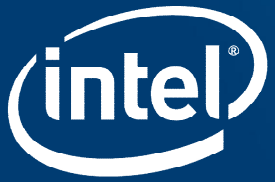
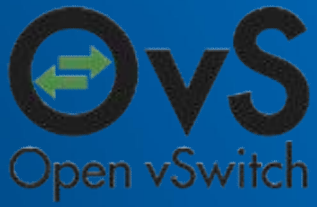




OVS, DPDK and Software Dataplane Acceleration

OVS Fall Conference

November 17th, 2015

Who are we?

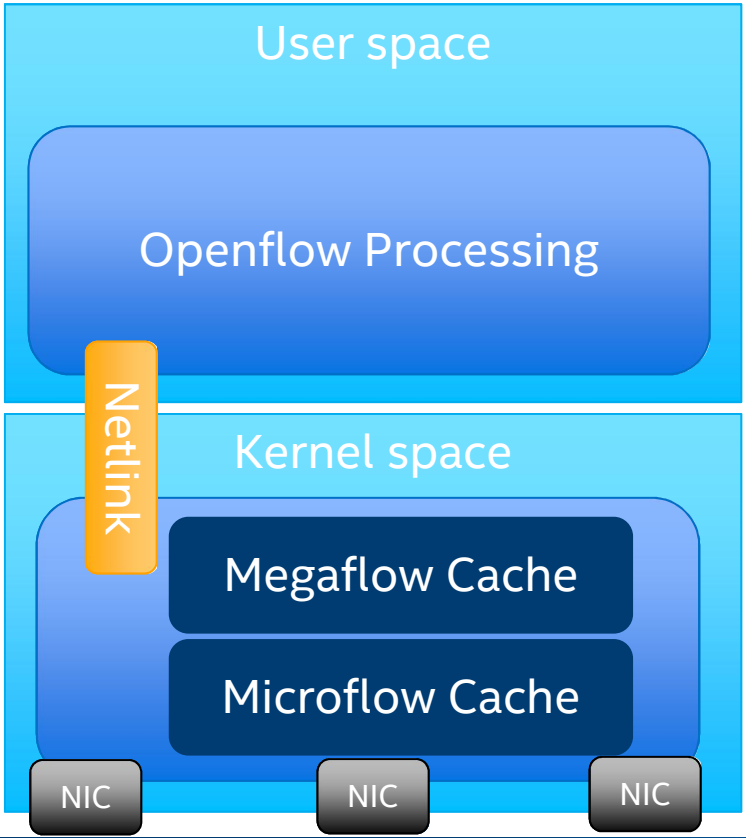
Thomas F. Herbert

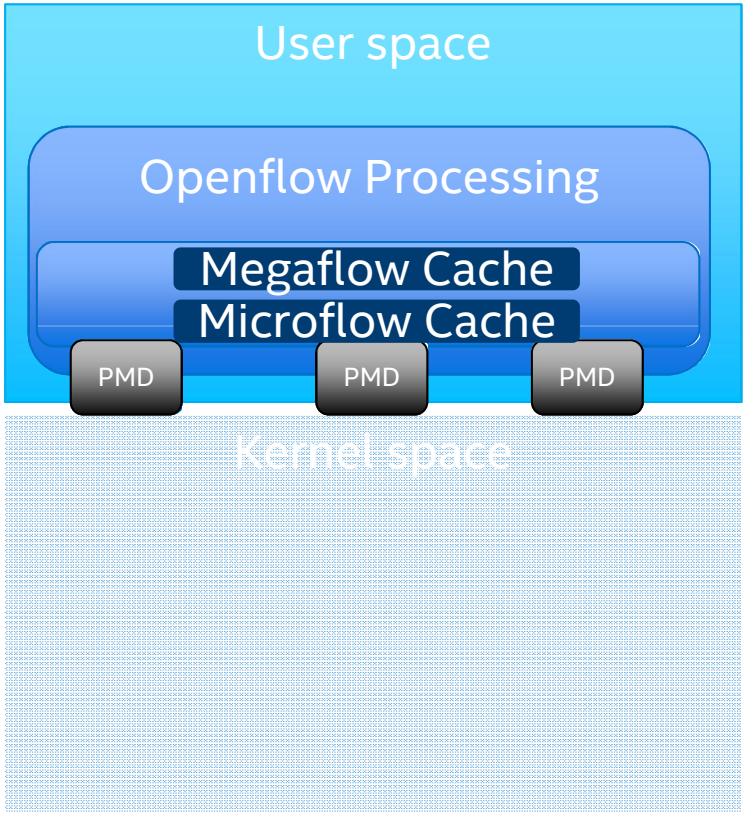
- Red Hat
- therbert@redhat.com

Kevin Traynor

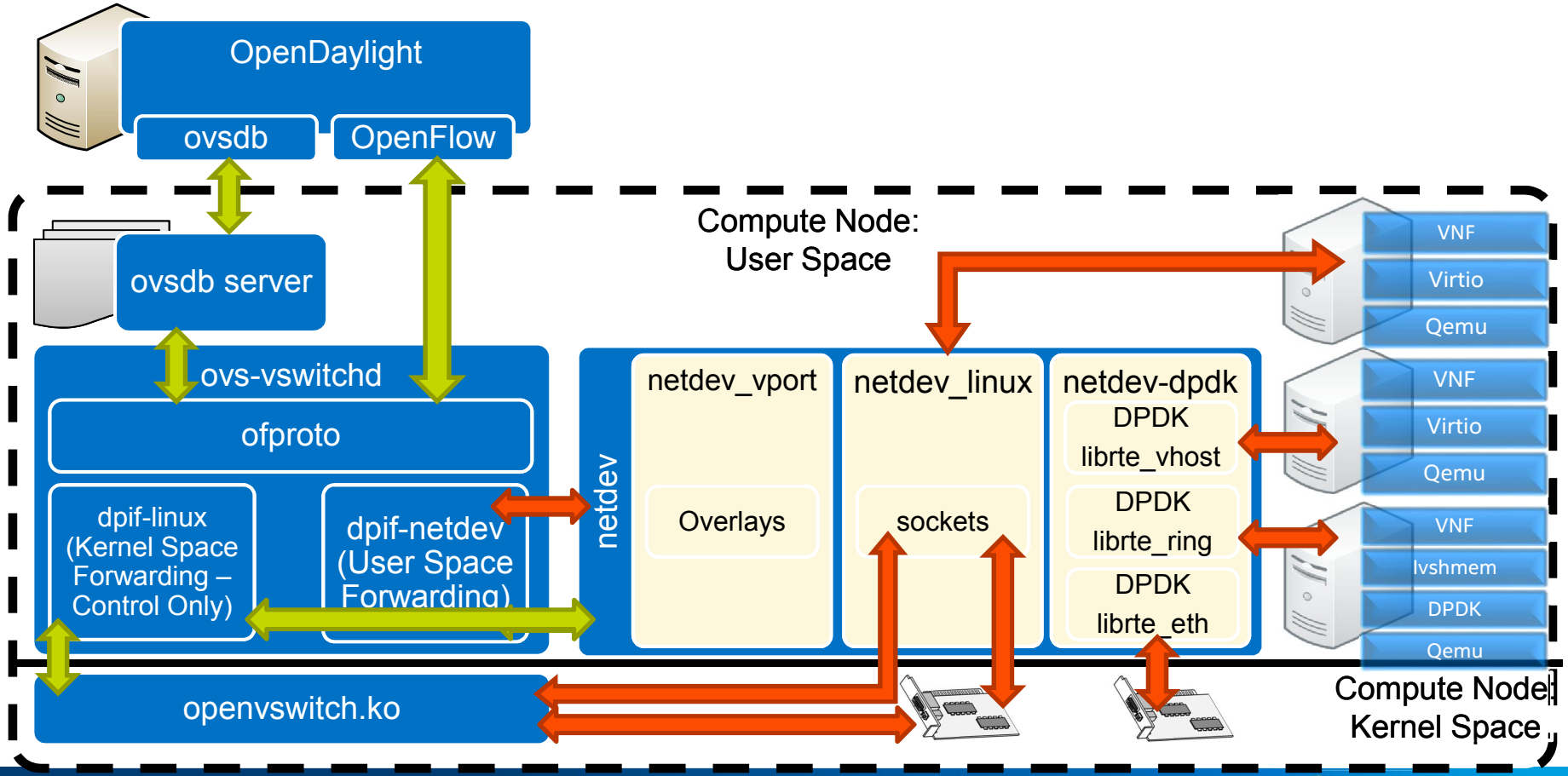
- Intel
- kevin.traynor@intel.com

Open vSwitch with DPDK





OVS Architectural Evolution



The Netdev Interface Makes it Possible

- Transparency of Data Plane
- Netdev – API Between Data Plane and OVS
  - Generic network device API independent from data plane implementation.
  - Similar to network driver interface in BSD
  - Netdev Abstracts forwarding of packets in data plane
- Conceptually like any Network device driver
  - With Start, Stop, Private Data Area, Queue Management
- Struct netdev Holds the interface Specific Function Pointers
  - Includes the generic part followed by private part for use by driver.
  - Constructur for netdev provider
  - Dpdk Creates dpdk personality of struct netdev
    - Multiple rx queues Managed by OVS

