Improve Performance of Kube-proxy and GTP-U using VPP

Hongjun Ni (hongjun.ni@intel.com)

Danny Zhou (danny.zhou@intel.com)

Johnson Li (johnson.li@intel.com)

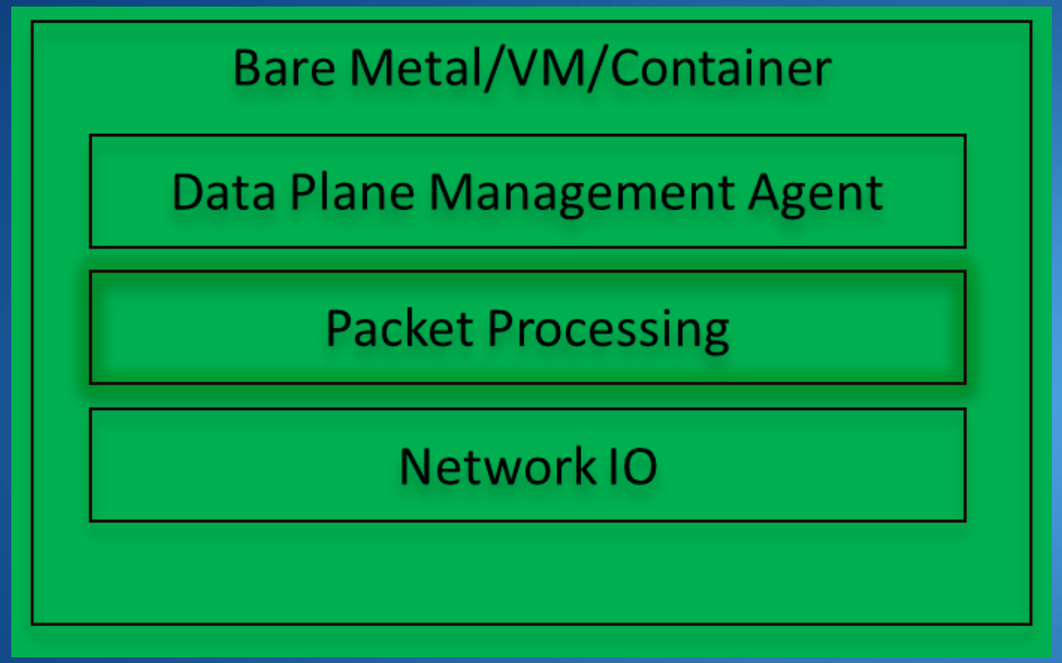
Network Platform Group, DCG, Intel

Acknowledgement: Jianfeng Tan

Agenda

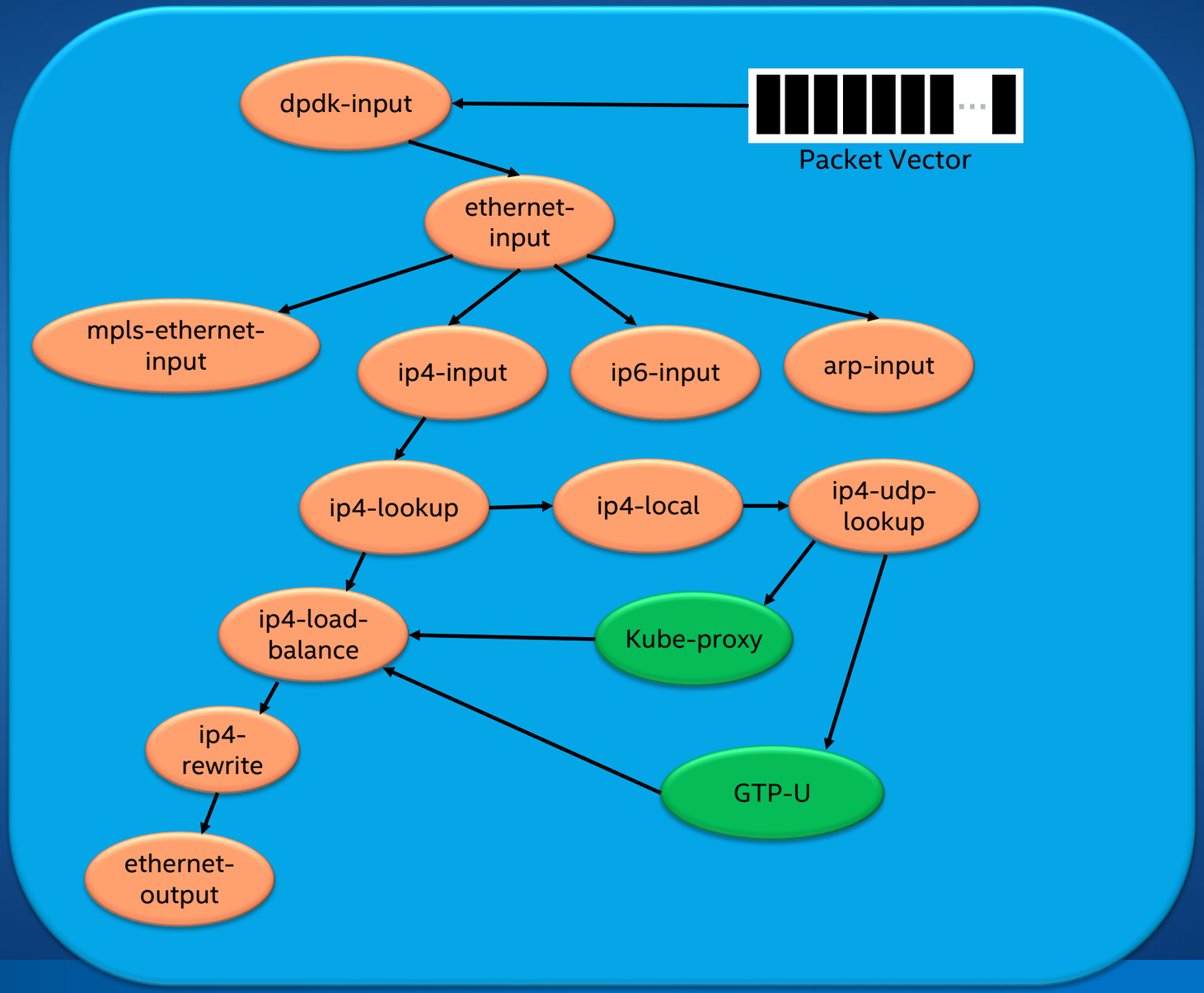
- Enabling high performance kube-proxy
- Six ways to improve the performance of GTP-U
- Key takeaway

Introducing Vector Packet Processor



- Packet processing development platform
- Runs on commodity CPUs and leverages DPDK
- Runs as a Linux user-space application

