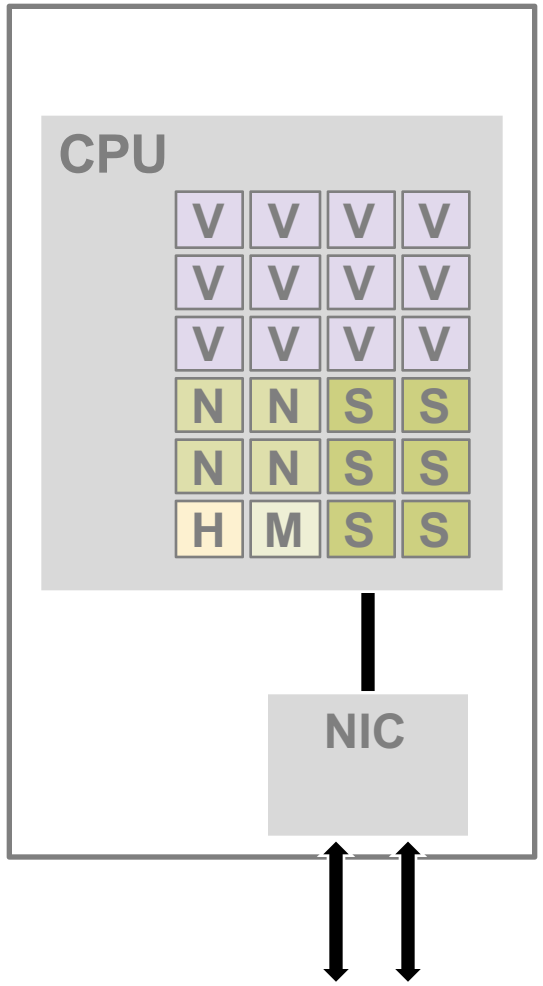
The Core Cost of Doing Business

Netdev 0x13
Prague, Czech Republic

Don Wallwork
Andy Gospodarek

Compute and Connectivity Division (CCX)

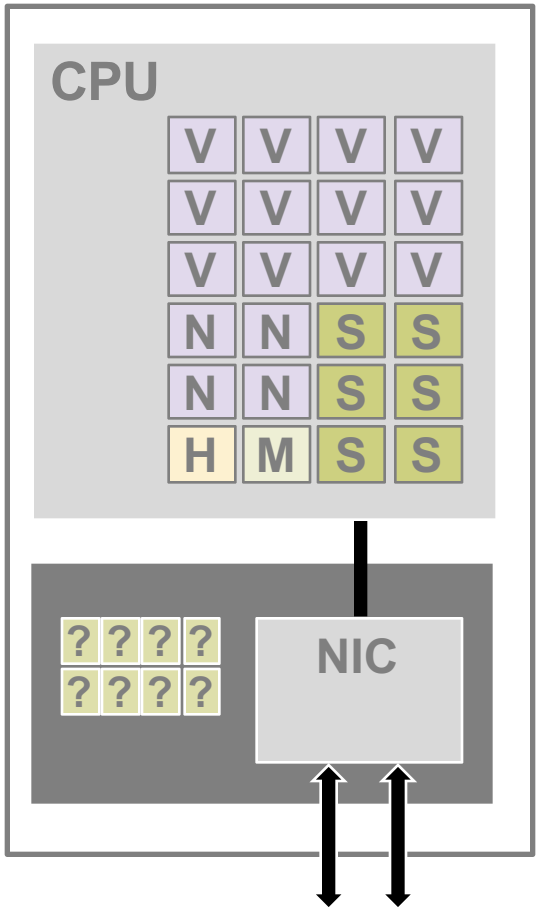
Network Services Running on a Server



BROADCOM

Migrate Services off Compute Server?

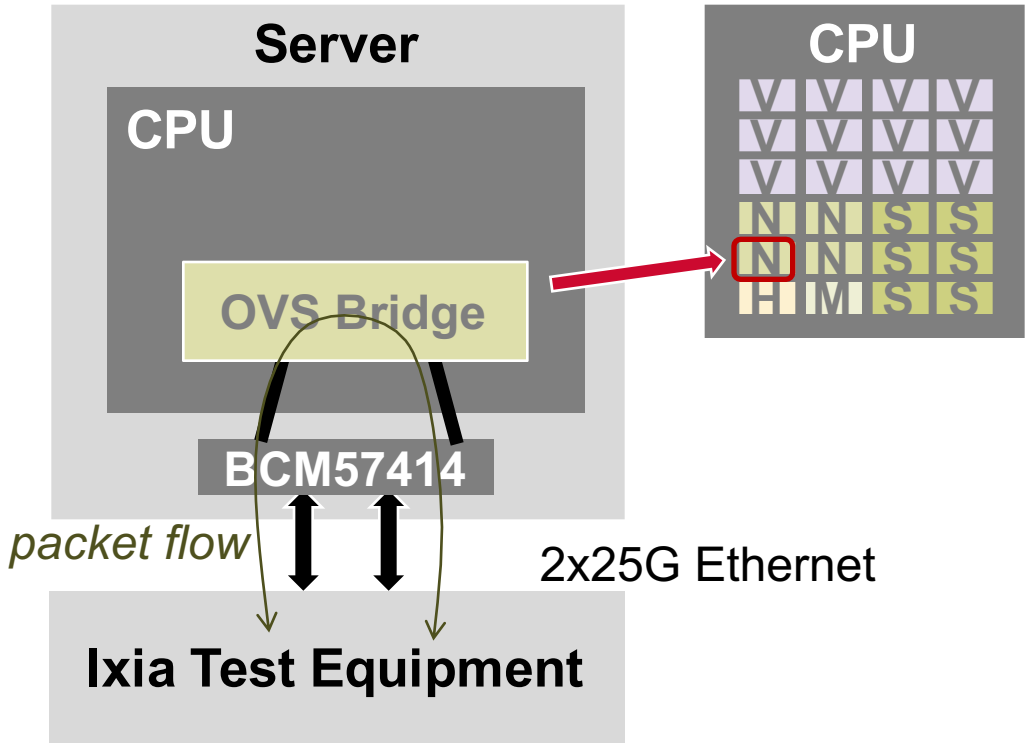
What can the hardware do?



