Hardware Accelerated L7 Monitoring at 100 Gbps

Introduction

- Network monitoring shifting to L7 processing
  - To keep pace with modern threats (Heartbleed)
- 100 Gbps networks available and deployed
  - 150 Mpps

➔ **L7 flow monitoring at 100 Gbps**

  - Performance of hardware, flexibility of software

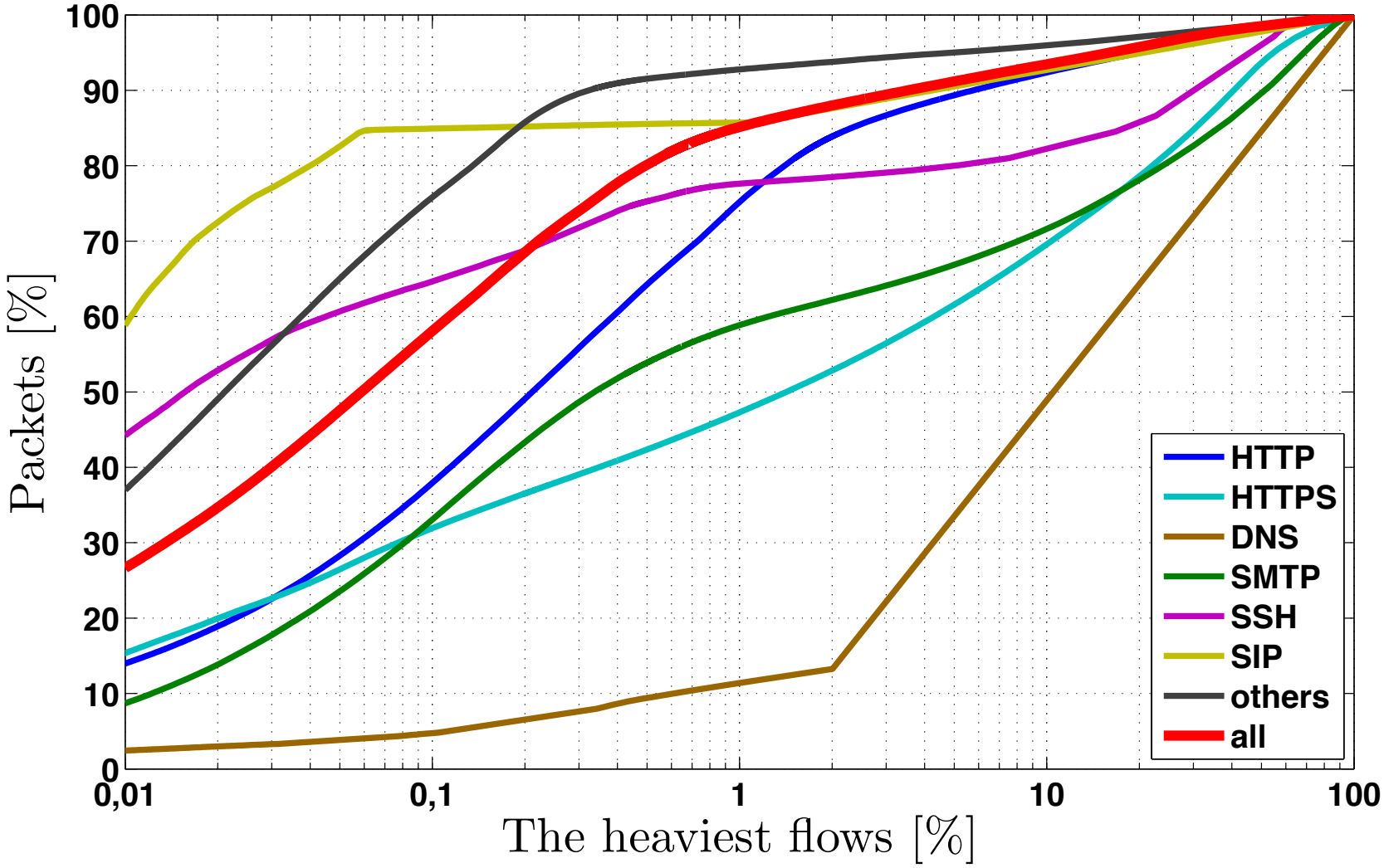
Pitfalls of 100G monitoring

- CPU limitation at 10-20 Mpps/core (basic NetFlow monitoring)
  - Statistics only up to transport layer (TCP, UDP)
  - L7 is more complex
- Complete hardware-based implementation of monitoring
  - Hardware processing of L7 is still research topic
  - Low flexibility
- Can we find a trade-off between these two?

Observation

- Network traffic has heavy-tail distribution
  - Large portion of traffic conveyed by small number of heavy flows
- Most of the traffic is irrelevant for L7 analysis
- By offloading NetFlow monitoring of small % of flows to **HW accelerator**, large % of packets are not sent to SW