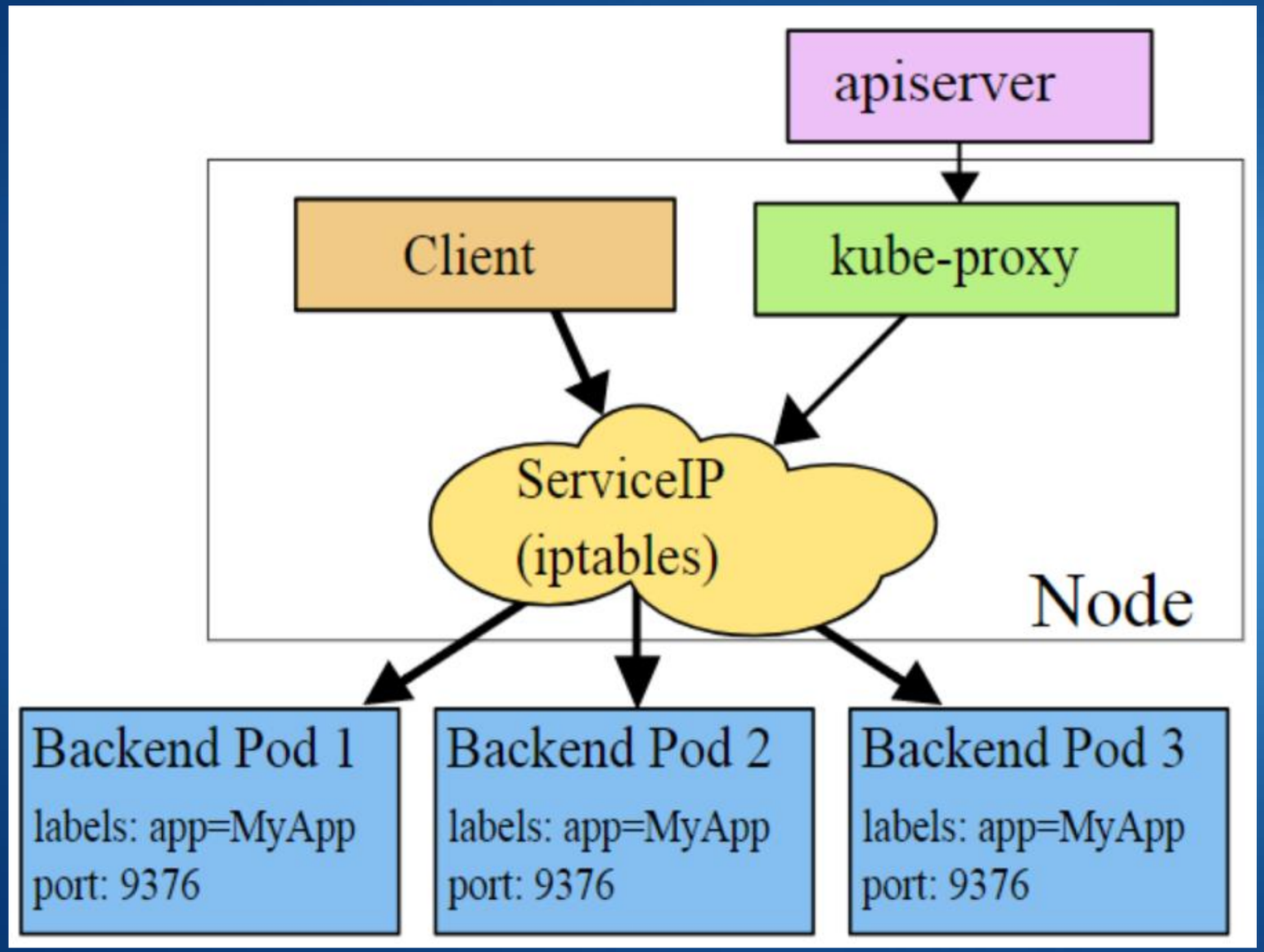
VPP Graph Node and Plugin Framework



Enabling high performance kube-proxy

- Introducing kube-proxy
- Current pain point
- Proposed kube-proxy solution
- Kube-proxy dataplane
- GoVPP
- Performance

Introducing kube-proxy



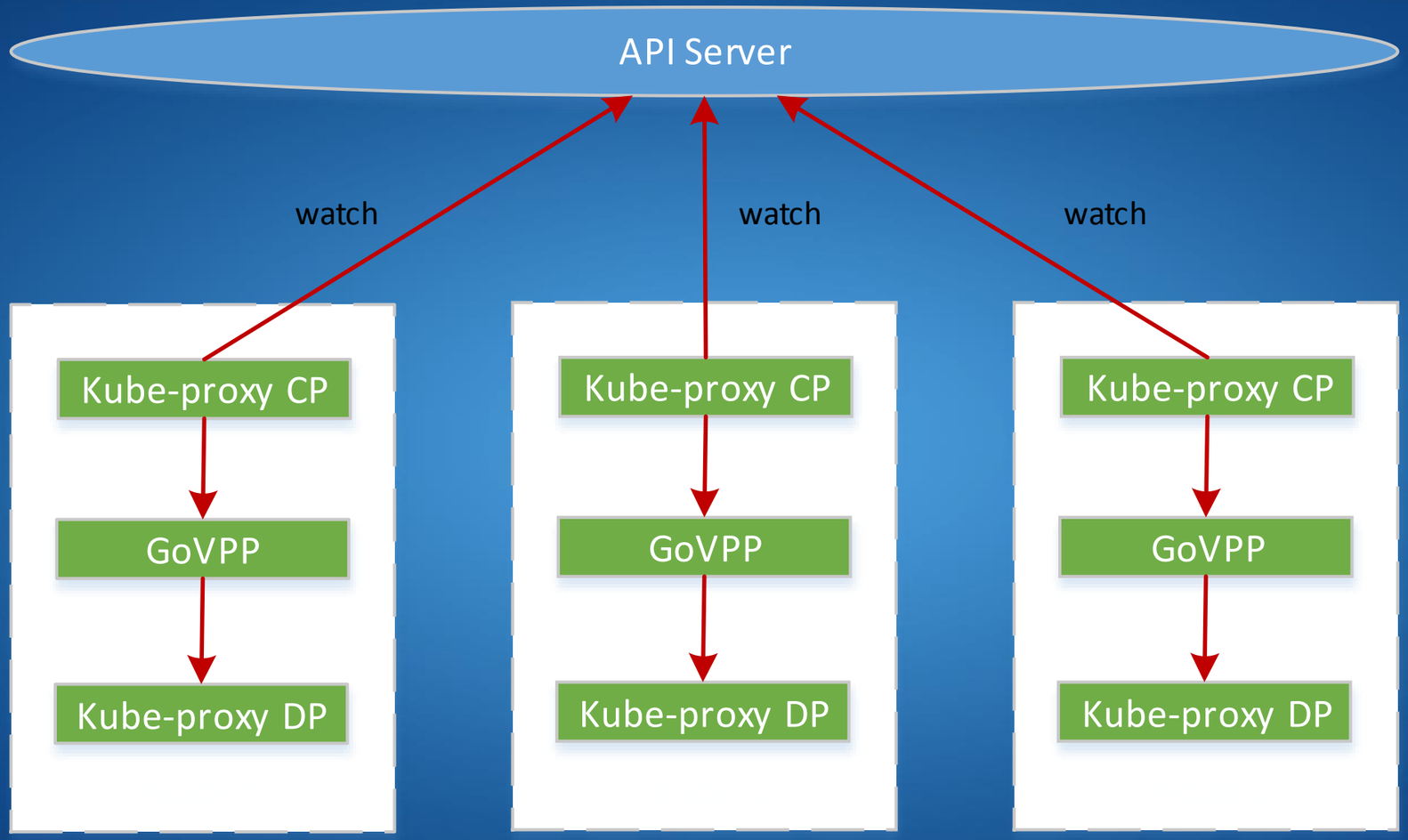
- Watches addition and removal of Service and Endpoints.
- Installs iptables rules
- Captures traffic and select Pod
- Redirects traffic

Reference: https://kubernetes.io/docs/concepts/services-networking/service/

Current Pain Point

- Supports userspace and iptables
- Uses kernel iptables NAT
- Performance degrades when service/endpoint pairs increase

