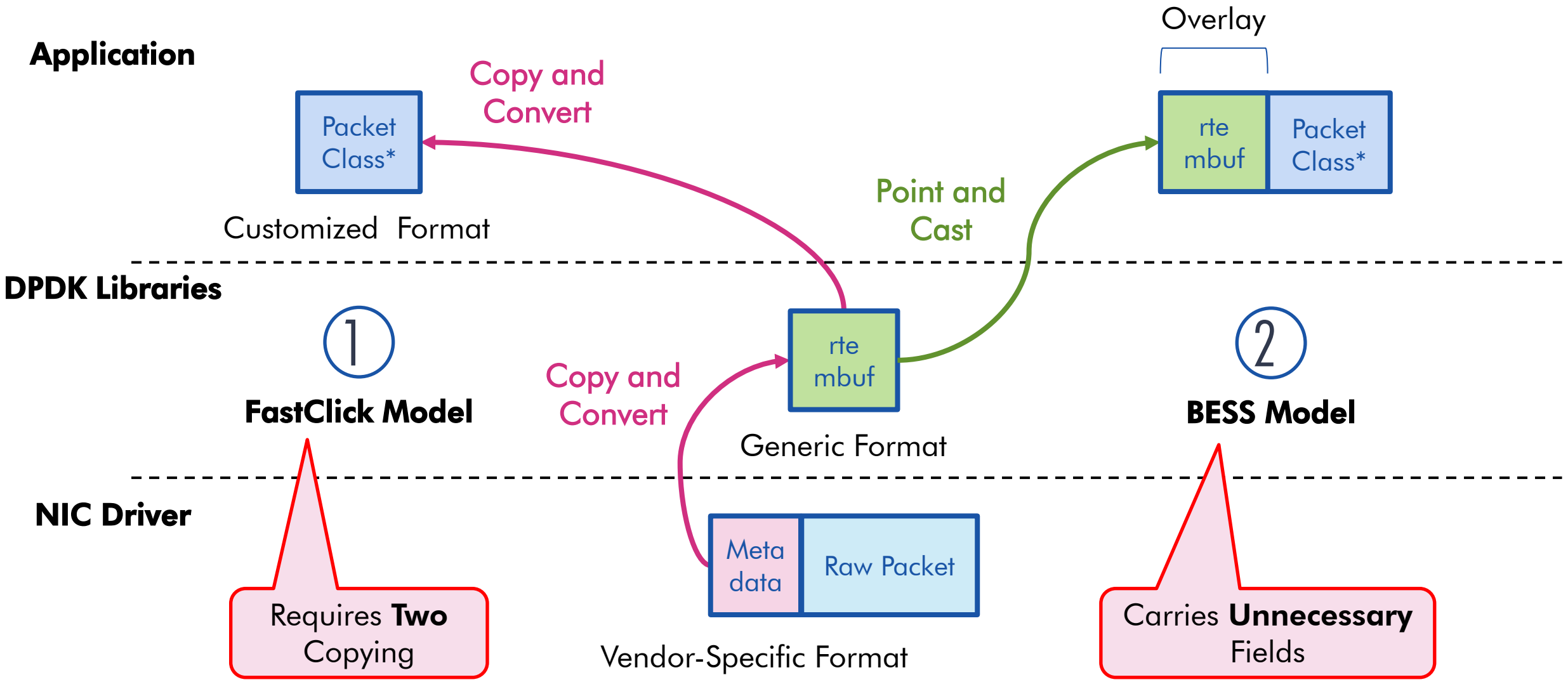


* Network Function Virtualizations (NFV)

