ESnet
ENERGY SCIENCES NETWORK

DPDK as an Offload Engine for P4 SmartNIC Applications

Chris Cummings
Network Automation Software Engineer
Sean Cummings
Student Assistant

Energy Sciences Network (ESnet)
Lawrence Berkeley National Laboratory
U.S. Department of Energy

DPDK Summit

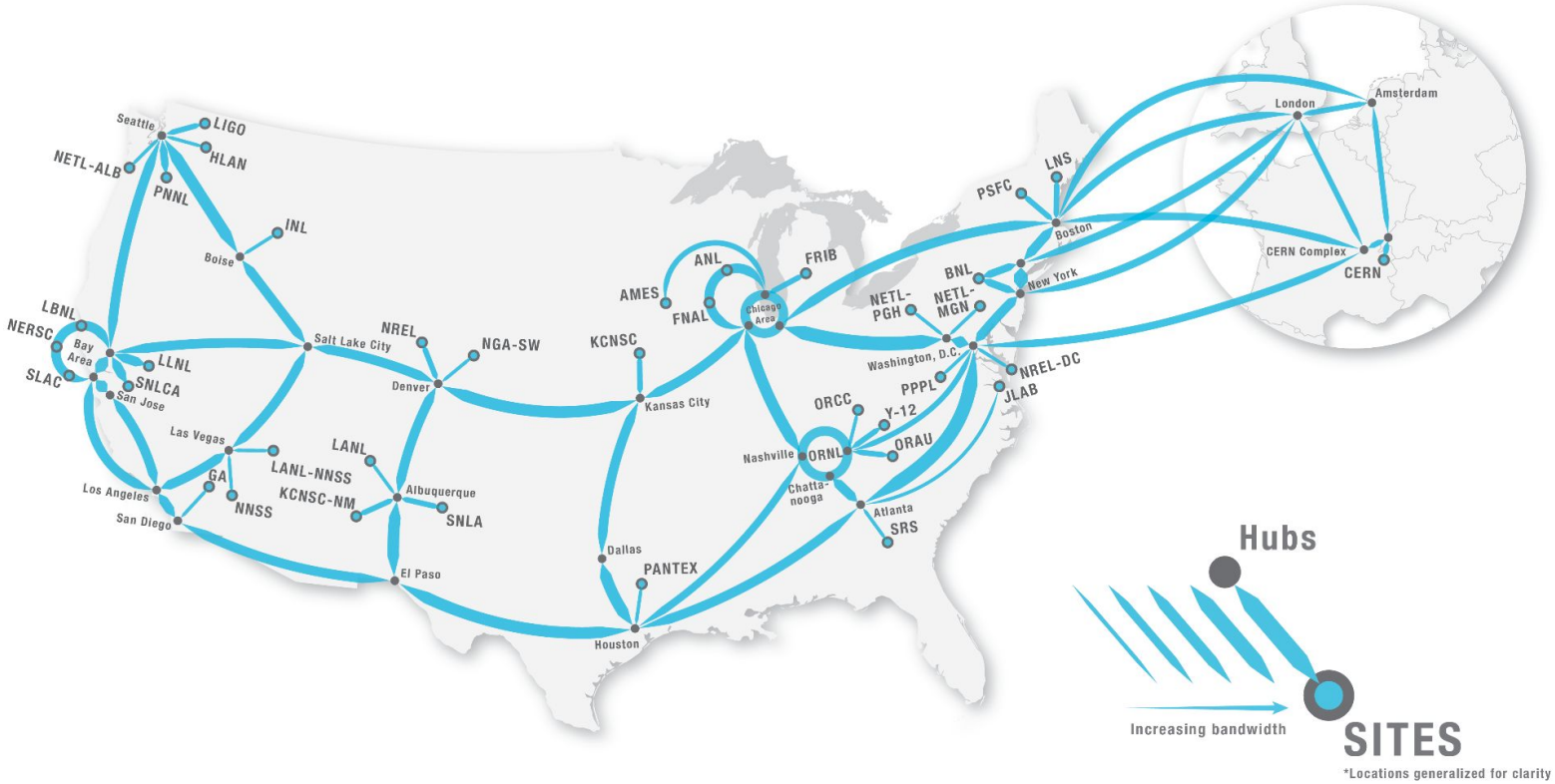
September 2023

U.S. DEPARTMENT OF ENERGY
Office of Science

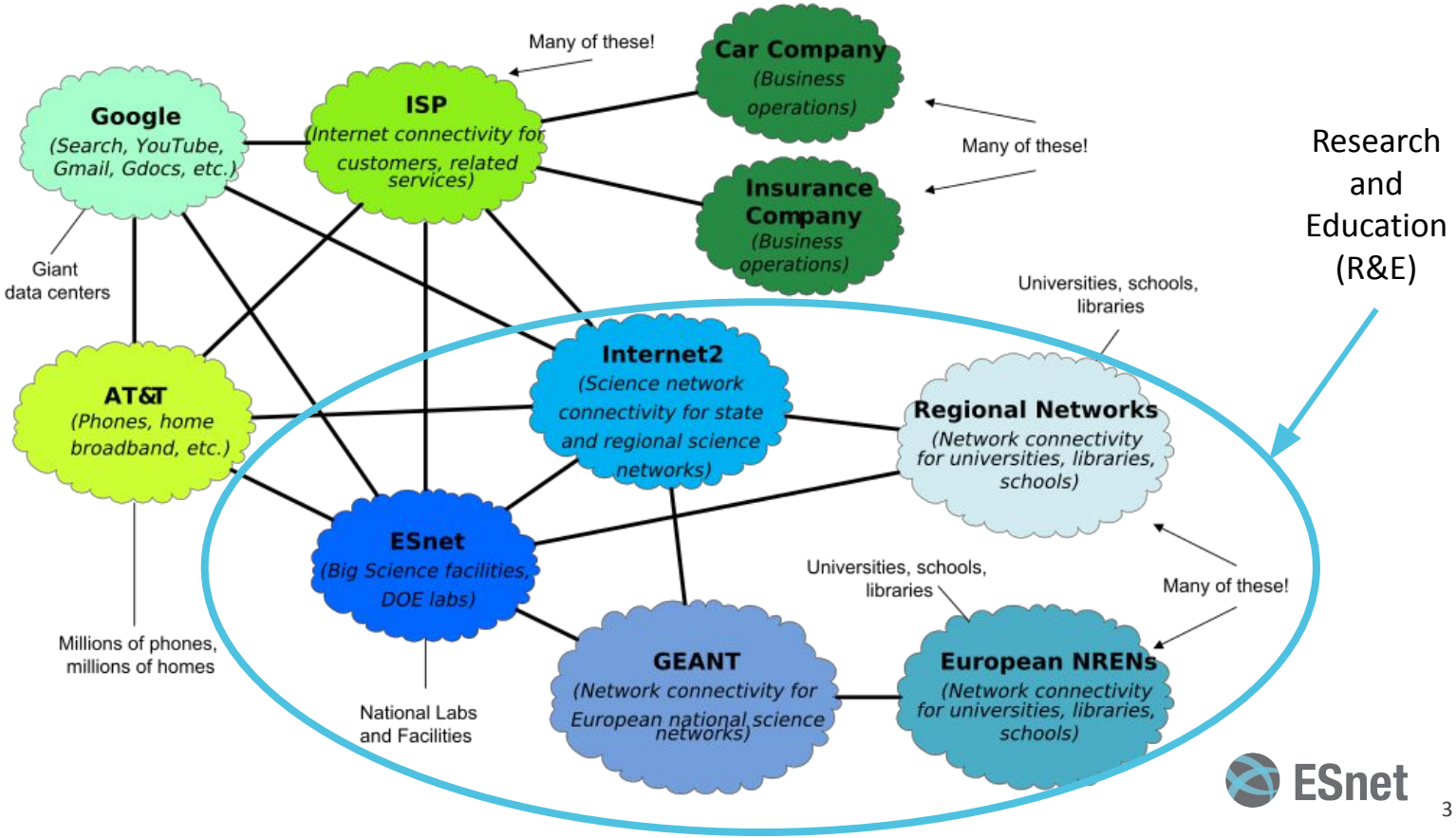
BERKELEY LAB

ESnet6

What is ESnet?



What is an R&E Network?



