DPDK PACKET CAPTURE

THE NEXT GENERATION

Stephen Hemminger

DPDK Userspace Summit – Bordeaux 2019

History

- Libpcap
  - Lawrence Berkeley Lab 1998
  - Better than SunOS
  - Filtering BPF
- DPDK
  - rte_pdump

Earlier work

	Model	Capture Format	URL
rte_pdump	Secondary	pcap	*http://dpdk.org/git/dpdk*
dpdkcap	Primary	pcap	*https://github.com/dpdkcap/dpdkcap.git*
Libpcap dpdk	Primary	pcap/pcapng	*https://github.com/the-tcpdump-group/libpcap.git*
dpdk-pcapng	Secondary	pcapng	*https://github.com/shemminger/dpdk-pcapng.git*

Libpcap Issues

- Libpcap security – 141 CVE’s
  - Decoding packets in C is hard
- Limited DPDK native support
- Pcap file format
  - Timestamp limitations
  - No meta data
  - Single interface

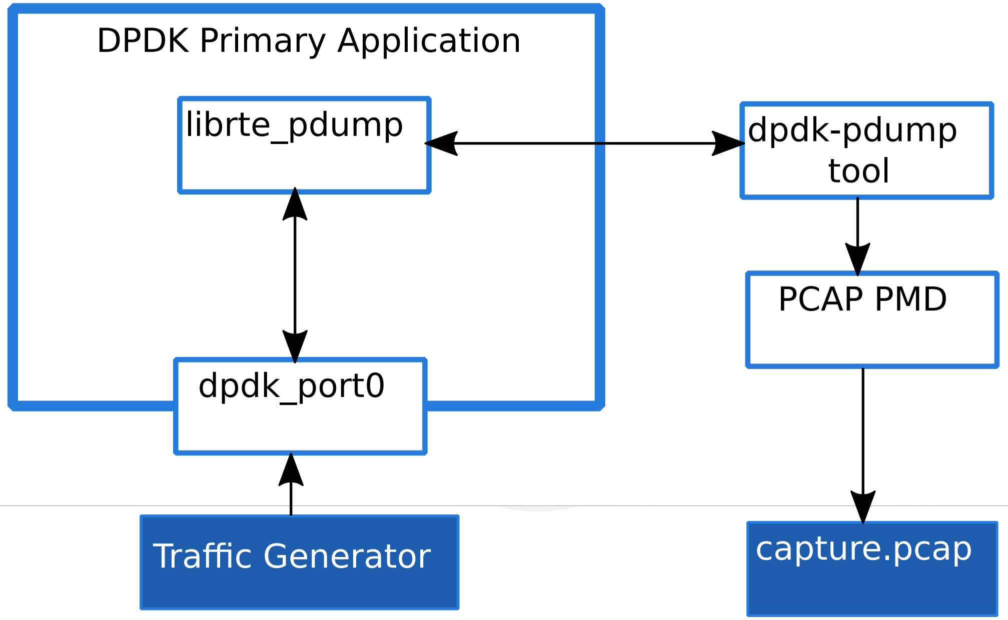
Pcapng

- Evolving standard
- Used by Wireshark/tshark
- TCPdump - read/only
  - https://github.com/pcapng/pcapng

```
Section Header
|
+- Interface Description
|   +- Simple Packet
|   +- Enhanced Packet
|   +- Interface Statistics
|
+- Name Resolution
```