Metadata Is not Specialized for the Network Function

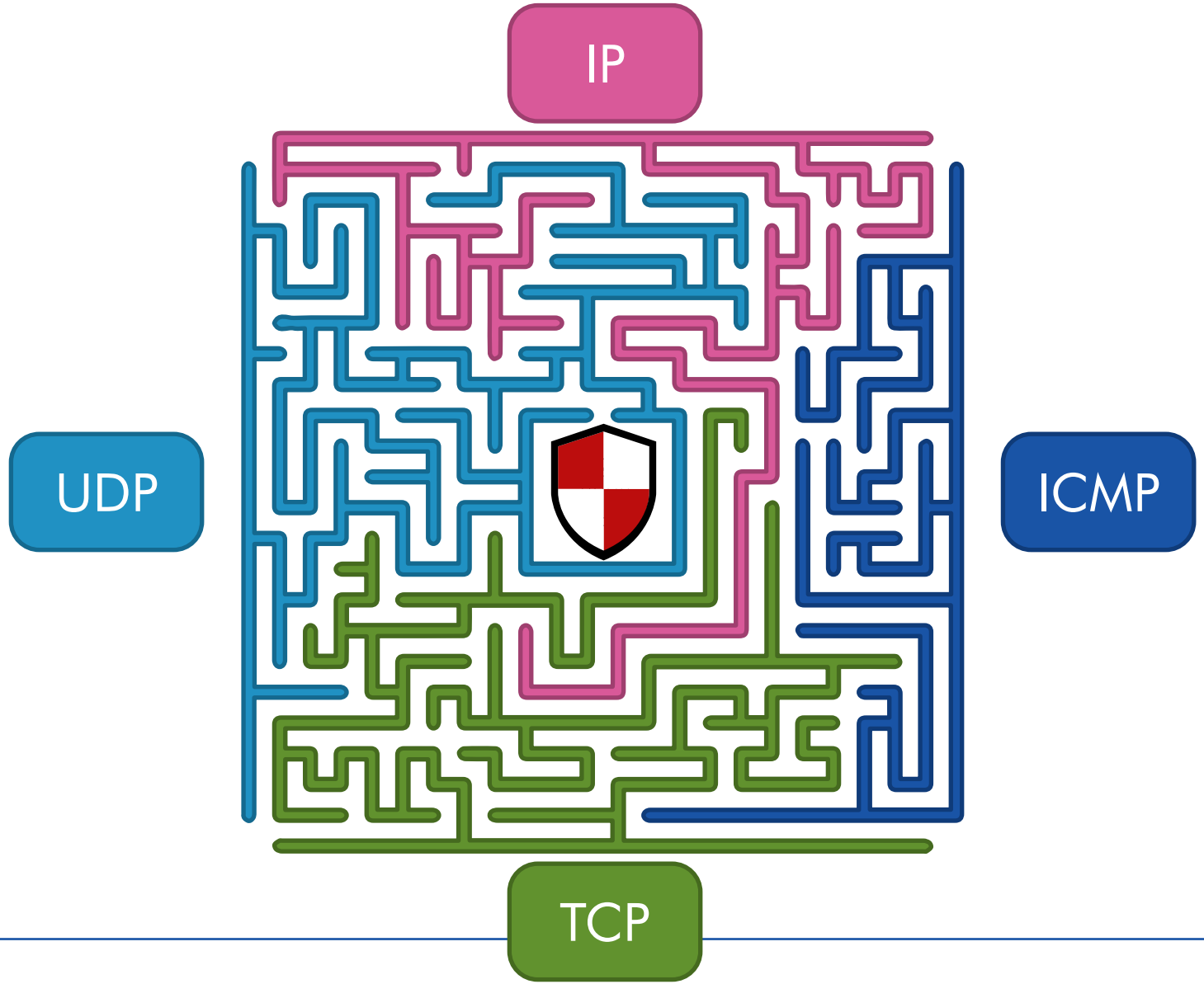


Inefficient Metadata Management

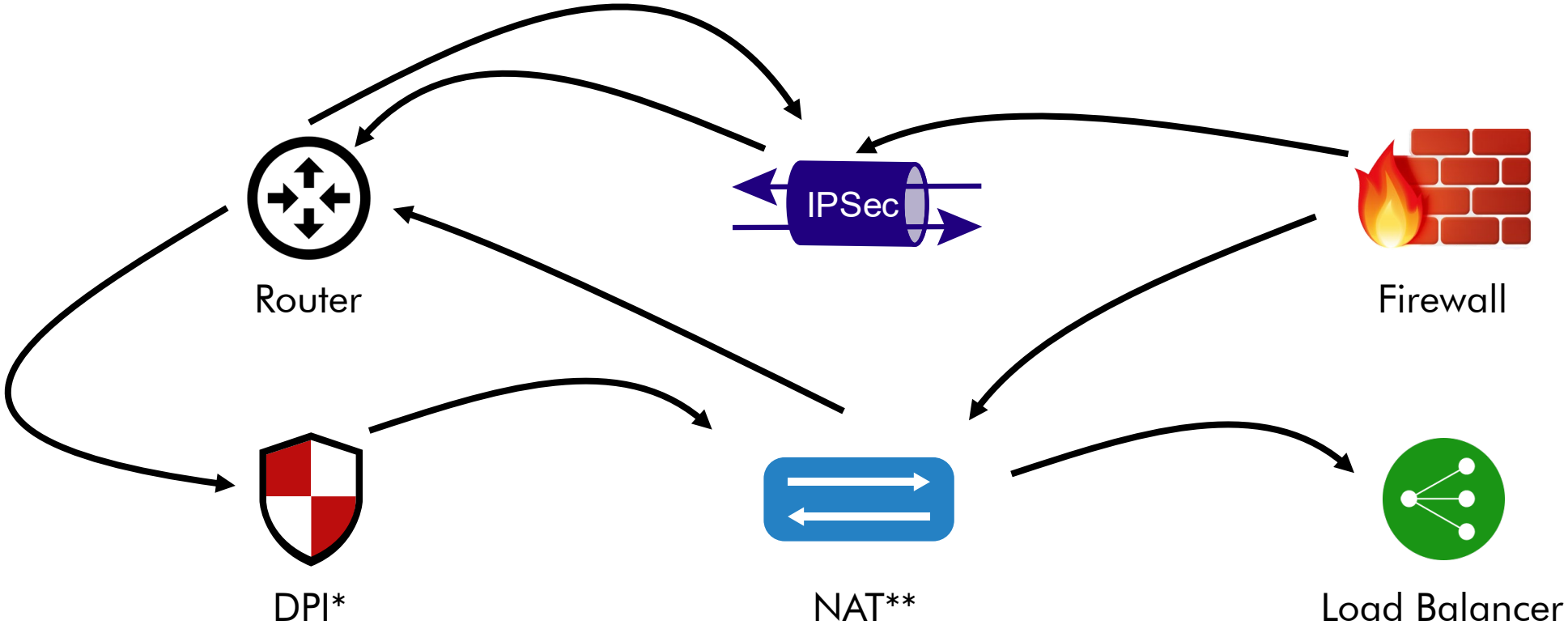


* An Example Data Structure (used in FastClick)

Network Functions are a maze



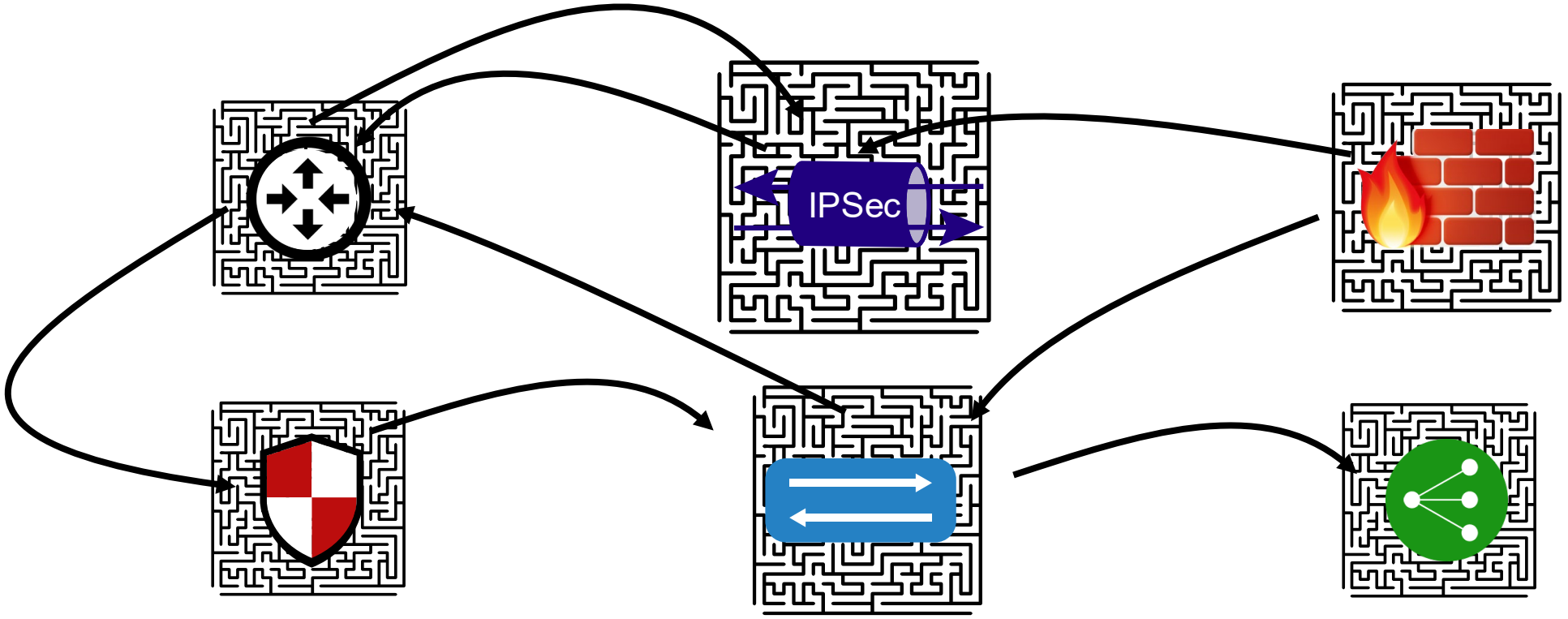
Modular Packet Processing Frameworks Implement a Chain of Network Functions