Open vSwitch with DPDK

- DPDK – Data Plane Development Kit
  - About 4 Years Old
  - First Integrated with OVS from 2013
  - Fast Packet Forwarding
  - Poll Mode Drivers
  - Uses Commodity Hardware
  - Multiple Threads and Cores
- Up to 12X Speed Improvement for small packets
- Over 15mpps Forwarding
  - Small Packets
- Disadvantages WRT Linux Kernel
- Linux Data Plane Has
  - Complete TCP/IP Stack
  - 20 years of development
  - Rich Debugging Options
    - Promiscuous IFs
  - Access to Wide Variety of Network IF's and VF's
  - Tunnels and Endpoints

Open vSwitch with DPDK User Perspective

- Who Uses OVS/DPDK
  - Open vSwitch
  - Open Stack/Neutron via ODL Plugin and OVS
  - OVS CLI Tools
  - Programmers - Application Developers of
    - Other Packet consumers, DPI, Classifiers
    - Infrastructure – Routers, Firewalls, Services
    - Other Packet consumers, DPI, Classifiers
- Typical Scenario
  - NF's in VMs or Containers
    - Service Function Chaining
  - Real World Performance
    - Challenge: DPDK Gains Over Wide Variety of Use Cases

Open vSwitch with DPDK Usability Story

- In the Beginning: My user Story starts in 2013
  - Inspired by Intel presentation of DPDK at ONS 2013
- Developing Network Threat Analyzer
  - Integrated Open vSwitch
  - Traffic shaping, threat blocking and mitigation
- Requirement: 10Gb without Adding $10K to $20K on custom HW Switch Fabric.
  - DPDK is the Answer?
  - How to prove the OVS/DPDK Claim?
- Started with DPDK 1.7.1 and OVS 2.3
- Scary: poor integration --Not integrated with OVS
  - Compilation issues, conflicting APIs. ABIs, OVS Versions
  - Three Confusing Forks:
    1. DPDK.org
    2. DPI Fork with custom API
    3. 01.org
- Came Together with DPDK 1.8
  - Integrated: Master Branch OVS
  - I Ran DPDK on guest with VirtIO/VMXnet3 saw 2.5X perf gain
  - Developed App using DPDK-ring to feed DPI
- Now: DPDK 2.1 with OVS: Much improved!

Improving Open vSwitch with DPDK

- Is DPDK really still Experimental?
  - Is it time for this patch?

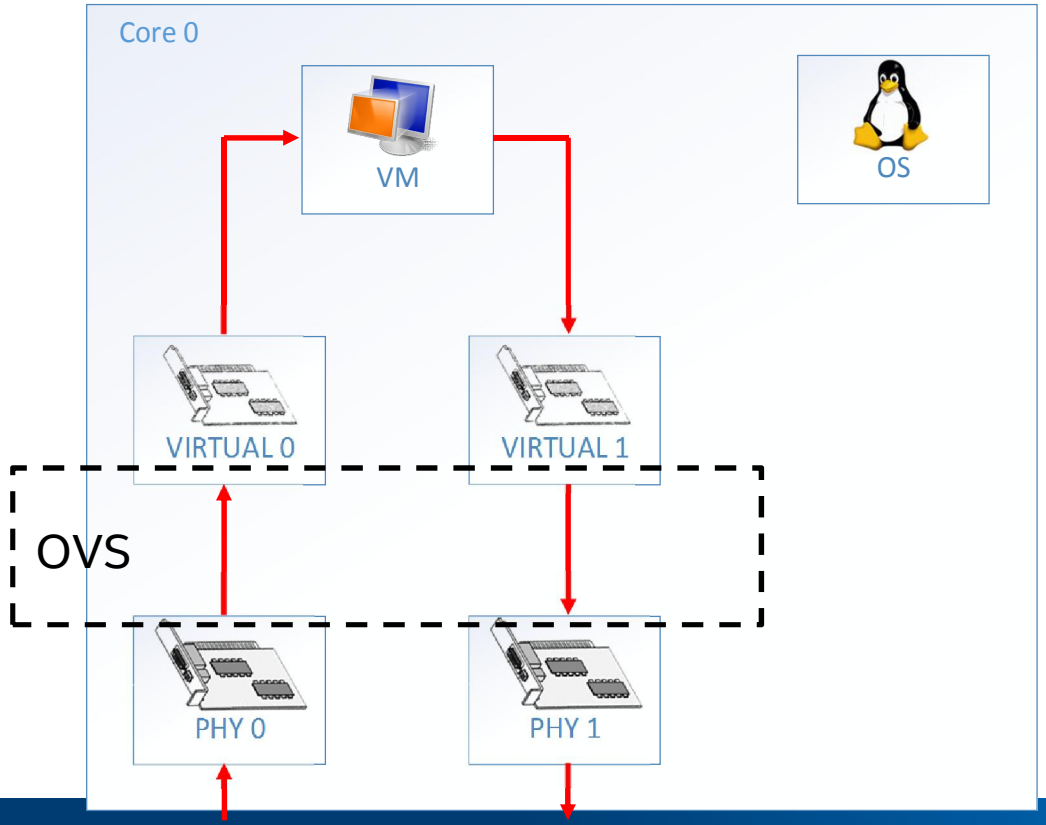
```
--- a/INSTALL.DPDK.md
+++ a/INSTALL.DPDK.md
@@ -5,8 +5,8 @@ Open vSwitch can use Intel® DPDK lib to operate entirely in
Userspace. This file explains how to install and use Open vSwitch in such a mode.

-The DPDK support of Open vSwitch is considered experimental.
-It has not been thoroughly tested.

This version of Open vSwitch should be built manually with `configure` and `make`.
```