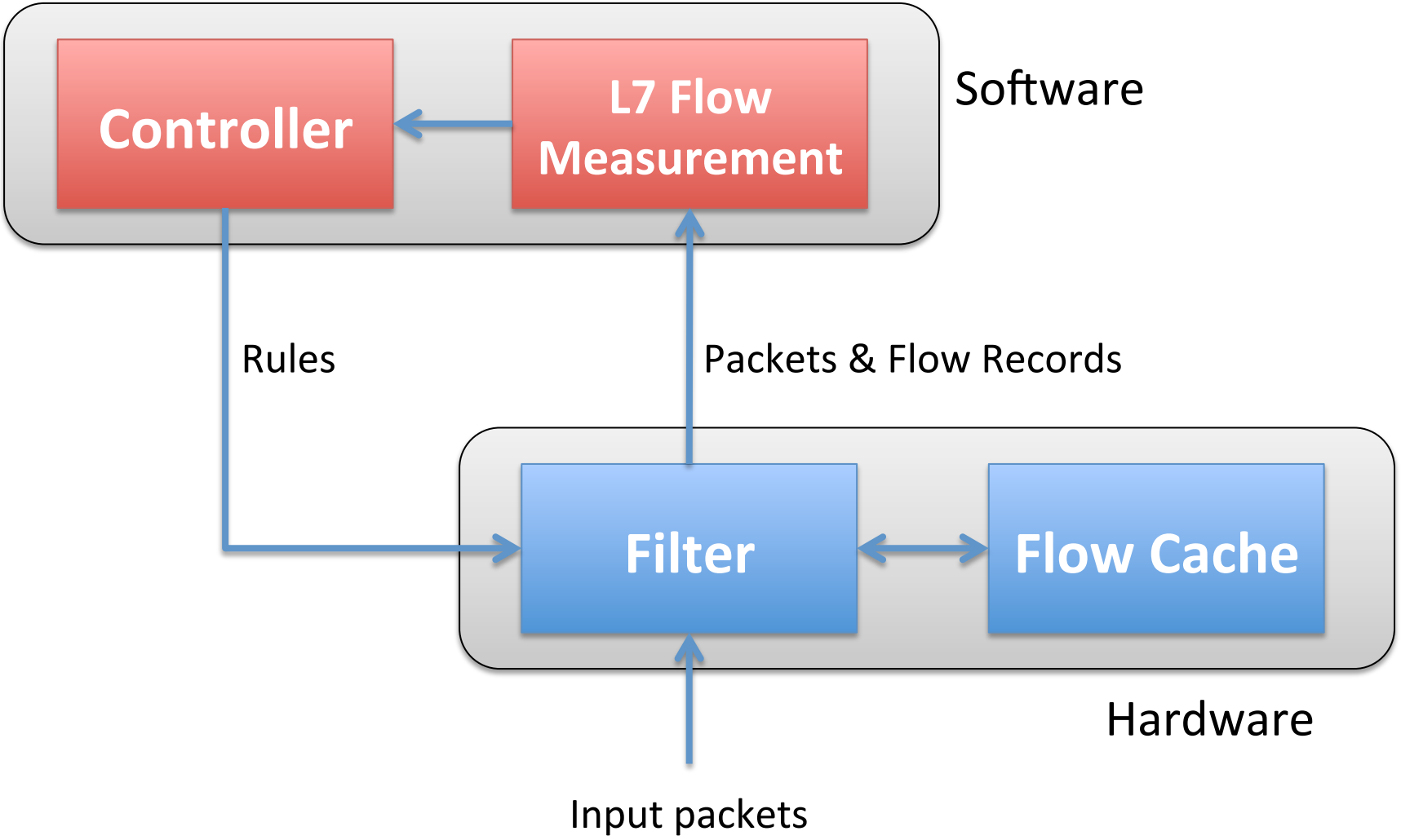
Heavy tail - graph



Design

- Hardware accelerator acts as advanced NIC
- New “unknown” flows sent to CPU
- **Software decides** what to do with each flow:
  - Software processing of interesting/suspicious traffic
  - Hardware NetFlow measurement of heavy and “uninteresting” flows