ESnet

Commercial ISPs vs. R&E Networks:

Normal ISP:[1]

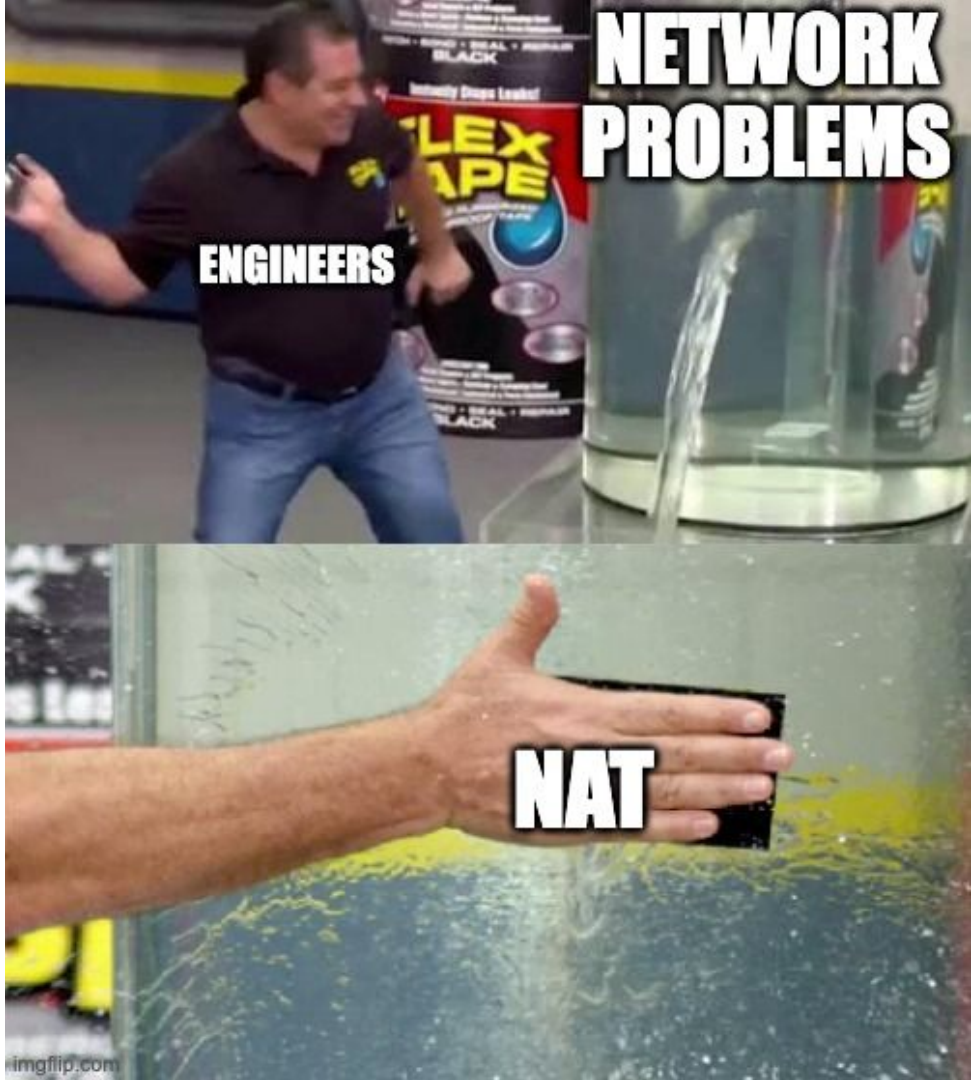
ESnet:[2]

ESnet

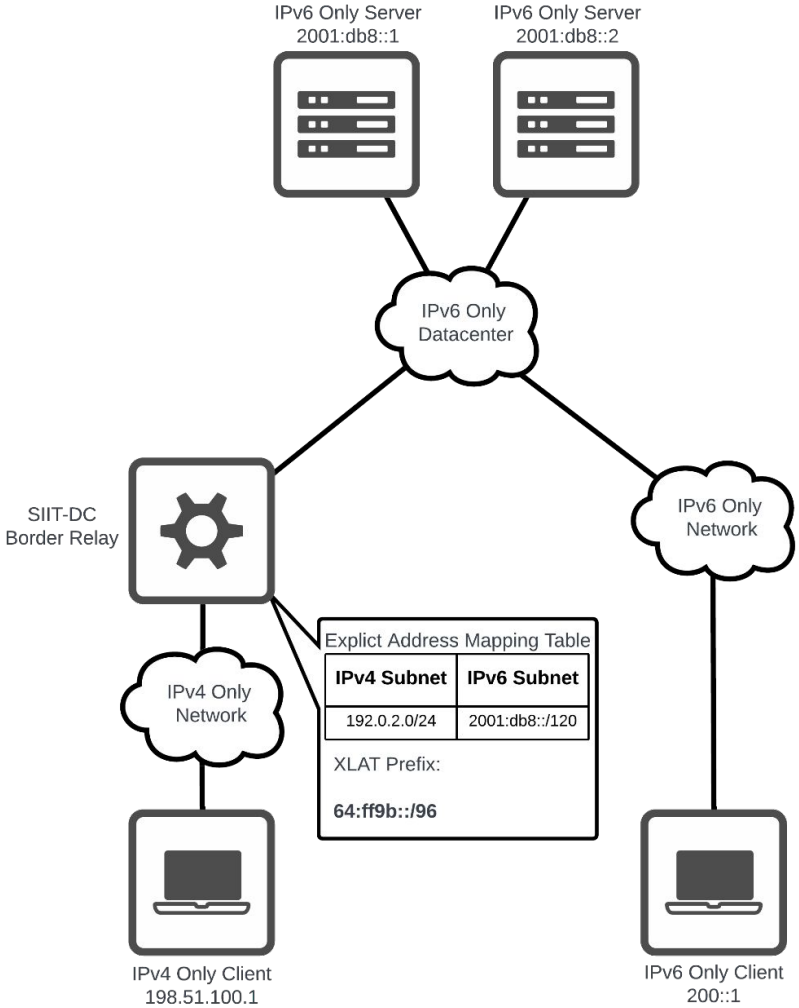
Project Background

- IPv6: The future is here, but not ***all*** the way here
- Some things can't be upgraded for IPv6 (expensive scientific instruments)
- Dual-stacking isn't necessarily the right solution anymore (OMB M-21-07)
- v4-only and v6-only segments still need to communicate… How?
- There is a tried-and-true solution…