- Issues with DPDK:
  - How to Improve?
    - This thread, http://openvswitch.org/pipermail/dev/2015-August/058814.html
  - Some Suggestions from Thread
    - Device management:
    - Udev/systemd – (Flavio Leitner)
      - Device creation, binding, destruction – handled by Host OS

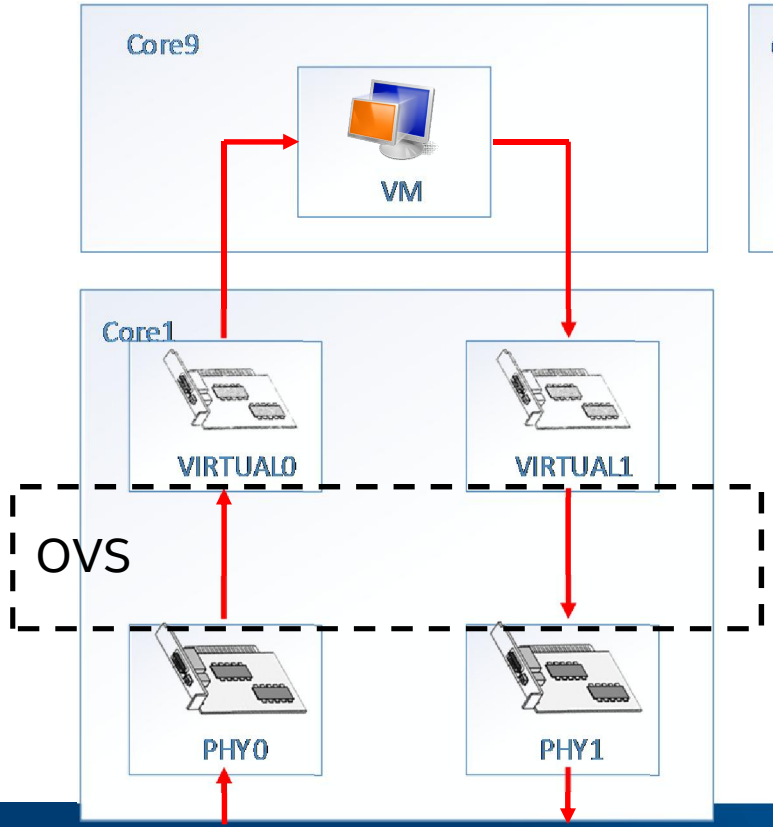
Improving DPDK in Open vSwitch

- How About Debugging?
  - TcpDump like capability
  - Use “Mirroring” of packets to pmd/libpcap or libpcap-ng
- How About Test Frameworks for DPDK
  - Add CI for Data Plane Testing
- Vsperf Project in OPNFV
- Support Only One type of vhost device
  - Drop Vhost – Cuse
- Better Documentation
  - Recent Patch to INSTALL.DPDK.md
- Education About Optimization Cache and Numa
  - From lstopo to Optimum DPDK
- Device Management
  - Driverctl utility for Fedora

