State of the art and requirements for the new WRS-4

Maciej Lipinski
CERN

New WRS-4 Workshop
2020-06-26

Setting requirements for the WRS-4

1. Analyse input from WR community (wr-dev list and 10th Workshop)
2. Discuss with CERN IT the main features of L2 Ethernet switches at CERN
3. Analyse and compare features of popular switches & time providers, focus:
   a. Power supply
   b. Fans
   c. Physical interfaces
   d. Price
4. Estimate future needs in terms of FPGA resources

Basic principles for setting the requirements

- Allow open source (minimise vendor-locking)
- Use standards if/when possible
- Address inputs from WR community
- Adjust to current trends for L2 switches (and time providers)
- Be modular/extensible
- Provide drop-in replacement for WRS-3
  (min 18 x 1Gb ports, RJ45 management, 10MHz and 1PPS in/out)
- Ensure reasonable price

Main areas of requirements

- **Hardware features and interfaces (excluding main board)**
  - Data ports
  - Management I/Fs
  - Clocking I/Fs
  - Power supply
  - Fans
  - Enclosure
  - Extension interfaces
- **Main board (PCB)**
  - CPU
  - FPGA
  - Clocking circuitry
- **Gateware**
- **Software**

Power supplies / fans in a device can be:

- **redundant :** there is hot-spare component (N+1), if one components fails, the spare allows to maintain operation of the device
- **hot-swappable:** in case of redundant components, it is possible to replace one component during operation of the device
- **standard:** the connector/manage/diagnostics are standardized and so the components can be purchased from different vendors

Input from WR community - highlights

Details in section 2 in “WRS-4 Main Board” and section 3 in “Study on hardware features”

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - Management I/Fs:
  - Clocking I/Fs:
  - Power supply:
  - Fans:
  - Enclosure:
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in “WRS-4 Main Board” and section 3 in “Study on hardware features”

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - Clocking I/Fs:
  - Power supply:
  - Fans:
  - Enclosure:
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in “WRS-4 Main Board” and section 3 in “Study on hardware features”

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - Power supply:
  - Fans:
  - Enclosure:
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in “WRS-4 Main Board” and section 3 in “Study on hardware features”

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - Fans:
  - Enclosure:
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in "WRS-4 Main Board" and section 3 in "Study on hardware features"

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - Enclosure:
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in "WRS-4 Main Board" and section 3 in "Study on hardware features"

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - **Enclosure:** 1U rack size
  - Extension I/F:
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in "WRS-4 Main Board" and section 3 in "Study on hardware features"

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - **Enclosure:** 1U rack size
  - **Extension I/F:** internal connector for extension module (e.g. oscillator board, DDS board for lpGBT)
- **Main board (PCB)**
  - CPU:
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

LpGBT:

- Project called "Low Power GigaBit Transceiver"
- provides LpGBT ASIC - a radiation tolerant serializer/deserializer ASIC
- provides LpGBT FPGA - FPGA cores for the non-rad-tol back-end counterpart of the ASIC
- offers a set of encoding/decoding schemas for transfer of high-bandwidth data
- allows transferring reference frequency (clock signal)

The new WR switch could be used as the no-rad-tol back-end for distribution of data and clock reference to the radiation areas.

Input from WR community - highlights

Details in section 2 in "WRS-4 Main Board" and section 3 in "Study on hardware features"

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - **Enclosure:** 1U rack size
  - **Extension I/F:** internal connector for extension module (e.g. oscillator board, DDS board for lpGBT)
- **Main board (PCB)**
  - **CPU:** at least ARM dual core with 1.2GHz with DDR4, QSPI, SD, eMMC
  - FPGA:
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in “WRS-4 Main Board” and section 3 in “Study on hardware features”

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply:** redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - **Enclosure:** 1U rack size
  - **Extension I/F:** internal connector for extension module (e.g. oscillator board, DDS board for lpGBT)
- **Main board (PCB)**
  - **CPU:** at least ARM dual core with 1.2GHz with DDR4, QSPI, SD, eMMC
  - **FPGA:** overdimension wrt. resources
  - Clocking circuitry:
- **Gateware**
- **Software**

Input from WR community - highlights

Details in section 2 in "WRS-4 Main Board" and section 3 in "Study on hardware features"

