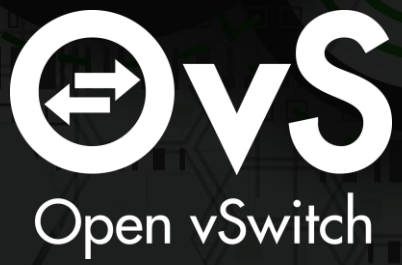
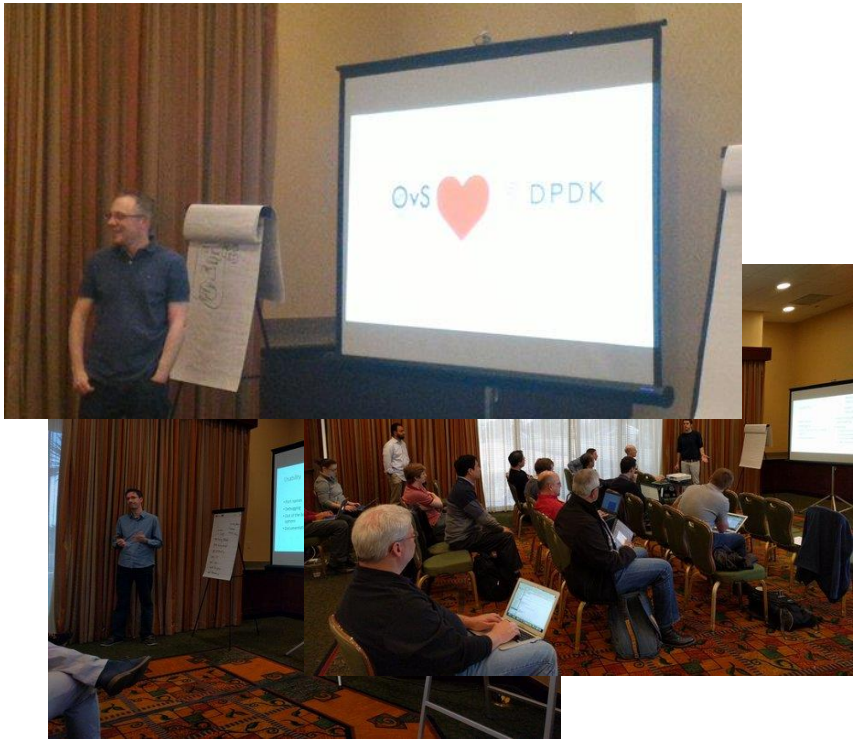


Aaron Conole, Red Hat
Robin Giller, Intel
Bhanuprakash Bodireddy, Intel

OvS-DPDK Usability Improvements for Real-World Applications

This time last year…

In a conference room not too far away



Branch: master ▾ ovs / INSTALL.DPDK.rst Find file

mvbpolito doc: fix bad link to dpdk advance installation guide f799f7a 2 c

2 contributors

605 lines (445 sloc) 19.9 KB Raw Blame History

Open vSwitch with DPDK

This document describes how to build and install Open vSwitch using a DPDK datapath. Open vSwitch can use the DPDK library to operate entirely in userspace.

Warning

The DPDK support of Open vSwitch is considered 'experimental'.

