DPDK Summit Bangalore - 2019

Technical Board and Roadmap

Hemant Agrawal
Jerin Jacob

Linux Foundation Project - Governing Board

DPDK
DATA PLANE DEVELOPMENT KIT

Chair
Jim St. Leger

Vice-Chair
Bob Monkman

Maen Suleiman

Board Members

Steven Nurenberg

Per-Erik Möckelind

James Hendergart

Erez Scop

Jaswinder Singh

Rashid

Songmin Yan

Silver Board Rep
George Zhao

Tech Board Rep
Konstantin Ananyev

Linux Foundation Project Manager
Trishan de Lanerolle

DPDK Technical Board History

- DPDK has a technical board to allow trusted contributors to the project to make technical decisions on behalf of the whole community
- The 9 current members of techboard@dpdk.org are:
  - Bruce Richardson
  - Ferruh Yigit
  - Hemant Agrawal
  - Jerin Jacob
  - Konstantin Ananyev *
  - Maxime Coquelin
  - Olivier Matz
  - Stephen Hemminger
  - Thomas Monjalon

* Current Tech board representative in DPDK Governing board (Quarterly Rotation basis)

What Does DPDK Tech-Board Do?

The Tech-board meets every two weeks (approx.) for the purposes of:

- Approving the inclusion of new functionality, libraries or device classes, into DPDK
- Approving any new sub-trees or staging trees for DPDK project work
- Approving the appointment of committers/maintainers for those trees
- Approving any subprojects on DPDK
  - Once created – subprojects get their own technical board
- Providing technical input on contributions when requested
- *Making technical decisions when consensus cannot be reached on-list*

Release process

- 4 major releases per year (.02, .05, .08, .11)
- 1 new LTS branch per year (17.11, 18.11, ….)

- Maintainer’s role
  - answer questions
  - make sure patches are reviewed
- Committer's role
  - easy: apply patches
  - medium: ask for review to trusted people
  - hard: reject patches when they are not mature enough

- local patch → patchwork → sub-sub-tree → sub-tree → main tree → stable tree

Repositories – maintainers (https://git.dpdk.org/)

- Main trees
  - dpdk.git *Ferruh, Thomas*
  - dpdk-stable.git *Luca, Kevin*
- Tools
  - dpdk-web.git *Thomas*
  - dpdk-ci.git *Thomas*
  - dts.git *Yong*
  - stable-scripts.git *Luca*
- Apps
  - pktgen-dpdk.git *Keith*
  - *nff-go*
  - spp.git *Ferruh*
  - *dpdk-burst-replay*

- Sub-trees
  - dpdk-next-crypto.git *Akhil*
  - dpdk-next-eventdev.git *Jerin*
  - dpdk-next-net.git *Ferruh*
    - dpdk-next-net-intel.git *Qi*
    - dpdk-next-net-mlx.git *Shahaf*
    - dpdk-next-virtio.git *Maxime*
    - dpdk-next-pipeline.git *Cristian*
  - dpdk-next-qos.git *Cristian*

- Draft trees (temporary)
  - dpdk-draft-cli.git *Keith*
  - dpdk-draft-windows.git

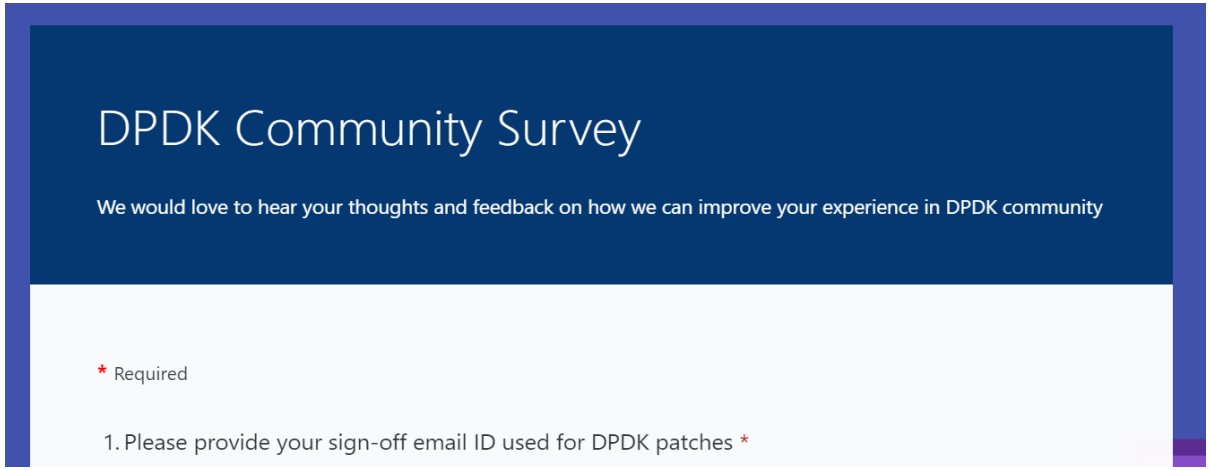
➢ Main tree also available at: https://github.com/DPDK