BROADCOM

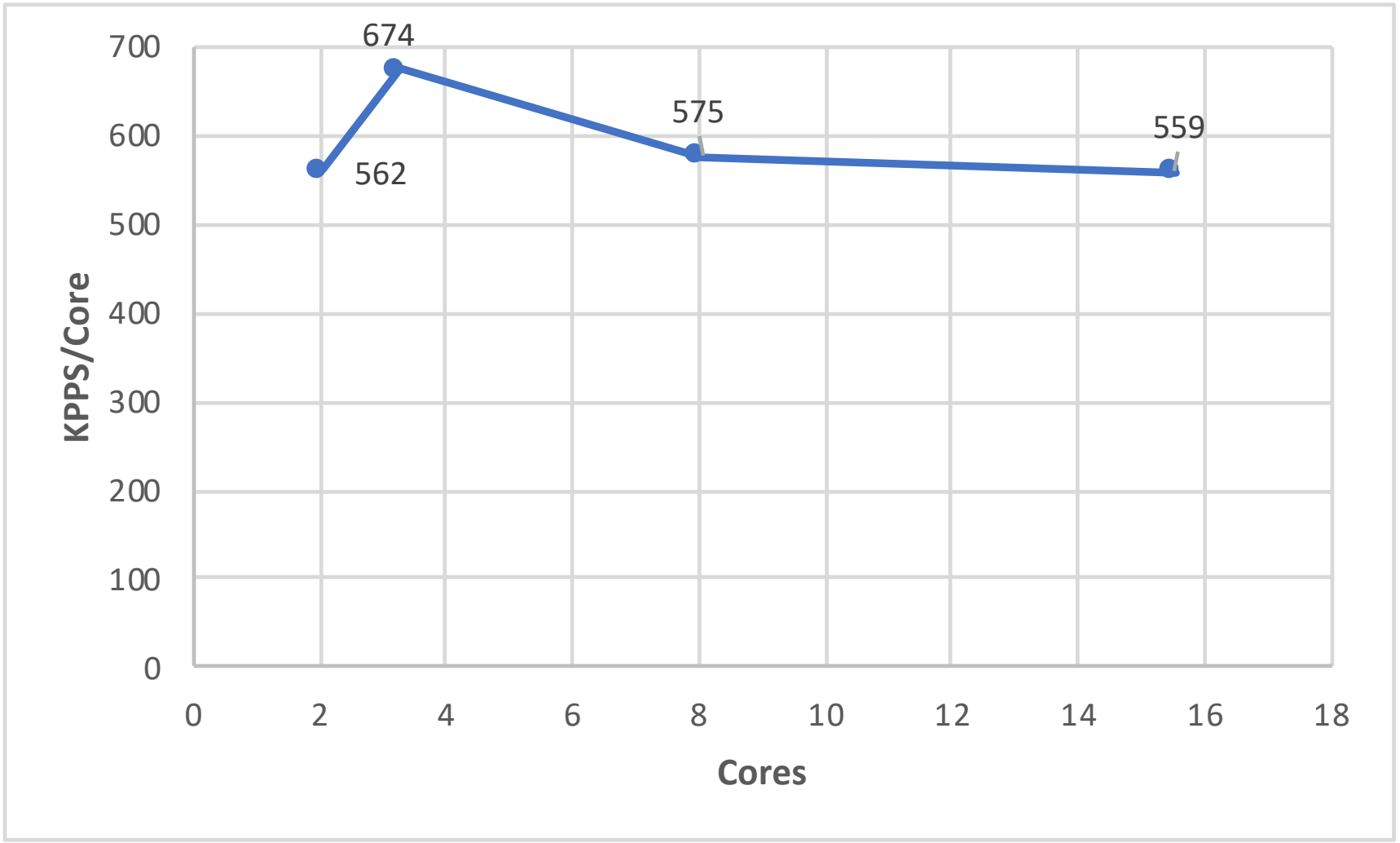
OVS Performance on Single Xeon Core

- **Data collected using Dell R730**
  - Dual socket Intel Xeon E5-2697v4 @2.6GHz
  - Broadcom BCM57414 2x25Gbps NIC
  - RHEL 7.6
    - kernel-3.10.0-957.1.3.el7
    - openvswitch-2.9.0-83.el7fdp.1
- **Layer 3 forwarding, no network overlay**