Build requirements

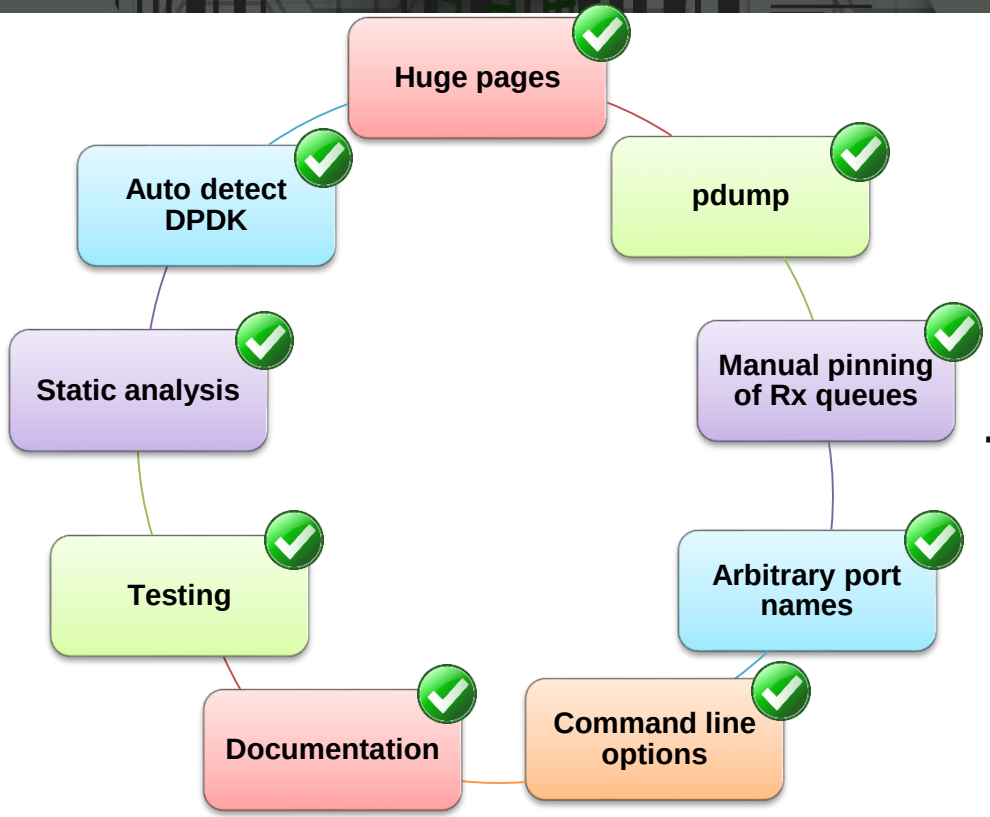
In addition to the requirements described in the installation guide, building Open vSwitch with DPDK will require the following:

- DPDK 16.07
- A DPDK supported NIC

  Only required when physical ports are in use

- A suitable kernel

Usability checklist



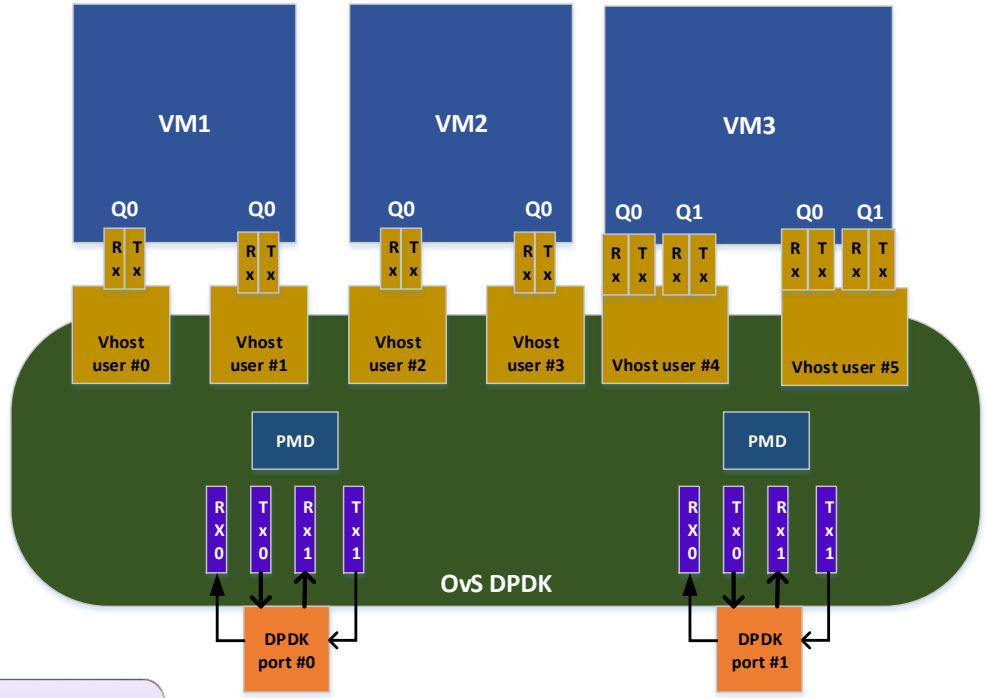
The DPDK support of OvS is considered 'excellent'

