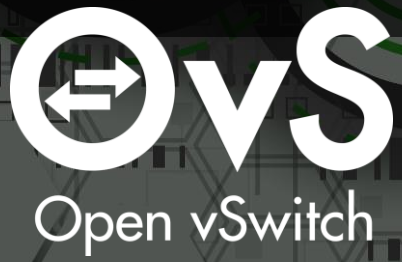


John Hurley

Netronome

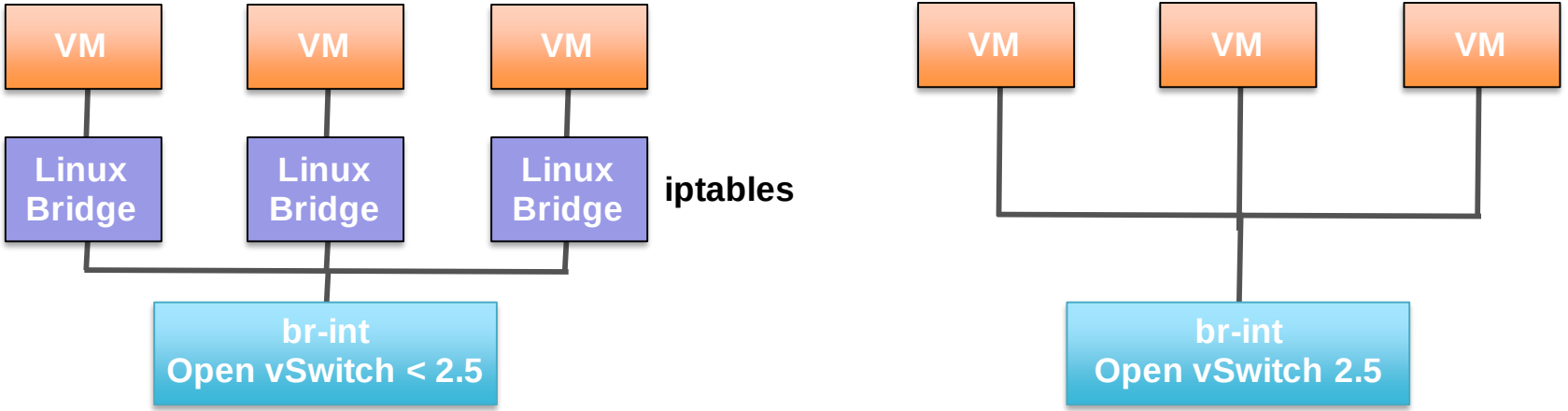
Evolving Stateful Firewalling: OVS+iptables, OVS+Conntrack, and Conntrack Acceleration

Overview

- Background to Conntrack within OVS/OVN
- Description of offloading Conntrack to SmartNIC
- Performance measurements for offload
- ‘Assume Established’ to improve performance
- Experimental/Initial results of proposed technique