BROADCOM

OVS Throughput per Core

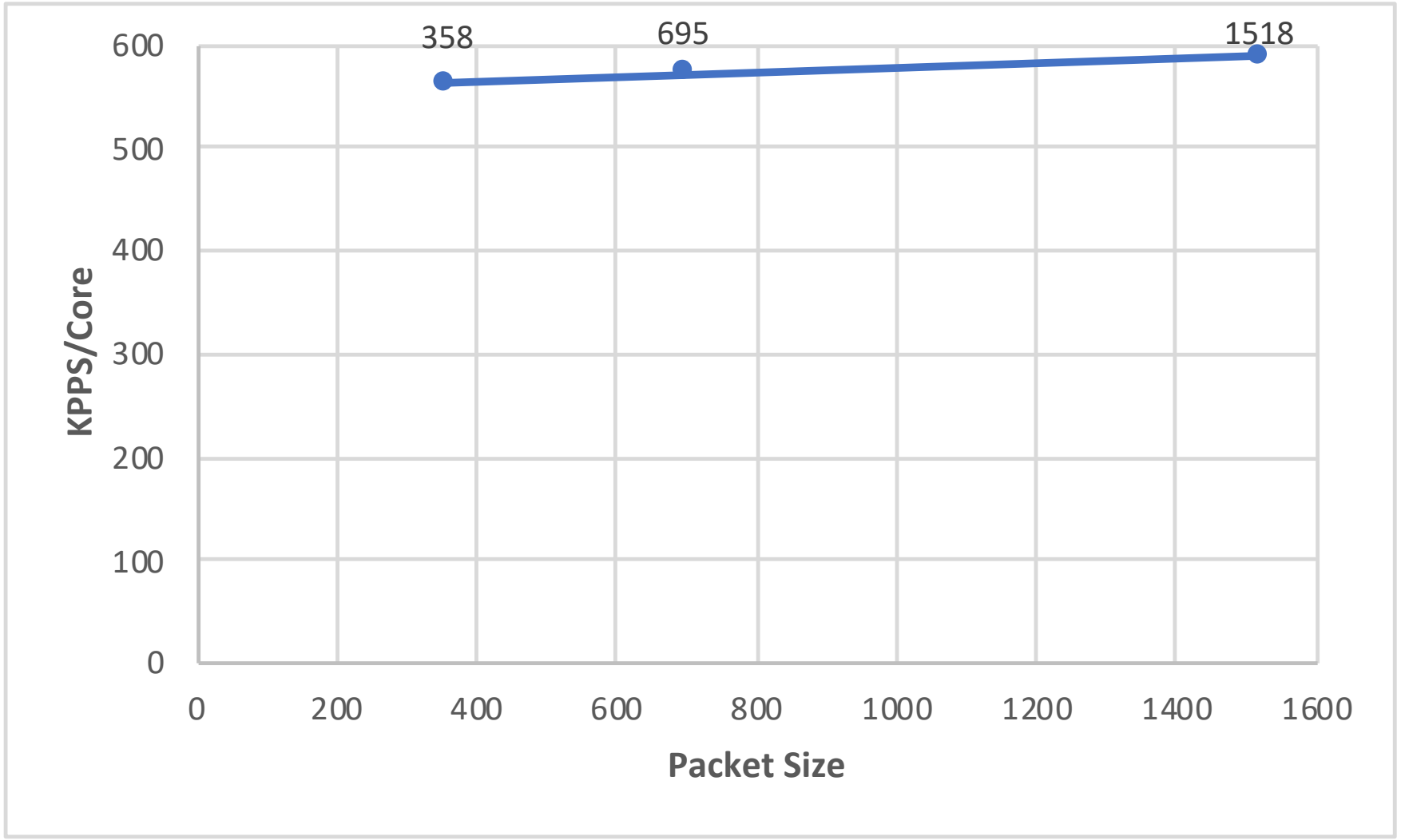


2.6GHz/560kpps = 4600 cycles/packet !!

Nearly linear scaling

BROADCOM

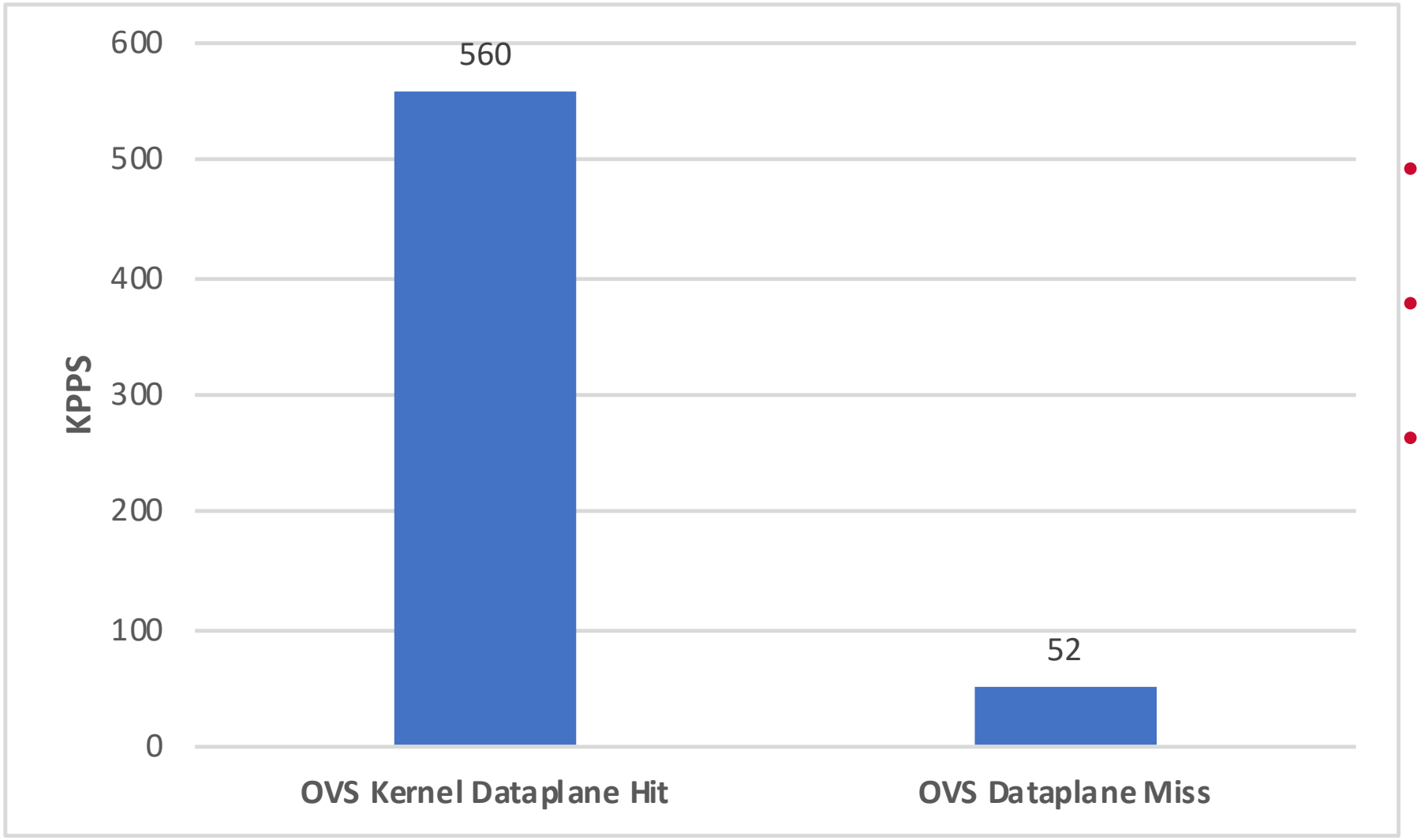
What about other packet sizes?