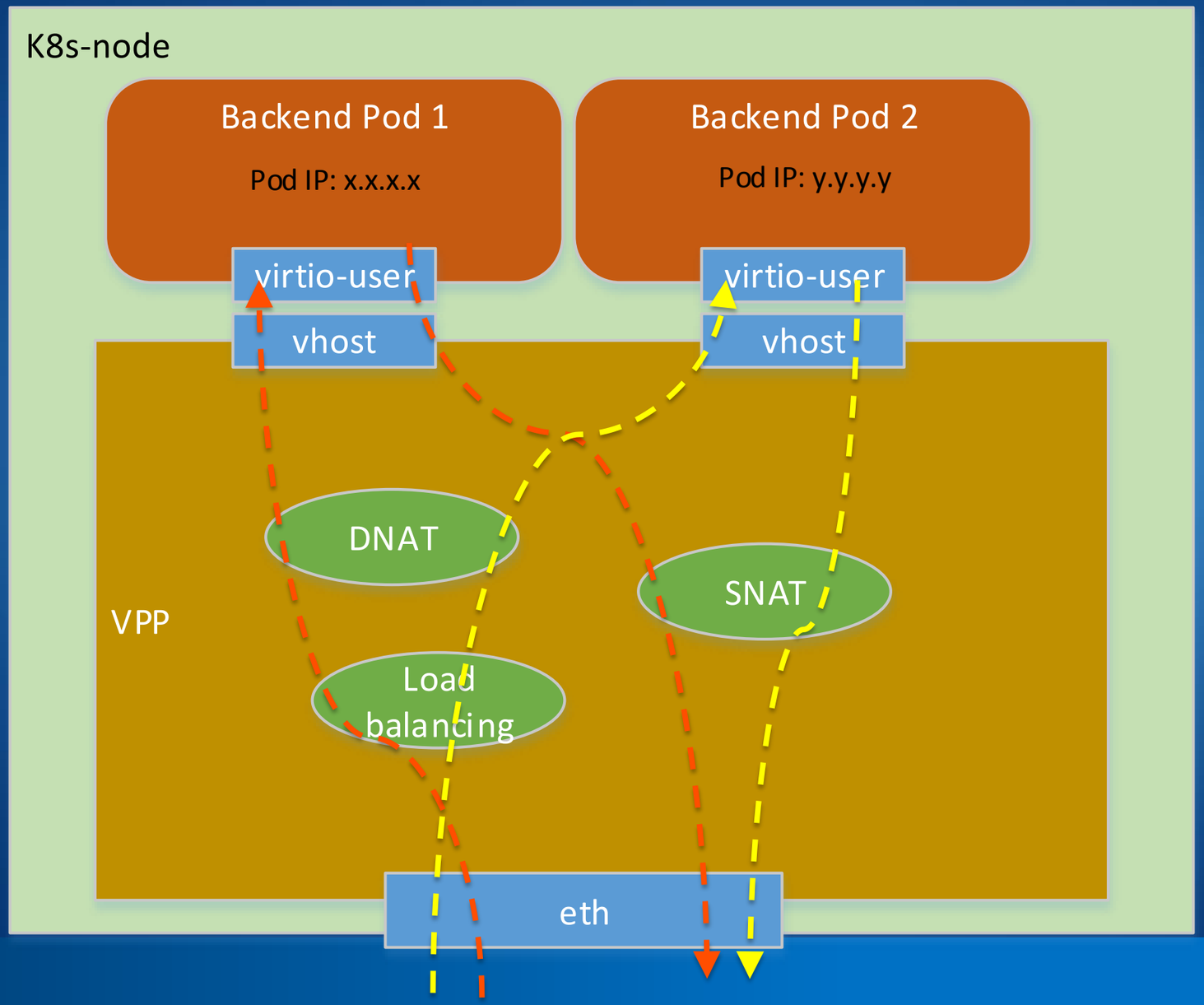
Proposed kube-proxy solution



CP: Control Plane

DP: Data Plane

Kube-proxy Data Plane



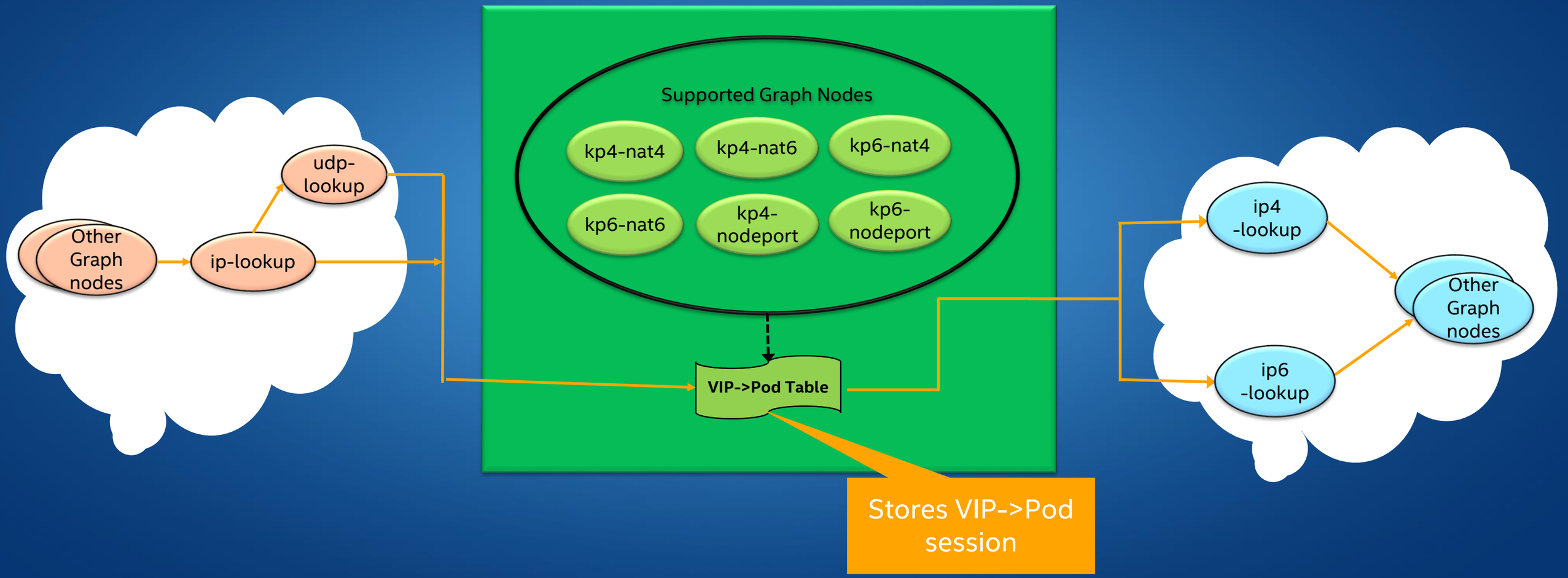
- Distribute traffic evenly
- Per flow stick to Pod
- Three service types:

  (1). ClusterIP: Port

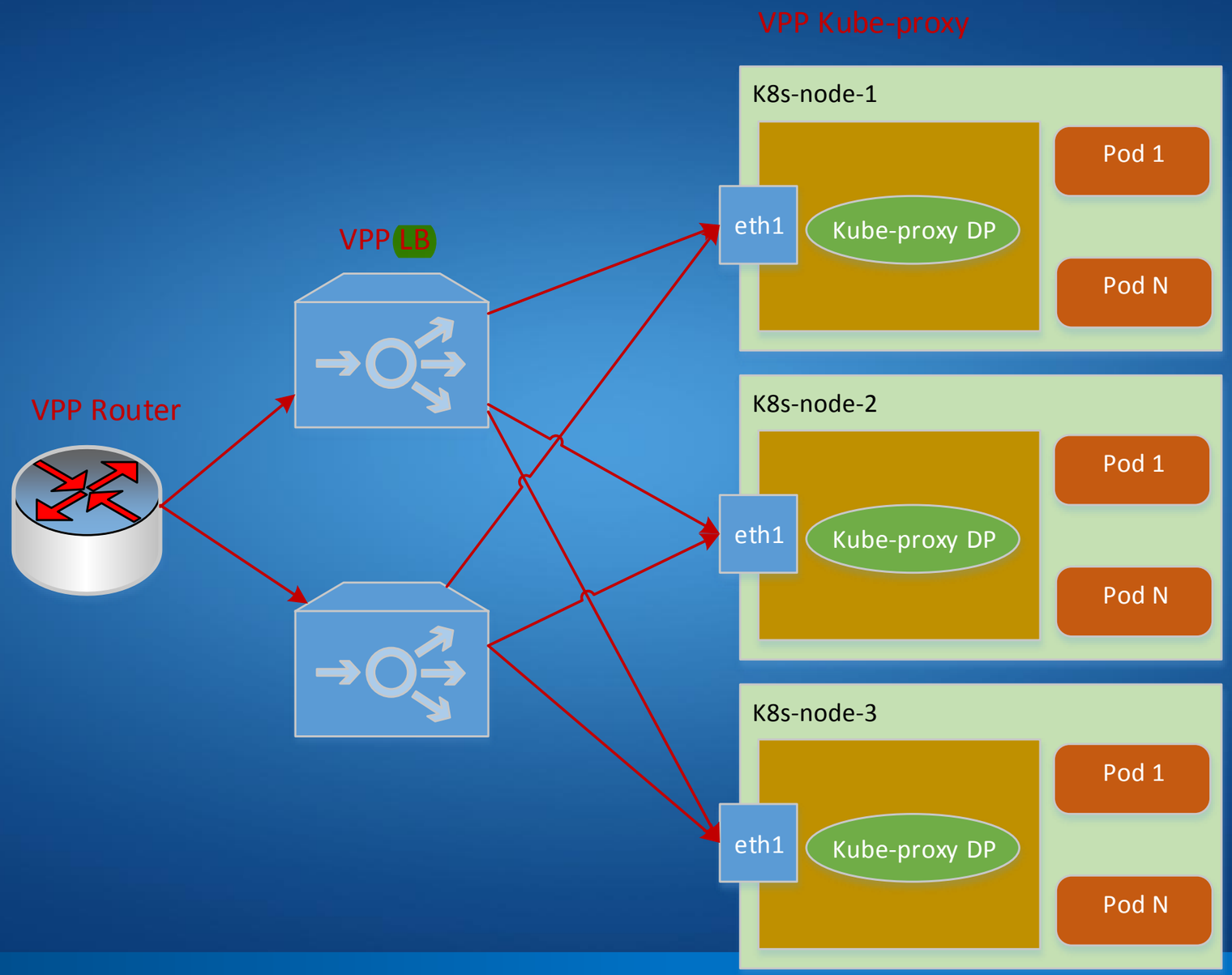
  (2). NodeIP: NodePort

  (3). External LoadBalancer

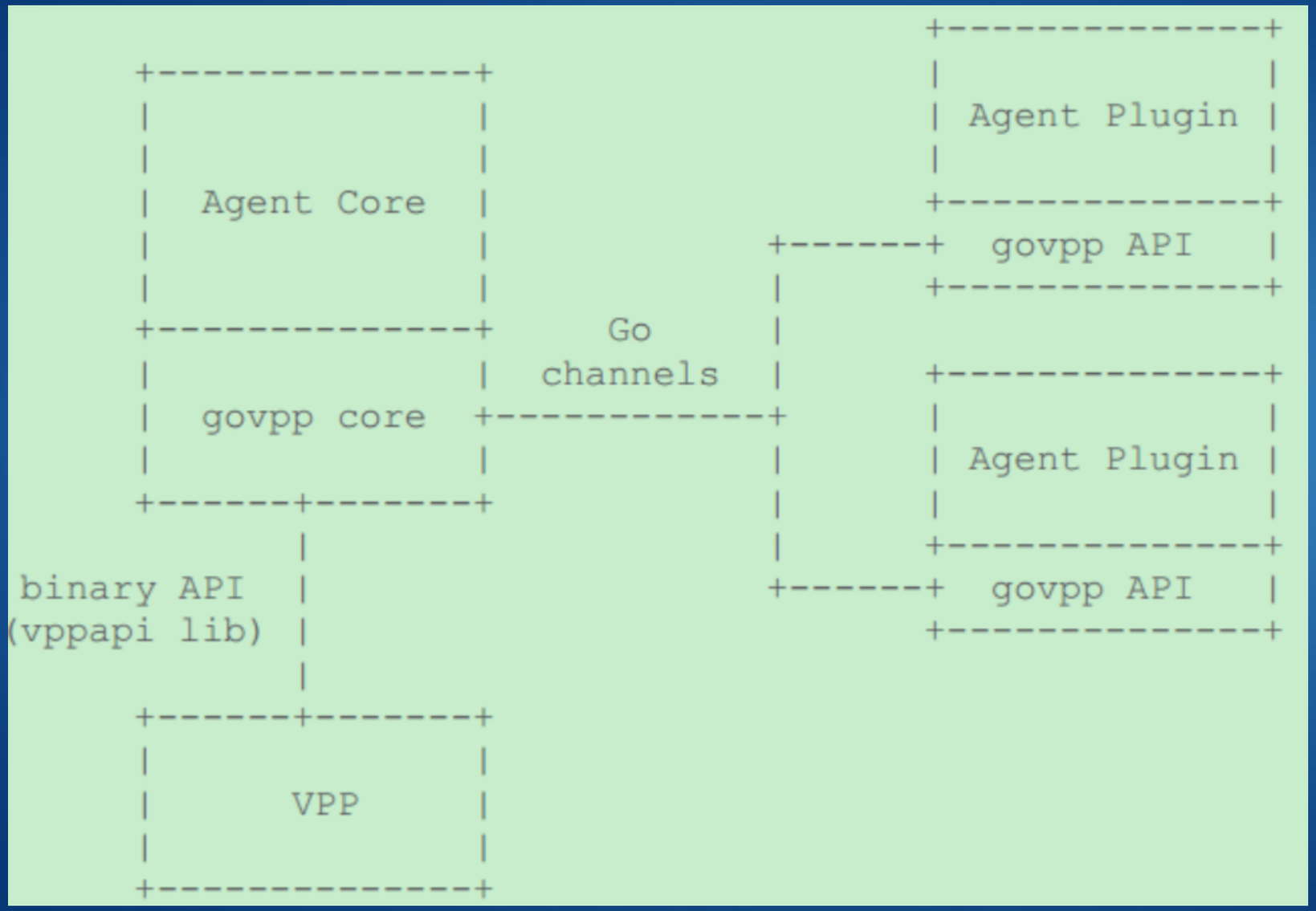
Kube-proxy Data Plane Internals



Kube-proxy: External LoadBalancer



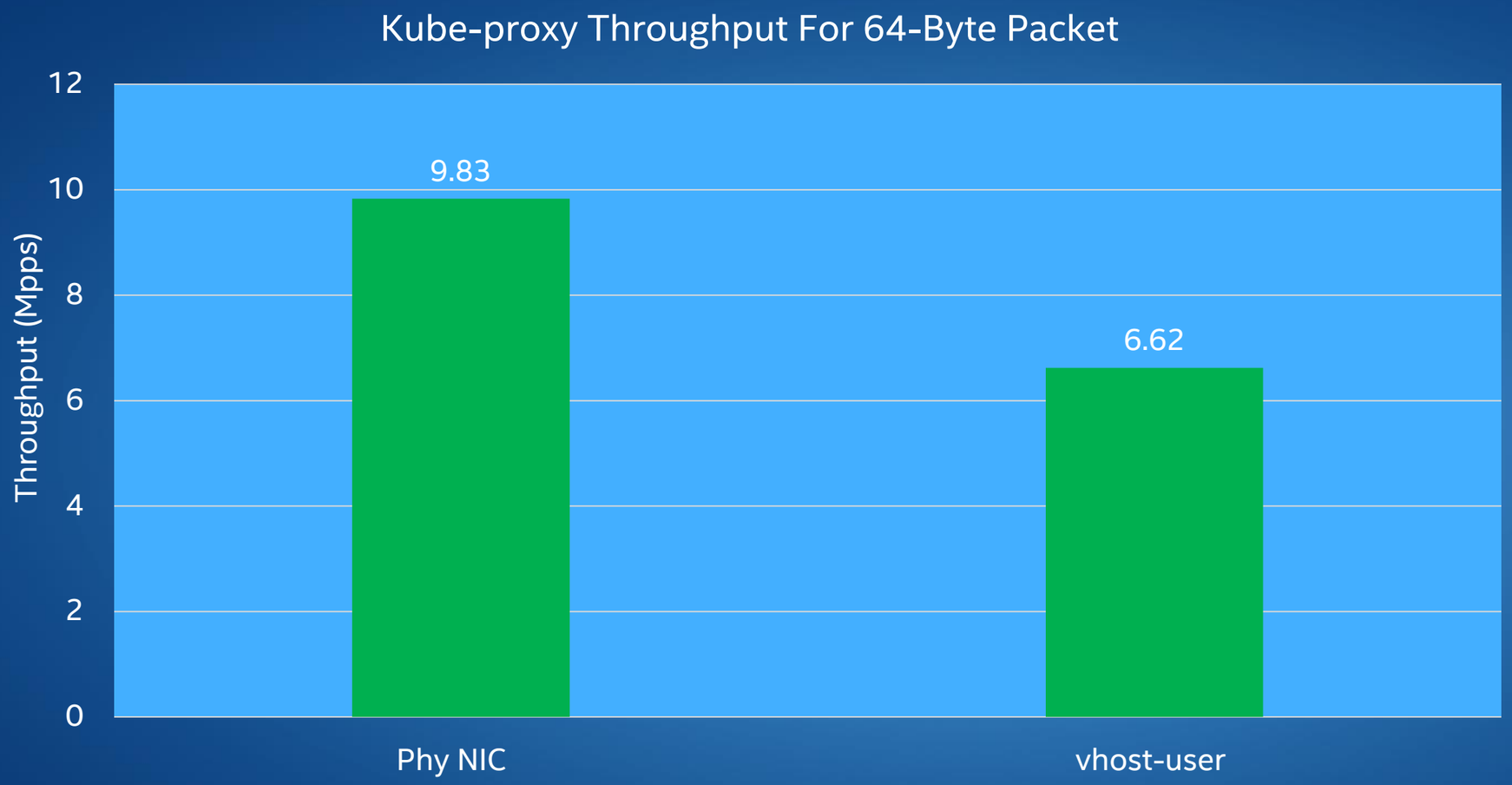
GoVPP



- Golang toolset for VPP management
- VPP binary API (JSON) → Go Structure
- Handle 250,000 binary API requests per second

