Feedback & discussion

New WRS-4 Workshop
2020-06-26
Basics

● References:
  ○ “Hardware study” - refers to “Study on the new hardware features for the WRS-4”
  ○ “Main board arch” - refers to “Study on the WRS-4 main board (Hardware Architecture)"

● Feedback:
  ○ Tracked with issues in wr-switch-hw-v4 project: https://ohwr.org/project/wr-switch-hw-v4/issues

● First: gather all feedback without discussion (next page)

● Second: discuss in the order of importance
Feedback

- **Ports**
  - At least 20 ports, ideally 24 ports (HDL support)
  - Use compact SFP (not for DWDM)
  - It should be possible to have 1Gb on 18 ports (could be with different bitstreams)
  - Switch on/off the laser of the SFP

- **Clocking circuit**
  - Check determinism of HMC7044, feed 10MHz directly to HMC7044
  - RF noise from the PSU
  - Use PPS rising edge as reference instead of 10MHz - TDC in FPGA, precision requirement between rising edge of 1PPS and 10MHz: ~100ps

- **Interfaces**
  - PPS in should be 5V TTL compatible, PPS out also TTL, always 50 ohms
  - 10MHz output should be AC coupled instead of LVCMOS 3.3
  - To gain space space in the front panel, squeeze the Aux/Abscal connectors - they are rarely used...
  - Do not use CP2105 dual USB, use FTDI FT2232HL
  - Do not use USB-C instead of mini USB type B.
  - Locator LED on the back of the switch
  - LEDs - color-blind friendly
  - Bad idea to have the button to flash directly from USB
  - Have buttons next to LCD to change the content
  - LCD - overkill and takes space
  - SMA connecter preferred
  - UFL connecter - very bad choice, long lifetime, plain PCB mountable SMA connector, MCX, or MCX would be much better MMCX rather

- **Expansion board**
  - Mimic the pinout of FMC in the expansion slot. Trying to put both the FMC and edge-connector footprints on the PCB in parallel could result in stubs
  - Provide proper airflow/cooling for the expansion board - cater for worst case scenario of power/heat needs
  - High speed interface for data (link aggregation)
  - ...

- **Memories**
  - Add high bandwidth external packet storage
  - A selection criteria for storage could also be TBW (TeraByte Written), e.g an SSD will be reliable and it has S.M.A.R.T for monitoring
  - To speed up configuration of the FPGA one should use Dual QSPI. Is 256 MB QSPI enough?
  - ...

- **Other**
  - Continued supply of WR switches... - drop-in replacement
  - Filter on the inlet of airflow
  - Dust cover unused connectors
  - External temperature sensor to know environmental conditions
  - Have simple version of the switch, no features, reliable → oriented for data-center
Number of ports - comment

Section “3.5 Interfaces in “Main board arch”

Comment: The chosen Zynq Ultrascale+ device (XCZU17EG-1FFVC1760E) has 48 GTH/GTY transceivers. It is a waste to stick to 18 ports. Many commercial switches have 24 ports minimum. If it isn’t taken into account now it will never be upgraded later (See INT-2 table 1 of [1]). There is front panel space for 3 blocks of 2x4 SFPs = 24 ports when SMAs are stacked using SMA feed-through cable assemblies that connect to SMAs on the main PCB (see for example [5], other example use in MultiSFP Crate [6] or HPSEC [7]). KM3NeT would love to have at least 20 ports!

Remark regarding connectors:

- Industrialization is more complex
- Delays between cables might be difficult to ensure
- Increase costs
- Increase difference between multi-vendor
**Number of ports - considerations**

**Answer:** Adding more ports has many consequences and needs to be well-justified to be added to the basic version of the switch.

**Considerations regarding adding more ports:**
- Front panel - do we want to
  - Keep the same layout (e.g. management port and clk in/out on the left)?
  - Keep the same clk in/out + abscal?
- Clocking:
  - Clk in/out, expansion board and clocking circuit must be on the same side
- Airflow
  - Limited by PSU, the cooling is more challenging
- Ethernet traffic forwarding:
  - for WRS-4v1 (1Gbps ports): forwarding efficiency with more than 18 ports will be much worse than for the current WRS-3
  - implementing Ethernet Traffic forwarding for 18 x 1Gbps was a challenge, doing it for 10Gbps and even more ports increases this challenge (i.e. man-hours/costs)
  - The size of FPGA was estimated for 18 ports. If we add ports, we use the FPGA resource overhead for traffic forwarding rather than features
- Increased cost

**Some ideas:**
- Move management ports to the right of the switch
- Make the switch longer to allow easier airflow “behind the PSU”
- IN_3 (aux) and OUT_3 (62.5MHz) - remove or make optional with SMA feed-through cable
- In any case, complexity and cost increase - acceptable?
Comment:

- HMC7044, did you actually check phase determinism? If not, verify with an EVAL-HMC7044 evaluation board.
- Phase determinism is critical. When you operate in master- or grand master mode your WR CLK (generated by the HMC7044) must be deterministic phase aligned with your external reference. To my opinion the 10MHz in (i.e. Sine or Digital CLK AUX IN SMA) can be connected to the HMC7044 as reference clock, just like the "optional 10MHz from te expansion board". So, no need for a LMX2594. And a as a bonus: you get a very clean WR CLK (without routing through the phase noisy FPGA).
- The (super clean) external 10 MHz should be used to create a clean 62.5/125 MHz WR_CLK without FPGA intervention to avoid additional phase noise. So the external 10MHz reference should be input to the HMC7044.