    - Which make up most of the traffic!
- Software plugins (C code with simple API)
  - No need to modify/know HW architecture

Scheme

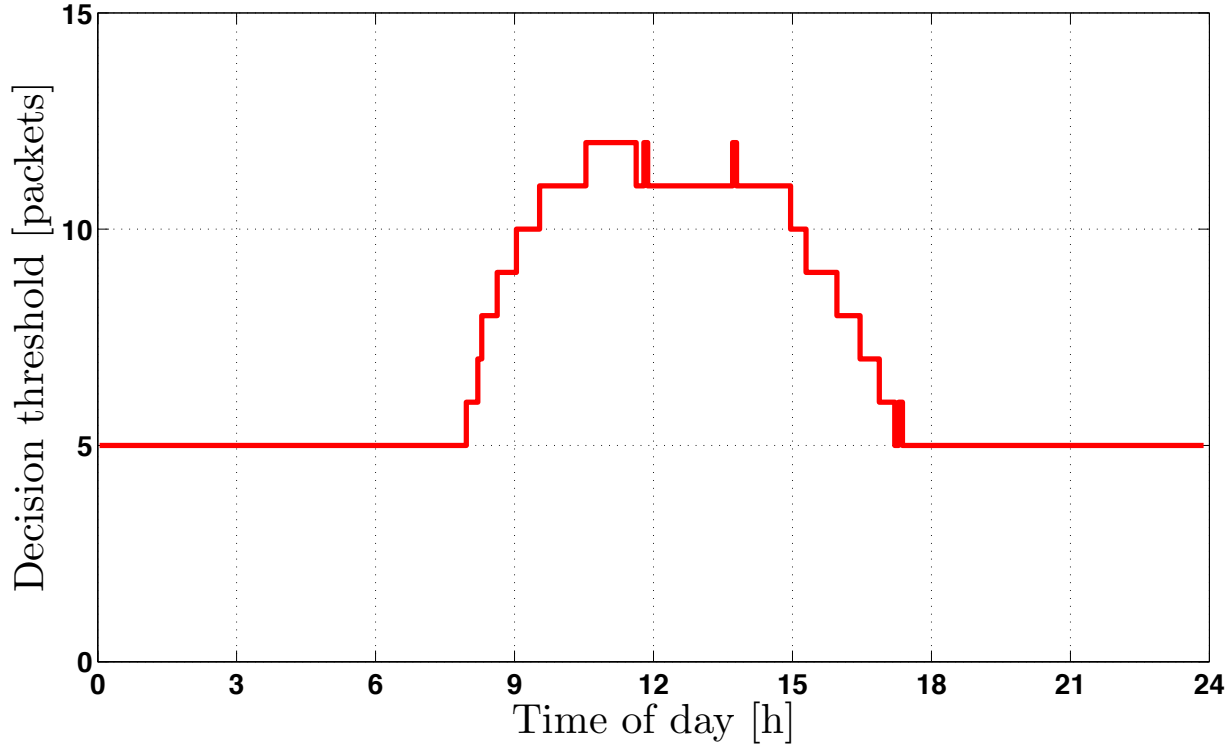

Controller
L7 Flow Measurement
Software
Rules
Packets & Flow Records
Filter
Flow Cache
Hardware
Input packets