Tools

- Patches → Patchwork
  - http://dpdk.org/dev/patchwork
  - per-patch community CI integration
- Performance → Community Lab
  - Provide a monitoring dashboard
- Bugs → Bugzilla
  - http://dpdk.org/tracker
- Static Code Analysis → Coverity
  - https://scan.coverity.com/projects/dpdk-data-plane-development-kit
- Features → Roadmap
  - dev@dpdk.org
  - http://dpdk.org/dev/roadmap

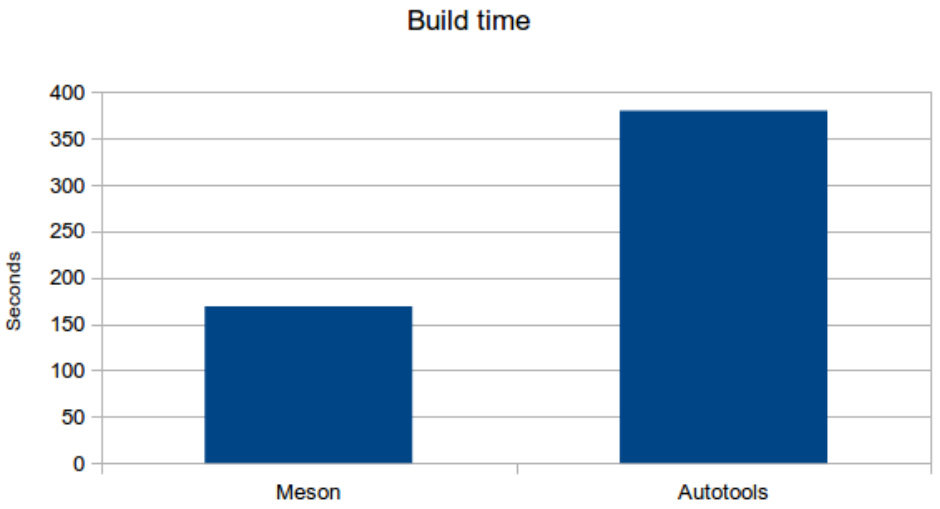
DPDK Survey: Live for your feedback

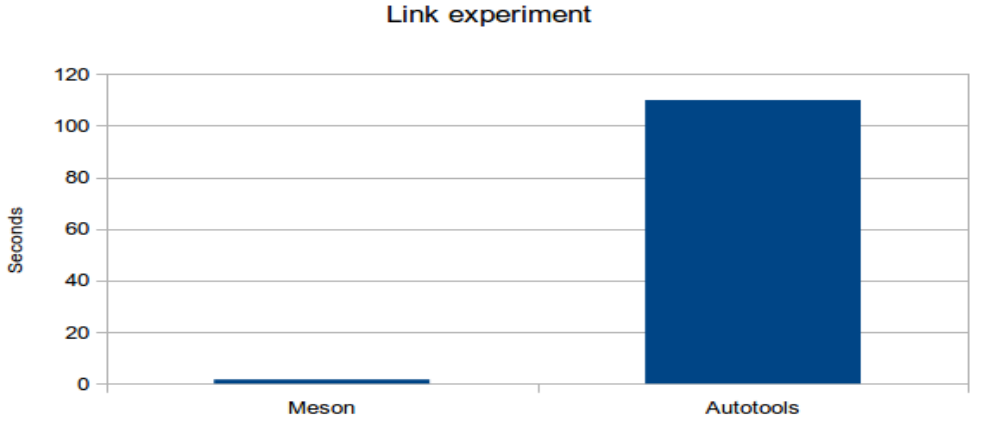
http://goo.gl/qZUgxG
Or,
https://forms.office.com/Pages/ResponsePage.aspx?id=eVlO89lXqkqtTbEipmlYTcwgJ8psxytOnArCkHeSZSZUREdIN09QOEVRSUJWN0l2TzNYUTk5STVJRC4u