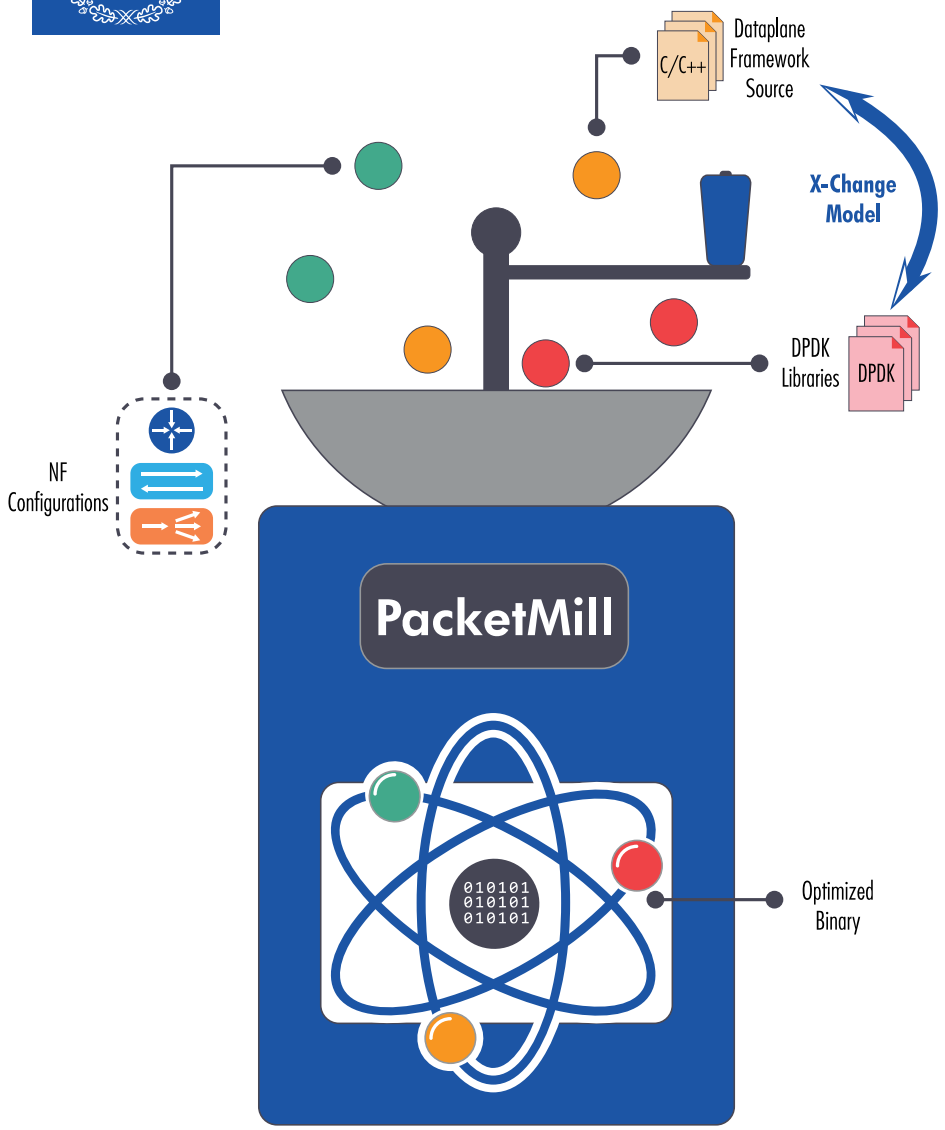


* Deep Packet Inspection (DPI)

** Network Address Translator (NAT)

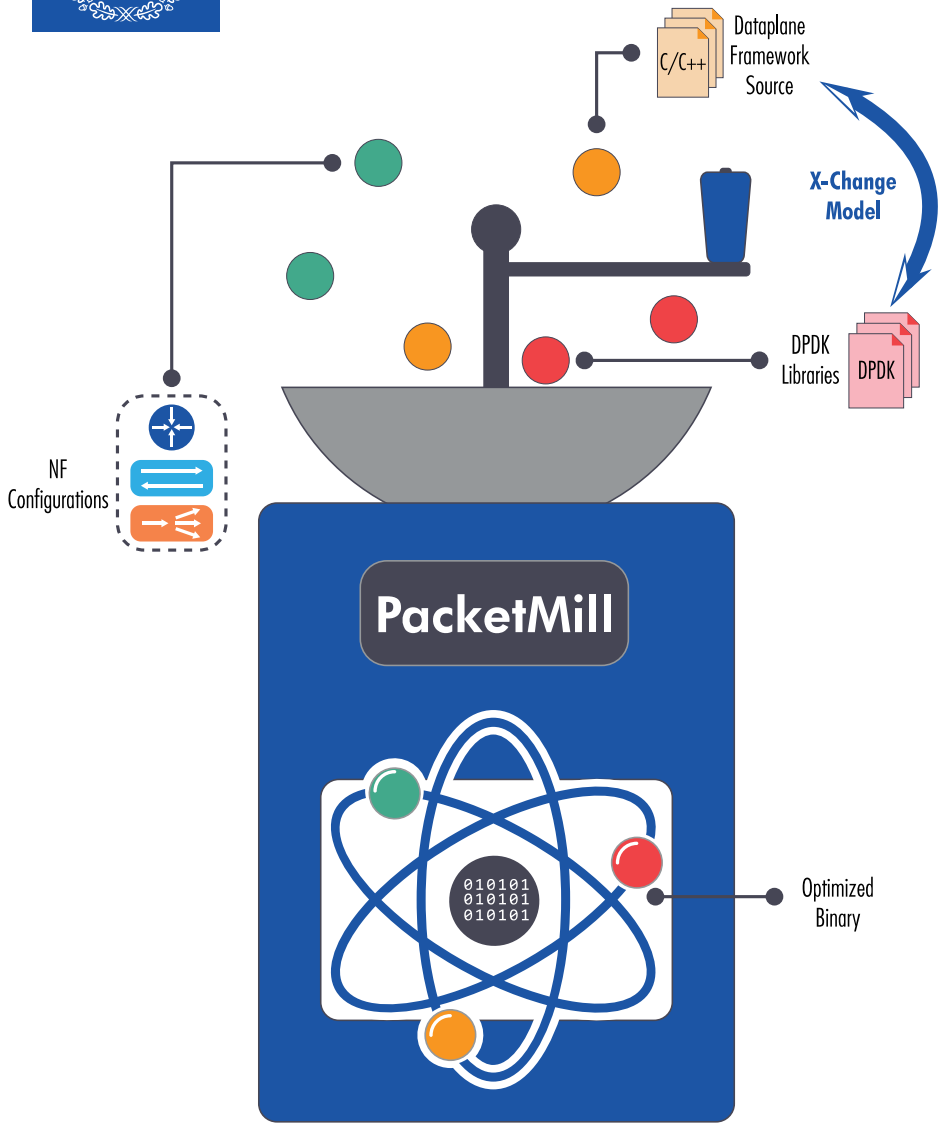
Modular Packet Processing Frameworks Create a Chain of Mazes





PacketMill mitigates these problems, increasing the performance and efficiency of the current software & hardware when processing packets

PacketMill



A metadata management model called **X-Change** that enables DPDK-based applications to use customized data structures instead of *rte_mbuf*