- **Hardware features and interfaces (excluding main board)**
  - **Data ports:** min 18 ports, ideally more, 1 & 10Gb, min 2 ports with 10Gb
  - **Management I/Fs:** SFP & Ethernet RJ45, UART RJ45, mini USB type B, RS232, power button, OLED/LCD, location LED
  - **Clocking I/Fs:** 1PPS in/out, 10MHz in/out, AUX in/out, Abscal out
  - **Power supply**: redundant, hot-swappable (optional), sufficient current for DWDM, optimal air flow direction, standard
  - **Fans:** redundant, hot-swappable, changeable air-flow, advanced control/diag, consider fanless
  - **Enclosure:** 1U rack size
  - **Extension I/F**: internal connector for extension module (e.g. oscillator board, DDS board for lpGBT)
- **Main board (PCB)**
  - **CPU:** at least ARM dual core with 1.2GHz with DDR4, QSPI, SD, eMMC
  - **FPGA:** overdimension wrt. resources,
  - **Clocking circuitry:** VCXOs with low phase, allow extensions, include LJD board improvements, 10 & 5 MHz inputs to GM
- **Gateware**
- **Software**

Features of L2 switches and time providers

Details in section 2 in "Study on hardware features"

Power Supply:

- Redundancy - all switches
- Rear location - all switches
- Hot-swappable - majority of switches
- Internal fan - majority of switches, airflow direction inline with enclosure fan
- Standard - none of switches

Fans:

Only few switches were fanless. For the switches with fans:

- Redundancy - all switches 2 or more fans
- Fixed - few switches, fans located in the right part of the device
- Hot-swappable - majority of switches, located in the rear of the device
- Airflow direction - all allow different direction as a function of part number chosen
- Standard - none of switches

Features of L2 switches and time providers

Details in section 2 in “Study on hardware features”

Interfaces:

- Female USB type A - very popular for firmware or configuration update
- Ethernet RJ45 - the most popular for serial communication
- Mini USB and USB C - also used for serial communication
- Ethernet RJ45 port - the most common for management
- SFP port - also used for management
- Reset button - commonly in the front panel interfaces.
- Power on/off button - not present
- LCD screen - not present in switches, present in time providers
- Timing connectors - in few switches, different types (SMA, SMB, BNC)

Discussion with CERN IT

Notes available here

- There is no standard for power supplies/fans/management interfaces
- The power supplies are bought for a given vendor and for a give model
- Standardized power supplies are available for servers
- For most of the removable fans, one can choose the direction of airflow
- For non-removable fans, typically the airflow is front-to-side or side-to-side
- For cooling PHYs/SFPs in front panel, the best airflow is front-to-back
- Most common airflow in data centers: front to back
- Management interfaces:
  - Serial communication: miniUSB or RJ45, RS232 connector is not used in new switches (only old)
  - USB: in many switches, there is a USB connector used to update firmware (upgrade or recovery)
- Most switches have small reset button in front, no power reset

Choice of FPGA

- Considered **Xilinx Zynq/Kintex** and **Intel Arria 10**
- Worked closely with Avnet (Xilinx) and Intel for almost 1 year (March 2019 to February 2020)
- With the help of CERN procurement, we negotiated very competitive prices from both

Xilinx Zynq Ultrascale+	Intel Arria 10
● Extensive experience at CERN/HT & 7Sols ● Cortex-A53 is 64bit – no potential problems with end of Linux support for 32bit ● Support offered ● LpGBT core was tested on Kintex Ultrascale (GTH) ● Better price	● Little experience ● Cortex-A9 is 32bit – potential problems with end of Linux support for 32bit ● Support offered

Final choice: **Xilinx Zynq Ultrascale+ XCZU17EG-1FFVC1760E**
at a lower price* than the Virtex-6 XC6VLX240T used in the current WRS-3

*The final step

Estimation of required FPGA resources

1. **Estimation of resources needed for 18-port 10GbE switch**
   (see WRS-v4-resource-utilization.pdf):
   a. HDL of WRS-3 was ported to the candidate FPGA: Xilinx Ultrascale+
   b. All redundancy features (not in release) were instantiated/enabled
   c. 10GbE PHYs were generated and data-paths extended accordingly
   d. Result: 210k LUT, 198k FF, 517 BRAM (post-implementation report)
2. **Criteria for new FPGA choice:**
   The above worst case scenario should take less than 70% of FPGA resources (ideally 50-60%).
3. **Estimated required resources:**
   350k LUTs, **600k FFs** and **860 BRAM** blocks and minimum **18 ports (min 10GbE)**

Summary

- Inputs described in this presentation were used to prepare two documents:
  - “Study on the new hardware features for the WRS-4”
  - “Study on the WRS-4 main board (Hardware Architecture)”
- The documents will be presented in the next two presentations

Summary

- Inputs described in this presentation were used to prepare two documents:
  - “Study on the new hardware features for the WRS-4”
  - “Study on the WRS-4 main board (Hardware Architecture)”
- The documents will be presented in the next two presentations

Questions?