Consistently around 4600 cycles/packet

BROADCOM

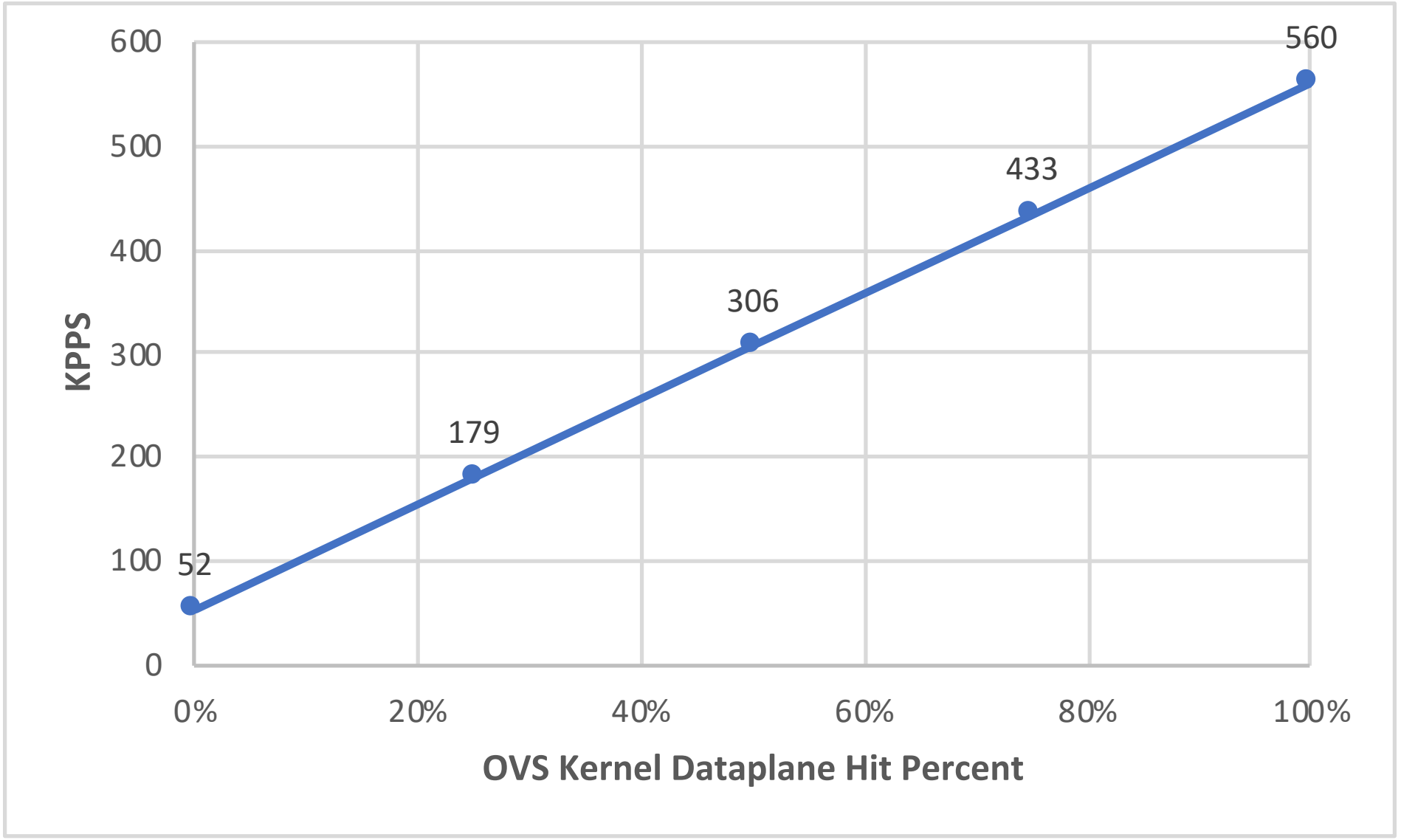
What if the flow isn't already known?