Build System migration – Meson (http://mesonbuild.com)

- DPDK build system is currently based on “make” by default
  - Static configuration only - no dynamic configuration at build time
  - Custom build system with many complicated makefiles
- New build system using “meson” was introduced.
  - Open source build system meant to be both extremely fast, and, even more importantly, as user friendly as possible.
  - Very popular with open-source projects, e.g. Xorg, system
- Meson advantage:
  - support for detecting dependencies on the system, compiler features, including functions, defines.
  - faster builds using ninja - especially in the delta or nothing-has-changed case.
  - pkg-config support
  - dependencies in each lib can be moved back to being tracked in the libs files themselves, not up a level





Ref. http://mesonbuild.com/ARM-performance-test.html#arm-performance-test for building glib

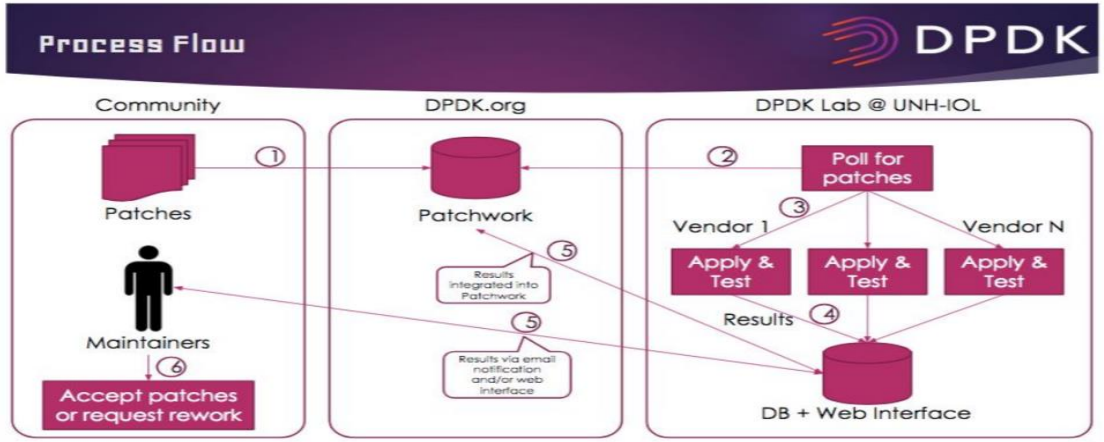
DPDK will migrate to meson in 19.08 and discontinue Makefile in 20.02

DPDK Community Lab

University of New Hampshire InterOperability Laboratory

DPDK DATA PLANE DEVELOPMENT KIT

- Objective and scope
  - Identify any regression in DPDK performance/function.
  - Identify any regression in the performance of DPDK-enabled application.
  - Demonstrate any new feature performance of DPDK.
  - May be used as a training or demo lab for DPDK events.
- Host: University of New Hampshire InterOperability Lab (UNH-IOL)
- Existing Test Hardware in Lab
  - Intel Ethernet Converged Network Adapter 82599ES 10 Gbps
    •
  - Intel Ethernet Converged Network Adapter XL710-QDA2 40 Gbps
  - Mellanox ConnectX-5 100 Gbps
  - Mellanox ConnectX-4 Lx 25 Gbps
  - Mellanox ConnectX-4 Lx 40 Gbps
  - NXP LS2088ARDB in the process of setting up with the Lab



http://core.dpdk.org/lab/

Context	Check	Description
ci/intel-Performance-Testing	success	Performance Testing PASS
ci/mellanox-Performance-Testing	success	Performance Testing PASS
ci/Intel-compilation	success	Compilation OK
ci/checkpatch	warning	coding style issues