Monitoring hardware

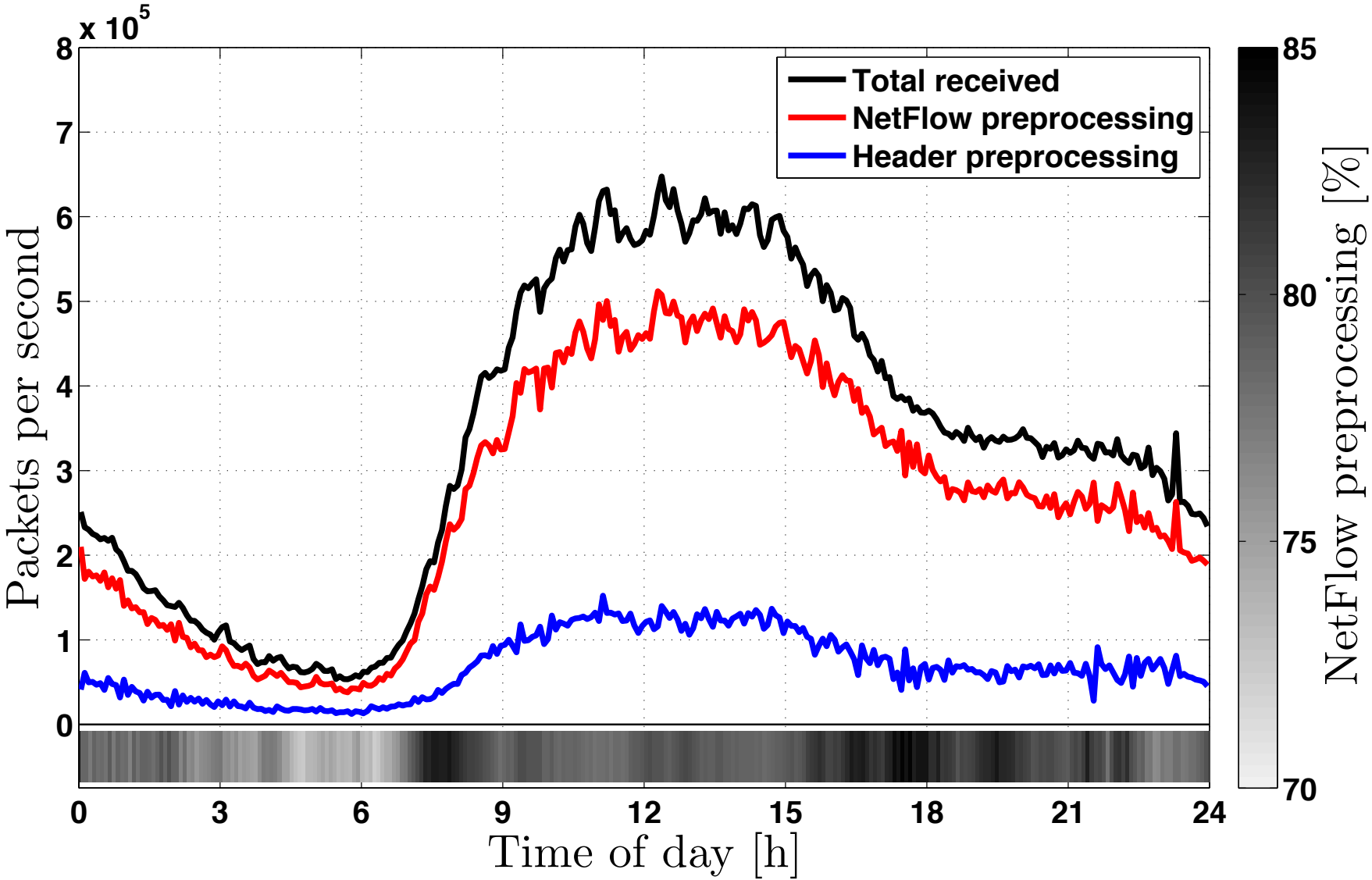
- Cards with Virtex-7 FPGAs
- COMBO-80G:
  - 2x QSFP+ optical module
  - PCI Express 3rd generation x8
- COMBO-100G:
  - CFP2 optical module
  - PCI Express 3rd generation x16
    - Up to 128 Gb/s