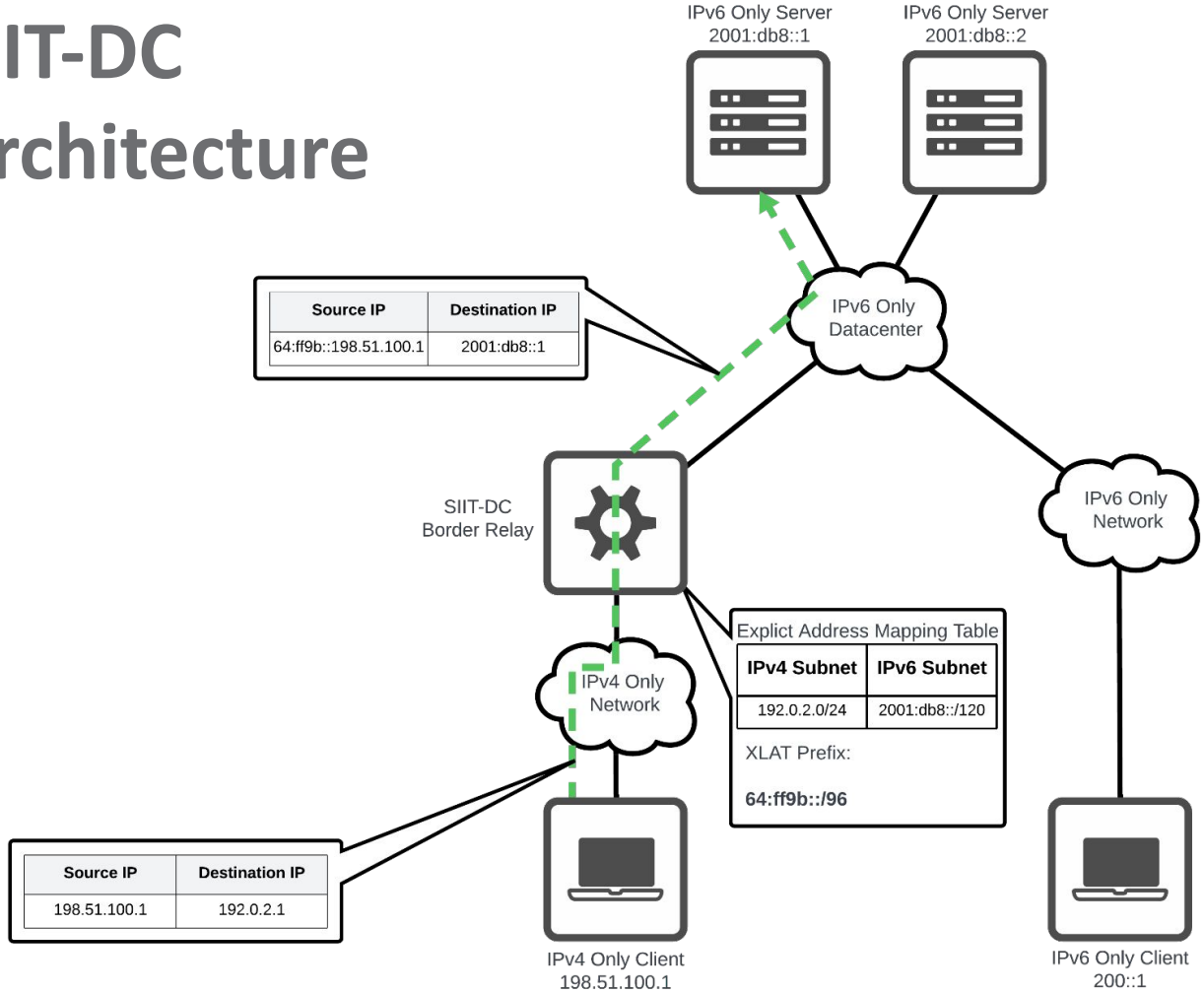
ESnet


ENGINEERS
NETWORK PROBLEMS
NAT
imgflip.com

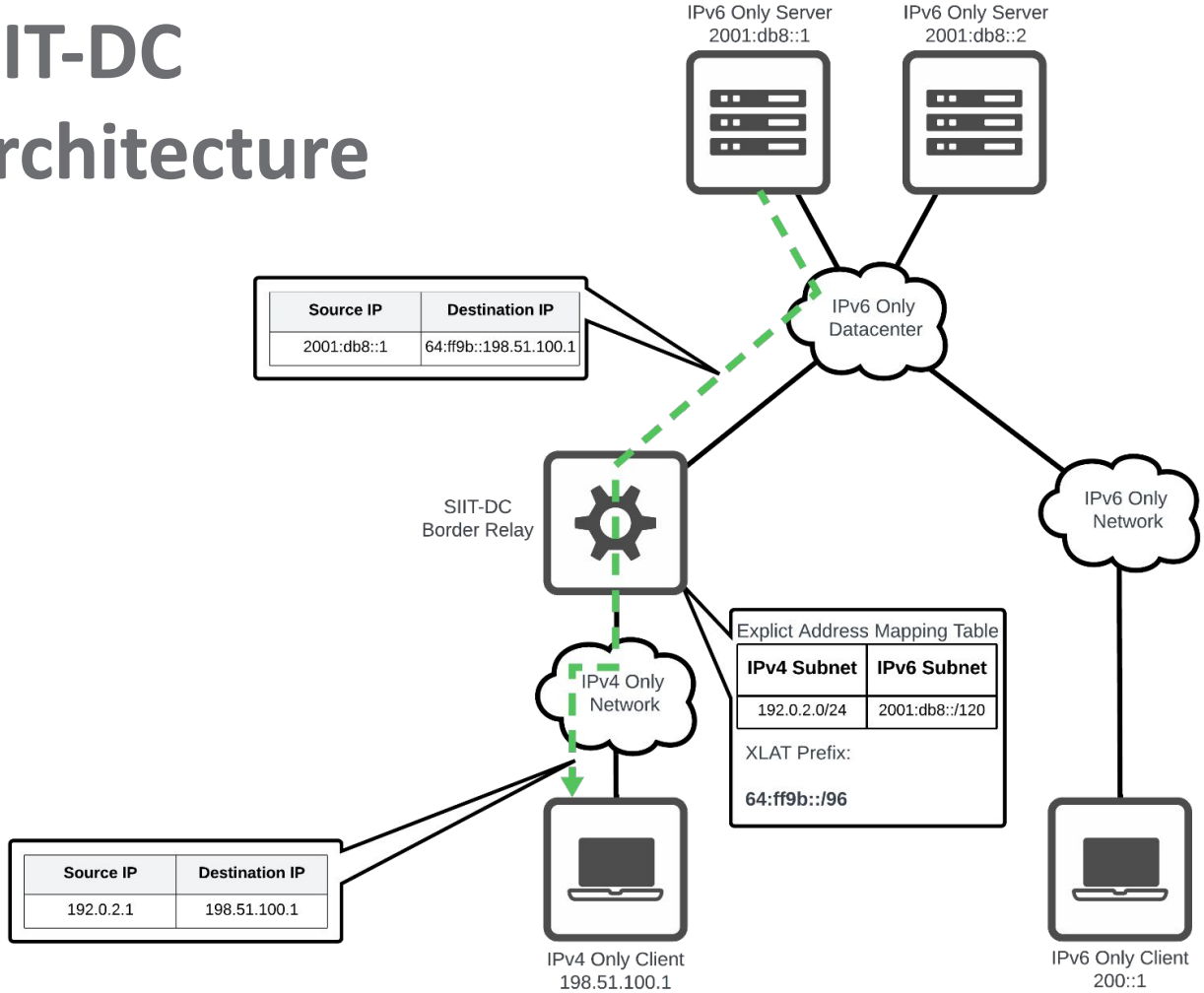
SIIT-DC Architecture



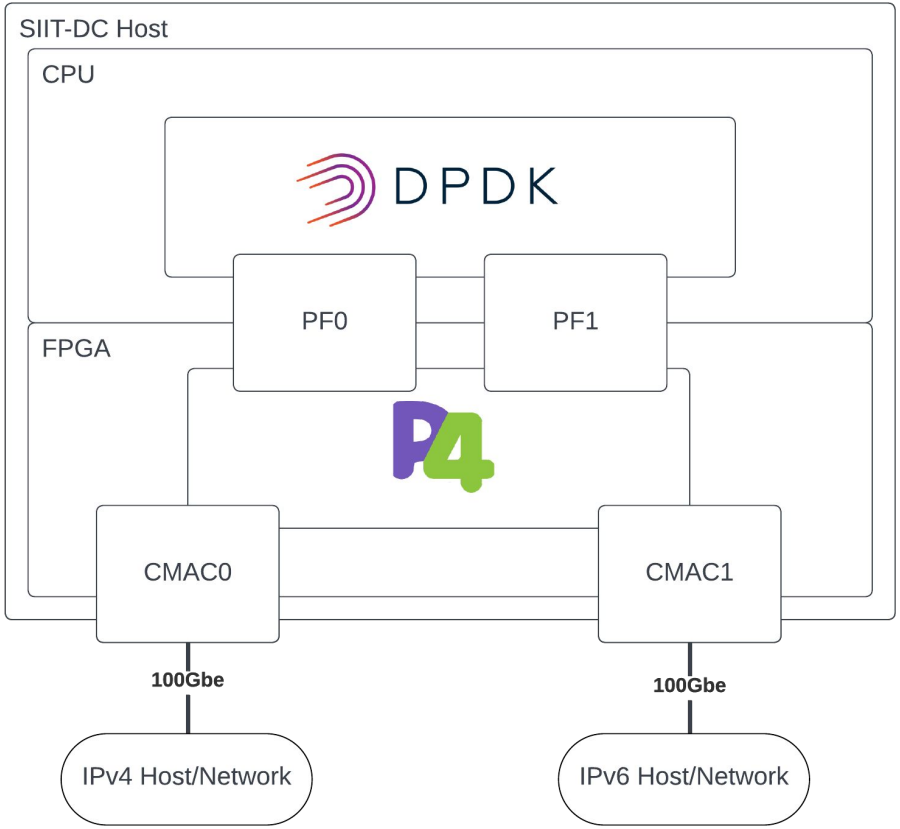
SIIT-DC Architecture



SIIT-DC Architecture



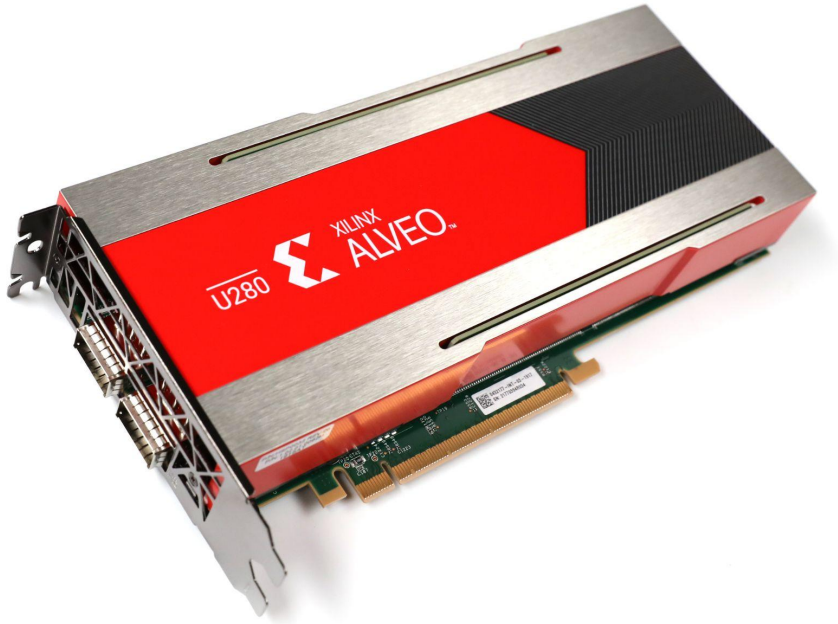
Architecture



ESnet

Hardware

- Xilinx Alveo U280 FPGA[3]
- 2x 100GbE Ethernet CMAC interfaces
- 8x PCIe Gen4 Lanes
- Intel Xeon E5-2670 CPU in Host Server



