P4 and IO Visor for Building Datacenter Infrastructure Components

Brenden Blanco
Architect - PLUMgrid
Nov. 18, 2015

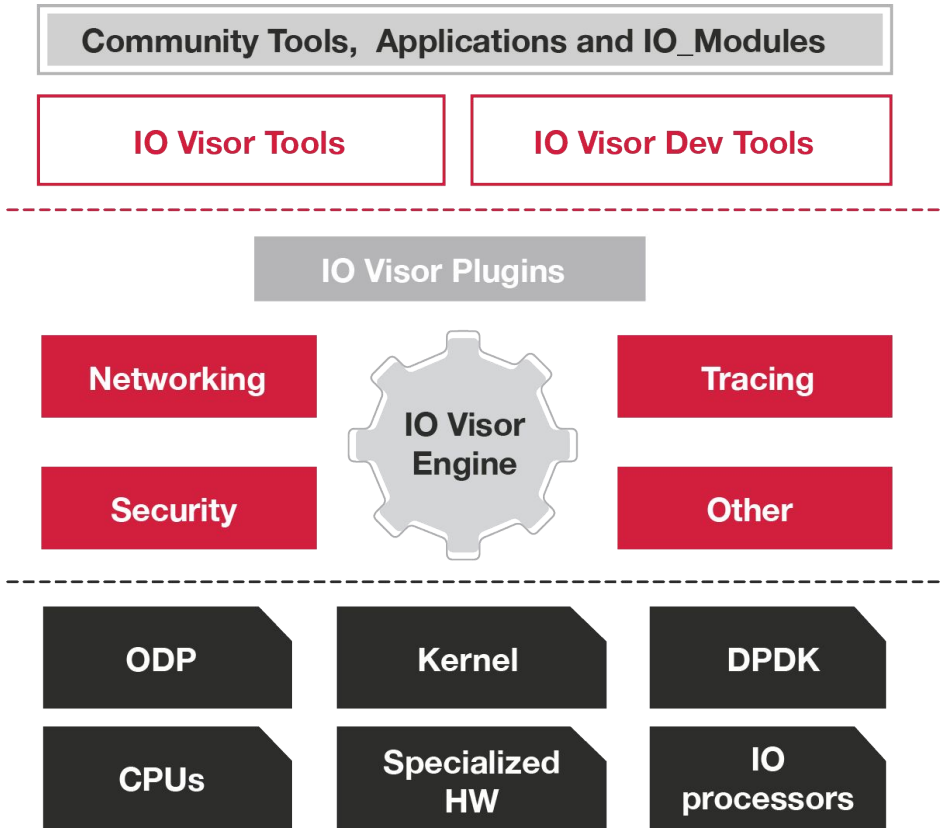
Founding Members

THE LINUX FOUNDATION

Motivation

- Modern flexible datacenter is leveraging programmability in all levels of network hierarchy
- Applications at the network edge create demand for high velocity feature implementation
- Linux based hypervisors and container hosts are fertile ground for quickly developing such features
- but...
- Linux kernel programming is hard

IO Visor Project, What is in it?



- A set of development tools, **IO Visor Dev Tools**
- A set of **IO Visor Tools** for management and operations of the IO Visor Engine
- A set of Applications, Tools and open **IO Modules** build on top of the IO Visor framework
- A set of possible use cases & applications like **Networking, Security, Tracing & others**