No need to translate

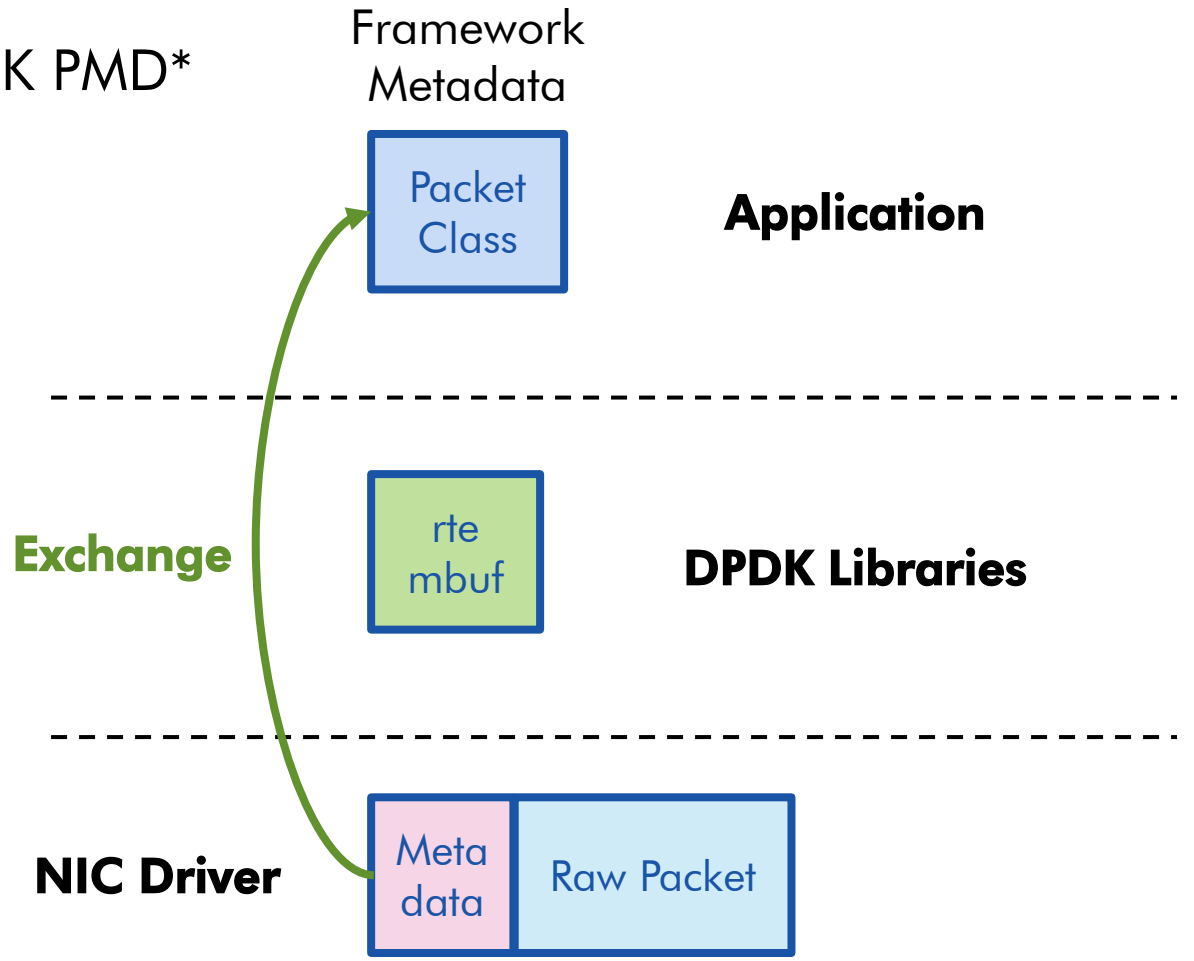
X–Change Prevents any Extra Operations

- Exchanges data structures with DPDK PMD*

Others:

- Uses fewer in-flight buffers
- Avoid allocating/releasing buffers

Check Our Paper



* Poll Mode Driver (PMD): user space DPDK drivers

X-Change Uses Conversion Functions

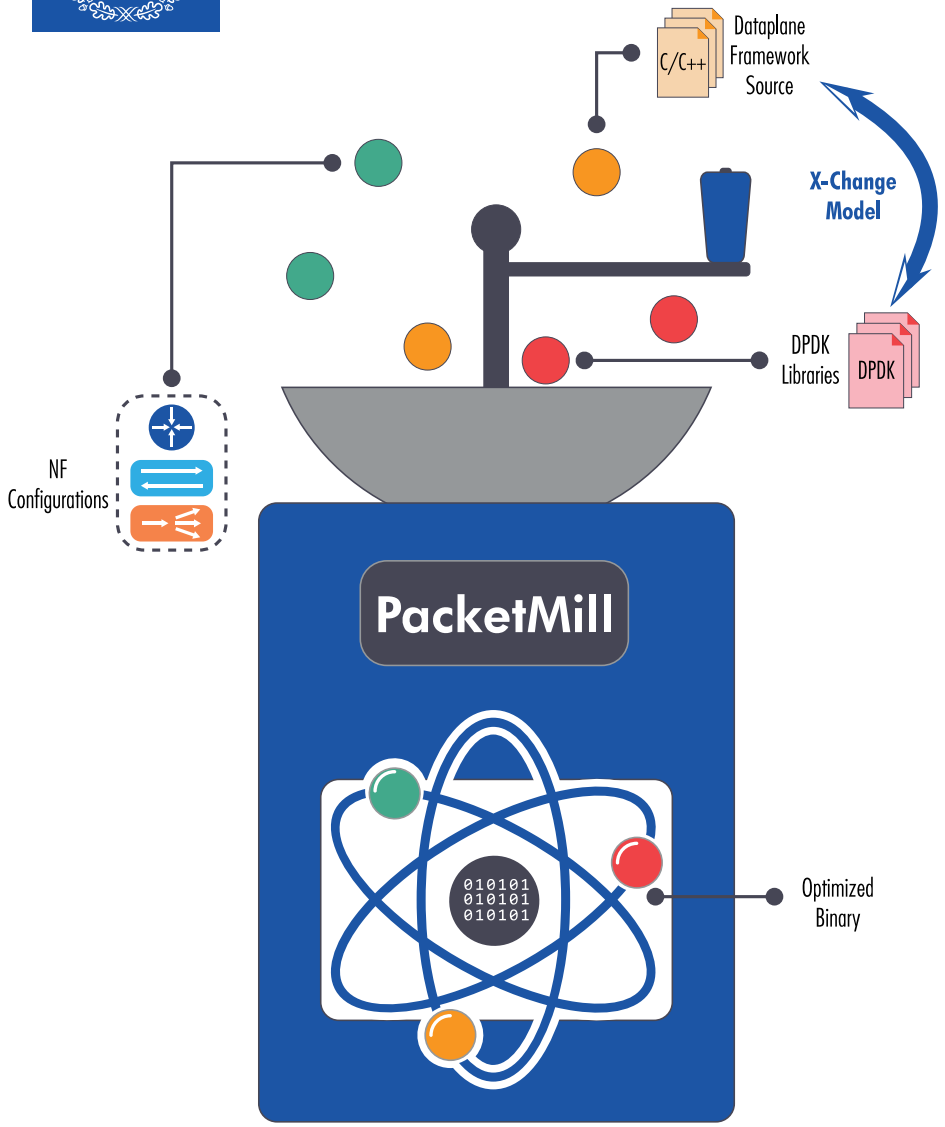
X-Change uses **LTO**** to inline the functions

```
/* X-Change Implementation for Default DPDK */
void xchg_set_vlan_tci(struct xchg* pkt, uint16_t vlan_tci) {
        ((struct rte_mbuf*)pkt)->vlan_tci = vlan_tci;
}
/* X-Change Implementation for Custom Buffers */
void xchg_set_vlan_tci(struct xchg* pkt, uint16_t vlan_tci)
{
        SET_VLAN_ANNO((Packet*)pkt, vlan_tci);
}
```


* Mellanox PMD

** Link Time Optimization (LTO)

PacketMill



A metadata management model called **X-Change** that enables DPDK-based applications to use customized data structures instead of *rte_mbuf*