Heavy flow identification

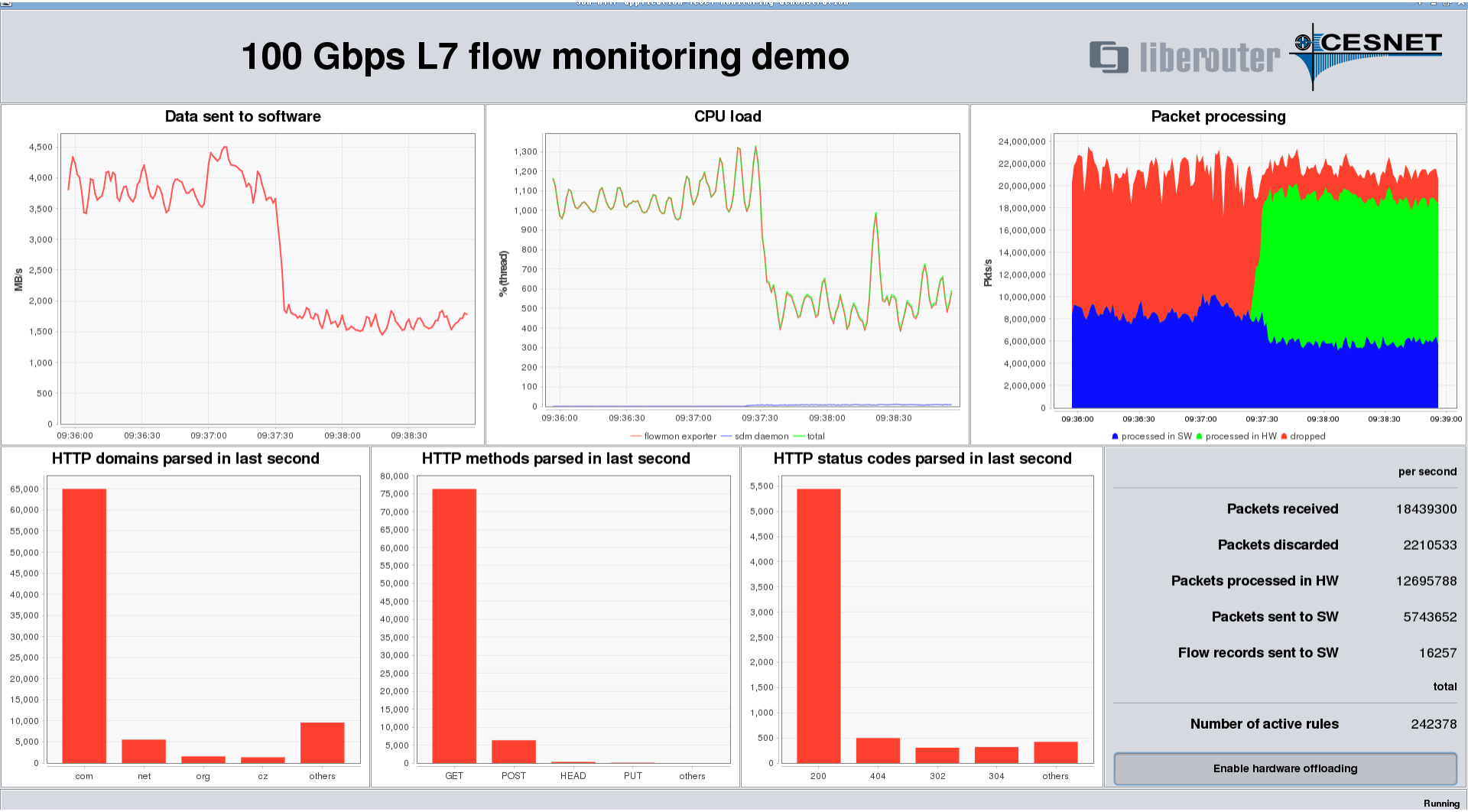
- HW flow cache size is limited, we can't offload all flows
- Flow with more than **X** packets is likely to be heavy
- Adaptive setting of **X** to keep HW flow cache utilized



