PDP : A Flexible and Programmable Data Plane

Massimo Gallo et al.
Introduction

Network Function evolution
Introduction
Network Function evolution

Can we build a fully modular and customizable stack (L2-L7) for NF deployment without sacrificing performance?
Click Router

Overview

- Elements that initiate packet flow have a **task** scheduled periodically
- *FromDevice* and *ToDevice* call push and pull within their tasks
PDP A Flexible and programmable data plane

Outline

ClickNF (CliMB)

vNS

FlowMon-DPDK
ClickNF: a Modular Stack for Custom Network Functions

Usenix ATC ’18 & HotMiddleboxes ’16 (& CCR 46 ‘16)
Click Router
Features & Limitations

- **Modularity**
- **Programmability**
- **Multithreading**
- **Kernel/User Space**
- **No Transport**
- **No Network Acceleration**
- **No Blocking I/O**
- **No Application (Modularity)**

Nokia Bell Labs
ClickNF Design

- Monolithic Network stacks:
  - Lack of flexibility
  - Hard to customize and debug

- ClickNF, decompose it into several simple elements:
  - Unfettered access to the network stack
  - Module reuse
  - Easier cross-layer optimization
ClickNF
1. High-Speed Packet I/O

DPDK MODULE
- 10/40 GbE cards
- Polling to fetch packets
- Wrap into Click packet and push

HARDWARE OFFLOADING
- Hardware flow control
- IP checksum offloading
- TCP checksum offloading

MULTICORE
- Receive Side Scaling
- Dedicated per-core data structure
ClickNF

2. Click TCP

• Full compliance with standards
• Support for:
  – TCP options
  – Congestion control
  – RTT estimation
• Composed of 60+ elements
• Elements modify TCP Control Block stored in TCPInfo
• Packet annotations and flags
ClickNF
3. Blocking I/O

- Click natively supports nonblocking I/O
  - Complex logic for applications
- Blocking I/O
  - Simpler, equivalent performance
  - Blocking for socket I/O multiplexing
- ClickNF introduces **blocking tasks**
  - If I/O request cannot be completed, task **yields** CPU
  - Task resumes → illusion of continuity
  - Backwards compatible
ClickNF
4. Application modularity

• ClickNF provides 4 building blocks:
  – TCPEpollServer
  – TCPEpollClient
  – SSLServer
  – SSLClient
• Application layer elements use packet annotations to control lower layers
• Simplify NF programming and code reuse
ClickNF

4. Application modularity

• ClickNF provides 4 building blocks:
  – TCPEpollServer
  – TCPEpollClient
  – SSLServer
  – SSLClient

• Application layer elements use packet annotations to control lower layers

• Simplify NF programming and code reuse
ClickNF

Experimental results

• **Hardware:**
  – 3 machines Intel Xeon® 40-core E5-2660 v3 2.60GHz, 64 GB RAM
  – Intel Xeon® 82599ES NICs with two 10 GbE interfaces

• **Software:**
  – Ubuntu 16.10 (GNU/Linux 4.4.0-51-generic x86 64)
  – Click 2.1
  – DPDK 17.02, DPDK Traffic generator

• **Tests:**
  – Microbenchmarks (Packet I/O, modularity, HW offloading, Performance)
  – Applications (HTTP Cache/Server, SOCKS4 proxy)
ClickNF
Performance

• Small TCP connection (64 bytes packet).
  Echo request <-> Echo reply
  – Slightly better than state of the art user-space stack
CliMBOS: A Modular NFV Cloud Backend for the Internet of Things

NMTS ’18 & Demo @ Nokia open days ‘17
• Current IoT backends relies on cloud-based aggregation model
• A better approach would be to have user-centric IoT backends

• We need a large number of NFs while ensuring fast spawn, isolation and high performance
• Not feasible with today’s NFV solutions, i.e. full fledged VM/containers
• We propose CliMBOS that allows high-density and VM-isolated deployment of network functions with both flexibility and performance
FlowMon-DPDK: Parsimonious per-flow software monitoring at line rate

TMA ’18 & Demo @ SIGCOMM ‘18
FlowMon-DPDK

Design

• Experimental network devices are usually validated using expensive dedicated HW or custom SW

• We designed FlowMon-DPDK
  – Accurate per-flow statistics
  – Limited resources
  – Deployed inside the device under test
vNS: a modular programmable virtual network switch

Poster @ SIGCOMM ‘18
Network stacks are typically implemented as monolithic packages:
- Inflexible, hard to extend and debug
- Agnostic to user needs and application requirements

Tenants can interfere with network stack behavior

Building on ClickNF, vNS Introduces programmability in the virtual network switch by allowing complete network stack delegation
Conclusion

• **ClickNF** - Usenix ATC ’18 & HotMiddleboxes ‘16
  – good scaling properties with increased flexibility
  – Open source [https://github.com/nokia/ClickNF](https://github.com/nokia/ClickNF)

• **CliMBOS** - NTMS ‘18 & DEMO @ Nokia Open days ‘17
  – Isolation through microkernel

• **FlowMon-DPDK** - TMA ’18 & Demo @ SIGCOMM ‘18
  – Per-flow statistics
  – High performance with minimal amount of resources

• **vNS** - Poster @ SIGCOMM ‘18
  – Network stack delegation
  – Ongoing