Reference: https://wiki.fd.io/images/f/fa/GoVPP-intro.pdf

Performance



- Linux iptables perf: < 400 kpps
- Scaling

Test Case	Input packet size (bytes)	Output packet size (bytes)
Load balance + DNAT + Routing	64	64

Reference: https://people.netfilter.org/kadlec/nftest.pdf

intel

Six Ways to Improve Performance of GTP-U

- Cache table lookup result
- Bypass second ip-input
- Bypass first route lookup
- Dual-loop and Quad-loop
- Packet prefetching
- Bypass second route lookup

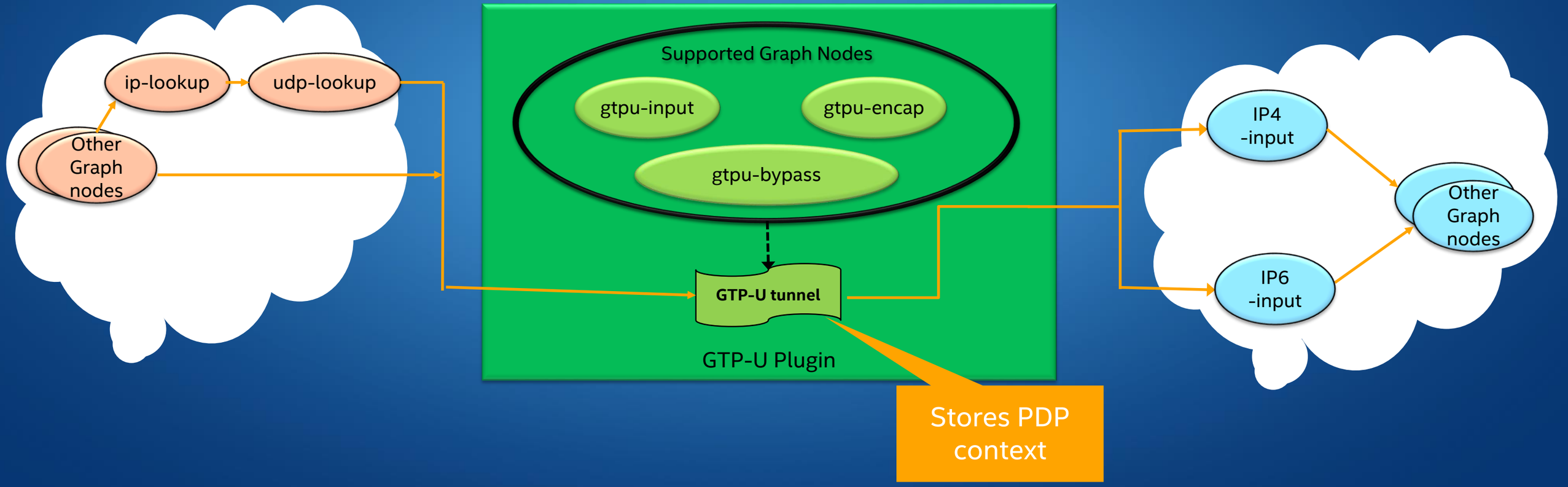
☒缓存表查找结果
☒绕过第二个IP输入
☒绕过第一条路线
☒双回路和四回路
☒数据包预取
☒绕过第二条路线

GTP-U Plugin Internals

- Typical GTP-U packet processed by VPP and GTP-U plugin

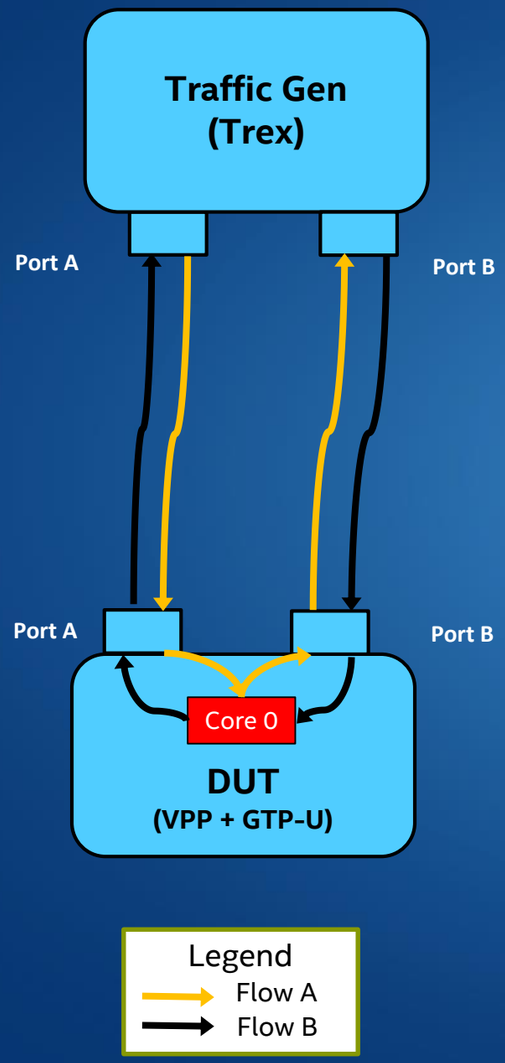
Outer MAC header	Outer IP header	UDP header	GTP-U header	Inner IP header	L4 header	Payload

- GTP-U Plugin



Device Under Test for Performance

- Network Topology



- Hardware Configuration

CPU	Intel(R) Xeon(R) CPU E5-2699 v3 @ 2.30GHz
DIMM	2133 MHz, 64GB Total
NIC	2x 82599ES 10-Gigabit SFI/SFP+ Network Connection
PacketGen	Ixia* 10 Gigabit Ethernet Traffic Generator

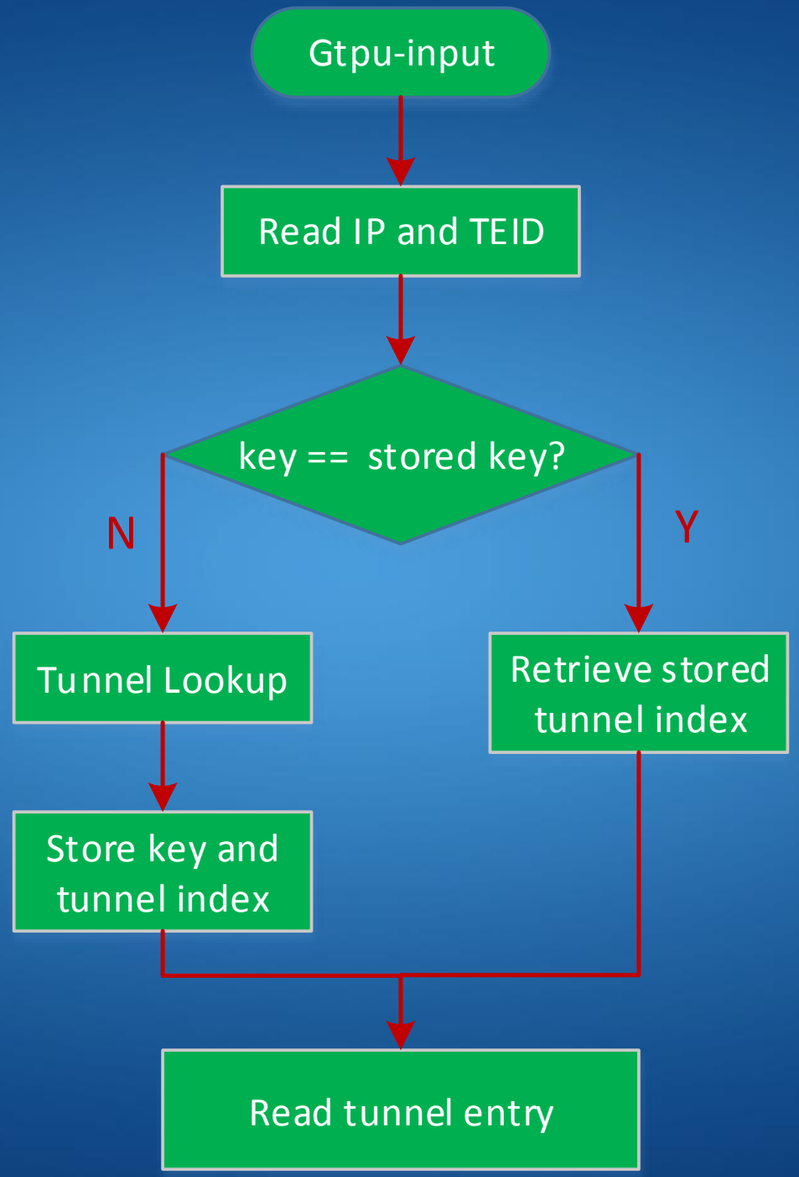
- Software Configuration

OS	Ubuntu 16.04.2 LTS
Kernel	Linux version 4.4.0-62-generic
DPDK	17.08
VPP	17.10-rc0
GTP-U	17.10-rc0