Effect of offload

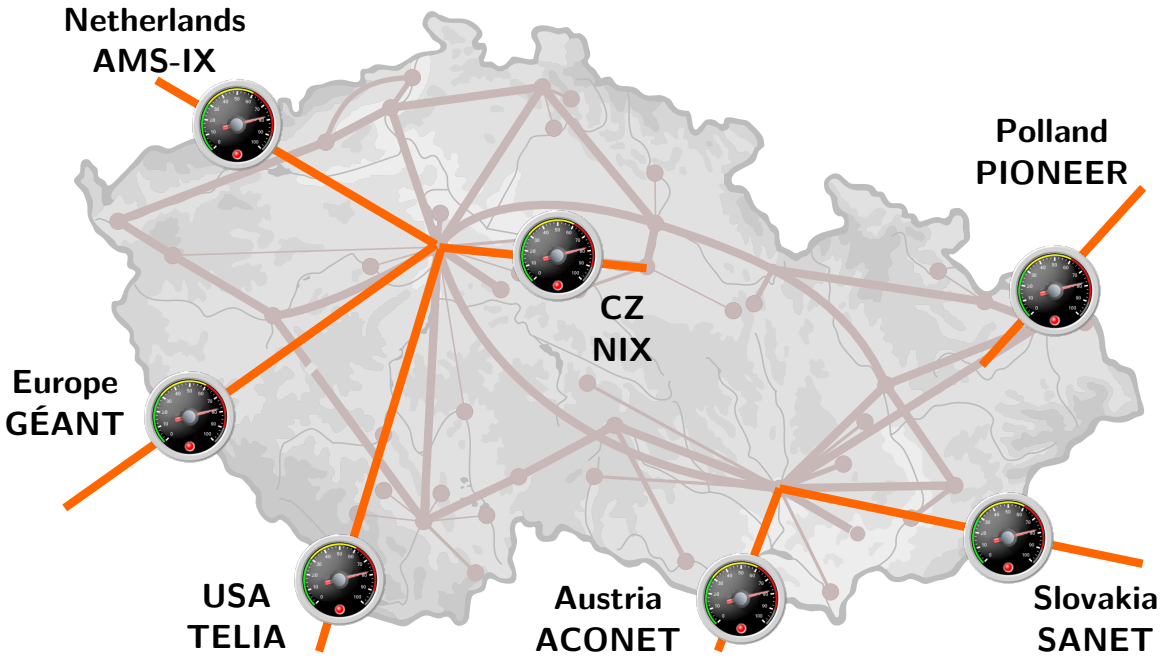


Bits-n-Bites live demo: HTTP



Our testbed

- National research and education network (400k users)
- Observation points at all external lines (+ datacenter)
- >200 GB of NetFlow data daily



Conclusion

- 100G traffic monitoring
  - **Flexible** through software plugins
  - **High-speed** with hardware offload of most traffic
- **Tight control feedback loop** between software controller and hardware accelerator
  - *Software Defined*, but not OpenFlow
- Implementation for 80G and 100G cards
- Deployment in CESNET network

Thank you