Tuning – Core Affinities

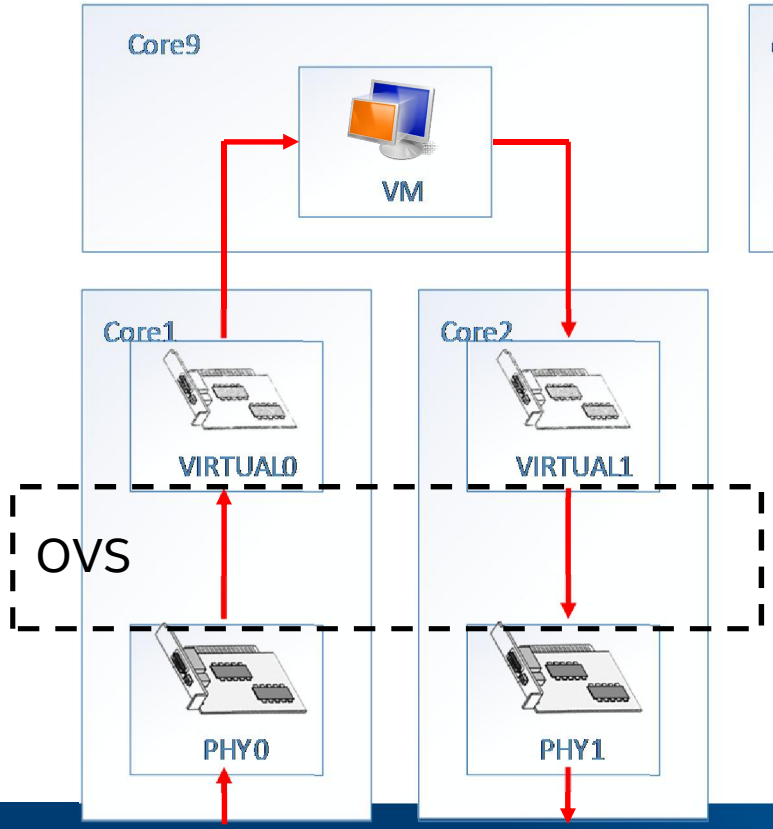


Tuning – Core Affinities



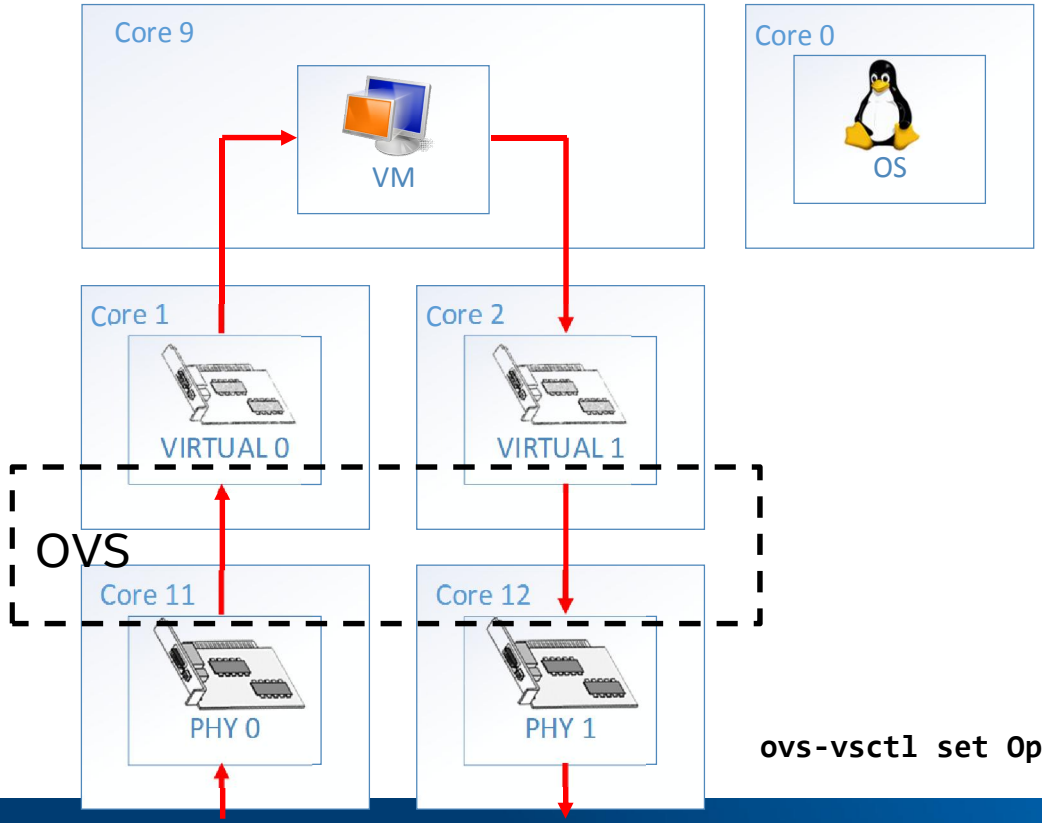
```
ovs-vsctl set Open_vSwitch . other_config:pmd-cpu-mask=2
```