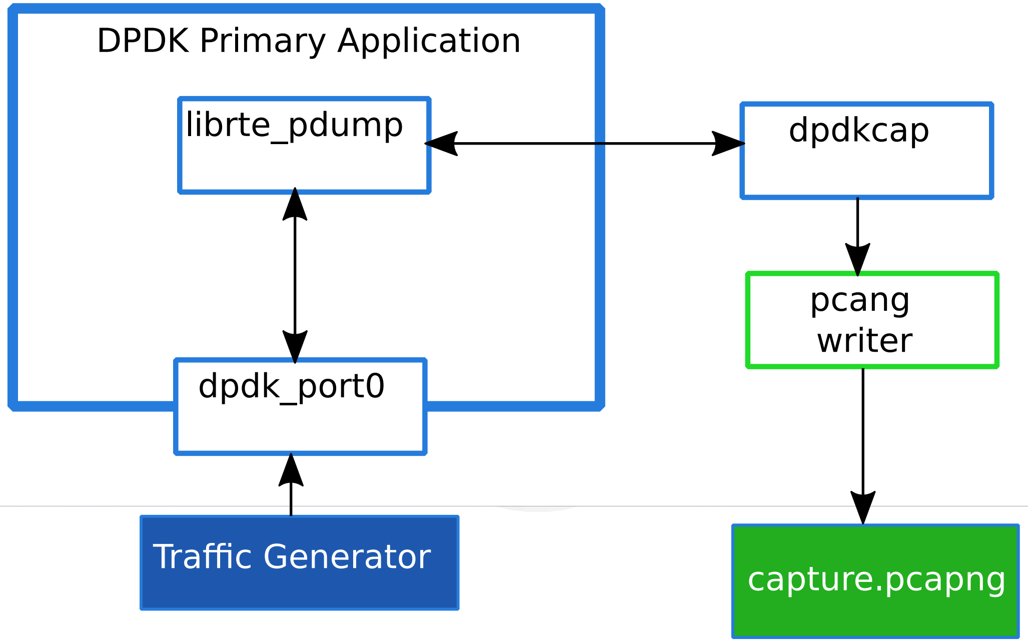
pdump

- rte_pdump
  - Rxtx callback hooks
  - Copy packet to new mbuf
  - Ring to secondary
- Pdump tool
  - Inject to instance of PCAP PMD



Pcapng Application

- Existing librte_pdump
  - Backward compatible enhancements
- New secondary process
  - Pcapng writer library



```
# dpdk-pcapng -h
Usage: dpdk-pcapng [options] ...

Interface:
  -i <interface>         name or port index of interface
  -D                     print list of interfaces and exit
Stop condition:
  -c <packet count>      stop after N packets (default: infinite)
Output file:
  -w <filename>          name of file to save (default: tempfile)
  -g                     enable group read access of output file
  -n                     use pcapng format instead of pcap (default)
Miscellaneous
  -N <packet limit>      maximum number of packets buffered (default:
2048)
  -q                     don't report packet capture counts
  -v                     print version information and exit
  -h                     display this help and exit
```

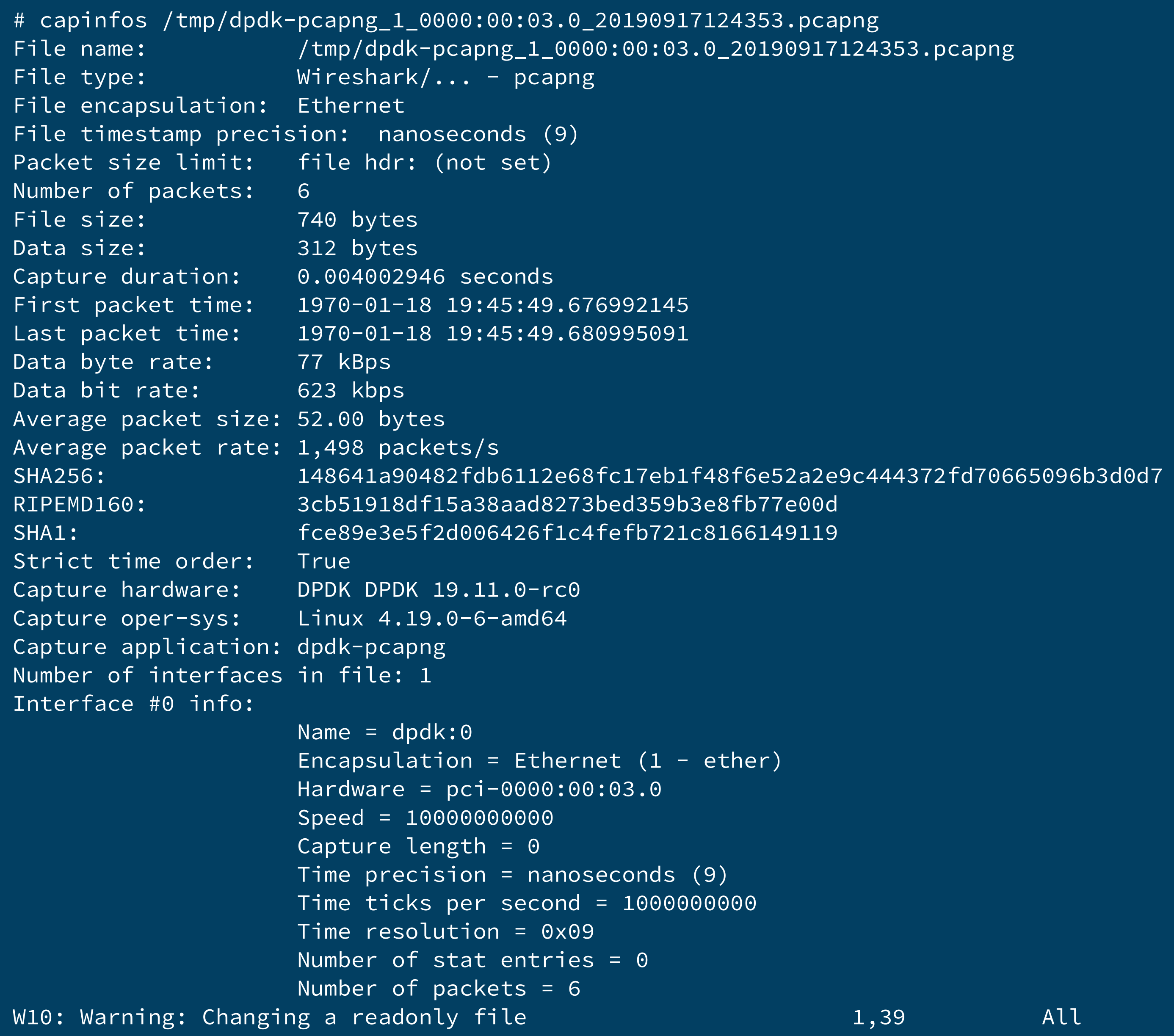
```
# dpdk-pcapng -D
0. 0000:00:03.0

# dpdk-pcapng -c 6
Packets captured: 6
Packets received/dropped on interface '0000:00:03.0': 6/0

# tshark -r /tmp/dpdk-pcapng_1_0000:00:03.0_20190917124353.pcapng
Running as user "root" and group "root". This could be dangerous.
    1 0.000000000 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
    2 0.000000002 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
    3 0.002017483 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
    4 0.002017485 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
    5 0.004002944 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
    6 0.004002946 fe:54:00:3b:29:82 → Spanning-tree-(for-bridges)_00 STP
52 Conf. Root = 32768/0/52:54:00:cc:30:31  Cost = 0  Port = 0x8002
```