Huge pages

Huge pages

- Huge pages need kernel and HW support
  - CONFIG_HUGETLBFS should be enabled
  - Architectures support multiple pages(4k, 8k, 64k, 256K, 1M, 4M, 16M, 256M, 1G)
  - Cpu flags(***pse*** – 2M, ***pdpe1gb*** – 1G huge pages)
- Gigantic pages (1GB pages)
- Persistent vs run-time allocation of Huge pages
- Performance

Manual pinning of rx queues




```
pmd thread numa_id 0 core_id 4:
        isolated : false
        port: dpdk0     queue-id: 0
        port: dpdk1     queue-id: 0
        port: dpdkvhostuser0    queue-id: 0
        port: dpdkvhostuser2    queue-id: 0
        port: dpdkvhostuser4    queue-id: 0
        port: dpdkvhostuser5    queue-id: 1
pmd thread numa_id 0 core_id 5:
        isolated : false
        port: dpdk0     queue-id: 1
        port: dpdk1     queue-id: 1
        port: dpdkvhostuser1    queue-id: 0
        port: dpdkvhostuser3    queue-id: 0
        port: dpdkvhostuser4    queue-id: 1
        port: dpdkvhostuser5    queue-id: 0
```

```
pmd thread numa_id 0 core_id 4:
        isolated : true
        port: dpdk0     queue-id: 0 1
        port: dpdkvhostuser0    queue-id: 0
        port: dpdkvhostuser3    queue-id: 0
        port: dpdkvhostuser2    queue-id: 0
        port: dpdkvhostuser1    queue-id: 0
pmd thread numa_id 0 core_id 5:
        isolated : true
        port: dpdkvhostuser4    queue-id: 0 1
        port: dpdkvhostuser5    queue-id: 0 1
        port: dpdk1     queue-id: 0 1
```

Manual pinning of Rx queues

Arbitrary port names & Auto detect DPDK library

Arbitrary port names

$ ovs-vsctl add-port $BRIDGE dpdk0 – set interface dpdk0 type=dpdk

$ ovs-vsctl add-port $BRIDGE DPDKRX – set interface DPDKRX type=dpdk options:dpdk-devargs:0000:03:00.1

'dpdk-devargs' to indicate which DPDK port to associate with a given OVS port

Auto detect DPDK

$ export DPDK_BUILD=$DPDK_DIR/$DPDK_TARGET
$ ./configure --with-dpdk=$DPDK_BUILD

$ ./configure --with-dpdk

Auto discover DPDK library/headers if present in compiler search paths.

Documentation and static analysis

INSTALL.DPDK.md

- Install DPDK, OvS
- setup OvS

INSTALL.DPDK-ADVANCED.md

- System configuration
- Performance Tuning
- Test cases
- Vhost walkthrough

Documentation

Command Line:	make
Clang Version:	clang version 3.5.0 (tags/RELEASE_350/final)
Date:	Wed Nov 2 21:00:59 2016