Tuning – Core Affinities



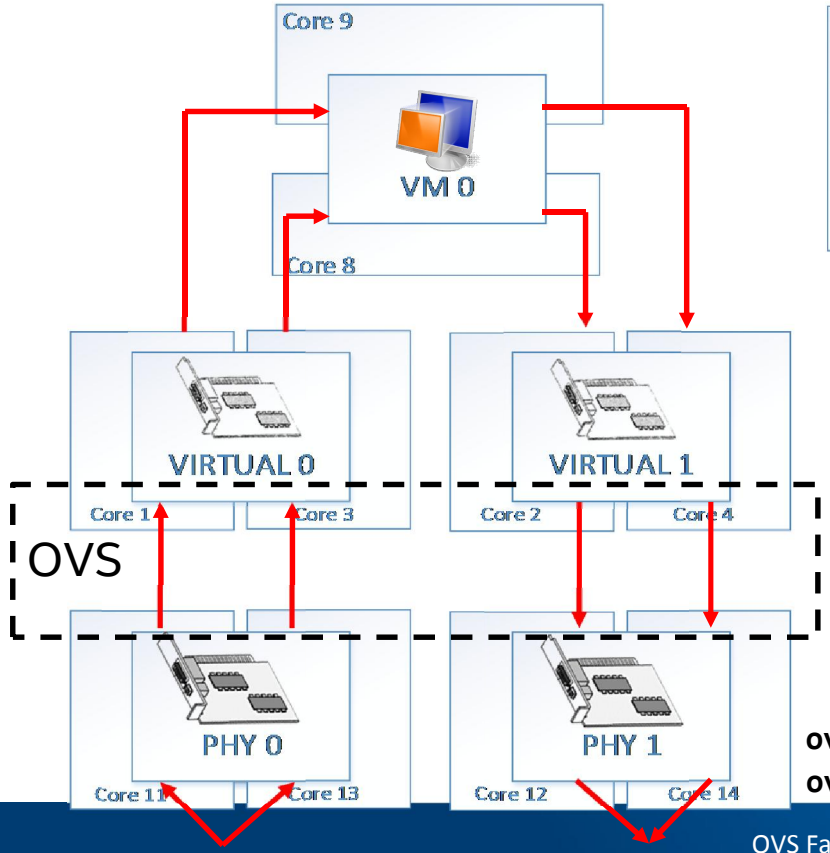
```
ovs-vsctl set Open_vSwitch . other_config:pmd-cpu-mask=6
```

Tuning – Core Affinities




```
ovs-vsctl set Open_vSwitch . other_config:pmd-cpu-mask=1806
```

Tuning – Multiqueue



```
ovs-vsctl set Open_vSwitch . other_config:pmd-cpu-mask=781E
ovs-vsctl set Open_vSwitch . other_config:n-dpdk-rxqs=2
```

Usability Guides

- Instructions for setup

  https://github.com/openvswitch/ovs/blob/master/INSTALL.DPDK.md#using-open-vswitch-with-dpdk

- Performance tuning

  https://github.com/openvswitch/ovs/blob/master/INSTALL.DPDK.md#performance-tuning

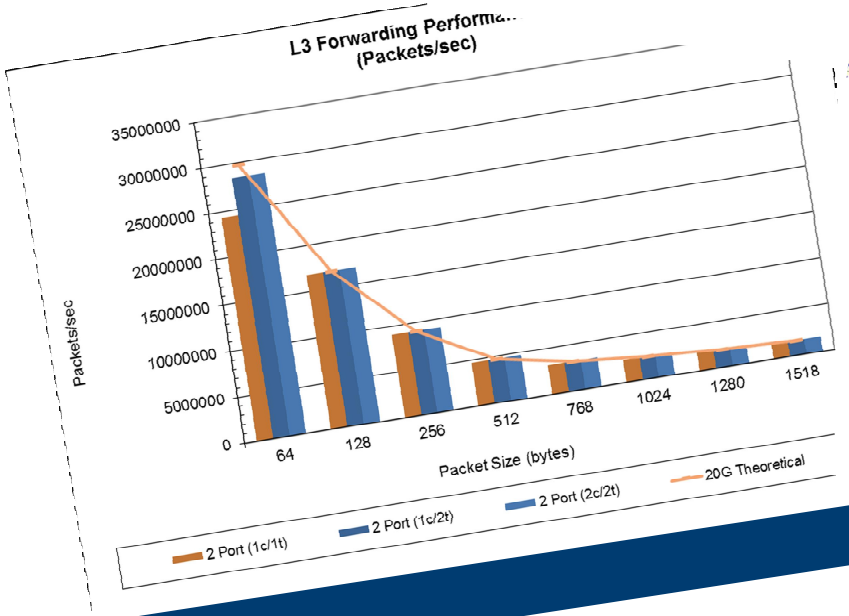
  https://download.01.org/packet-processing/ONPS1.5/Intel_ONP_Server_Release_1.5_Performance_Test_Report_Rev1.2.pdf