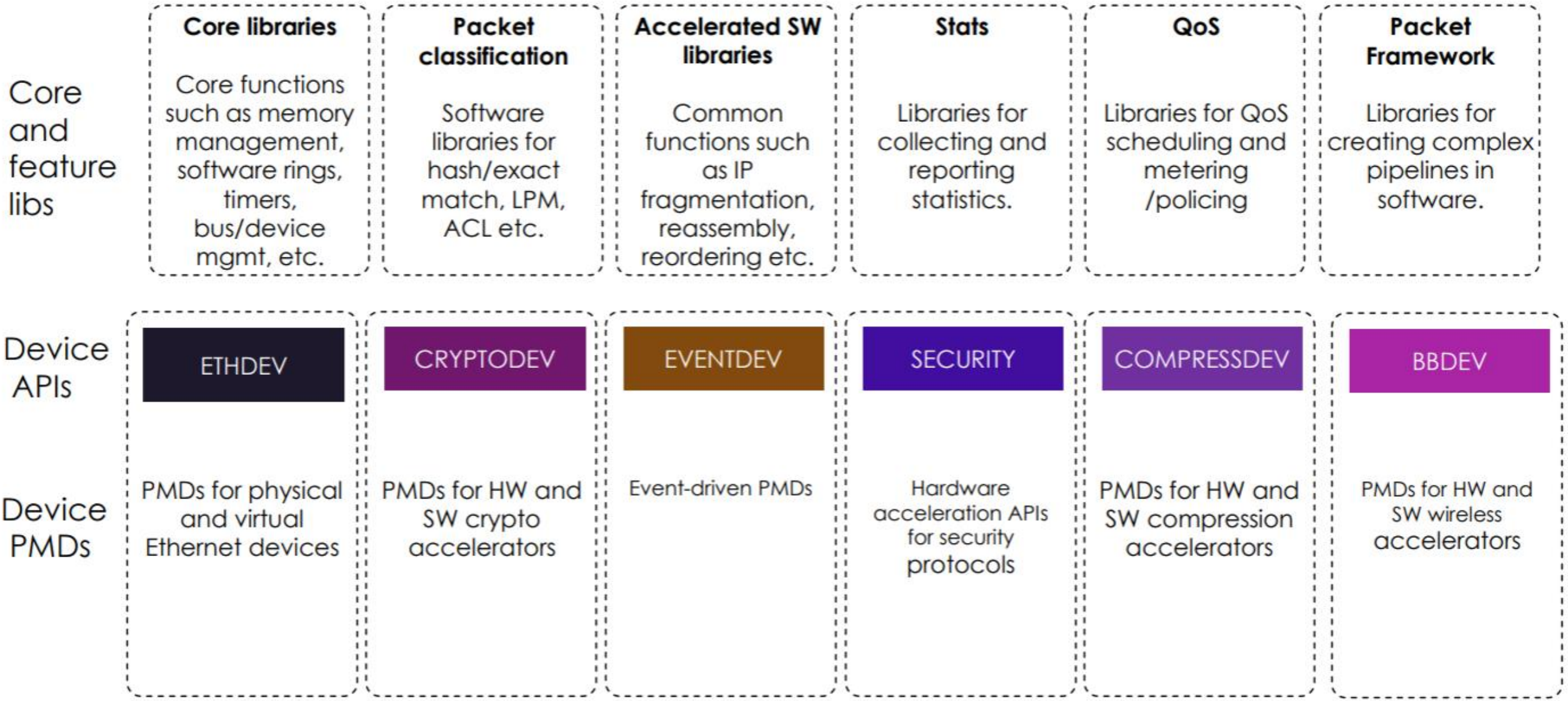
intel

Mellanox TECHNOLOGIES

NXP

DPDK libraries

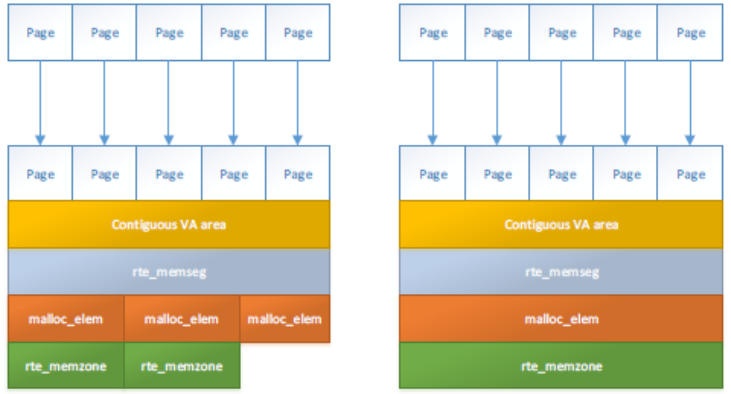
DPDK
DATA PLANE DEVELOPMENT KIT



Memory hotplug

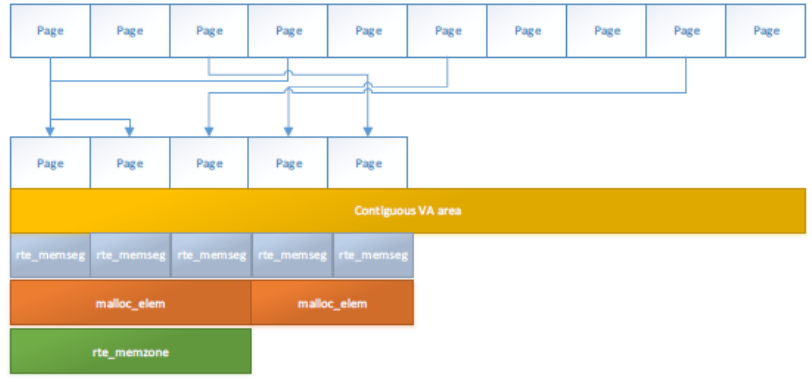