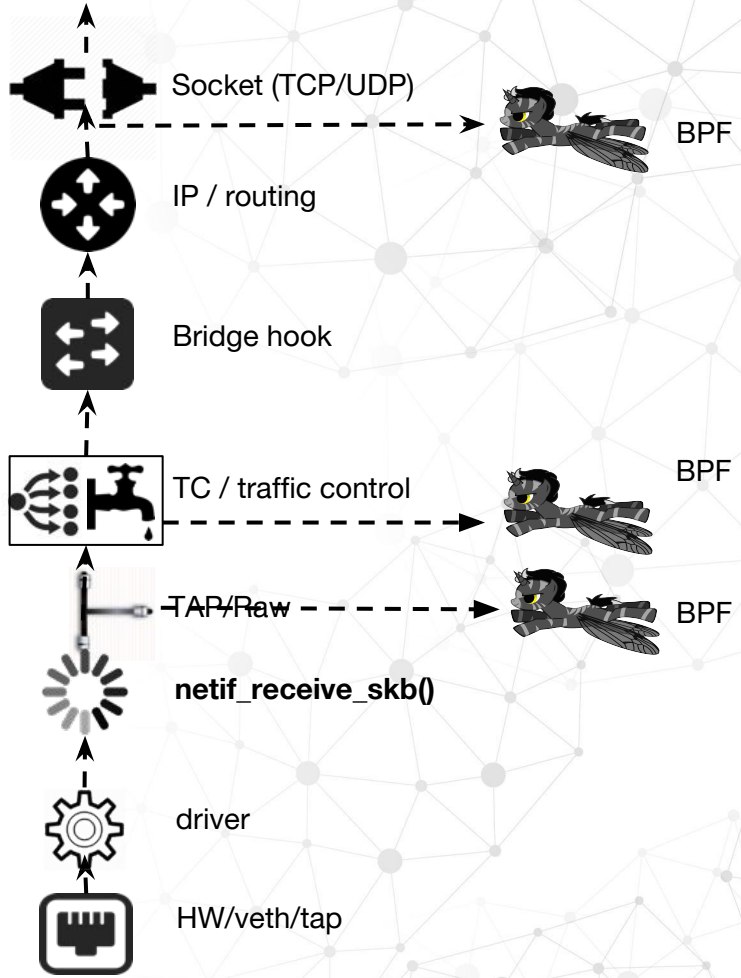
eBPF: Loading New Modules

BPF program written in C

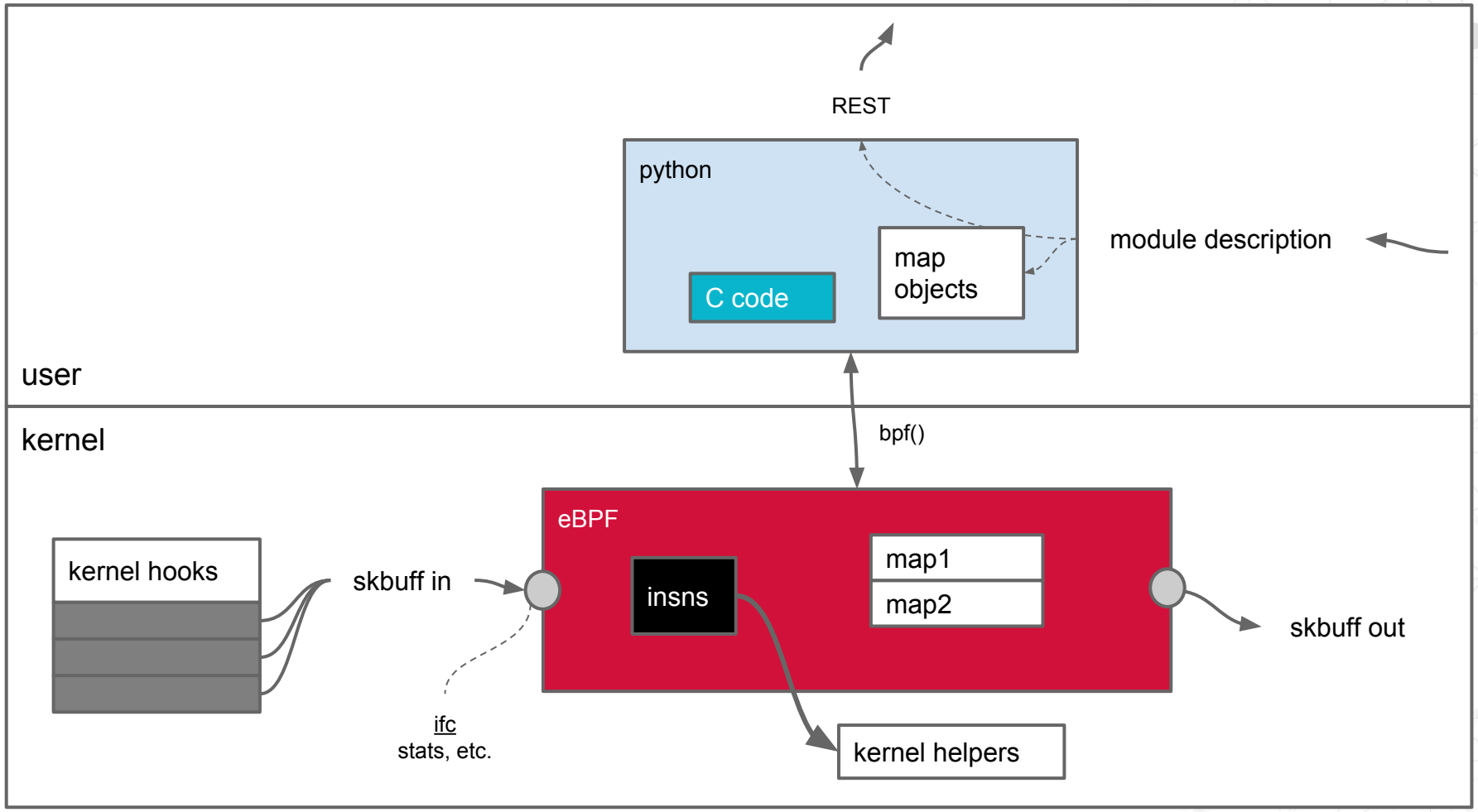
Translated into eBPF instructions (LLVM)