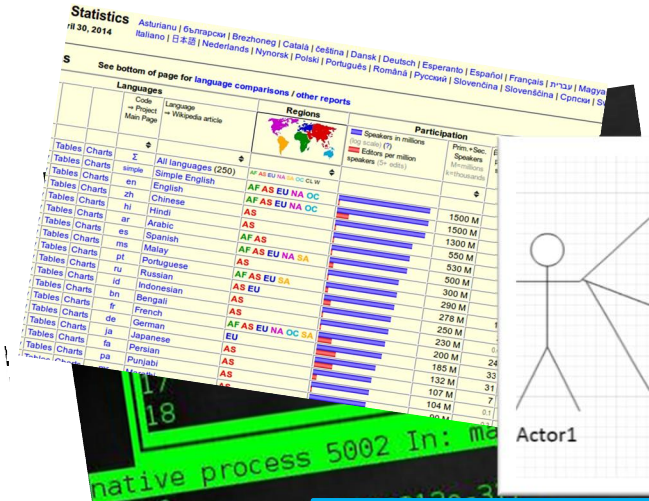
Availability: Distro Packages and Git

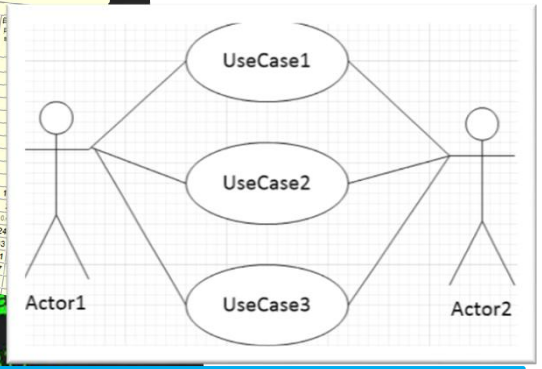
- ISV and OSV recognizing the progress of OVS with DPDK
- Centos7: 7.4: DPDK 2.1; 7.2: OVS 2.4
- Fedora: F23;F22 updates DPDK 2.0; F24: DPDK 2.1
- Fedora Copr repo for latest:https://copr.fedoraproject.org/coprs/pmatilai/dpdk/
- Red Hat OSP8...soon:
  - OVS/DPDK Integrated with Neutron
- Ubuntu: 15.10: OVS with DPDK package

- OVSNFV OPNFV Project planning deployment in future OPNFV releases
- https://01.org/packet-processing/intel%C2%AE-onp-servers
- git clone http://dpdk.org/git/dpdk
- git clone https://github.com/openvswitch/ovs.git