No need to translate

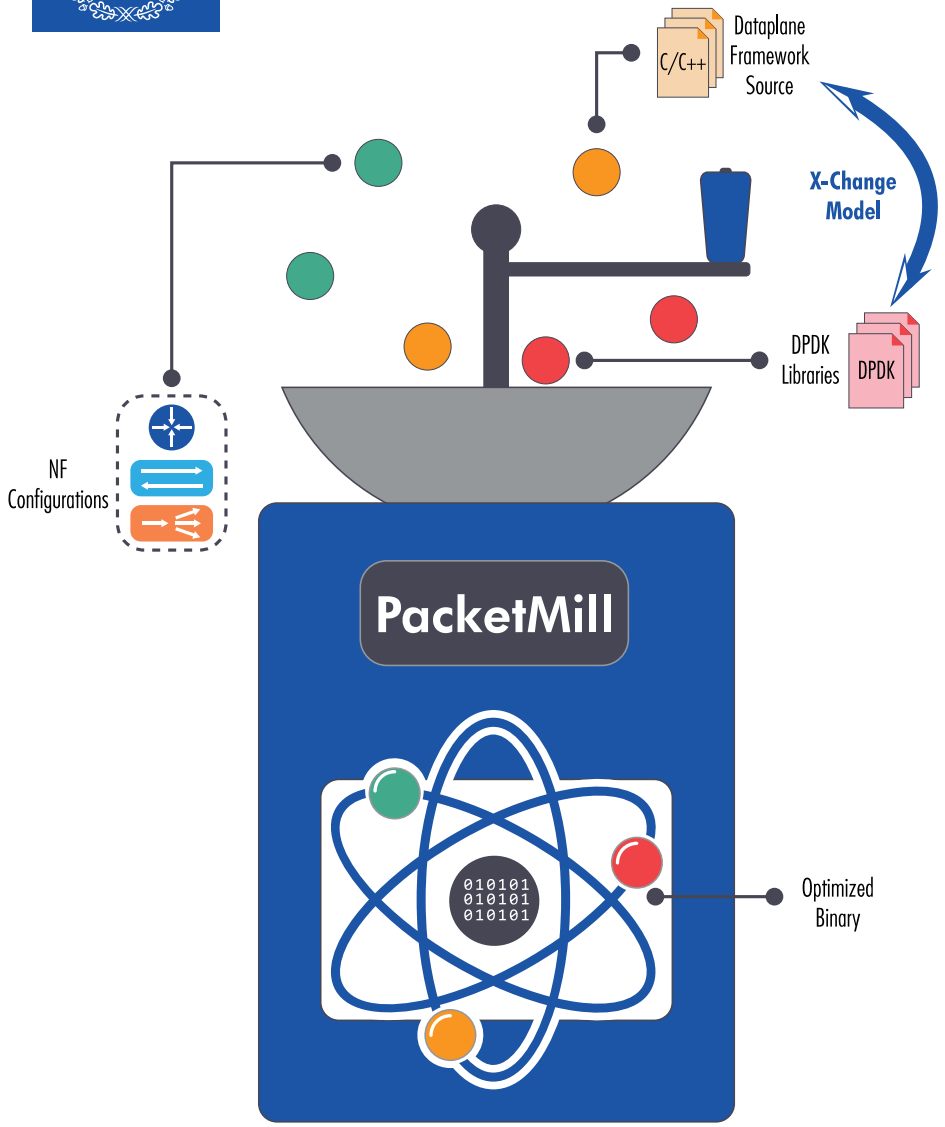
B Uses LLVM optimization passes to reorder data structures in the IR* (LLVM bit code)

Provides better order

- Tracks GetElementPtrInst (GEPI) Instructions
- Reorders the application-specific data structure
- Fix the GEPI Instructions

* Intermediate Representation (IR)

PacketMill



A A metadata management model called **X-Change** that enables DPDK-based applications to use customized data structures instead of *rte_mbuf*

No need to translate

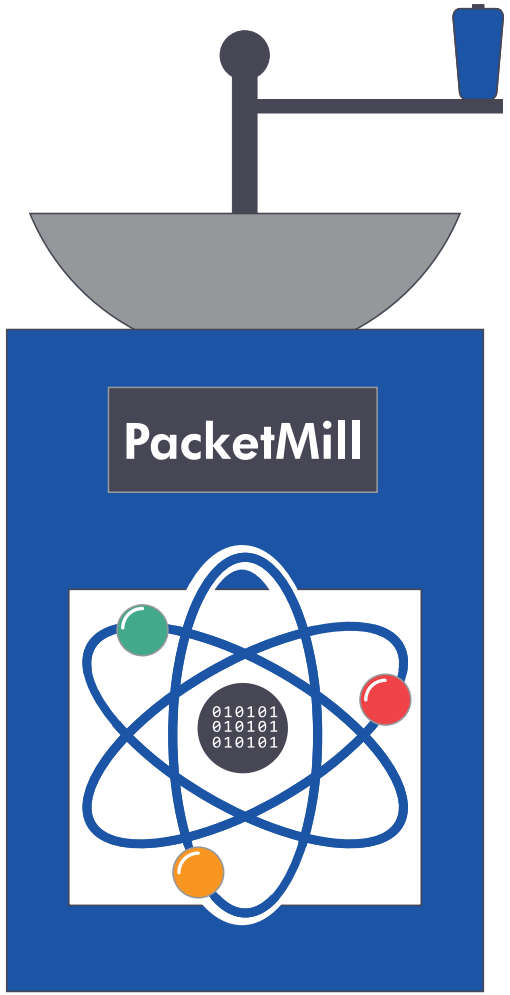
B Uses LLVM optimization passes to reorder data structures in the IR* (LLVM bit code)

Provides better order

C Uses/Embeds the available information in the configuration file to perform source-code modifications

Simplifies the maze

* Intermediate Representation (IR)



PacketMill generates a **customized** binary for a given chain of network function by performing whole-stack optimizations

Currently supports:

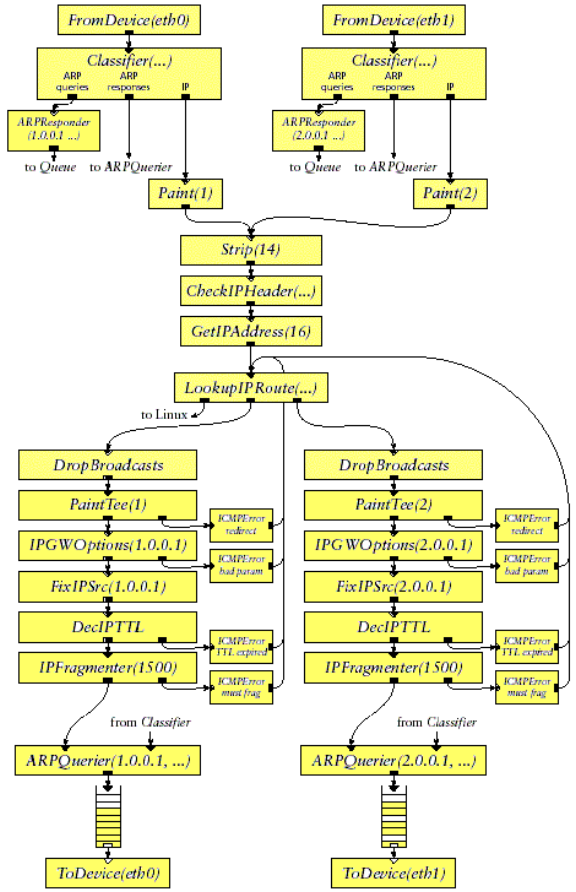
FastClick & Mellanox PMD (mlx5)