Loaded in kernel
JIT to native instructions

Hooked at different levels of Linux Networking Stack



Components of an IO Module



The eBPF Instruction Set - In Kernel

Instructions

- 10x 64bit registers
- small stack
- 1-8B load/store
- conditional jump
- arithmetic
- function call

Helper functions

- forward/clone/drop packet
- load/store packet data
- load/store packet metadata
- checksum (incremental)
- push/pop vlan
- map lookup/update/delete
  - hash, array, others in future

All upstreamed in Linux Kernel

P4 to eBPF using BCC

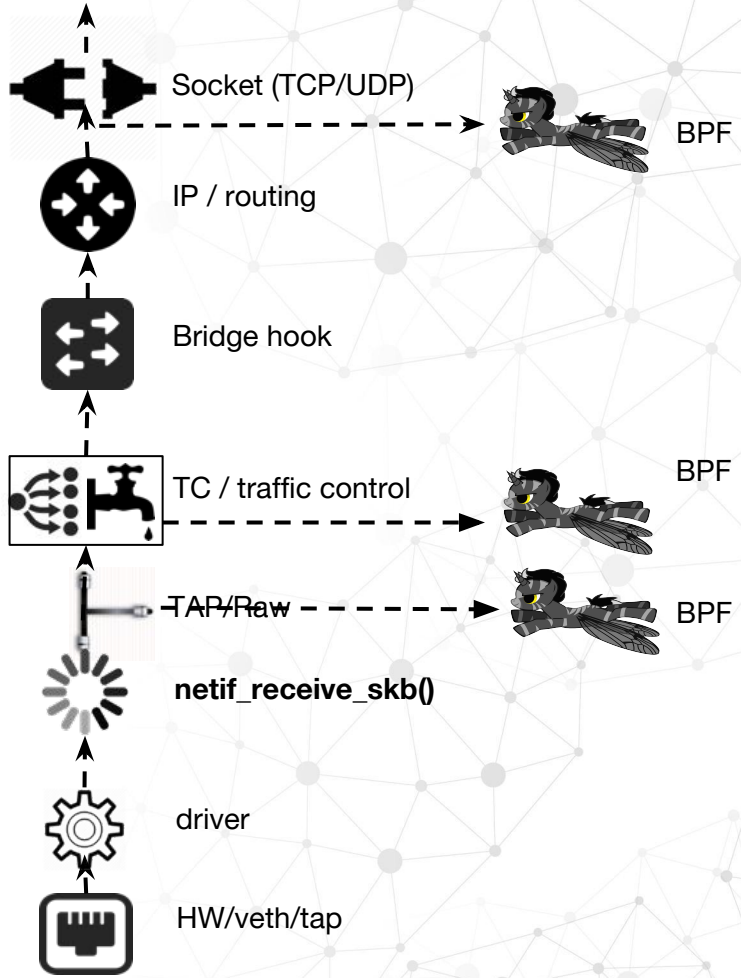
P4 Program

BPF program written in C

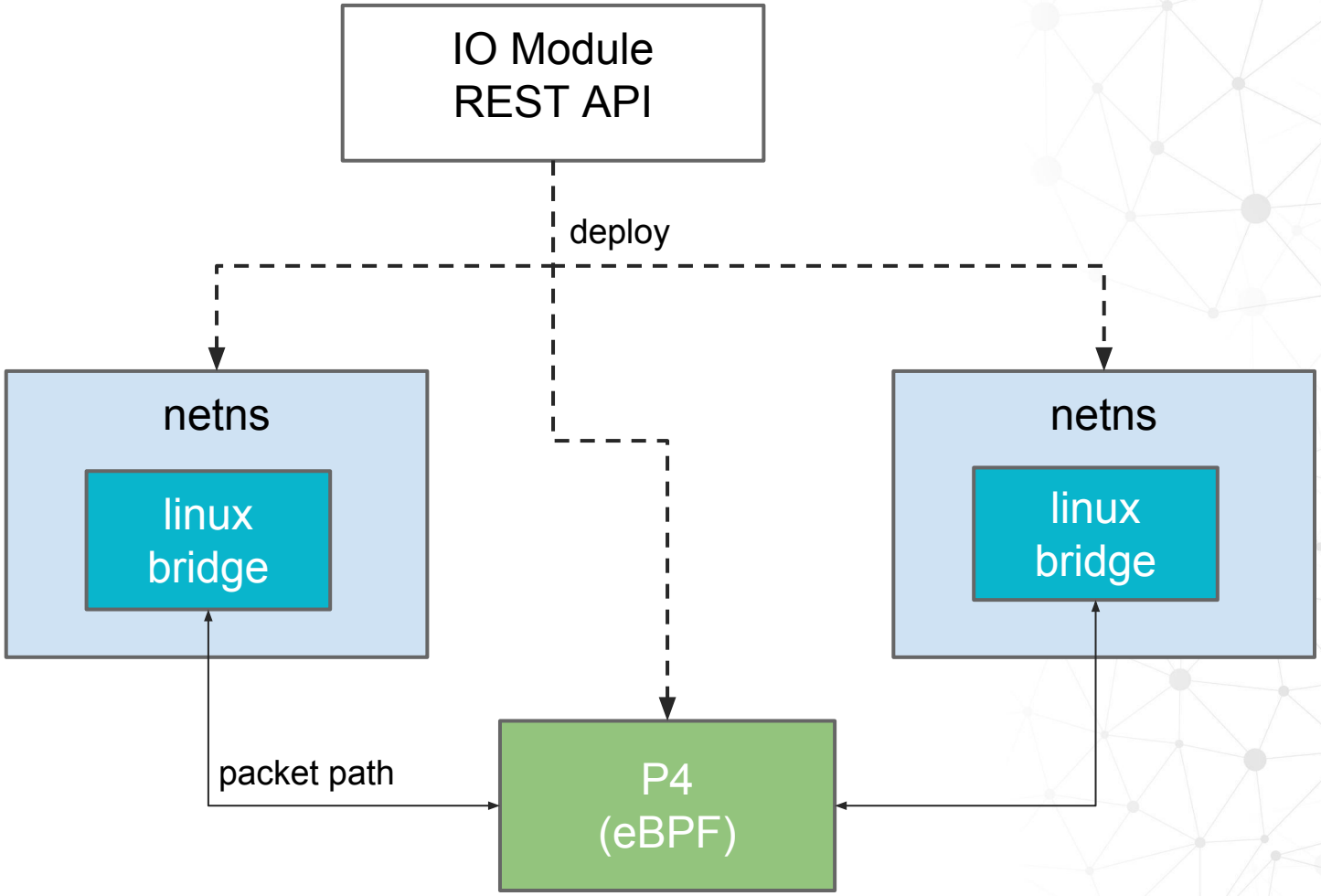
Translated into eBPF instructions (LLVM)

Loaded in kernel
JIT to native instructions

Hooked at different levels of Linux Networking Stack



Demo


IO Module
REST API
deploy
netns
linux
bridge
netns
linux
bridge
packet path
P4
(eBPF)

TODO

- Language features
  - Checksum
  - Add/Remove header
  - Ternary/LPM table support
  - Counters/Meters
  - Many other primitive actions
  - v1.1
- Remove C translation step
  - Direct P4 -> LLVM IR
- API for control plane table access
- Write applications!

Learn More and Contribute

- https://iovisor.org
- https://github.com/iovisor
- #iovisor irc.oftc.net
- @IOVisor

Questions