Conclusion

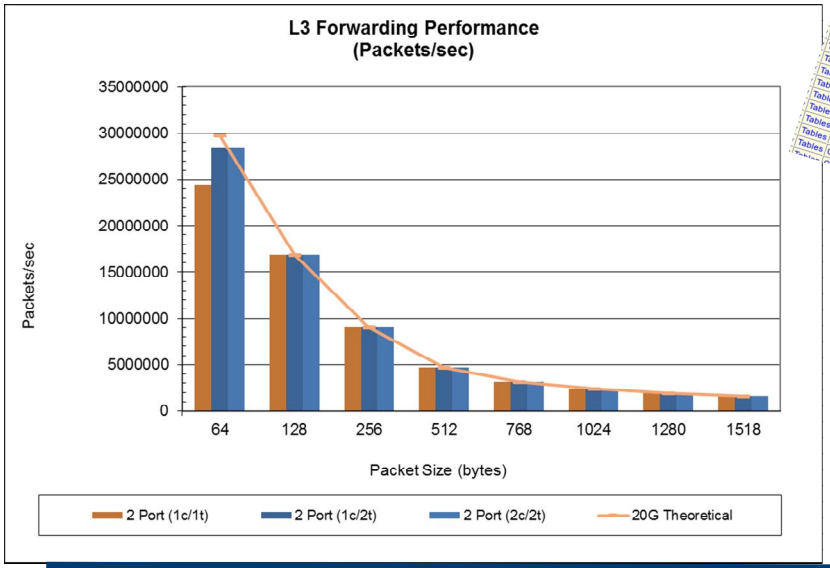


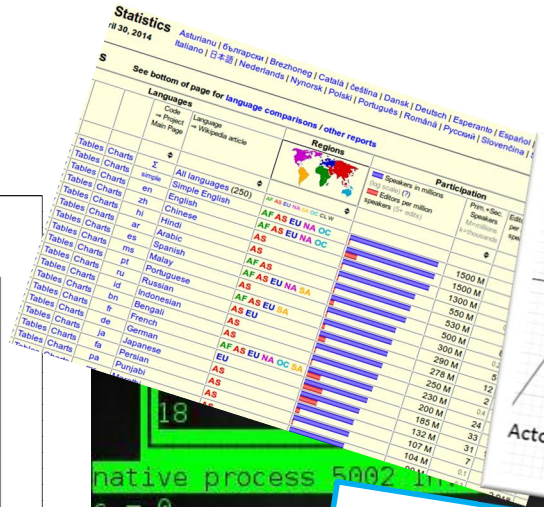


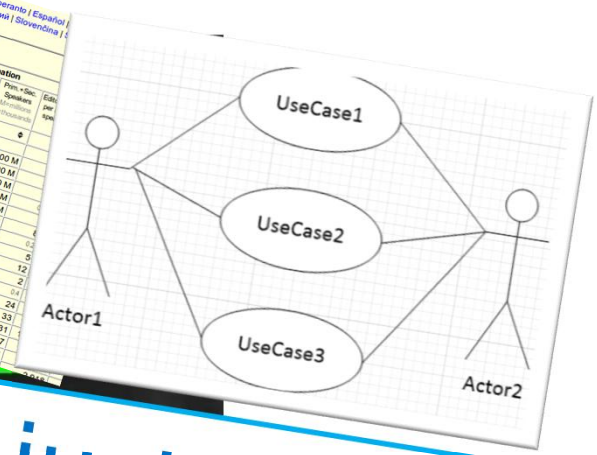


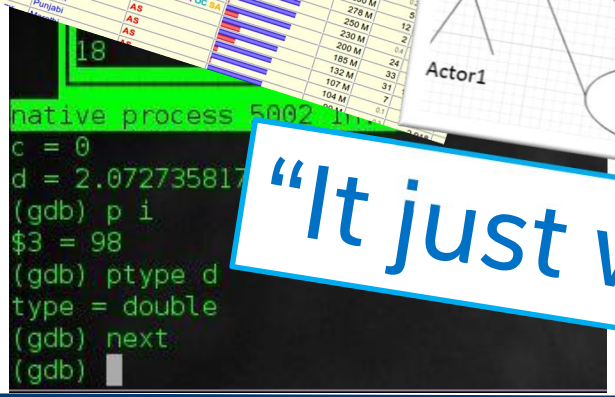
"It just works"

Conclusion



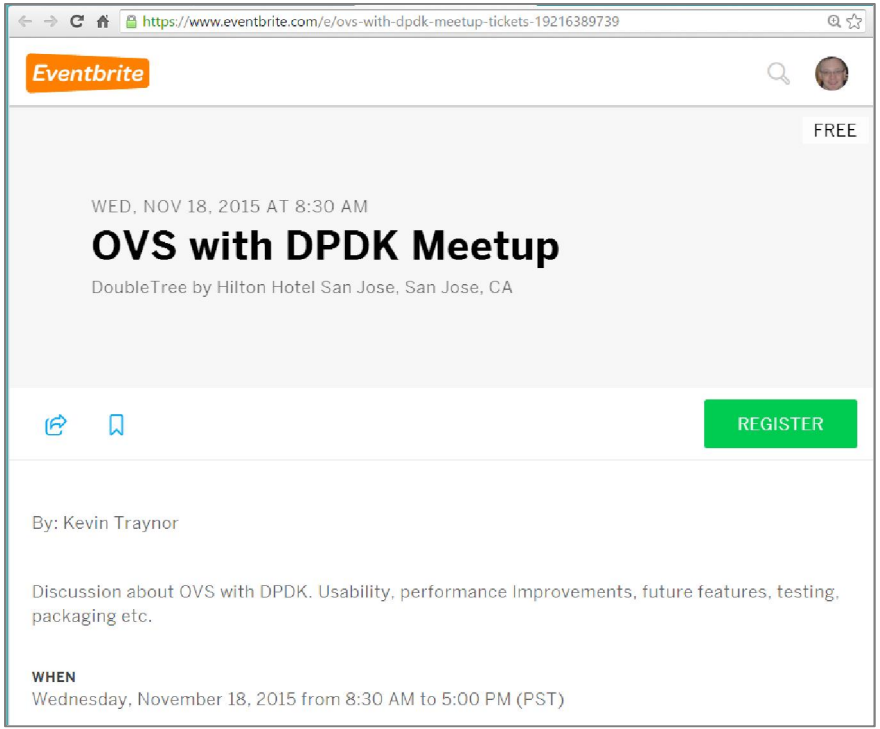






"It just works"

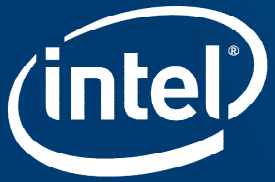
OVS with DPDK Meetup

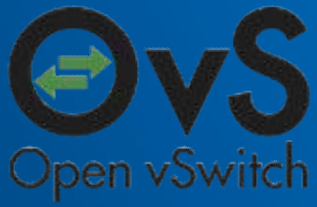


8:30 - 9:00	**Arrive**
9:00 - 10:45	**Usability Session** Discuss the issues highlighted in ML post here and ways to improve usability http://openvswitch.org/pipermail/dev/2015-August/058814.html
10:45 - 11:00	**Mid morning break**
11:00 - 12:30	**Features Session** Discuss status of current and potential userspace features e.g. vhost multiqueue, vhost-cuse, ivshmem, conn track, SFC, QoS etc.
12:30 - 1:30	**Lunch**
1:30 - 2:30	**Performance Session** Discuss ways to increase performance e.g. host-guest, guest-guest, tunneling, emc, dpcls etc.
2:30 - 2:45	**Mid afternoon break**
2:45 - 3:45	**Adoption/Community Session** Discuss ways to grow community and adoption e.g. Downstreaming, DPDK and QEMU community dependencies, further events
3:45...	**Open Mic/Wrap up**

https://www.eventbrite.com/e/ovs-with-dpdk-meetup-tickets-19216389739

Legal Disclaimer





Questions?