- **Kernel hit: 4600 cycles/packet**
- **Kernel miss: 50,000 cycles/packet**
- **Miss 10.8x more costly than hit**

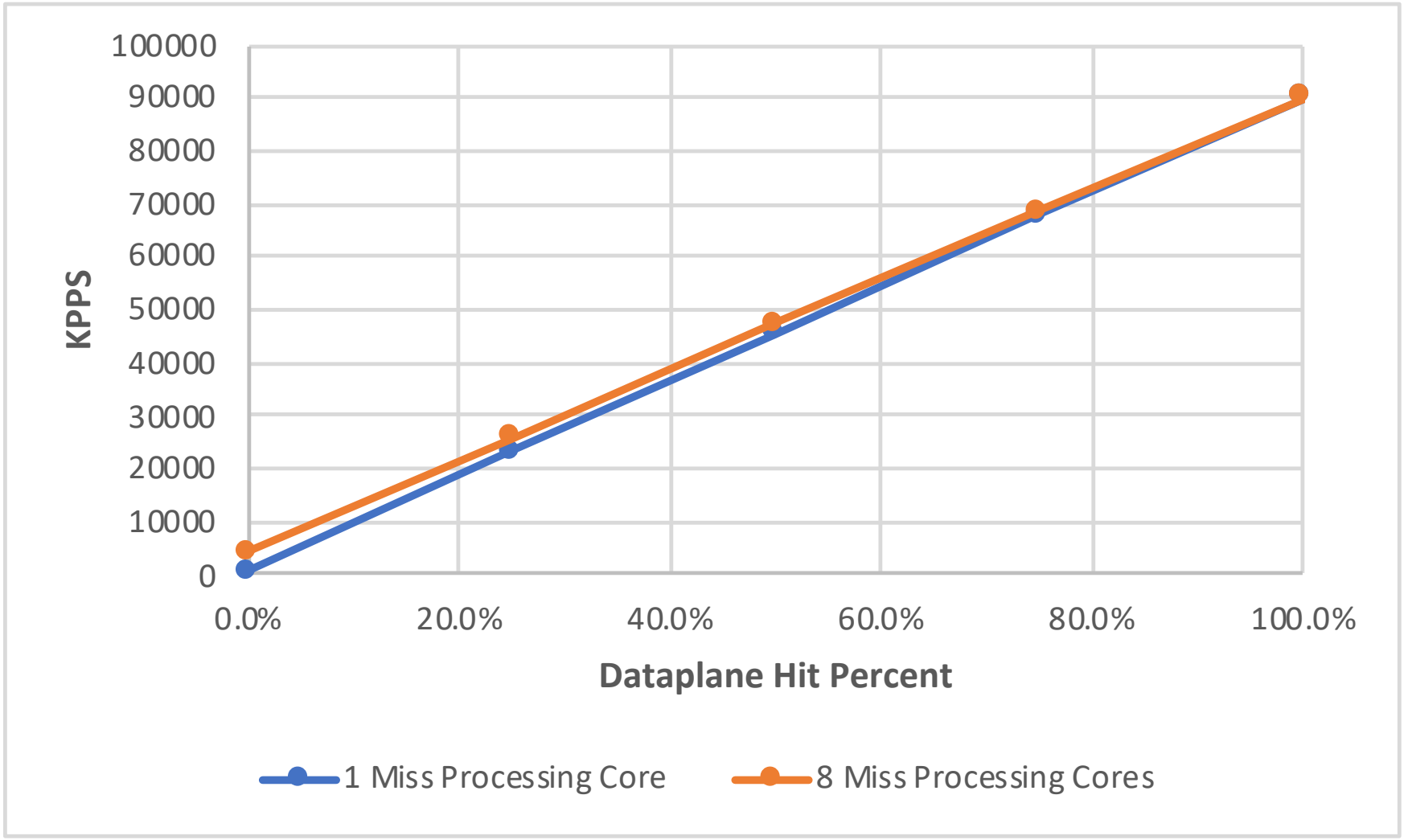
BROADCOM

What about a mix of known and unknown flows?