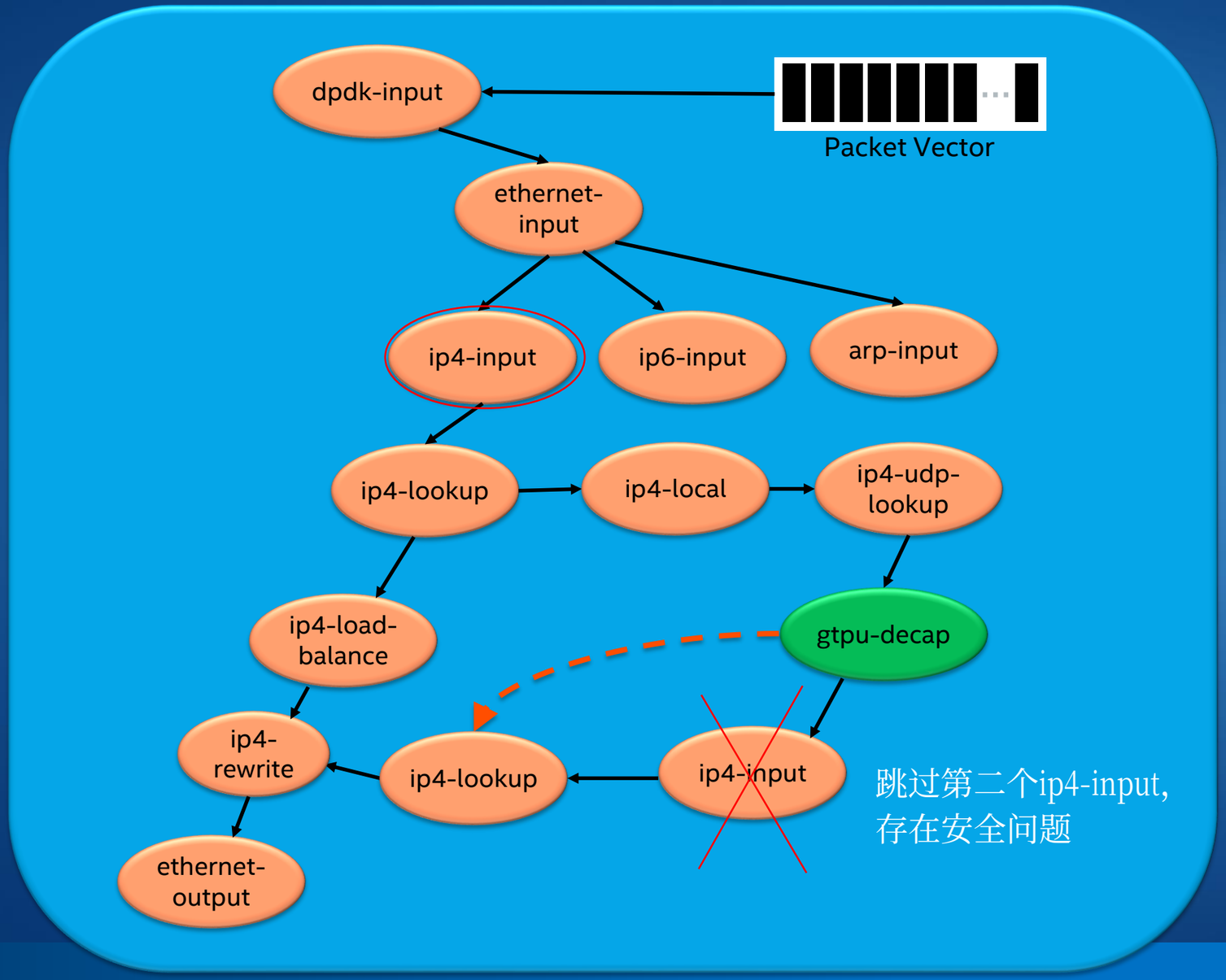
- BIOS Configuration

Enhanced Intel Speedstep	Enabled
Turbo Boost	Enabled
Processor C3	Disabled
Processor C6	Disabled
Hyper-Threading	Disabled
Intel VT-d	Enabled
CPU Power and Performance Policy	Performance
Memory Freq.	2133 MHz
Total Memory Size	64 GB
Memory RAS and Performance Configuration -> NUMA Optimized	ENABLED
QPI B/W	9.6 GT/s
MLC Streamer	ENABLED
MLC Spatial Prefetcher	ENABLED
DCU Data Prefetcher	ENABLED
DCU Instruction Prefetcher	ENABLED
Direct Cache Access (DCA)	ENABLED