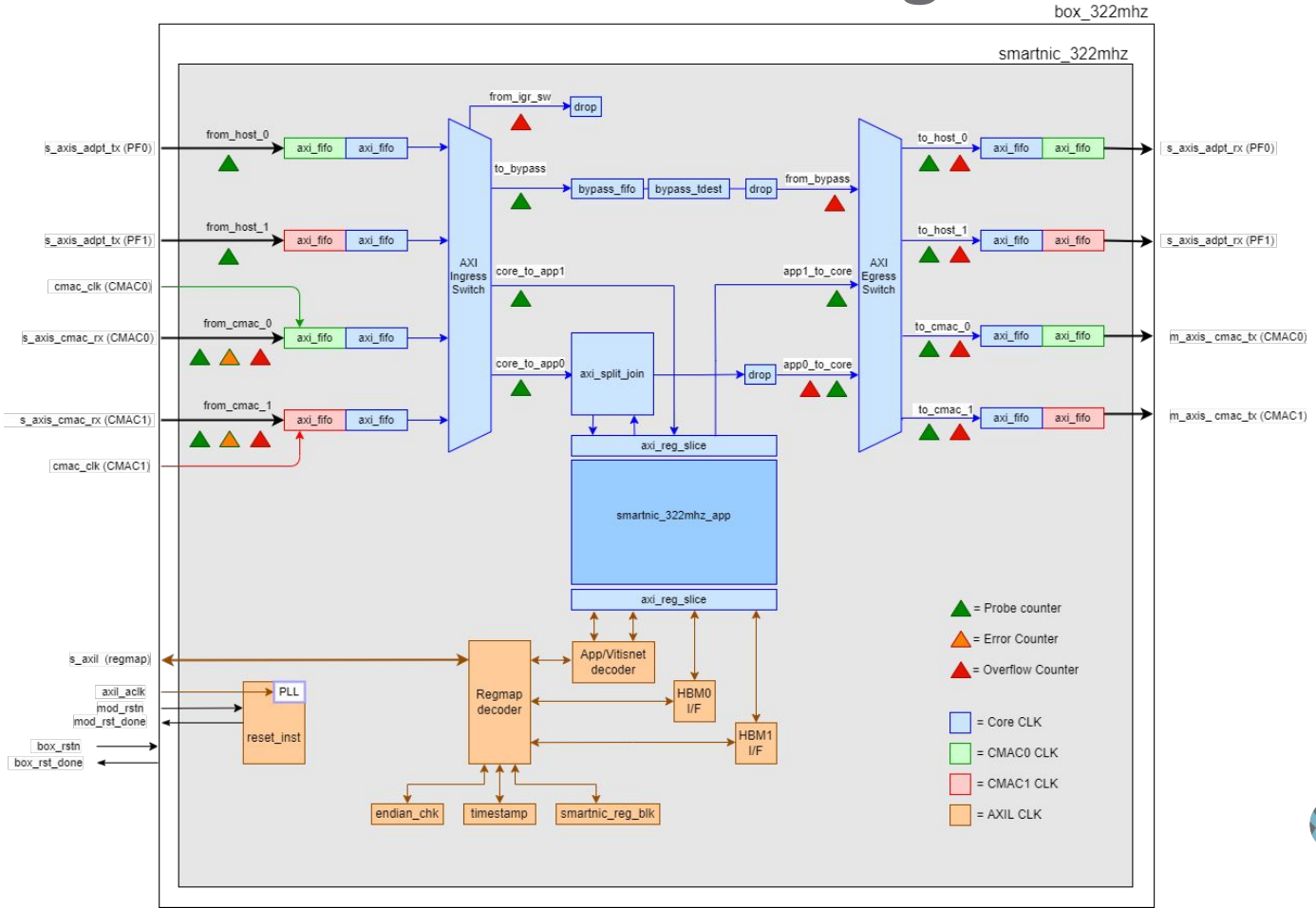
ESnet

What is ESnet SmartNIC?

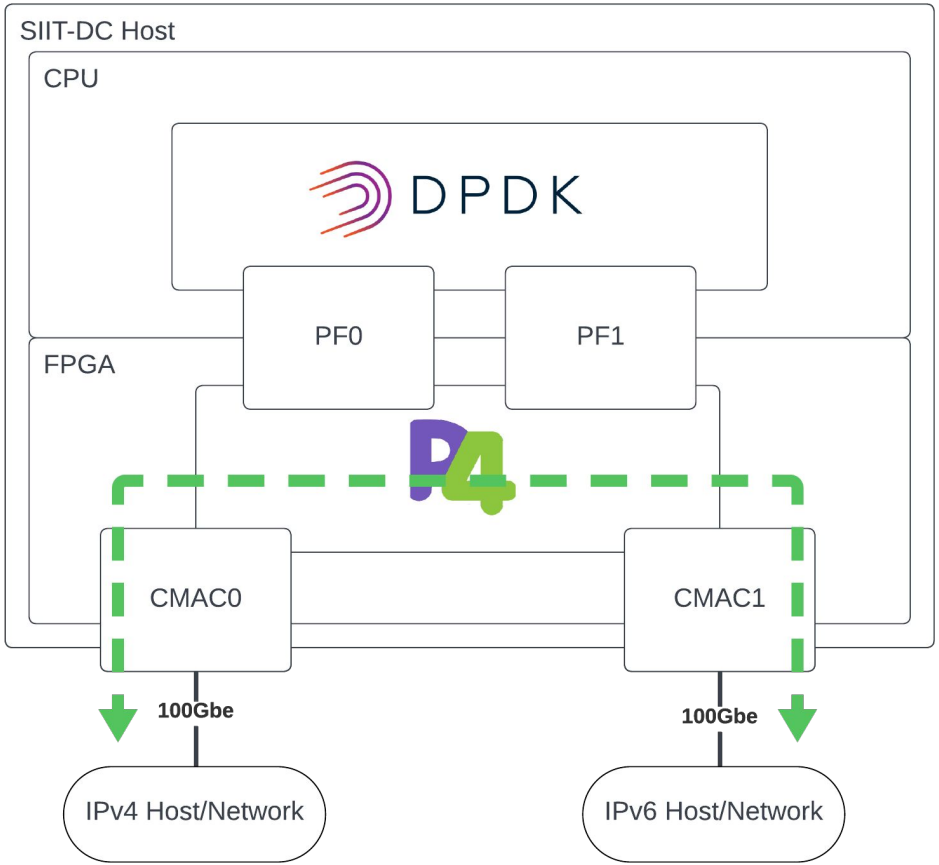
- Framework for developing FPGA-based applications
- Helps you deploy your FPGA-based applications
- Uses docker compose
- Repos Links:
  - esnet-smartnic-fw
  - esnet-smartnic-hw