Legacy DPDK Memory Architecture

- VA layout follows PA layout
- VA and PA layout is fixed



18.05+ DPDK Memory Architecture

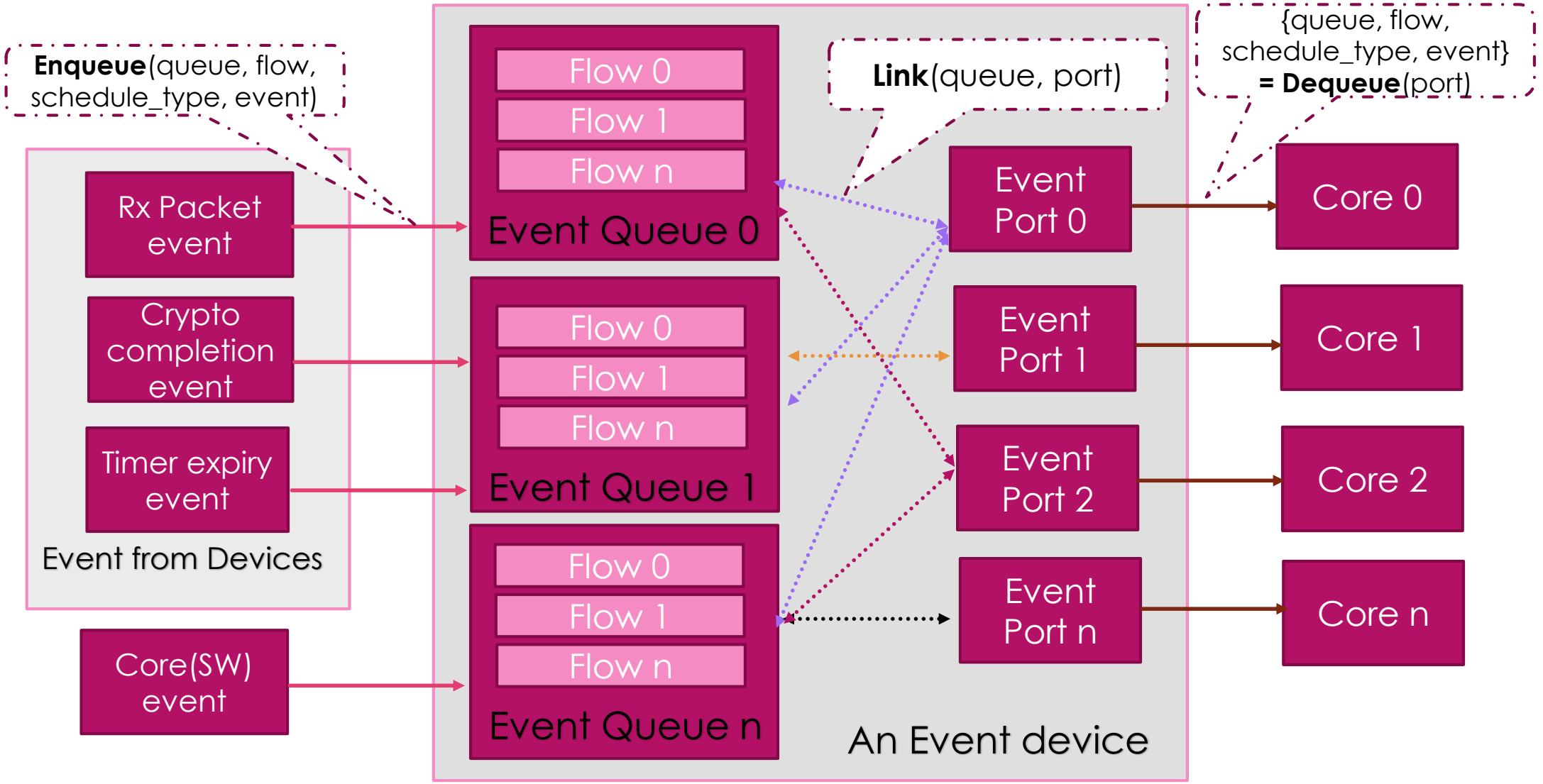
- VA layout is independent from PA layout
- VA layout is fixed, PA layout is not