PacketMill Workflow

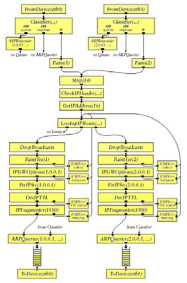
Describing the Network Function

Input Config File

Click-based Router



PacketMill Workflow



Describing the Network Function

Input Config File

Modifies Source Code

Embeds the Constants and the Graph

Fast Click Source

+

XCHG* + DPDK

Whole-Program IR

Compiles with LTO*

IR Bitcode

Modifies IR Bitcode

Opt. IR

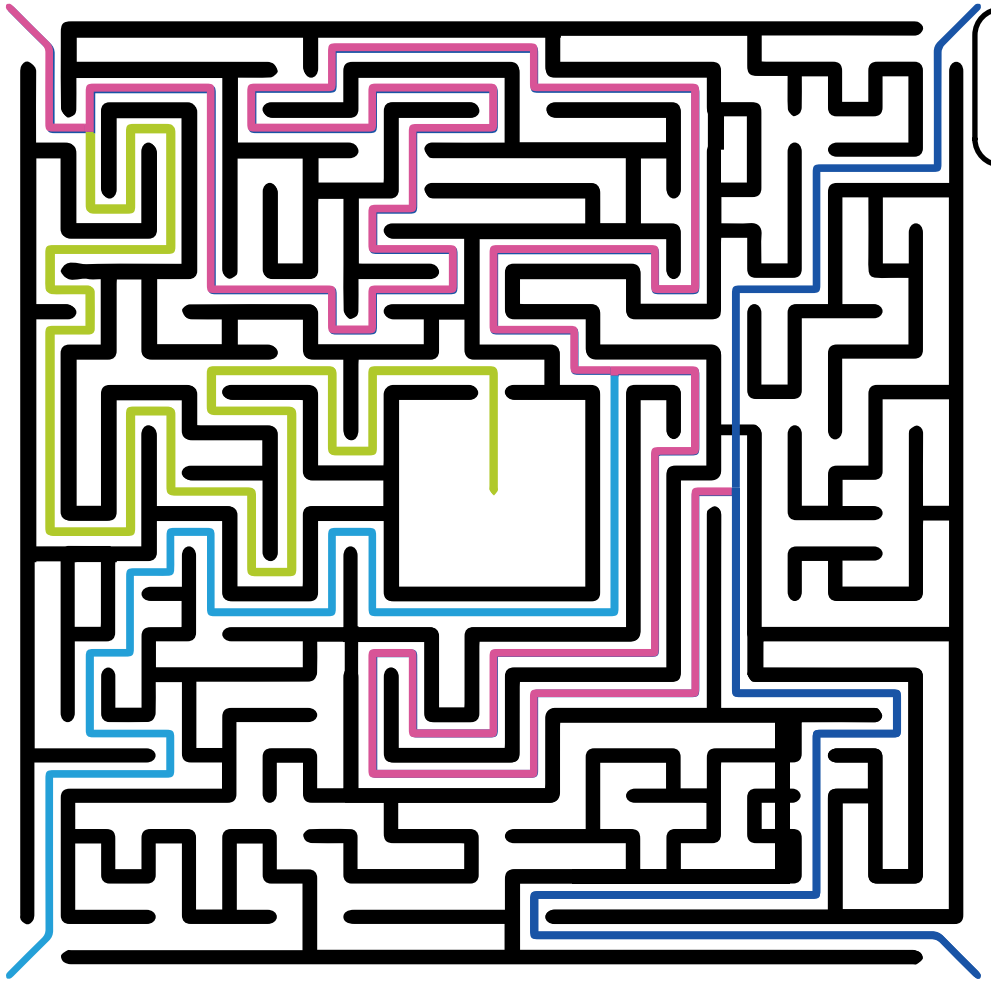
Link

Customized Binary

* X-Change

PacketMill Provides the Right Format & Order for Metadata and Minimizes the Framework Footprint

Better Cache Locality



Right Format
Better Order

NF Metadata
[Router]
[DPI]
Data
[NAT]
VLAN

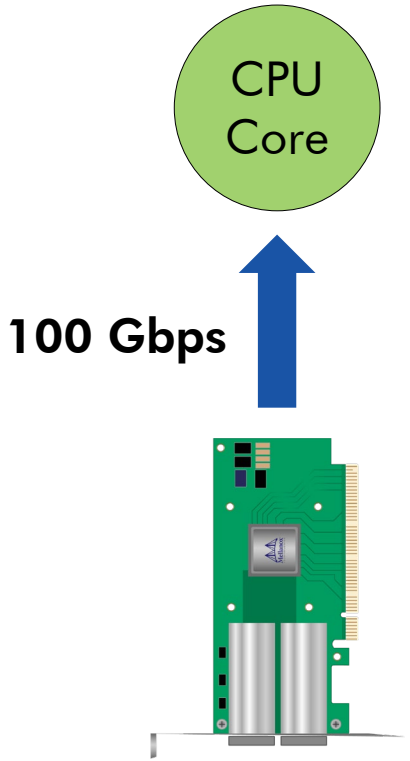
Packet

Evaluation

1. Impact of Code Optimizations
2. X-Change vs. Existing Metadata Management Models
3. Impact of Workload/Trace
4. Sophisticated Network Functions
5. Multicore Network Functions
6. PacketMill vs. State-of-the-Art Packet Processing Frameworks