NETRONOME

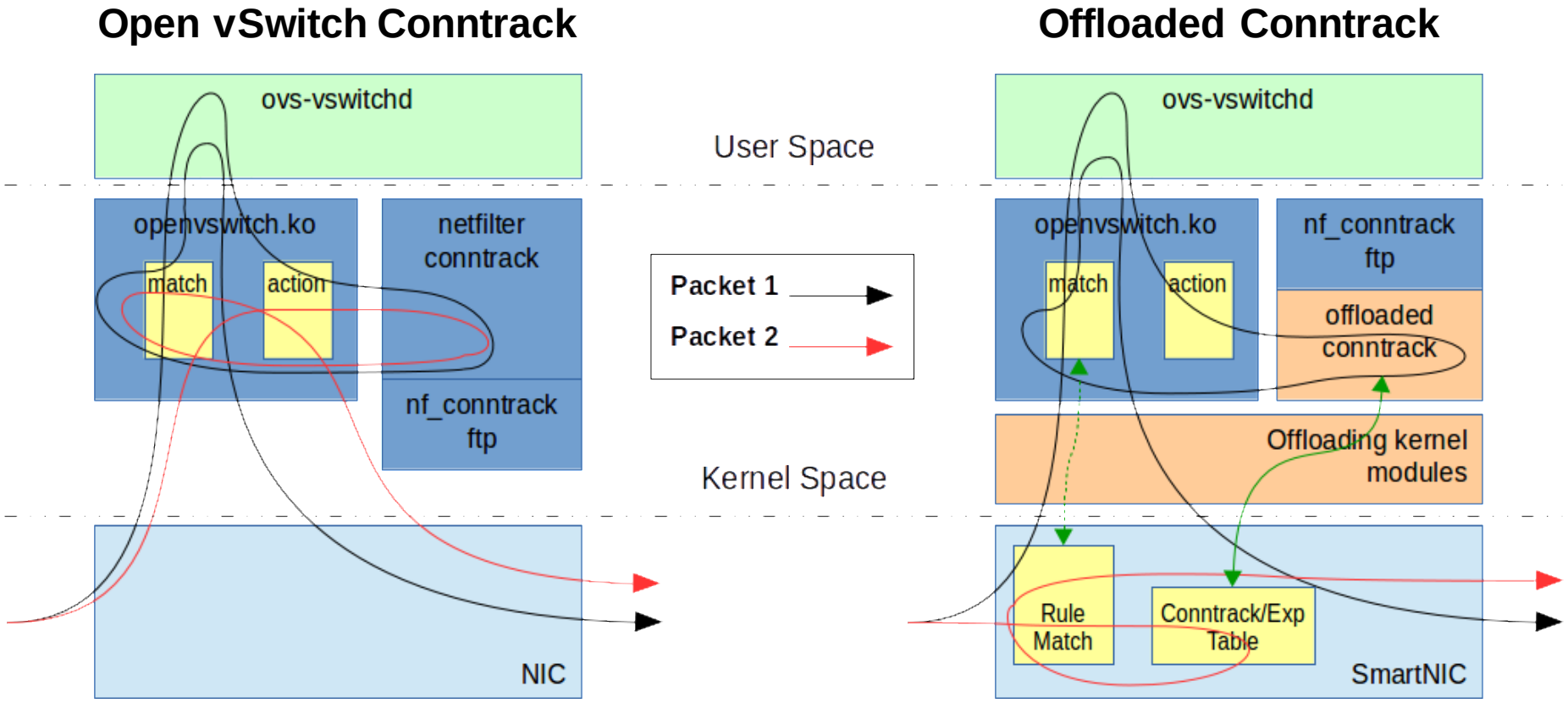
Conntrack within OVS/OVN



'Taking Security Groups to Ludicrous Speed with Open vSwitch' - Ajo, Graf, Lazzaro, Petit, OpenStack Summit 2015

- OVN allows L2/L3 overlay networks and **security groups**
- Security groups within OVN are based on Conntrack
- Current OVS handles security groups via Linux bridges …… messy and slower
- Solution: move security groups to OVS bridge