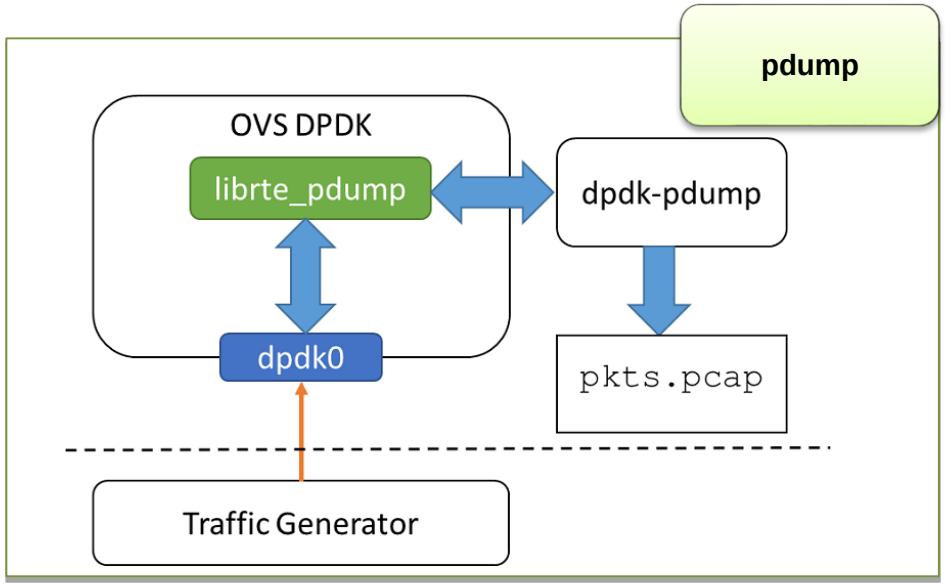
Bug Summary

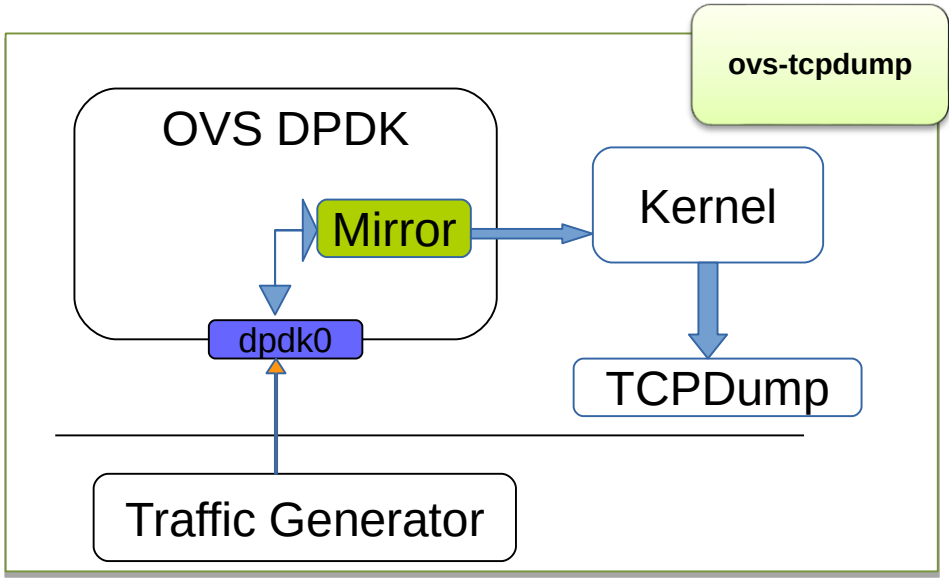
Bug Type	Quantity	Display?
All Bugs	**42**	☑
API		
Argument with 'nonnull' attribute passed null	13	☑
Dead store		
Dead assignment	1	☑
Logic error		
Assigned value is garbage or undefined	2	☑
Branch condition evaluates to a garbage value	1	☑
Dangerous variable-length array (VLA) declaration	1	☑

Static analysis

- **./boot.sh**
- **./configure CC=clang/ ./configure CC=gcc**
- **make clang-analyze**
- **Scan-view <results dir>**