- DPDK can now allocate hugepage memory as needed
- DPDK can also release memory that is unused
- Small page sizes and virtio are not enemies anymore!
- (18.08+) DPDK no longer requires a hugetlbfs mountpoint

- Memory is no longer guaranteed to be IOVA-contiguous
- what if you need IOVA-contiguous memory
- - you may actually not need it with VFIO
- - Ask for it - Memzone has a flag
- - Use legacy mode
- External memory is also supported.

event driven model – no more 100% polling



Latest DPDK releases (19.02 & 18.11)

- General:
  - ability to use externally allocated memory
  - new hotplug features, including multi-process and PCI failure handler
  - extendable table and lock-free r/w concurrency in hash library
  - traffic pattern aware power management
  - JSON interfaces for power policy example and new telemetry library
- Networking:
  - MAC swap, MPLS encapsulation and metadata matching in rte_flow API
  - new networking drivers for Aquantia Atlantic, Marvell Armada and NXP ENETC
  - postcopy live-migration in vhost-user
  - vDPA sample application
  - classification, metering and crypto in SoftNIC (using Packet Framework)
  - virtio packed ring (19.02)
- Cryptography:
  - new crypto drivers for Cavium OCTEON TX and NXP CAAM JR
  - PDCP in security library
  - library for Ipsec (19.02)
- Event mode:
  - eventdev Tx adapter, Eventdev Crypto adapter
  - new event driver by Ericsson: DSW (distributed software eventdev PMD)
- Applications:
  - noisy VNF forward mode in testpmd
  - FIPS validation application
  - IPSEC GW updated for rte_ipsec library

522	Intel (61)	301	Intel (54)
233	Mellanox (13)	66	Mellanox (10)
125	Cavium (18)	33	RedHat (6)
95	NXP (8)	26	NXP (6)
50	Solarflare (10)	24	Semihalf (2)
35	AT&T (2)	17	Microsoft (2)
32	OKTET Labs (4)	13	6WIND (3)
31	Semihalf (6)	11	Rami Rosen (1)
		9	Samsung (1)

Roadmap – to 19.05 and beyond!

- Initial windows support with meson and clang
- New device specification (devargs) syntax
- Documentation for device management
- DMA mapping API for external memory
- Relaxed memory ordering in spinlock and rwlock
- ticket-lock
- RCU library
- lock-free extended bucket in hash library
- TCP SEQ and ACK offload with rte_flow API and mlx5 implementation
- ICMP ping offload with rte_flow API and mlx5 implementation
- New mlx5 steering flow engine for high (millions/sec) insertion rate
- Failsafe support in multi-process
- AF_XDP poll mode driver
- baseband device turbo PMD
- NXP DPAA multi process application support
- NXP DPAA split Mac driver to split incoming traffic between kernel and DPDK.
- QAT compression PMD support for large scatter-gather lists
- QAT asymmetric crypto with support for modexp and modinv
- additional crypto/auth algorithms in the IPsec library
- Octeontx2 poll mode drivers

DPDK Adoption in Open Source Projects

DPDK
DATA PLANE DEVELOPMENT KIT

Open Source Projects Leveraging DPDK

ANS

BESS

Butterfly

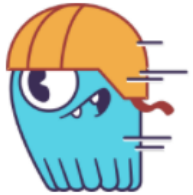
MoonGen

mTCP

Packet-journey

Pktgen-dpdk

Ruru



SPDK

YANFF


DPDK
DATA PLANE DEVELOPMENT KIT
Questions?