Life at a Networking Hardware Vendor

[Keeping up with the Joneses]

•••

Andy Gospodarek
Open Source Architect
Broadcom, Compute and Connectivity

Defined: Keeping up with the Joneses

- Comparing yourself with your neighbors (surname ‘Jones’ is used to reference a generic neighbor) as a benchmark for social class or acheivement
- In this context it would reference make sure networking hardware or software support for hardware matches or beats that of other vendors.

Broadcom makes lots of different networking hardware

This talk is not about wireless hardware

This talk is not about switching hardware

This talk is about NIC hardware

Broadcom’s goal is to sell as many NICs are possible

${VENDOR}’s goal is to sell as many NICs are possible

To sell the most hardware you *might* need to build the *best* hardware

Define *best*...

Highest Packets Per Second?

Lowest Power Consumption?

Lowest Price?

Most Offloads?

Least Offloads?

Most ARM/RISC-V cores?

Most programmable FPGA or NPU?

You also might need to make the best firmware and drivers?

Best is different for almost every [potential] customer

With 1500 byte packets, most NICs can send and receive at line-rate (10/25/40/50/100Gbps)

Some can handle line-rate traffic at smaller packet sizes

If you need line rate with 64 byte packets then you need to find a NIC that can handle it

Individual component costs are important

NICs that can offload work from server cores can justify a higher price

Spending more money on a NIC might save money other places

Look closely at the prices for processors as you scale the core density

Fixed Function Offload Evolution

Checksum Offload and TSO...

...GSO, LRO, Hardware GRO, UFO, RSS, XPS, RPS...

...Tunnel Encap/Decap...

...Flow Offload via Ntuple Filters or CLS Flower...

...TLS Offload...

...XDP/BPF Offload...

...Control Plane Offload

Seems unlikely that all offloads are being used at the same time

But vendors need to make sure they can support as many of those as possible

Unless users do not want to offload anything

Some just want the hardware
to get out of the way

Smart or *Dumb Nics*?

Snabb Switch creator would like to see a low-cost *Dumb NIC* with no offload features

General purpose processors on NICs

Gives users the chance to have a “server inside there server”

~~Turtles~~ Linux all the way down

Offload of control plane and dataplane to Smart NIC instead of using server cores

FRR on the NIC

Open vSwitch on the NIC

XDP/BPF maps and forwarding on the NIC

FRR + XDP for routing on the NIC

Speaking of programmable dataplanes...

FPGAs and NPUs fill the gap left by fixed-function devices

NPUs that allow offload of P4/XDP/BPF dataplane

FPGAs can do anything

Small Matter of Programming

Tough to justify FPGA development cost

Unless you can get your hardware or OS vendor to do it for you...

Best Firmware

Some hardware features are enabled by firmware

Firmware version impacts user experience

Firmware feels like a ‘black box’
even if open source

What makes a driver the *best* ?

Upstream is all that matters

Inbox is all that matters

Out of tree drivers are not going away

Does your driver support...

...all that your hardware supports

Checksum Offload, TSO, GSO, LRO, Hardware GRO, UFO, RSS, XPS, RFS, Tunnel Encap/Decap, Flow Offload, TLS Offload, XDP Offload...

Let’s not forget software dataplane support

XDP

DIM

DPDK poll mode driver

DPDK PMD vector support

DPDK rte_flow support

Kernel by-pass generally not preferred

~~DPDK poll mode driver~~

AF_XDP

AF_XDP is the new black

Seems simple to make the *best* NIC, right?

What should vendors set as their goal?

Minimize the number of instructions needed to process a packet

Offloading to hardware saves instructions

Optimizing drivers saves instructions

XDP saves instructions

AF_XDP saves instructions

DPDK saves instructions

No single hardware/firmware/driver combination works for everyone

Focus on everything?

Not realistic

What can we do to help users today?

What can we do to enable future users?

Obrigado!