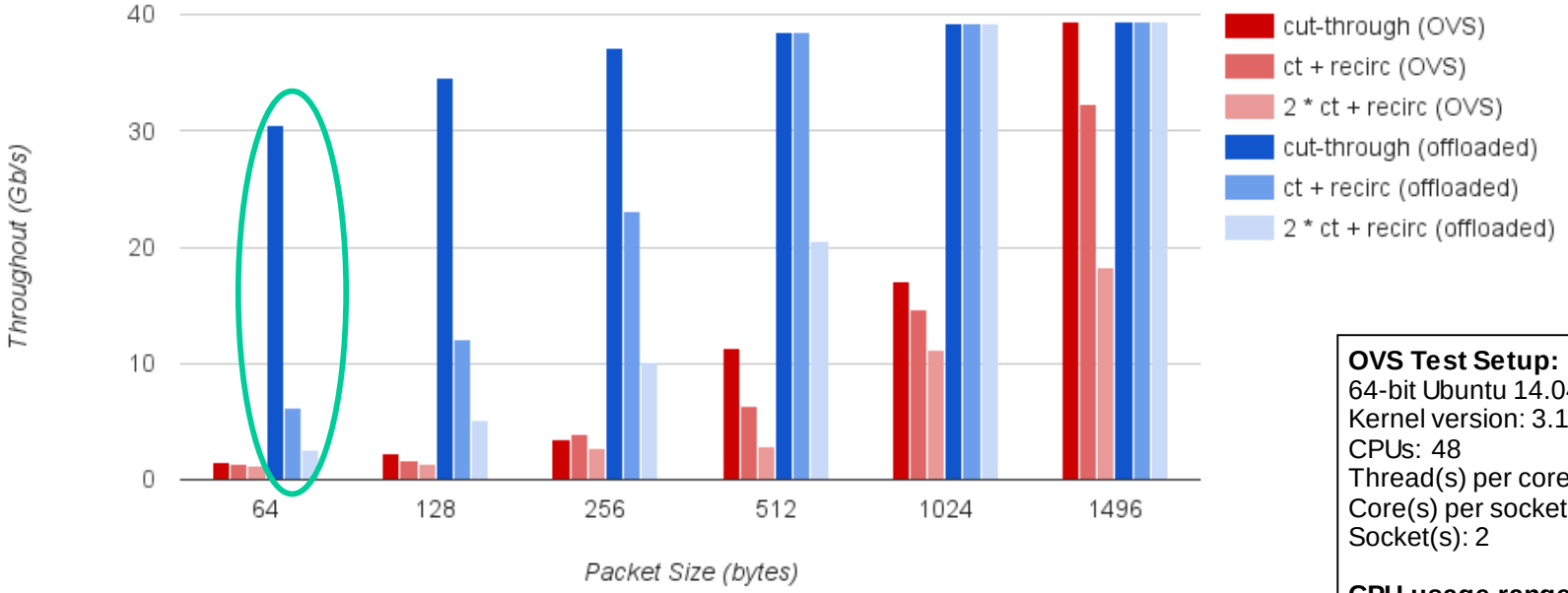
Conntrack Offload



Conntrack Offload versus Kernel OVS

10K Unique TCP Flows

Agilio LX
2x QSFP 40GbE
Netronome NFP-6480
8GB of DDR3DRAM @ 1866Mhz

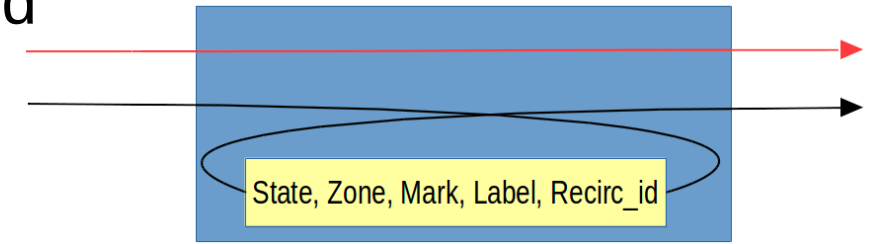


OVS Test Setup:
64-bit Ubuntu 14.04.3
Kernel version: 3.19.0-33-generic
CPUs: 48
Thread(s) per core: 2
Core(s) per socket: 12
Socket(s): 2

CPU usage ranged from 52 to 2%

‘Assume Established’ Proposal

- Assumptions/Observations… In majority of cases:
  - Packets will be in Established flow state
  - Conntrack metadata only key change after recirculation
- Cache each packet flow - exact match
- Assume packet is established
  - If it is then fastpath it
  - If not then do full processing
  - Merge out CT metadata



‘Assume Established’ Proposal

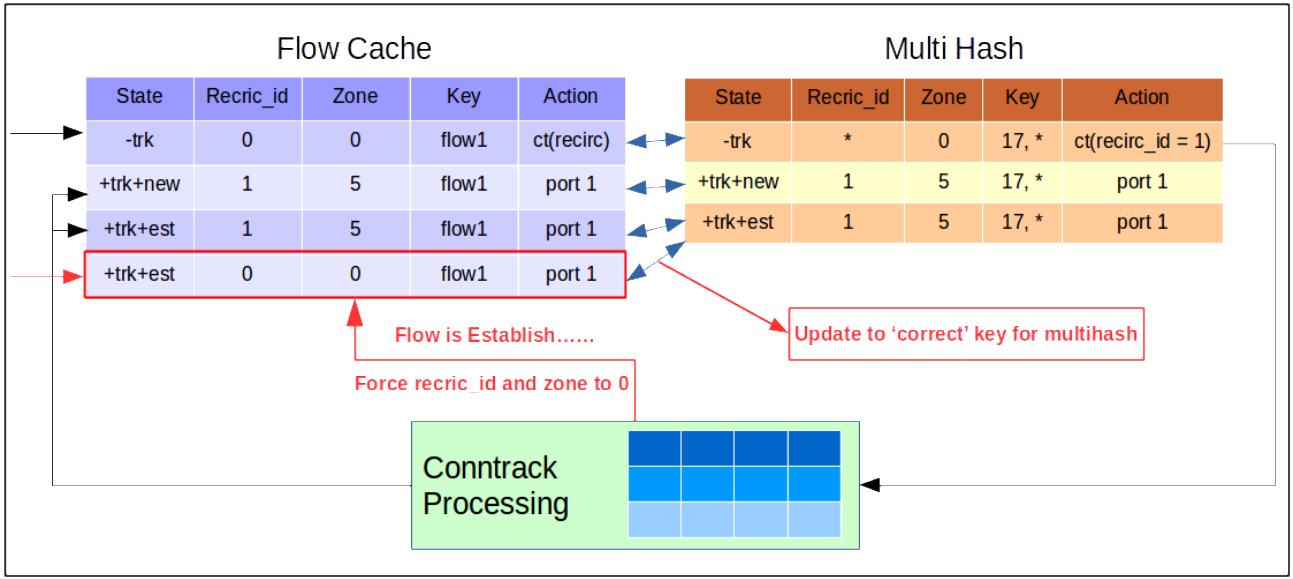
table=0,ct_state=-trk,udp,action=ct(commit,table=0,zone=5)
table=0,ct_state=+trk+new,ct_zone=5,udp,action=1
table=0,ct_state=+trk+est,ct_zone=5,udp,action=1