Answer: This is something to be discussed. Basically the safest design choice is to treat the external 10 MHz input like the recovered clock coming from the transceiver and use WR PLL to align it. Feeding it directly in HMC7044 will require:

1. phase determinism (and seems not deterministic also due to the divider settings that will be used)
2. a way to prevent the freezing of the entire design if the external clock start to become unreliable (and that will depend on how HMC7044 pll will behave in such conditions)
3. check of temp coefficient
4. Holdover implementation might be more difficult

It has an important disadvantage which is described in the document: if the 10MHz is lost, the system freezes.
Interfaces

Section “3.5 Interfaces” in “Main board arch”

Comment: PPS in should be 5V TTL compatible. Most 1 PPS sources are TTL out. No selectable termination, always 50 ohms!
Answer: OK but it would be nice if we have good feedback of 5V output buffer: part-to-part delay, rise time, etc..

Comment: About the 10 MHz output, I would go for an AC coupled output instead of LVCMOS 3.3. With the proposed clock scheme it’s also easy (just convert differential LVPECL to single-ended with a balun) and already robust against user mistakes (shorts, DC power supply,…) without additional components.
Answer: OK

Comment: The CP2105 dual USB seems to be vulnerable for ESD even when properly ESD protected on the PCB. For this reason we chose to abandon it for the next version of SPEC7 and move to FTDI FT2232HL (see also issue #16 of SPEC7 [10])
Answer: OK, thanks for the tip

Comment: To gain space space in the front panel, squeeze the Aux/Abscal connectors - they are rarely used…
Answer: To be considered

Section “4. Architecture proposal and its cost estimate” in “Main board arch”

Comment: Use USB-C instead of mini USB type B. USB-C is future proof (even when only the backward compatible slow USB2.0 interface is used).
Answer: OK?
Expansion board

**Comment:** Trying to put both the FMC and edge-connector footprints on the PCB in parallel could result in stubs which could affect the signal integrity. If we designed a dummy expansion board which is an FMC carrier itself, the problem would be gone.

**Answer:** OK

**Comment:** we should mimic the pinout of FMC in the expansion slot, e.g. we should be able to identify what expansion board is plugged using the standard FMC I2C mechanism.

**Answer:** OK

**Comment:** Provide proper airflow/cooling for the expansion board - cater for the worst case scenario of power/heat needs

**Answer:** OK
FPGA and memories

Section “3. Analysis of the main design choices” in “Main board arch”
Comment: A selection criteria for storage could also be TBW (TeraByte Written) e.g. when a running Linux distribution on the PS. Be aware of the amount of data that is transferred over time so for long-life and reliability one should select the proper type of storage. For example an SSD will be reliable and it also has S.M.A.R.T to monitor your disk. Note that the Zynq Ultrascale+ (XCZU17EG-1FFVC1760E) has a SATA interface.
Answer: OK, it can be an option, to be investigated

Section “3.2.1. ARM + FPGA on the same PCB” in “Main board arch”
Comment: Fast packet storage on chip in BRAM and/or Ultra-RAM. More high bandwidth packet storage might be needed to overcome issues like we had with WR switch v3.4 (see GSI report [8]) Add extra high bandwidth memory external to the FPGA.
Answer: To be considered, yet we intentionally have chosen FPGA with huge internal memory. From experience with WRS-3, external memory is a bottleneck. Adding high bandwidth storage in a prototype for evaluation is an option, however it increases complexity of PCB, thus costs.

Figure 8 in “3.5 Interfaces in “Main board arch”
Comment: To speed up configuration of the FPGA one should use Dual QSPI.
Is 256 MB QSPI enough?
Answer: OK ?
Comments on Power Supply and Fans

Section “2.1 Power Supply” in “Hardware study”
Comment: For critical applications (data centers or Internet exchange) a redundant (dual) hot-swappable PSU is either very desirable or even a must have.
Answer: sure

Section “2.2 Fans” in “Hardware study”
Comment: For critical applications (data centers or Internet exchange) hot-swappable fans are either very desirable or even a must have.
Answer: sure

Section “4.1 Power consumption estimation” in “Hardware study”
Sentence: "The previously commented performances increase the power consumption of each SFP+ transceiver up to 1.2W. Typically."
Comment: Several SFP+ modules specify at least 1.5Watt [3] some even 2.0 Watt [4]! The worst case power calculation should be adapted to these values.
Answer: OK, we will update the calculations, we overestimated the heat dissipation so we should be fine, power supply to be checked, noise to be investigated

Section “4.3 Power supply” in “Hardware study”
Comment: RF-Noise and EMC should also be a selection criteria for the Power Supply.
Answer: sure

Section “4.6 Quantity and shape of PCBs” in “Hardware study”
Comment: N4000-13SI substrate 10 GBps?
Answer: ????