BROADCOM

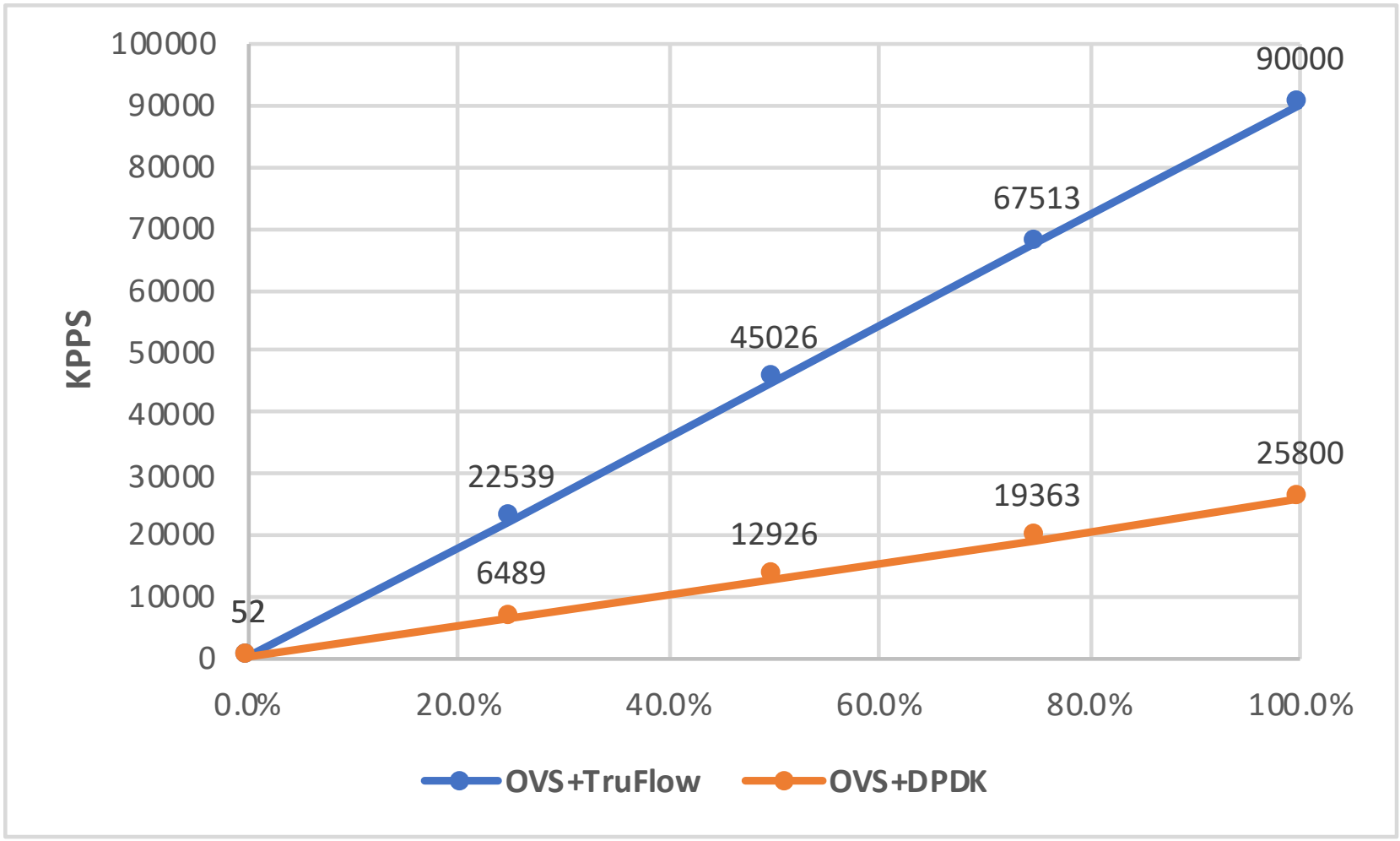
What about hardware acceleration?



- **Assumptions:**
  - 10x faster miss processing
  - 8 cores dedicated to miss processing
  - HW rate of 90MPPS
- **Hit rate dominates throughput**

BROADCOM

What about DPDK?



- **Kernel dataplane with HW acceleration outperforms DPDK**

BROADCOM

Amdhal's Law

$$S_{latency}(s) = \frac{1}{(1-p) + \frac{p}{s}}$$

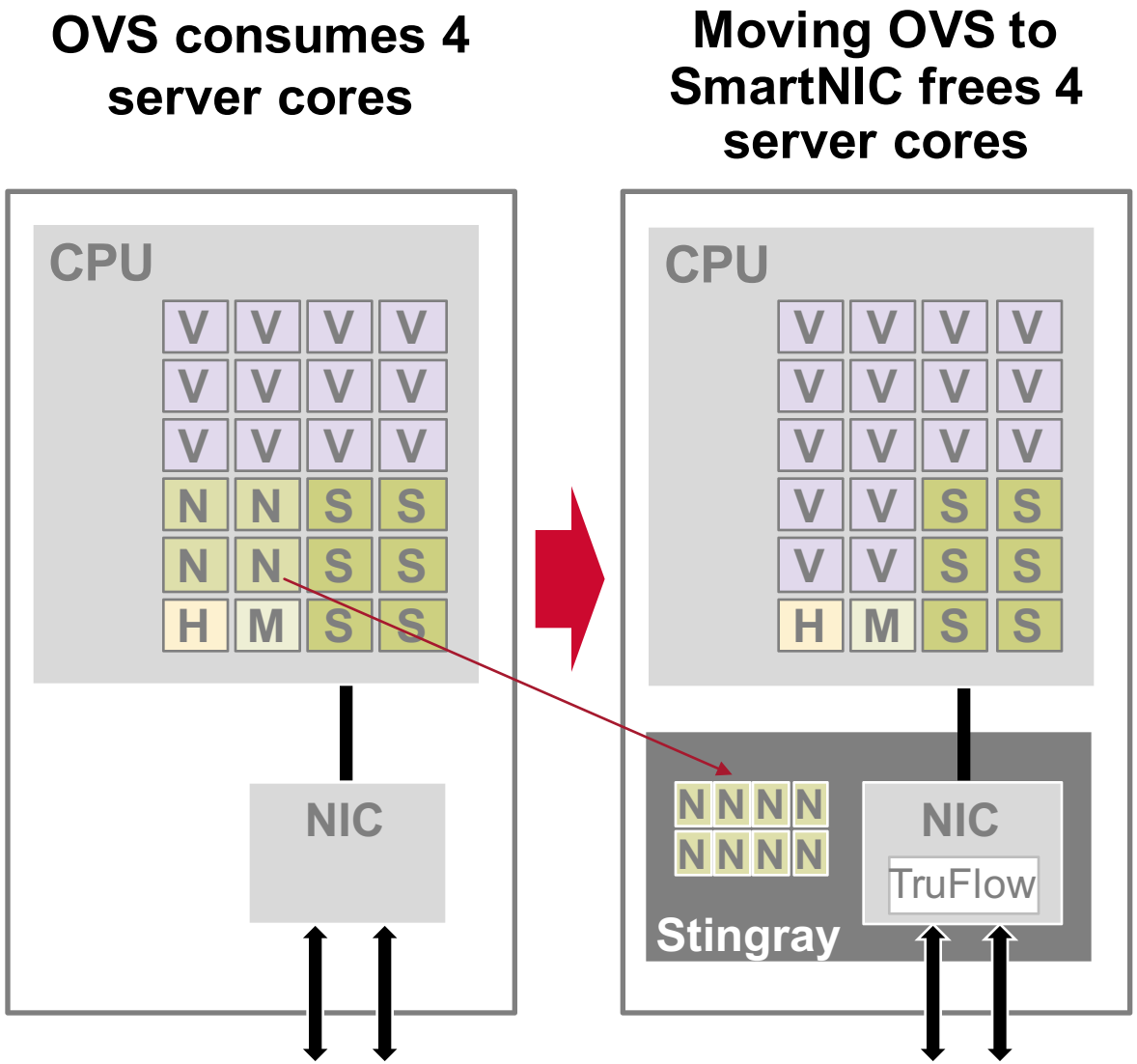
- $S_{latency}$ is the theoretical speedup of a task
- s is the speedup of the part of the task that is improved
- p is the proportion of the execution time that the original task occupied