Packet Tracing – Two approaches





OvS DPDK Start/Stop and library options

Command line options

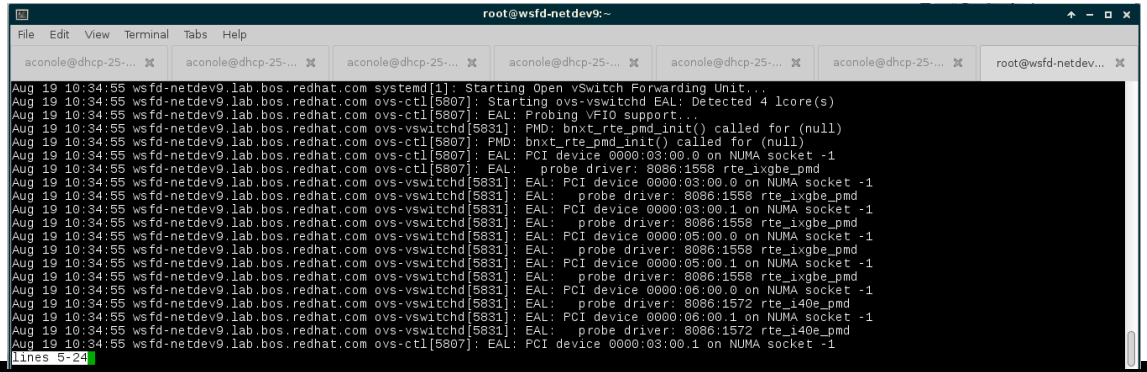
```
$ ovs-ctl  --no-ovs-vswitchd start
$ ovs-vsctl --no-wait set Open_vSwitch . other_config:dpdk-init=true
$ ovs-ctl --no-ovsdb-server start
```

(OR)

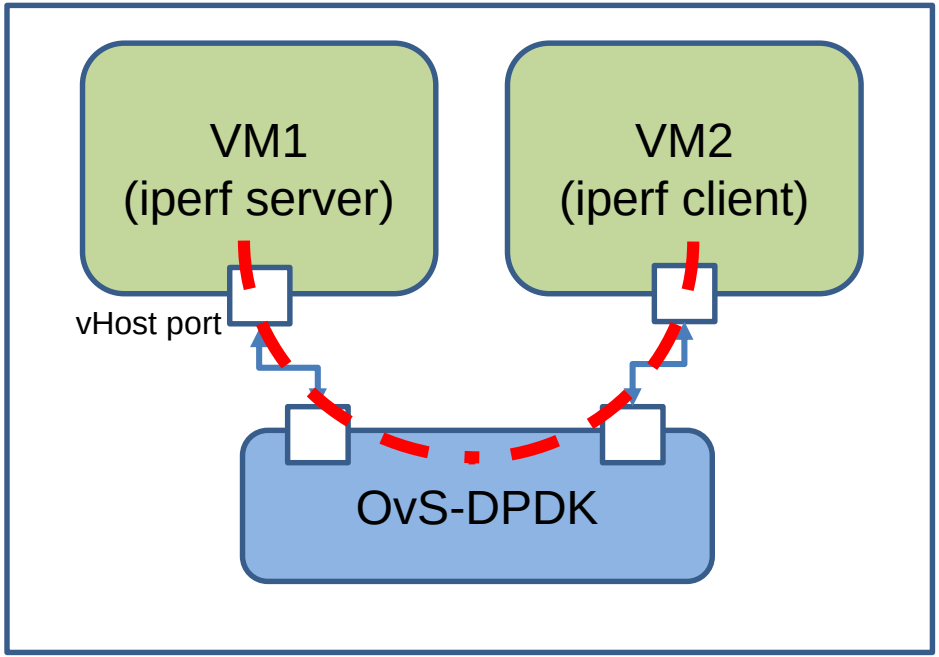
The RHEL systemd way

```
$ systemctl start ovsdb-server
$ ovs-vsctl --no-wait set Open_vSwitch . other_config:dpdk-init=true
$ systemctl start openvswitch
```

Values are set with **‘other_config:dpdk-extras’** or via a specific database key.