Cache table lookup result

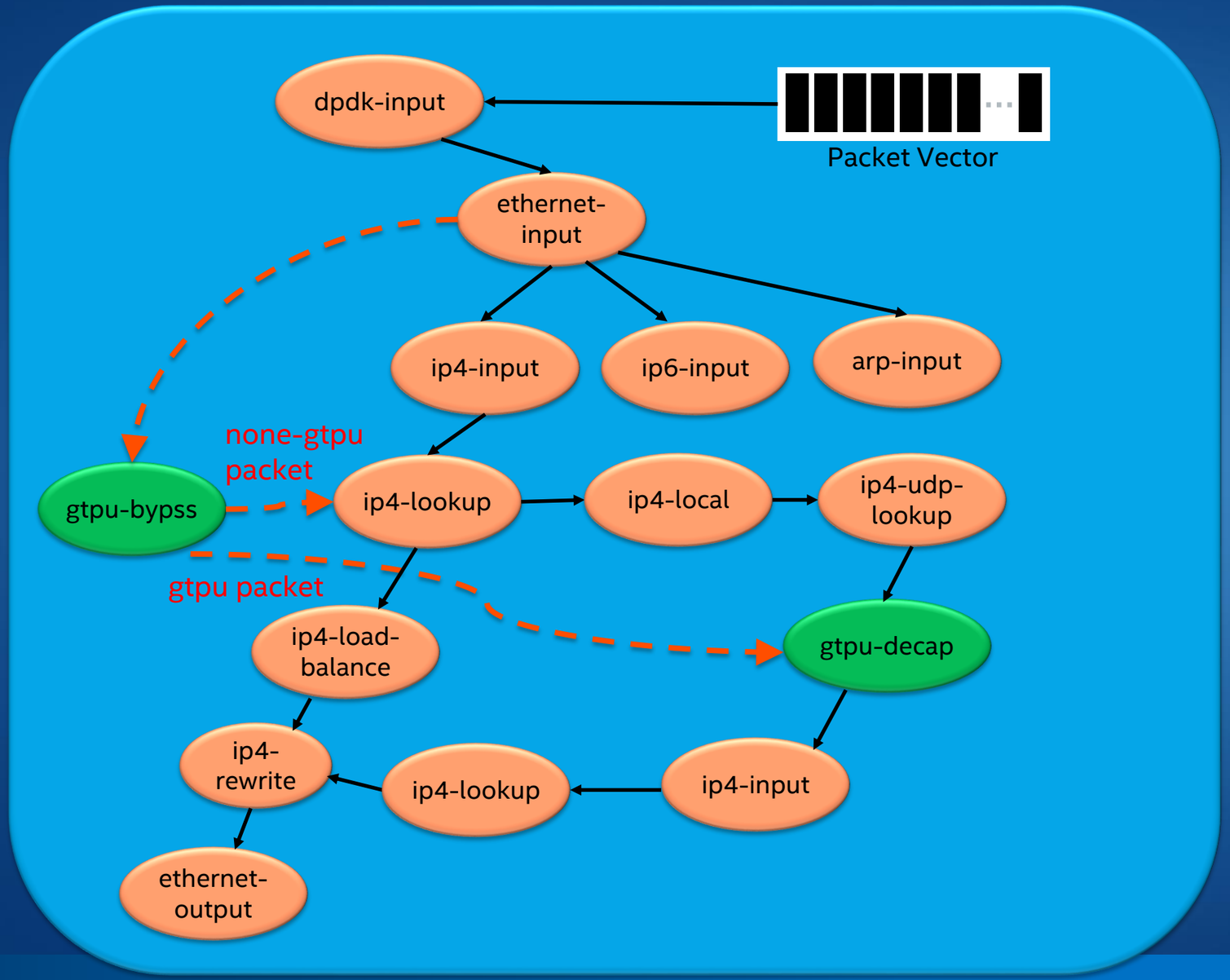


Bypass second ip-input



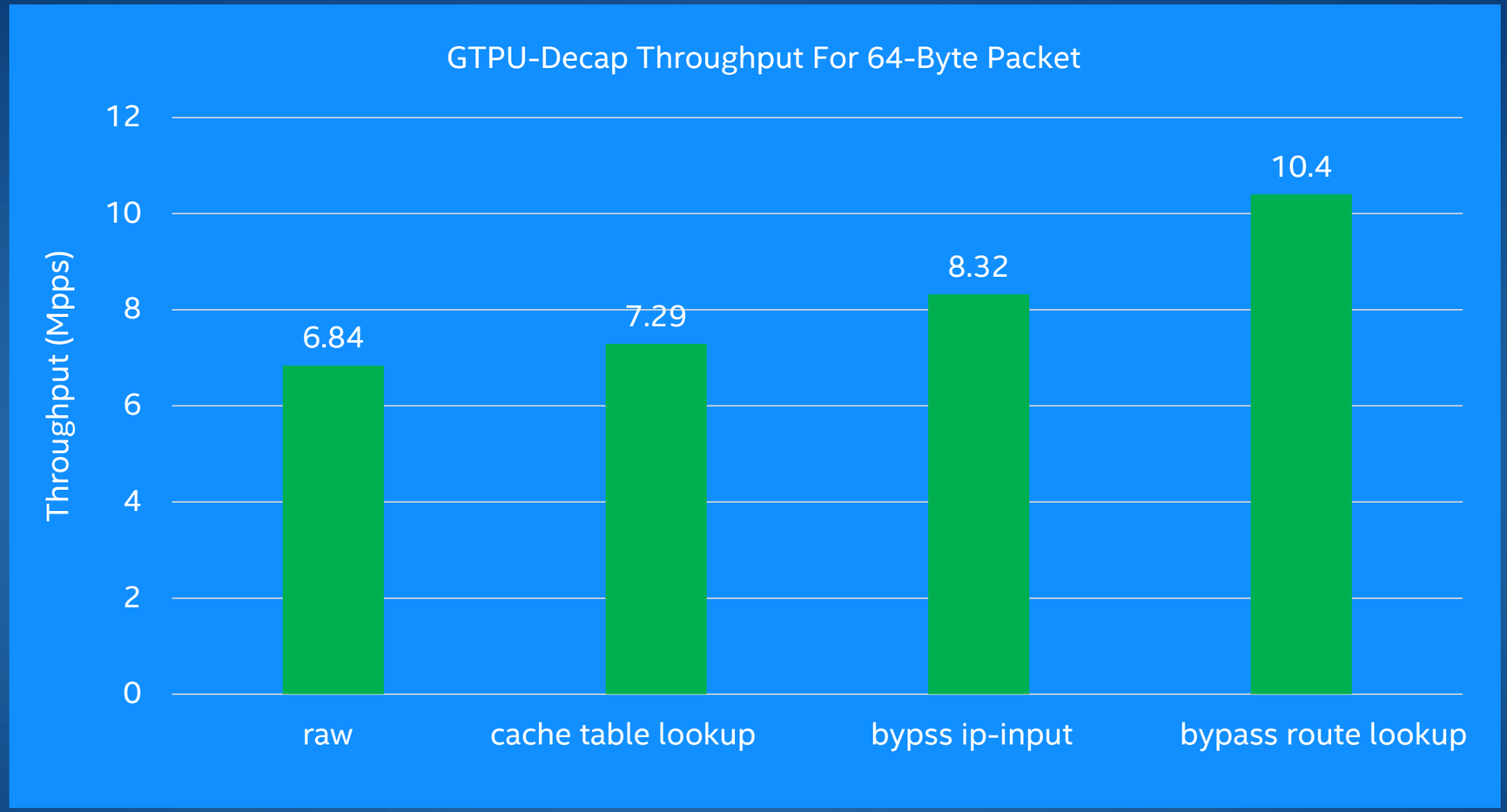
- Pros:
  Boost performance
- Cons :
  Security issue

Bypass first route lookup



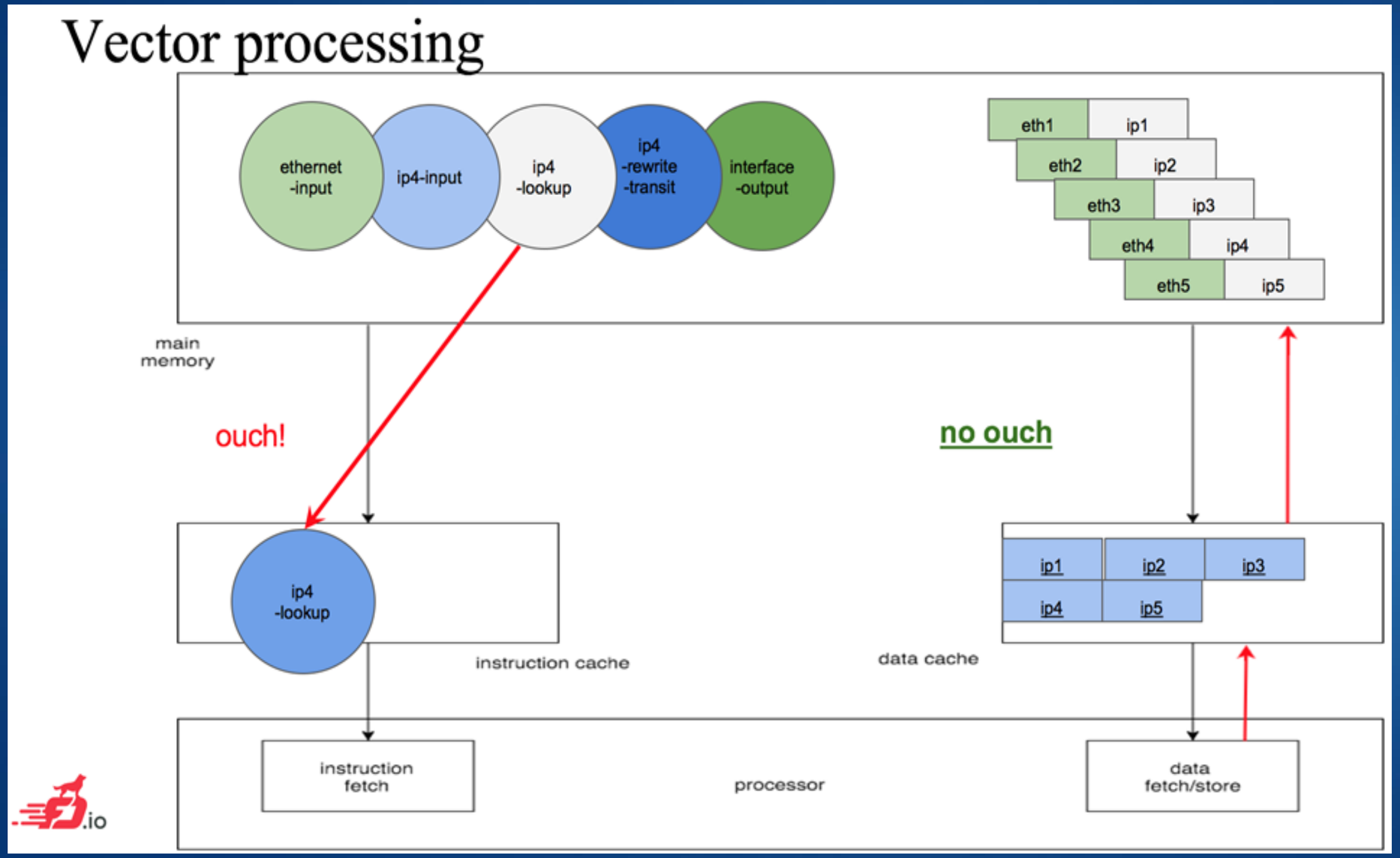
- gtpu packet:
  accelerate decap processing
- none-gtpu packet:
  an overhead with about 13 clocks per packet