Amdhal's Law Applied

$$S_{latency}(s) = \frac{1}{1 - 0.75 + \frac{0.75}{10}} = 3.1$$

- s = 10 (10x speedup)
- p = 0.75 (percentage of time previously spend on datapath miss processing)
- $S_{latency}$ = 3.1x speedup

Migrate Network Services off Compute Server



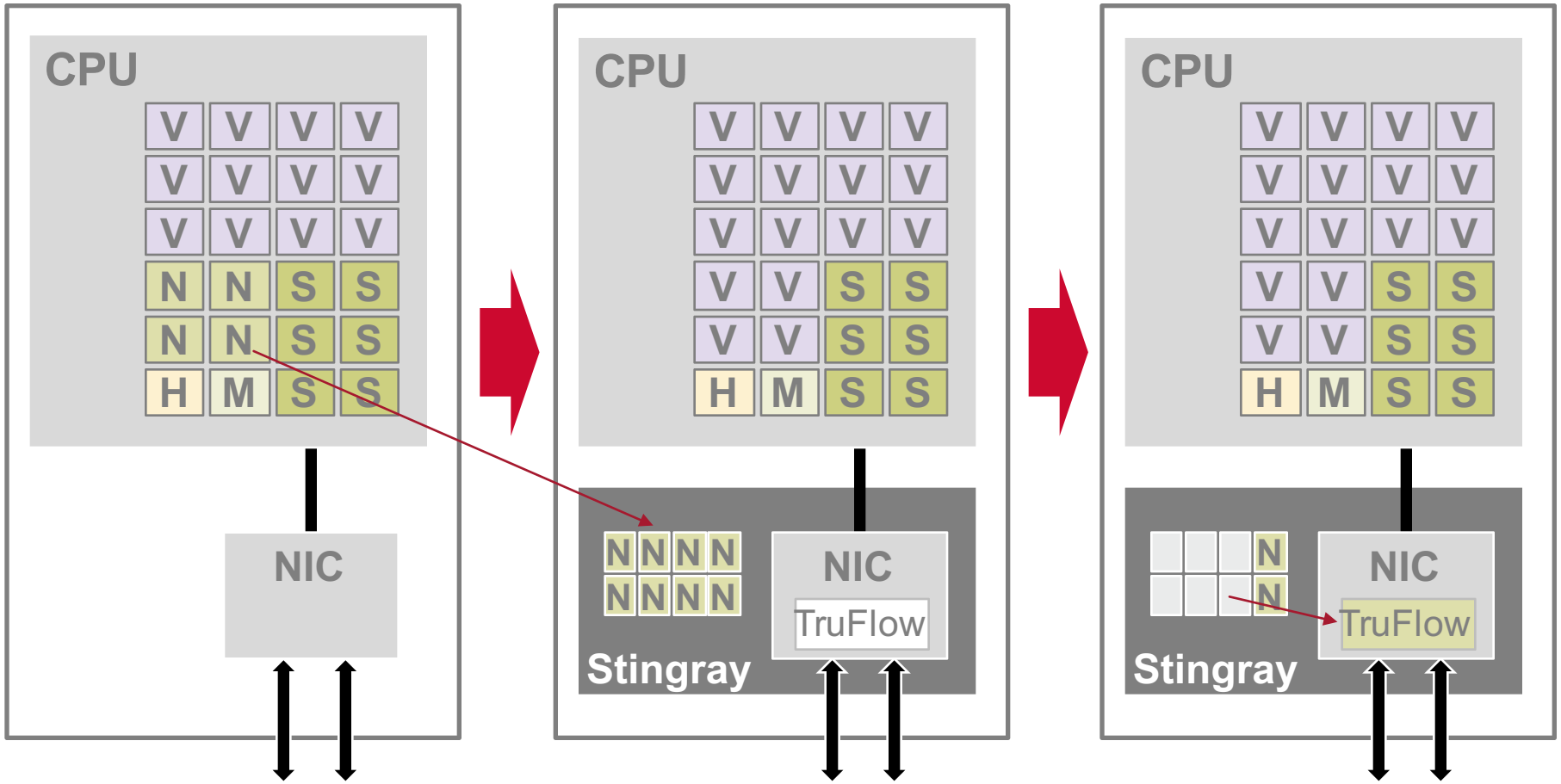
BROADCOM

Migrate Services off Compute Server

OVS consumes 4 server cores

Moving OVS to SmartNIC frees 4 server cores

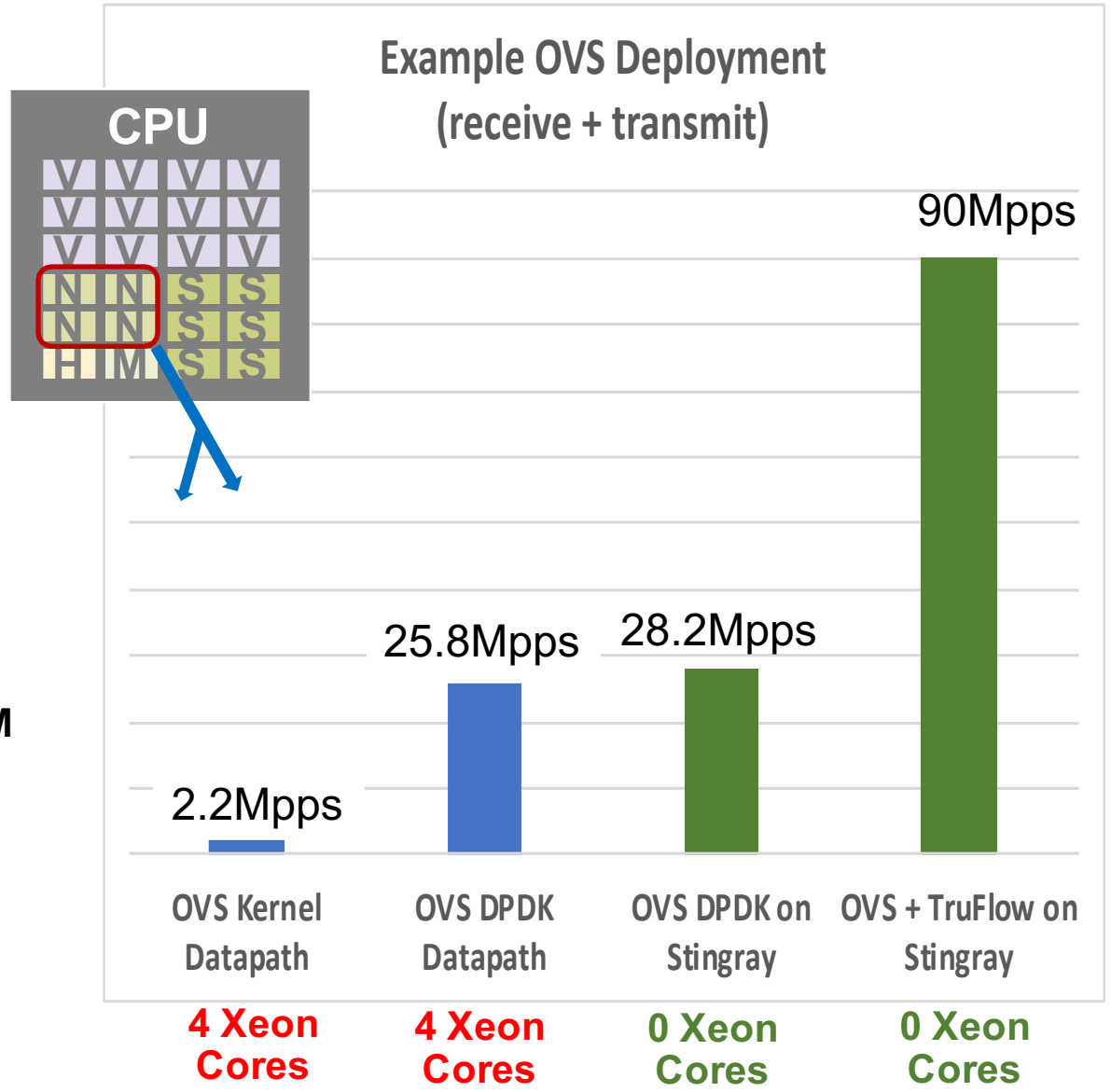
Utopia: Hardware acceleration of datapath in SmartNIC



BROADCOM

OVS Performance: 4x Xeon cores vs. Stingray

- **OVS performance on 4 Xeon cores**
  - Performance scales roughly linearly per core
  - Stingray 8x A72 cores offer similar performance
- **SmartNIC frees 4 Xeon cores for revenue generating applications**
  - Generate up to $4,000 additional revenue per year
- **Dedicated network processing HW accelerates performance → TruFlow™**
  - Increases PPS
  - Reduces power
  - Supports IPv6, VXLAN encap with no performance degradation



BROADCOM

Friends Don’t Let Friends Run OVS on Application Servers

- **Cost of flow setup and aging is high compared to kernel datapath processing**
- **Changing the number of cores processing new flows and aging old ones does not demonstrably change packet processing speed when compared to kernel or hardware datapath**
- **Moving OVS Control Plane and Dataplane to a Stingray/SmartNIC is a compelling option to save server resources**

Thank You