ESnet

ESnet SmartNIC HW block diagram

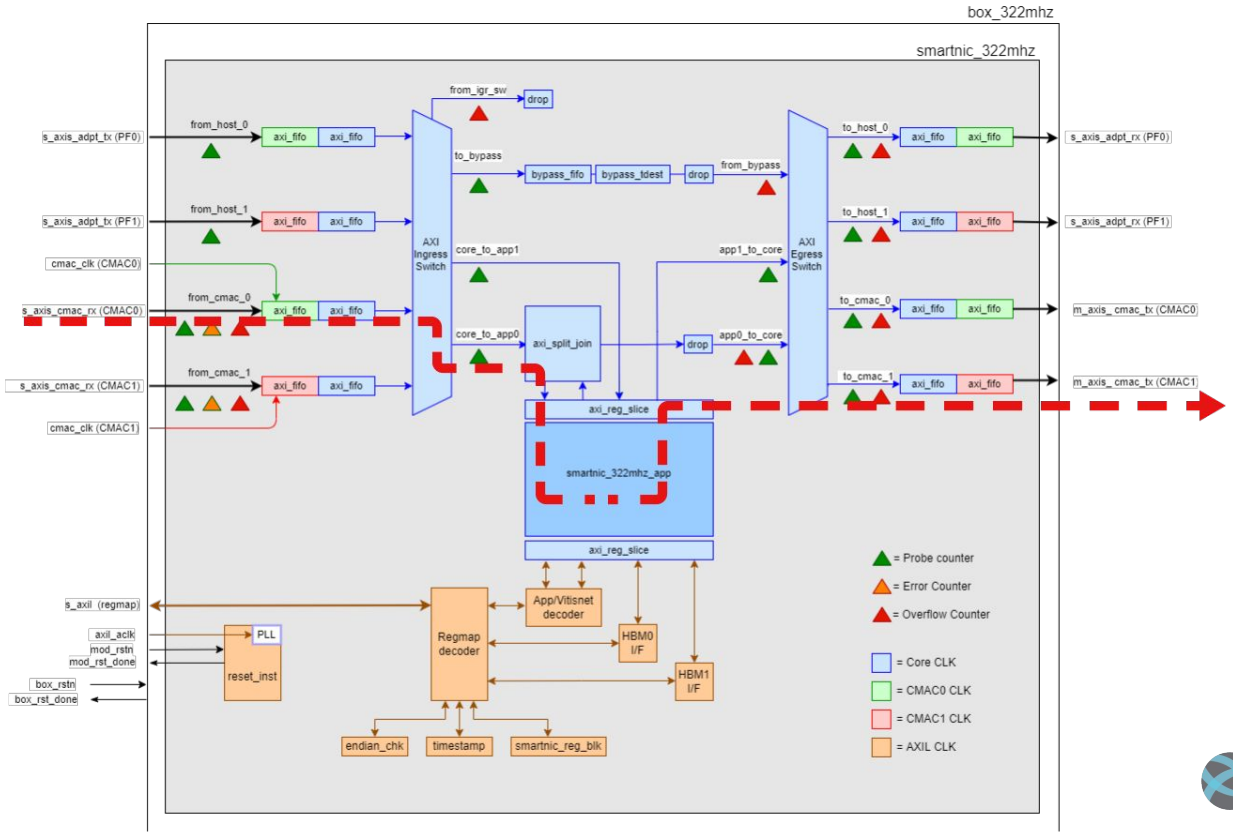


Fast Path Overview

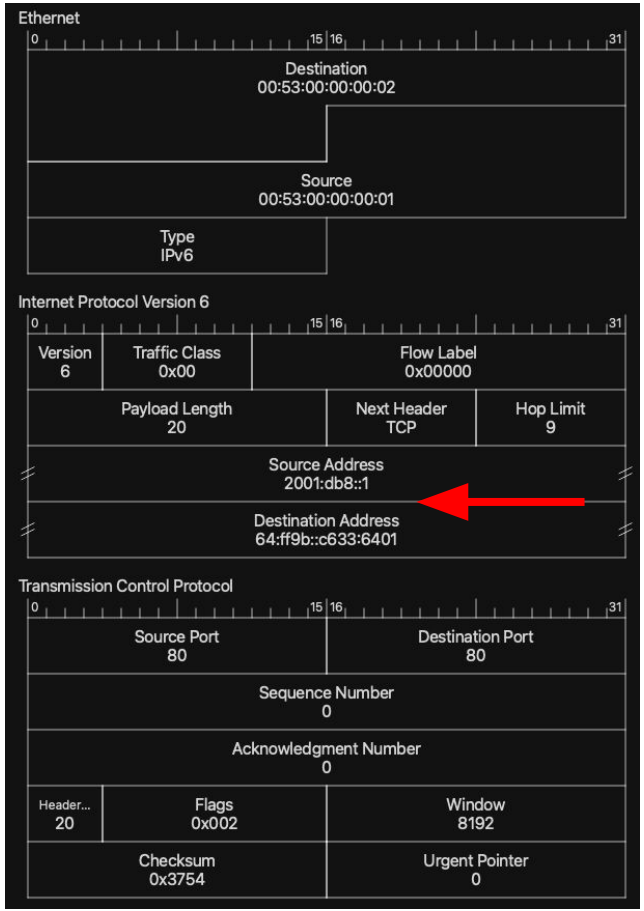


ESnet

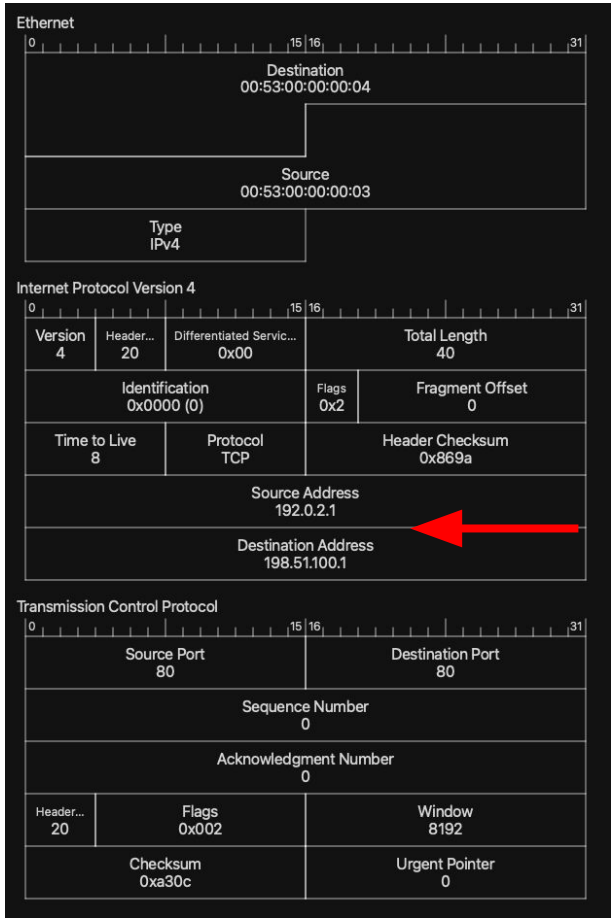
Fast Path Detail



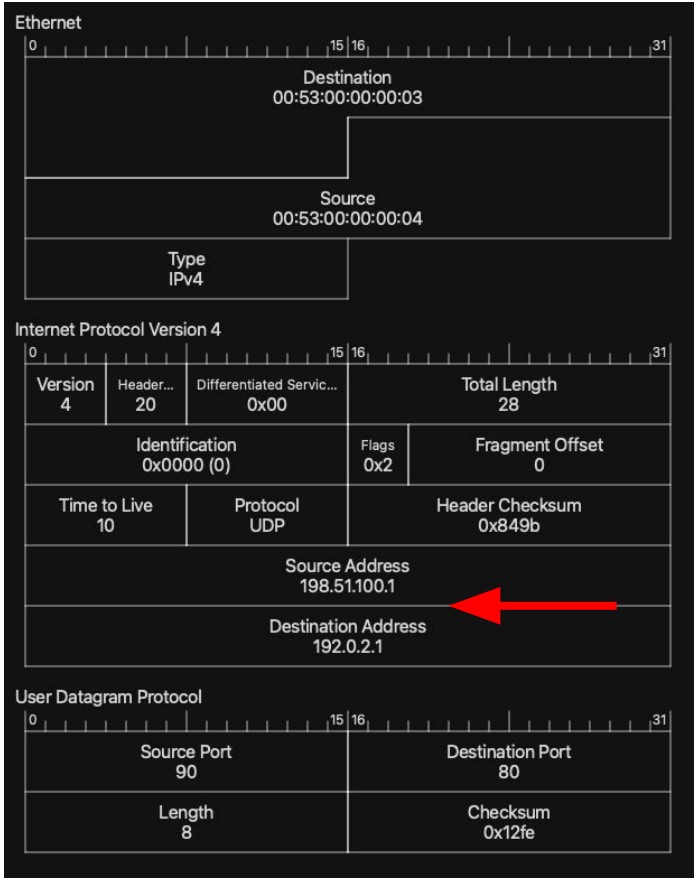
Untranslated IPv6/TCP



Translated IPv4/TCP

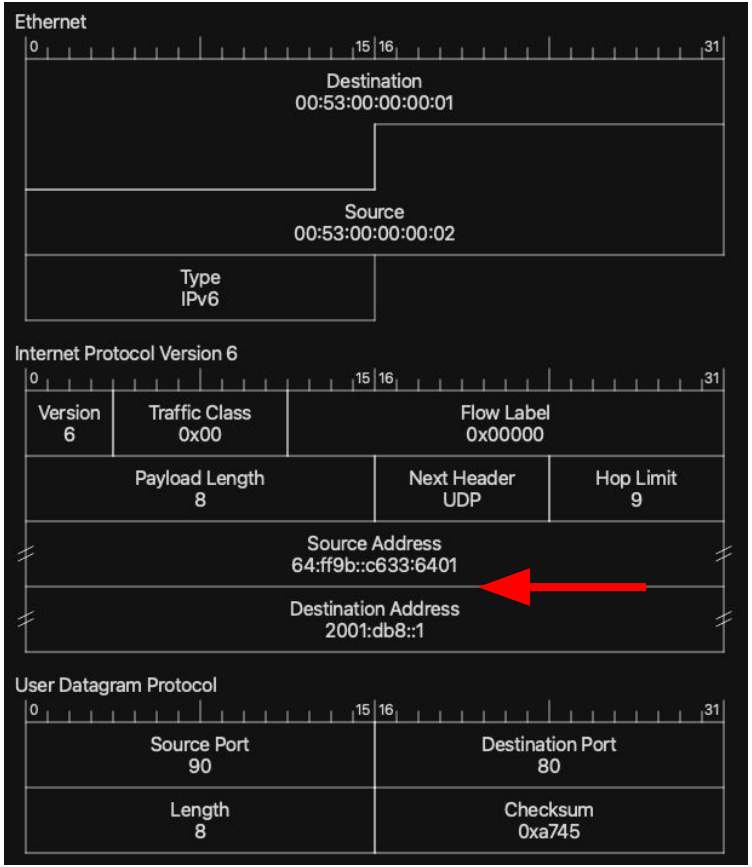