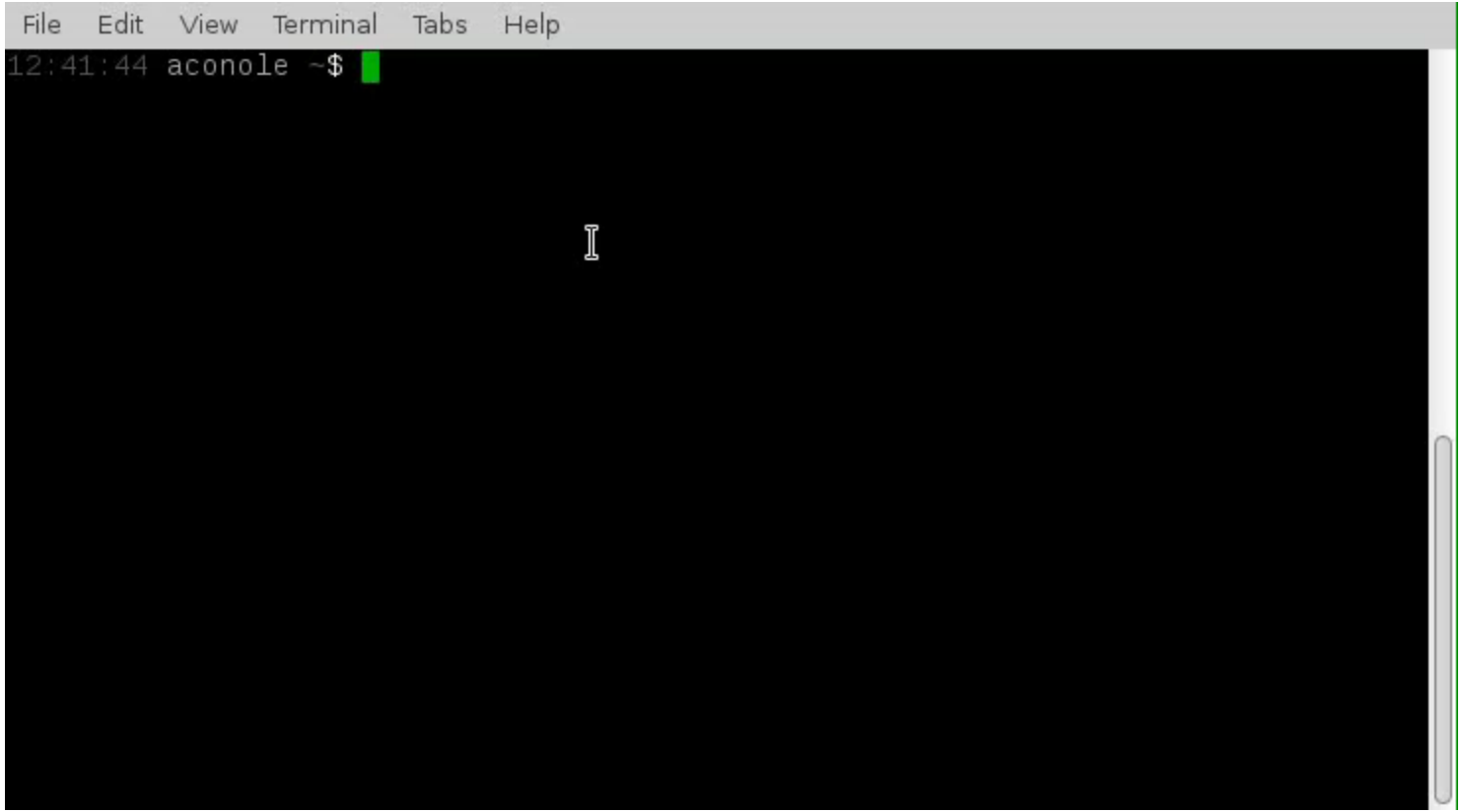


Demo



- Configuration
  - Host
  - VM
- Send traffic
- Packet dump

Demo



Summary

- Big focus on usability of OvS-DPDK and DPDK
- Many items addressed and merged in upstream project
- OvS-DPDK consumed by multiple Linux distributors and supported by installers
- Further discussion at OvS-DPDK design summit on Wednesday