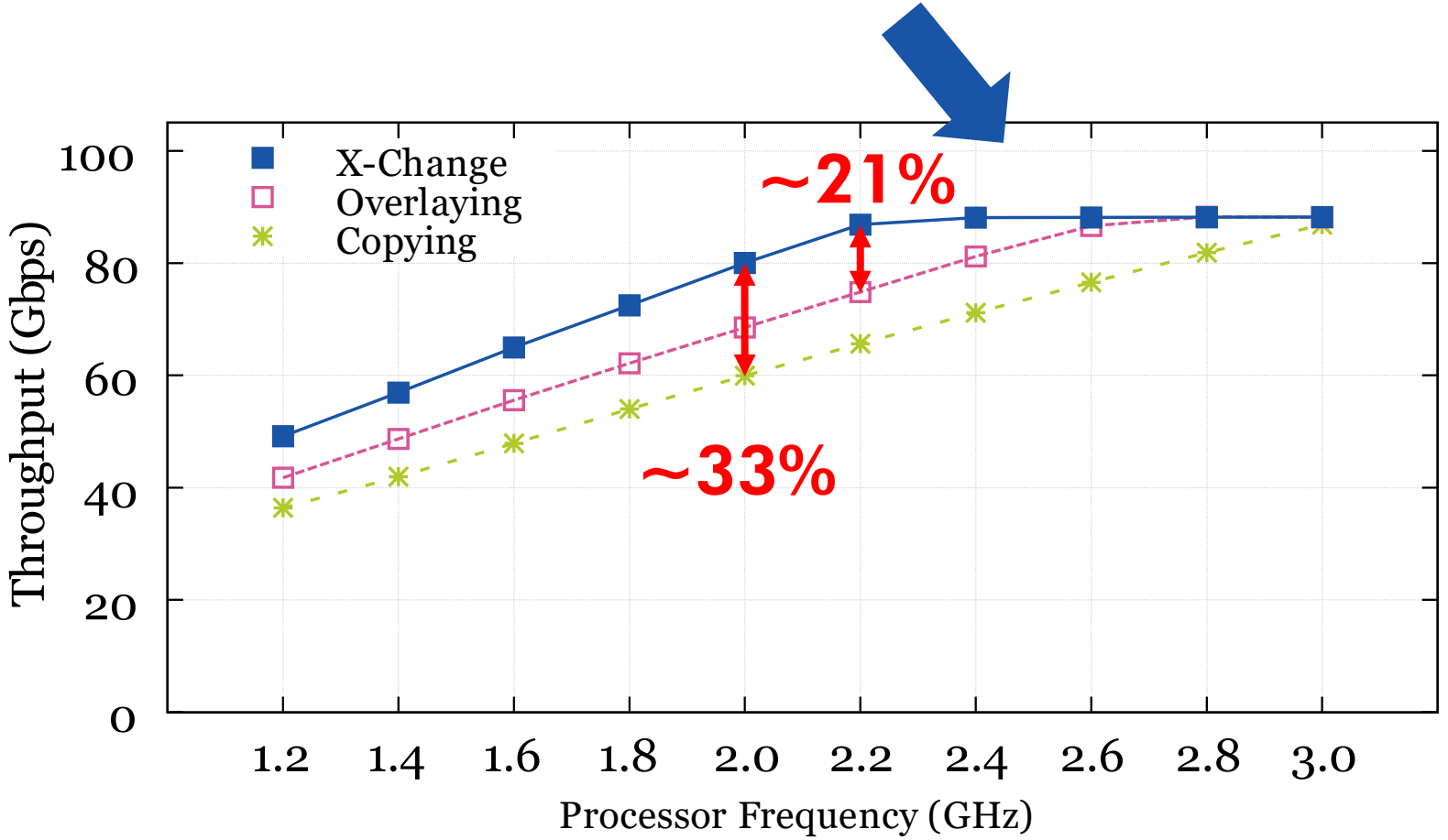
Check
Our Paper

X–Change vs. Existing Metadata Management Model



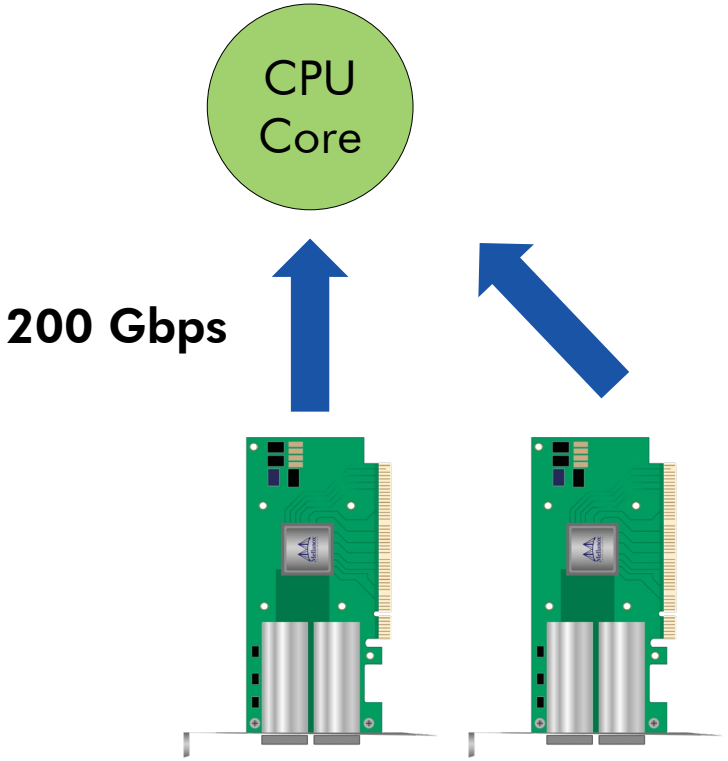
1x Mellanox Connect-X 5

Throughput stops increasing

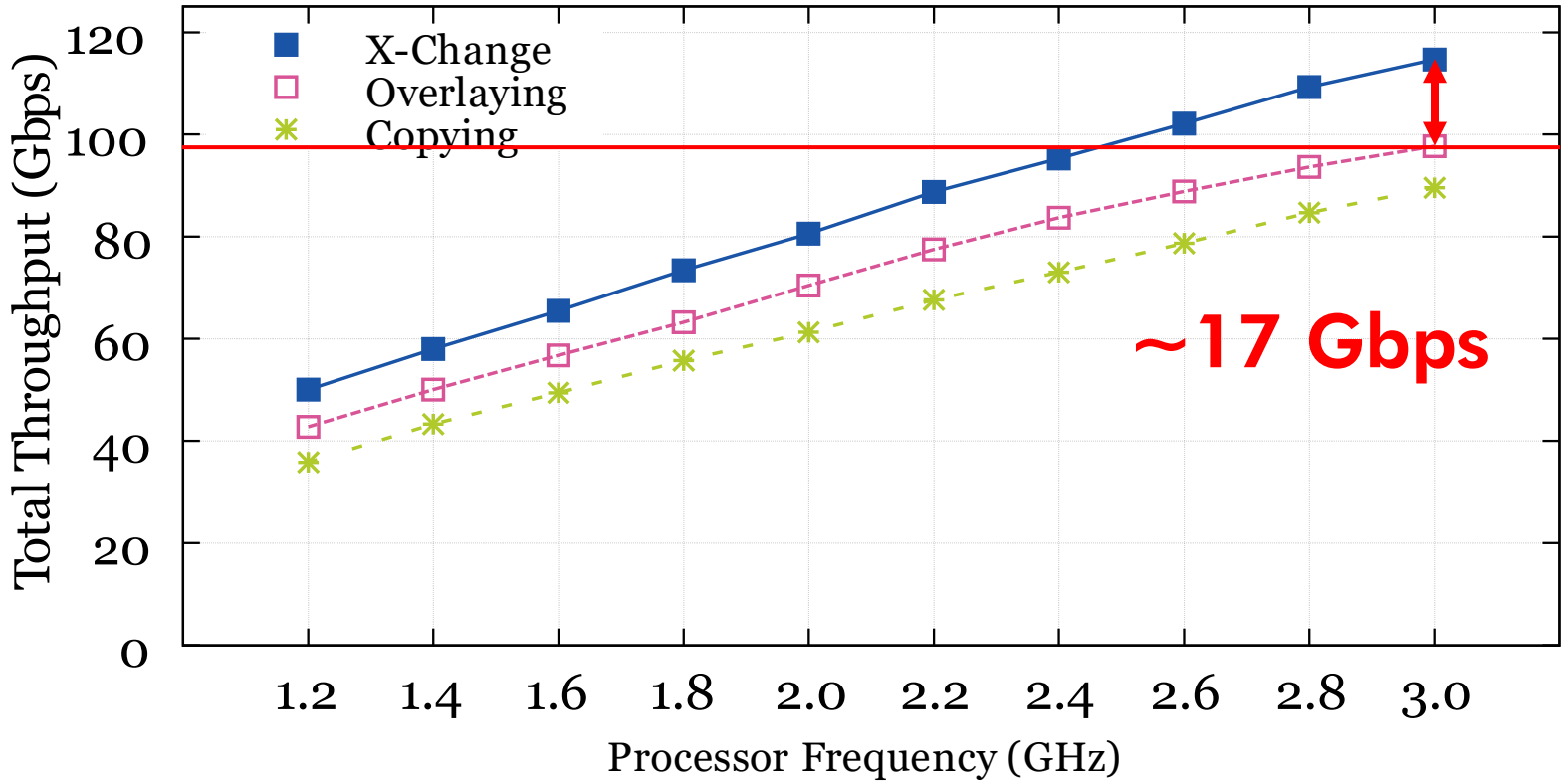


Higher is better

X–Change is the Only Model Capable of Forwarding Packets at >100 Gbps

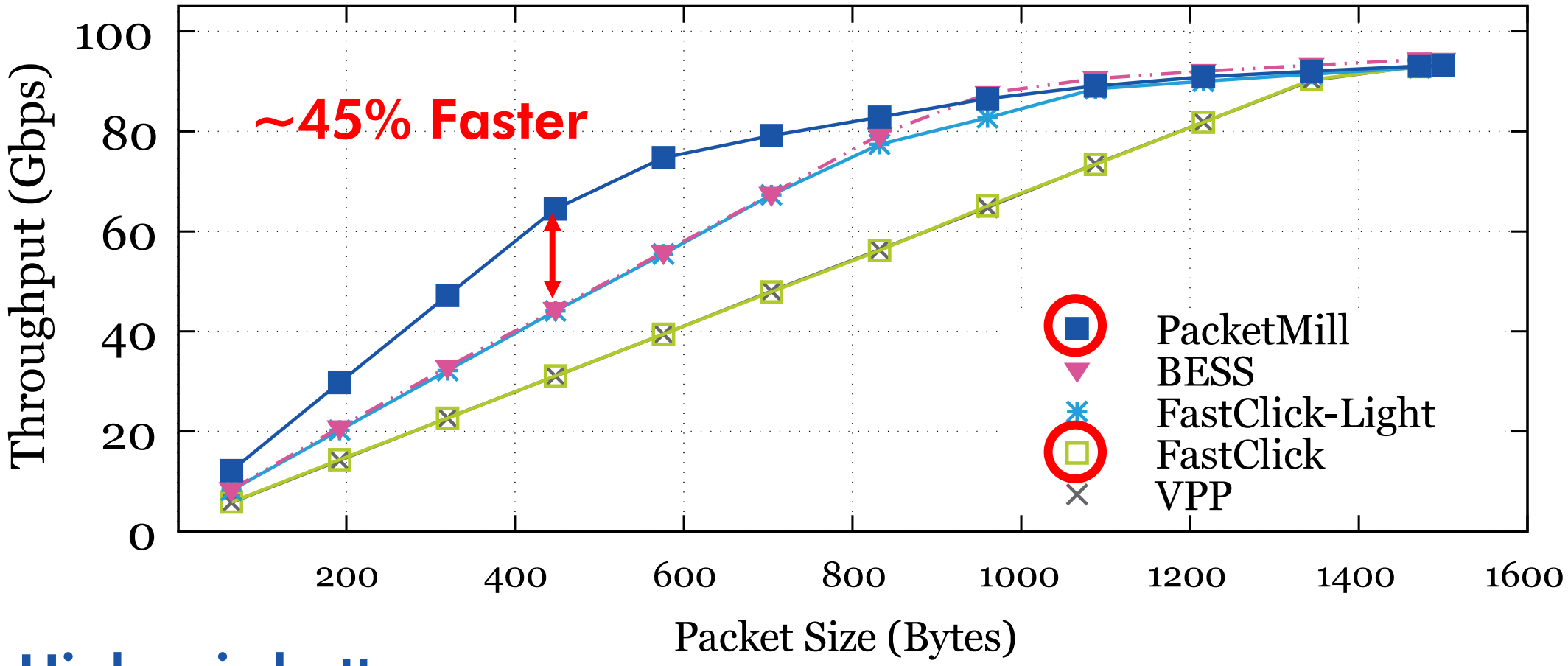


2x Mellanox Connect-X 5



Higher is better

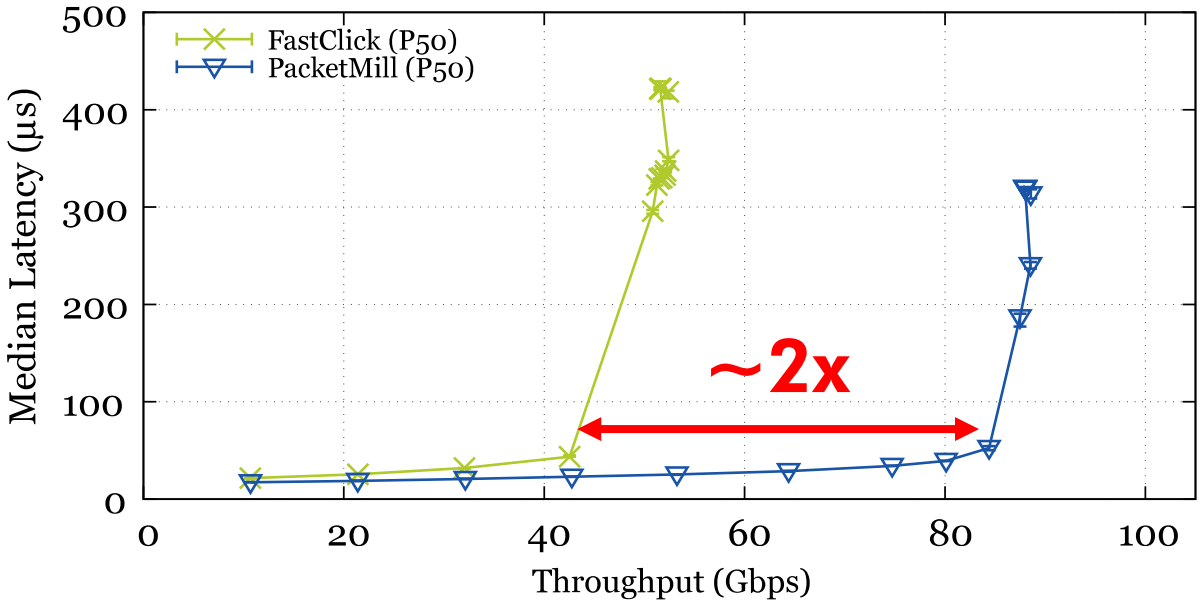
PacketMill Forwards Packets Faster than State-of-the-Art Frameworks

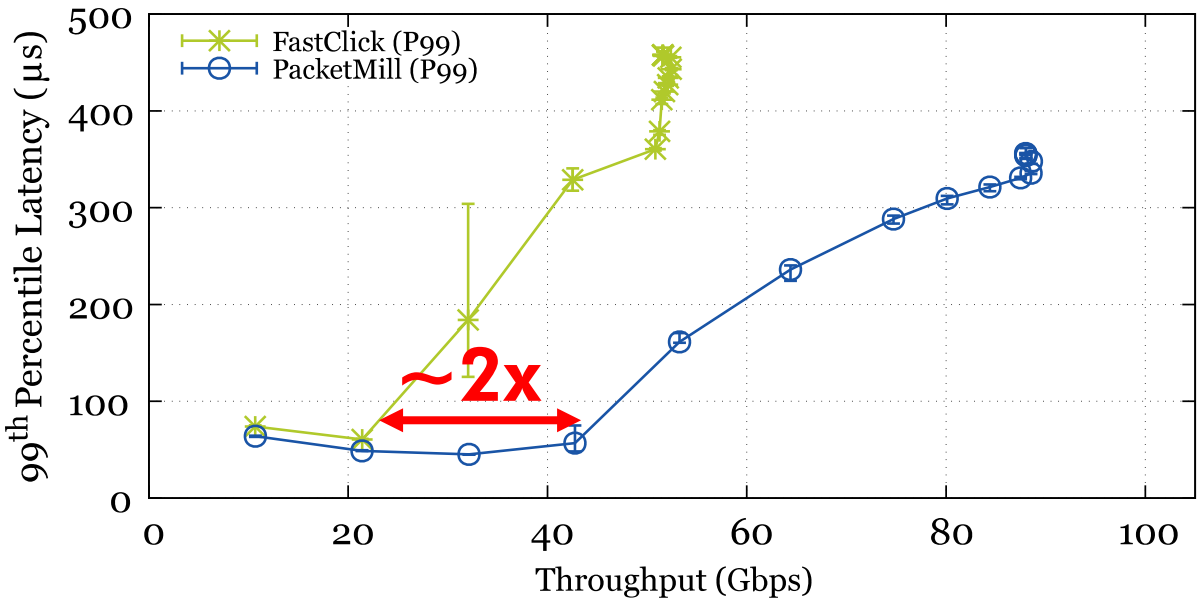


Higher is better

PacketMill Shifts the Knee of Throughput vs. Latency Curve

A router is forwarding a real campus trace with one core at different rates





Lower and More Right is Better

Conclusion

- Mitigating code inefficiencies and improving metadata management makes it possible to process packets on commodity hardware at higher rates.
- PacketMill achieves a better performance compared to other packet processing frameworks.
- PacketMill forwards at >100 Gbps with one core being fed with two NICs.
- Check out our paper for more information.

Our Paper

aliireza/packetmill

packetmill.io

European Research Council
Established by the European Commission
Swedish Foundation for Strategic Research
WASP Wallenberg Autonomous Systems Program

Thanks for watching

Do not hesitate to contact us if you have any questions.

farshin@kth.se and barbette@kth.se