Untranslated IPv4/UDP Packet

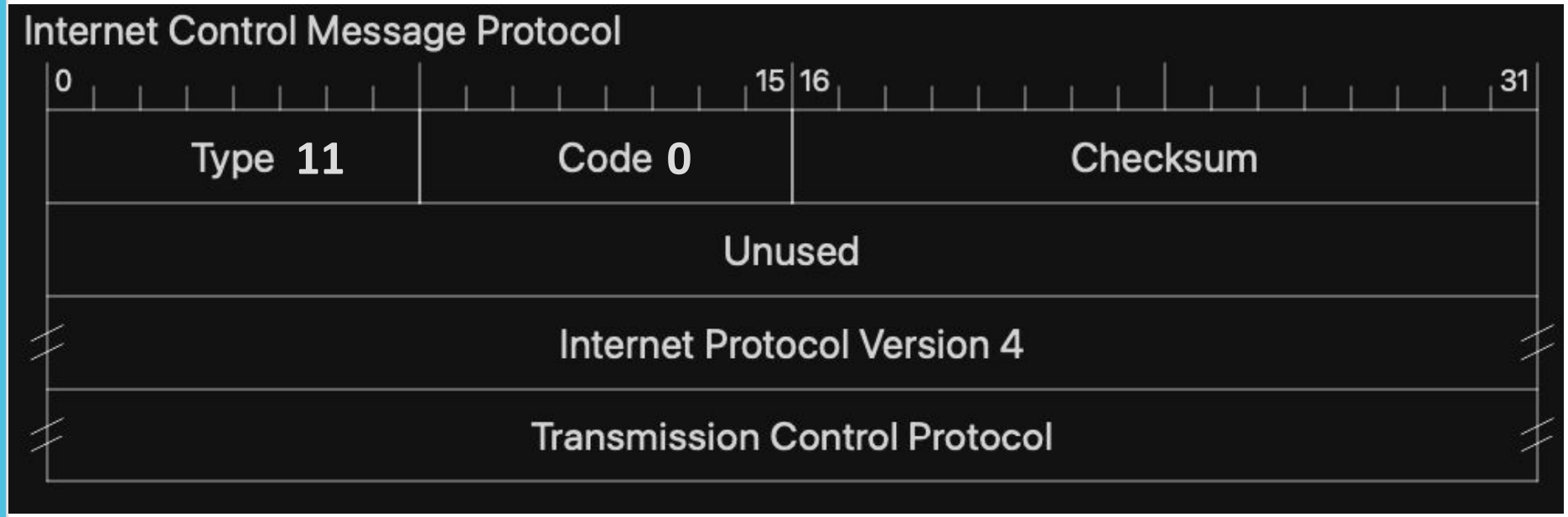


ESnet

Translated IPv6/UDP Packet



ICMPv4 Error



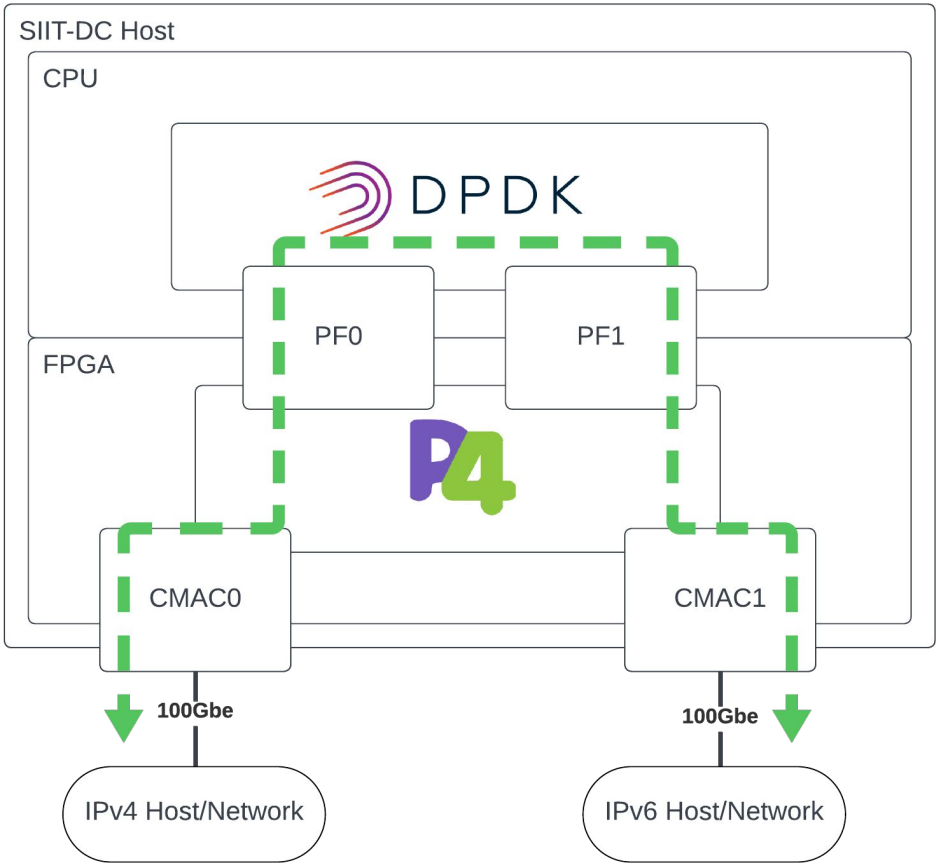
ICMPv6 Error (Variable length body!)

Time Exceeded

Bit offset	0–7	8–15	16–31
0	3	Code	Checksum
32	Unused		
64	Message body (Variable Size)		