Pcapng

- Section Header
  - OS, Hardware, Application
- Interface
  - Speed, Name, Description, ...
- Packet data
  - Timestamp (ns), Ifindex, flags, length (data, capture), ...



```
# capinfos /tmp/dpdk-pcapng_1_0000:00:03.0_20190917124353.pcapng
File name:           /tmp/dpdk-pcapng_1_0000:00:03.0_20190917124353.pcapng
File type:           Wireshark/... - pcapng
File encapsulation:  Ethernet
File timestamp precision:  nanoseconds (9)
Packet size limit:   file hdr: (not set)
Number of packets:   6
File size:           740 bytes
Data size:           312 bytes
Capture duration:    0.004002946 seconds
First packet time:   1970-01-18 19:45:49.676992145
Last packet time:    1970-01-18 19:45:49.680995091
Data byte rate:      77 kBps
Data bit rate:       623 kbps
Average packet size: 52.00 bytes
Average packet rate: 1,498 packets/s
SHA256:              148641a90482fdb6112e68fc17eb1f48f6e52a2e9c444372fd70665096b3d0d7
RIPEMD160:           3cb51918df15a38aad8273bed359b3e8fb77e00d
SHA1:                fce89e3e5f2d006426f1c4fefb721c8166149119
Strict time order:   True
Capture hardware:    DPDK DPDK 19.11.0-rc0
Capture oper-sys:    Linux 4.19.0-6-amd64
Capture application: dpdk-pcapng
Number of interfaces in file: 1
Interface #0 info:
                     Name = dpdk:0
                     Encapsulation = Ethernet (1 - ether)
                     Hardware = pci-0000:00:03.0
                     Speed = 10000000000
                     Capture length = 0
                     Time precision = nanoseconds (9)
                     Time ticks per second = 1000000000
                     Time resolution = 0x09
                     Number of stat entries = 0
                     Number of packets = 6
W10: Warning: Changing a readonly file                       1,39          All
```

rte_pdump enhancements

- Timestamp
  - Record time (tsc) when packet was captured
- Flags
  - Direction rx/tx

Future changes

- Cleanups
- Hotplug
- Multiple instances
- Mbuf ref count
- Snap length / Header only

Filtering

- Libpcap tools
  - classic BPF
- DPDK
  - Extended BPF

Summary

- Packet capture improvements
  - Command interface like wireshark (dumpcap)
  - Simplify architecture (no libpcap, no pcap PMD)
  - Pcapng output
    - Multiple devices
    - Timestamps
    - Metadata

Questions?

Stephen Hemminger

stephen@networkplumber.org