1. Set New packet to +Est, recirc_id=0

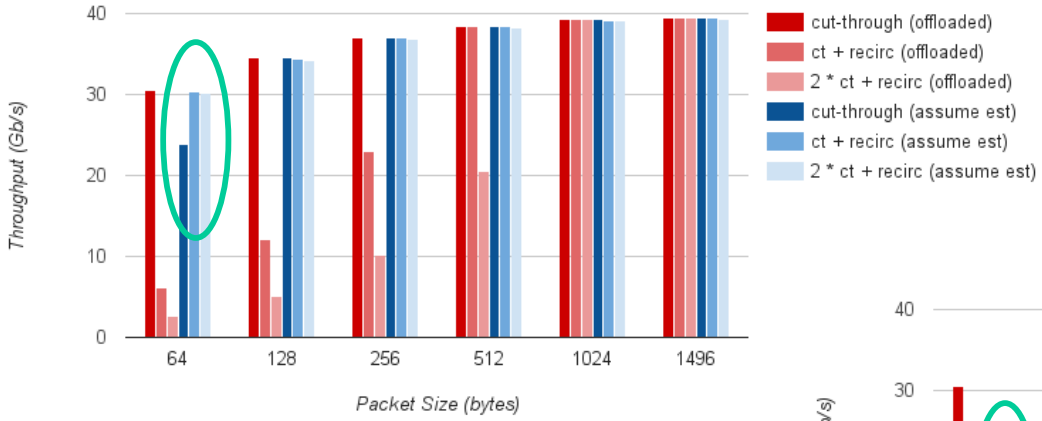
2. If not in FC then remove rule and retry with -trk state

3. When a Flow becomes established (and if it is permitted), force all metadata to 0 bar state and add to FC - update key for MH lookup

4. Packets from the same flow can now be ‘fast-pathed’

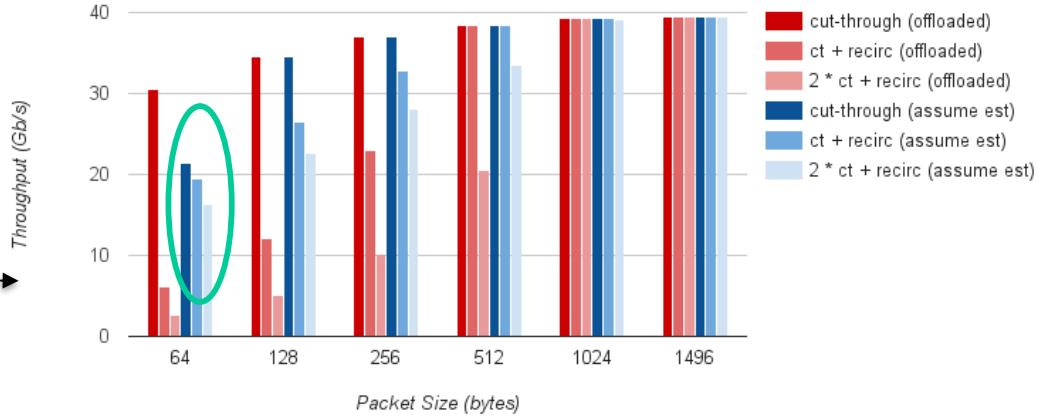


‘Assume Established’ - Initial Results



10k Unique TCP flows
All Established

20k Unique TCP flows
10k Established
10k Unidirectional
(1 in 3 packets will be NEW)



Future Work

- Full integration of TCP state tracking
- Further optimisations
- Corner case handling
- Sync with kernel Conntrack table?
- NAT - OVS 2.6