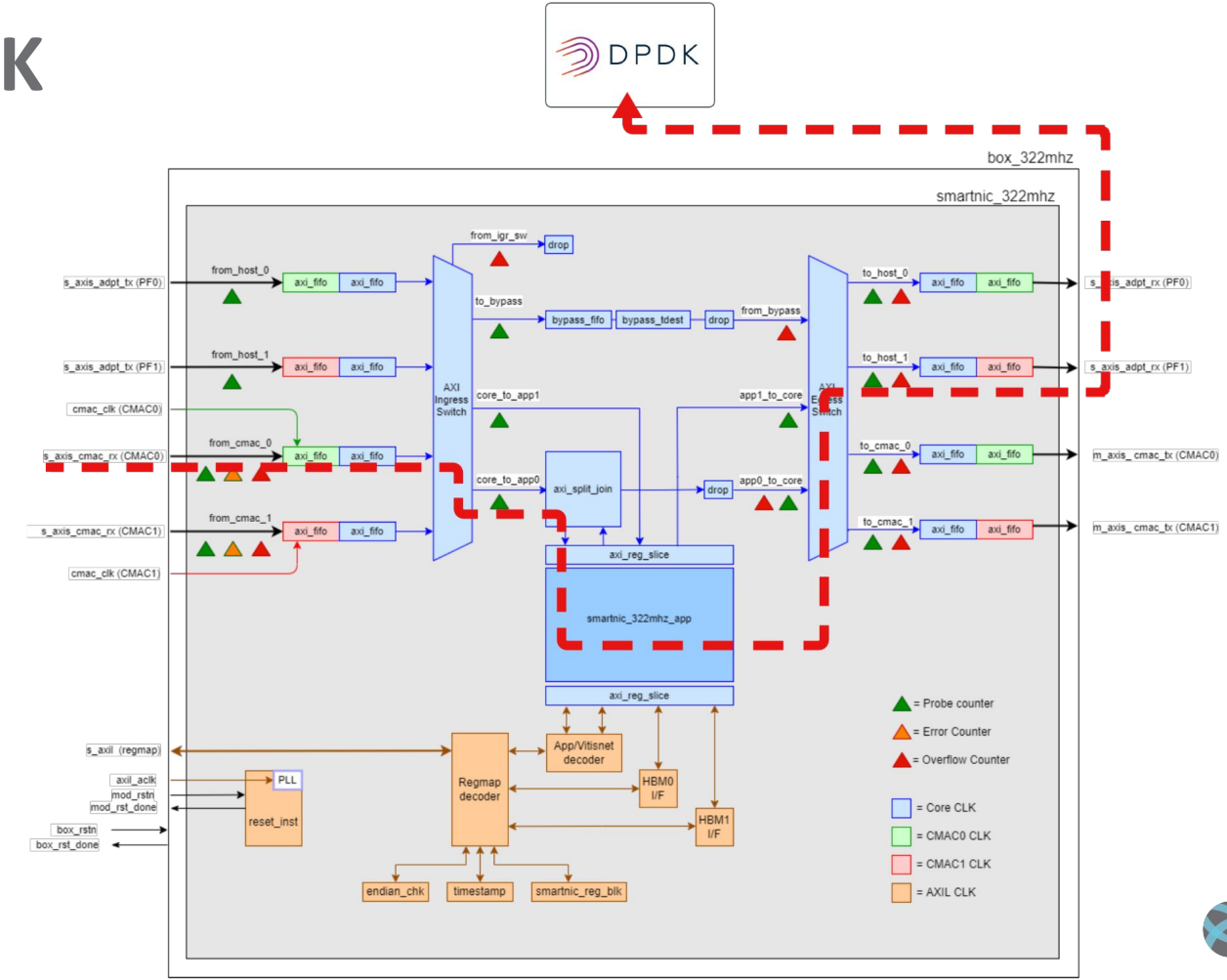
ESnet

“Slow” Path Overview



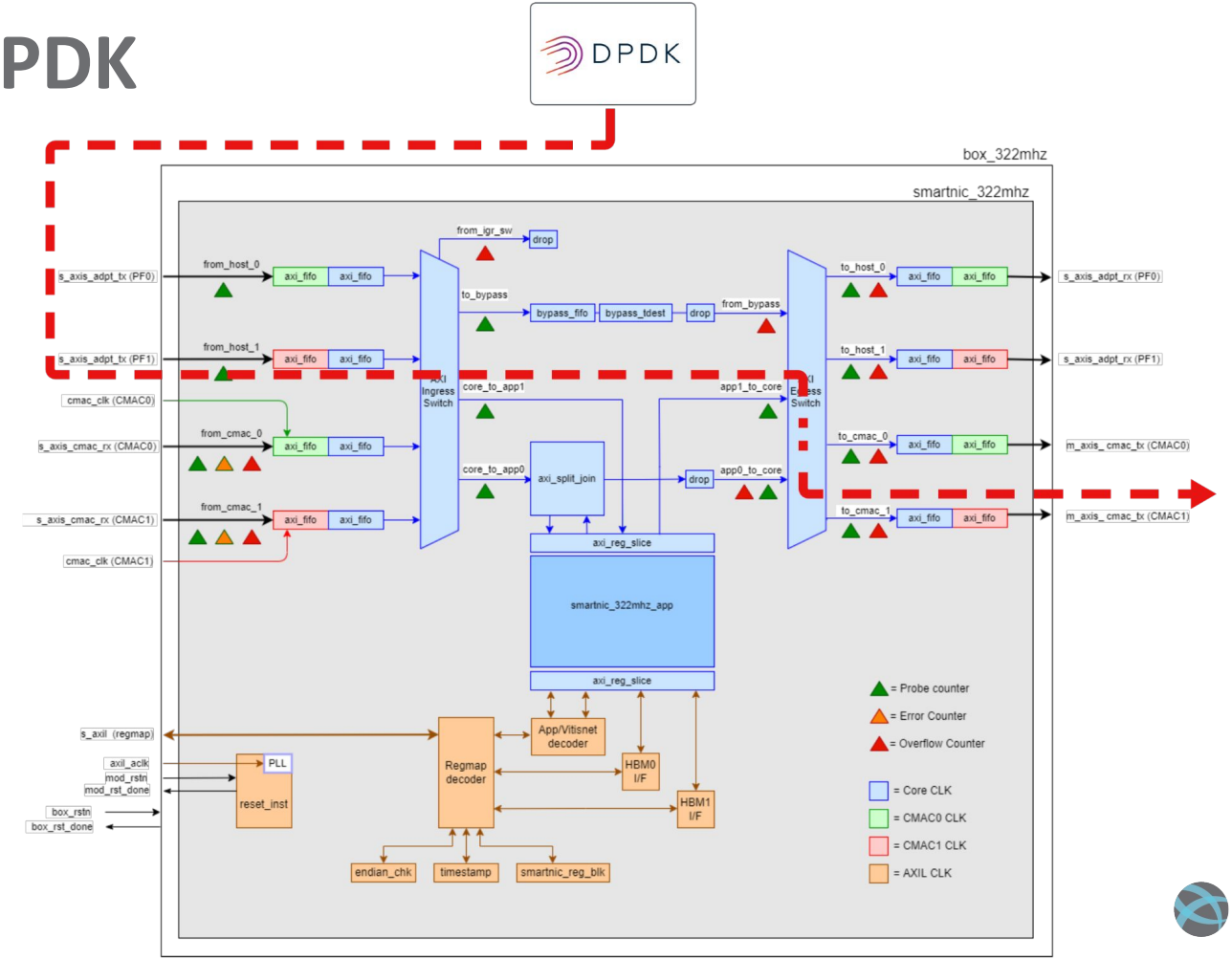
ESnet

To DPDK



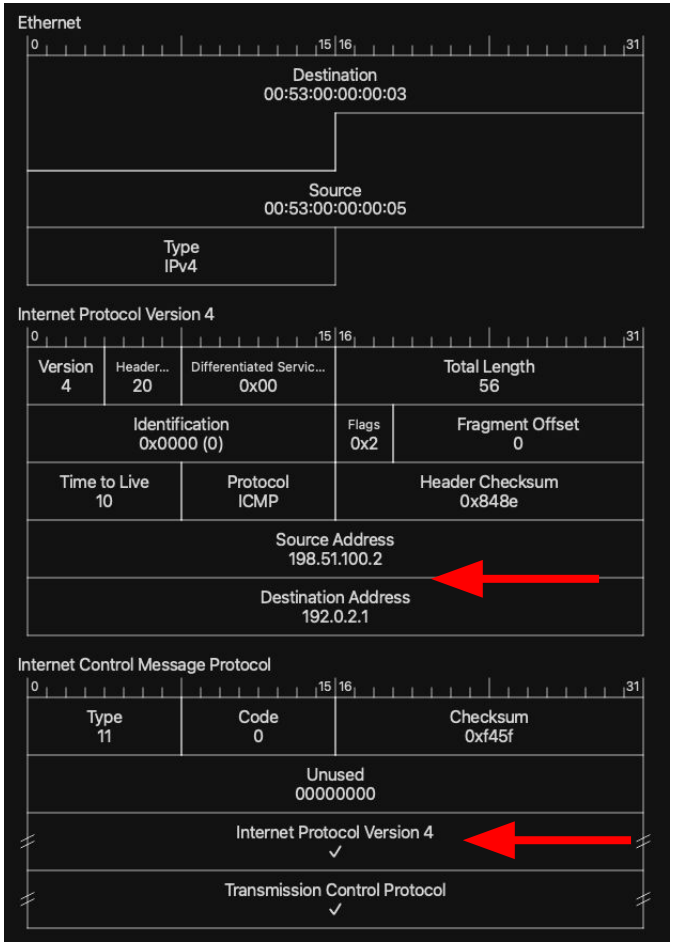
ESnet

From DPDK



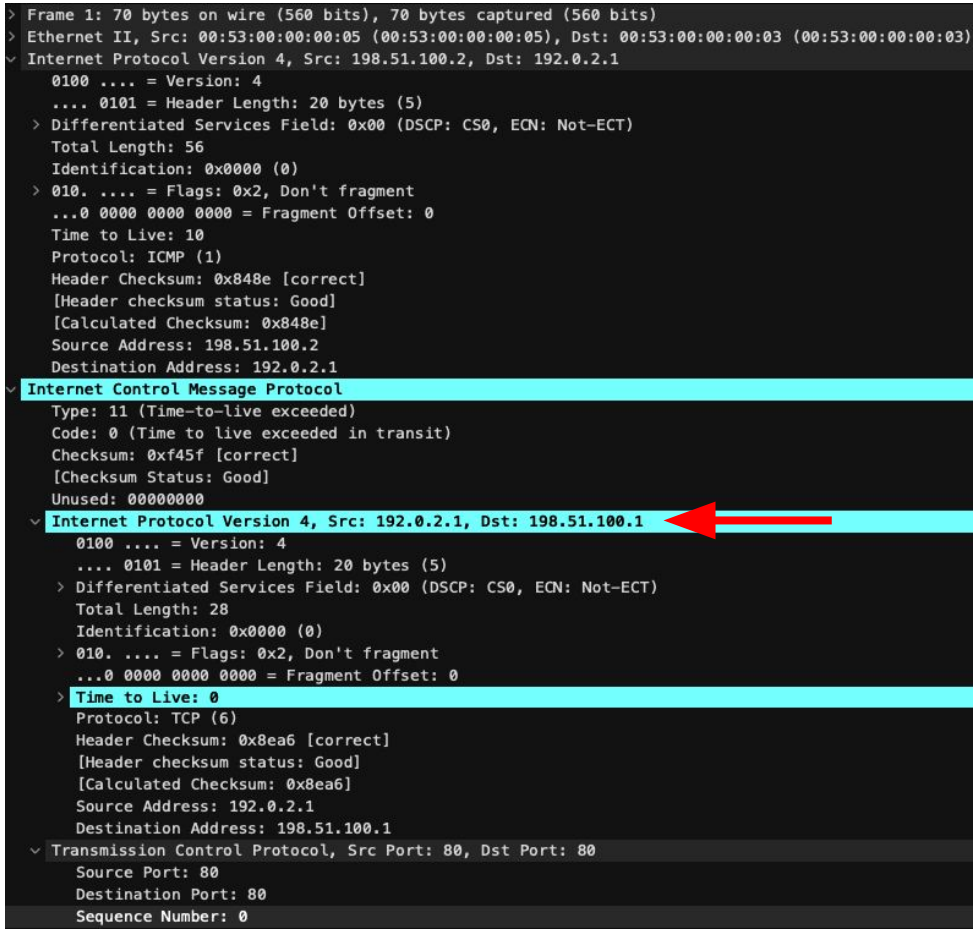
ESnet

Untranslated IPv4/ICMP4 Error



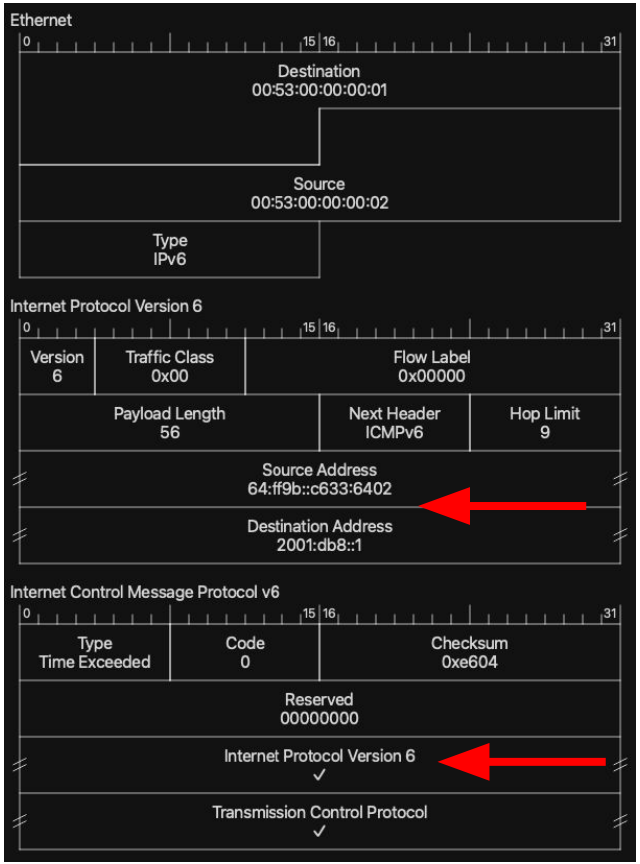
ESnet

Untranslated IPv4/ICMP4 Error



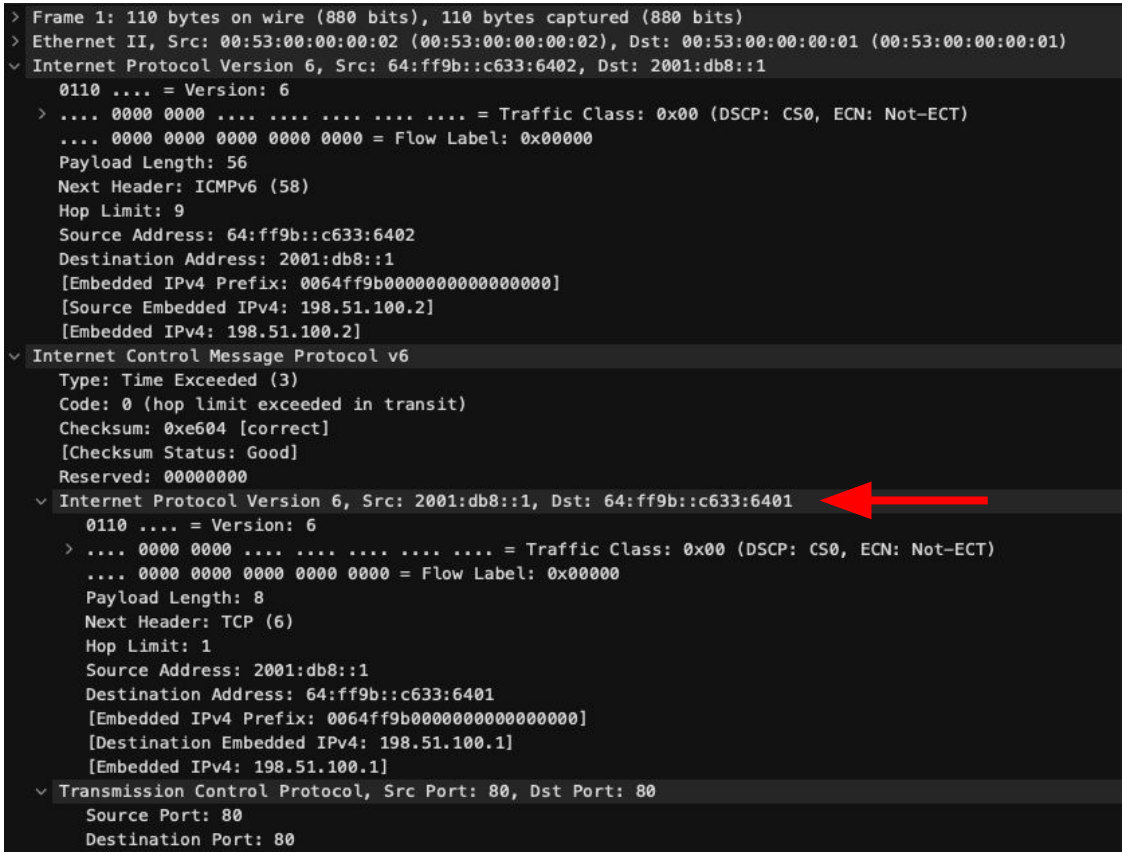
ESnet

Translated ICMPv6 Error



ESnet

Translated ICMPv6 Error



ESnet

Overcoming P4 Limitations

- P4 is limited by design - this is good!
- No loops or recursion
- Variable length packet processing is complex
- Reliance on externs for complex functionality
- Limited support for maintaining state

Offload possibilities

- Performing complex packet processing operations
- Prototyping externs without Verilog/VHDL
- Developing stateful applications
- Make quick table updates at runtime

Challenges Encountered

- Bit offsets are hard
- Compiling P4 app takes a long time (but bmv2 helps!)
- DPDK is very verbose compared to P4
- DPDK with C has less “syntactic sugar” compared to P4

Special Thanks!

Dr. Nik Sultana, IIT

Mohammad Firas Sada, IIT

Yatish Kumar, ESnet

Stacey Sheldon, ESnet

Scott Richmond, ESnet

Peter Bengough, ESnet

You!

ESnet

Questions?

- Contact Us:
  - Sean Cummings
    - scummings@hawk.iit.edu
    - https://www.linkedin.com/in/sean-cummings-6968061b4/
  - Chris Cummings
    - chriscummings@es.net
    - https://www.linkedin.com/in/chriscummingsak/

ESnet

References

- [1] Title: “Shinkansen N700 with Mount Fuji”, Author: tansaisuketti, Source: WikiMedia Commons, License: CC BY-SA 3.0
- [2] Title: “High Five”, Author: austrini, Source: WikiMedia Commons, License: CC BY 2.0
- [3] Title: “Alveo U280 Data Center Accelerator Card”, Author: AMD Xilinx, Source: AMD Xilix U280 Product Page