YVR18-212: Network switches and TSN configuration

Ilias Apalodimas
Ivan Khoronzhuk
Why

- An increasing number of vendors are including TSN support on their hardware
- Great diversity in the ecosystem for configuring devices up to now. Most vendors rely on out-of-tree modules and home-grown user-space tools.
- This becomes a maintenance nightmare
Issues

- Until today there is no standardized API to configure TSN
- Only CBS (Credit Based Shaper) configuration is available on TI cpsw and Intel i210
- Available solutions for switch configuration do not cover all the corner cases for different hardware across vendors
Example hardware

- Trust Architecture
- GPU and LCD Controller
- SEC
  - ARM® V8
  - 1 MB L2 Cache
- EnetC
- 4 Port Eth TSN Switch
- Integrated TSN Switch (Next-Generation AVB)
- Next-Generation Ethernet Controllers with TSN

Source:
Configuration

● Some ethernet interfaces already support CBS configuration via `iproute2`
● Ideally we’d like to use the same tools for configuring the switch
● `switchdev` (or alternatively DSA offer those possibilities)
Configuration

Kernel driver (vendor specific)

Netlink

Hardware configuration

User space configuration (common)

sw0p1 sw0p2 sw0p3 sw0p4 enp2s0

4-port switch
How

- Switchdev creates as many network interfaces on a Linux host as the number of available physical ports.
- Switchdev should be preferred over DSA when possible. DSA re-injects frames on the network stack after removing proprietary tags and introduces latency.
So

- `iproute2` can be used to configure each port individually.
- Configuration of a single switch port will be very similar to configuring an ethernet interface.
- Each vendor must add their driver but they will have a similar configuration from user space point of view.
TI driver

After applying https://patchwork.ozlabs.org/cover/929367/ we can configure the switch using iproute2/bridge utils instead of using TI proprietary tool and tc to configure CBS

- bridge vlan add dev sw0p1 vid 100 pvid untagged master
- tc qdisc add dev sw0p1 parent 100:4 cbs locredit -1335 ...
Future plans

- Upstream switchdev functionality for TI cpsw driver
- tapprio (on RFC) and ETF scheduler (merged) are valid options for TSN
- Extend tapprio, a software scheduler, to configure 802.1Qbv hardware features
  (https://www.spinics.net/lists/netdev/msg513149.html)
Future plans

- TSN is not only limited to configuration. Look into AF_XDP for offloading traffic to userspace.
Measurement approach

- hw timestamps for rx and tx packets
- hi resolution PHC counter
- PTP packets, both L2 and L4
- interclass/interchannel impact
- complete latency from app to wire
- latency in stack, packet scheduler and vlan layer, NIC driver
- network round trip latency
- even speed measured with hw ts
- shaping tests
- fast test with only one board
Layers latency impact

Layer: am572x, cpsw, ptpl2, itf speed = 1000Mbps,
packet = 512B, rate = 100.000pps, 412.80kbps,
Comple: min = 24.40us, max = 50.40us, mean = 26.00us mean+RMS = 27.97 +/- 1.15 us
Stack: min = 6.02us, max = 16.92us, mean = 10.90us mean+RMS = 7.36 +/- 0.45 us
Vlan: min = 2.44us, max = 7.97us, mean = 5.53us mean+RMS = 3.15 +/- 0.53 us
Driver: min = 15.56us, max = 27.63us, mean = 12.07us mean+RMS = 17.46 +/- 0.64 us
Interchannel impact investigation

- Based on interpacket gap
- As integrated characteristic, includes hw and sw impact
- No need to sync with system time, resolution 4ns
Interchannel impact investigation

HW Class B (par A) RMS = 0.08us

am572x, cpsw, class B (par A), ptp2, CBS, CPDMA, if speed = 1000Mbps, packet = 516B, rate = 4664.2pps, 19253.77kbps, min = 210.08us, max = 214.87us, p-t-p = 4.79us, mean + RMS = 214.40 + 0.08 us

HW Class B (par A and BE) RMS = 0.79us

am572x, cpsw, class B (par A + BE), ptp2, CBS, CPDMA, if speed = 1000Mbps, packet = 516B, rate = 4664.2pps, 19253.73kbps, min = 209.68us, max = 219.13us, p-t-p = 9.45us, mean + RMS = 214.40 + 0.79 us
TI cpsw sw vs hw CBS shaping
TI cpssw sw vs hw CBS shaping

Software CBS shaping

Hardware cpssw CBS shaping

Class A (Class B + BE in parallel)

Class B (Class A + BE in parallel)
Example of switch shaper tests

- AM572xx switch mode
  - PORT1 1000Mb
  - PORT2 100Mb
  - TSN talker class A 40 MBps

- BBB
  - PORT 100Mb
  - TSN talker class A 40 MBps

- Laptop
  - PORT 1000Mb
  - TSN listener (do nothing, just traffic indicator)

PHC timestamp (observe how it’s changed)
TI cpsw CBS configuration and test links

https://github.com/khorivan/plget
Documentation/networking/ti-cpsw.txt, since LK4.19-rc1
https://projects.linaro.org/secure/attachment/14172/cpsw_cbs_tests.pdf
https://git.linaro.org/people/ivan.khoronzhuk/tsn_latencies.git
https://git.linaro.org/people/ivan.khoronzhuk/tsn_kernel.git/log/?h=vlan_hw_ts_v2
Q&A

Questions?
Thank you!