GTPU-Decap Performance and Analysis



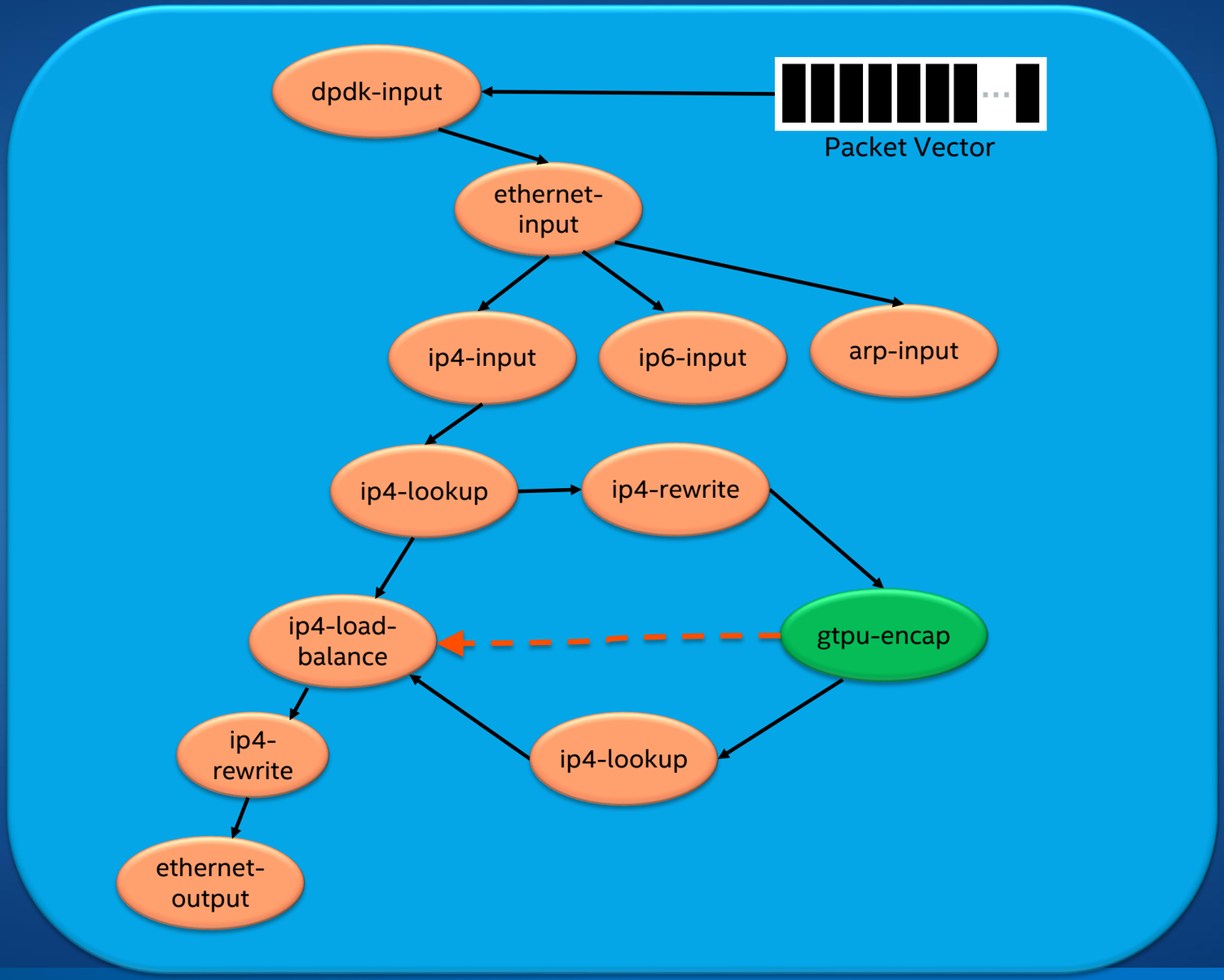
Test Case	Input packet size (bytes)	Output packet size (bytes)
Transport IP Routing + GTPU-Decap + Inner IP Routing	98	64

Dual-loop and Quad-loop, Packet Prefetching

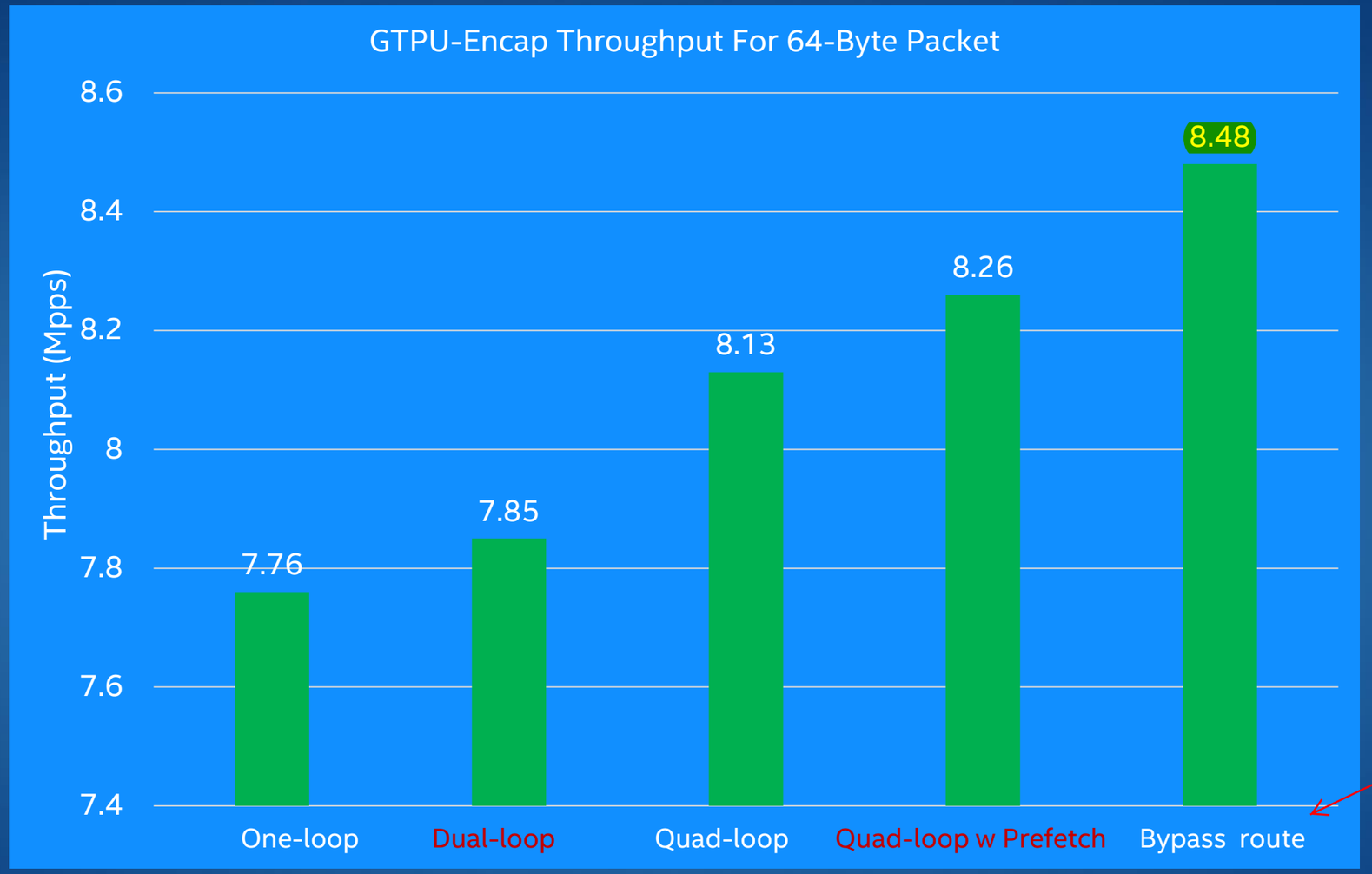


- Reduce read latency
- Process packets in parallel

Bypass second route lookup



GTPU-Encap Performance and Analysis



跳过路由

Test Case	Input packet size (bytes)	Output packet size (bytes)
Inner IP Routing + GTPU-Encap + Transport Routing	64	114

Key Takeaway

- Easy-to-use and flexible VPP plugin framework
- Kube-proxy plugin to boost DP's performance in Cloud environment
- Combination of a set of ways to optimize data plane performance
- Better platform for developing open source dataplane ingredients

Q&A