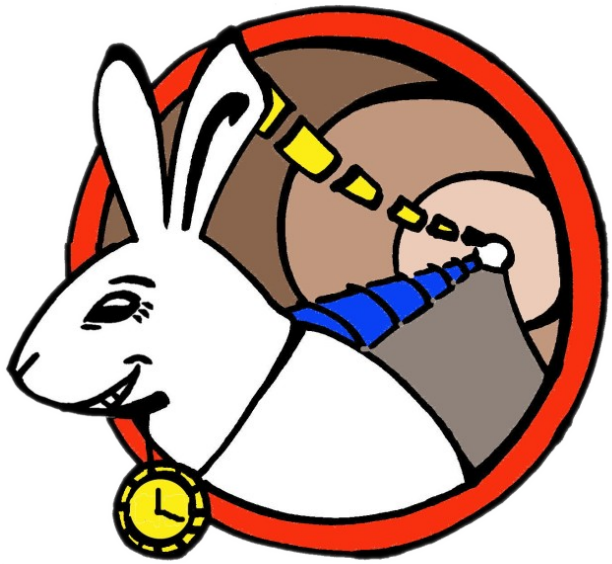
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH
ORGANISATION EUROPEENE POUR LA RECHERCHE NUCLEAIRE



White Rabbit Switch

Technical Specification

Version: 0.2
Date: November 24 of 2010.
Author: J. Gabriel Ramírez (Seven Solutions S.L.)

Acronyms List

AMC	Advanced Mezzanine Cards
CAM	Content Addressable Memory
CPU	Central Processor Unit
DDR	Double Data Rate
DRAM	Dynamic RAM
EMI	Electromagnetic Interference
EMC	Electromagnetic Compatibility
FCC	Federal Communications Commission
FPGA	Field Programmable Gate Array
GPIO	General Purpose Input/Output
GPS	Global Positioning System
ICD	Interface Control Document
IGMP	Internet Group Management Protocol
JTAG	Joint Test Action Group
LVDS	Low Voltage Differential Signaling
LVPECL	Low Voltage Positive Emiter-Coupled Logic
MAC	Media Access Control
MCH	MicroTCA Carrier Hub
MB	Mega Byte
Mbps	Megabits per Second
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor
PCB	Printed Circuit Board
PPS	Pulse Per Second
PSU	Power Supply Unit
PTP	Precise Time Protocol
RAM	Random Access Memory
SCB	Switching Core Board
SDRAM	Synchronous Dynamic RAM
SerDes	Serializer/Deserializer
SFP	Small Form-factor Pluggable
SMD	Surface Mounted Device
SMI	Switch Management Interface
SNMP	Simple Network Management Protocol
SPI	Serial Peripheral Interface
WRP	White Rabbit Protocol
WRS	White Rabbit Switch

Document Reference:

Ref	Title	Author	Version
[1]	White Rabbit Protocol Specification	White Rabbit Team	To be published
[2]	White Rabbit Switch ICD	White Rabbit Team	To be published
[3]	White Rabbit Switch Functional Specification	White Rabbit Team	0.0

Content Index

Index of Figures.

1 Introduction.

The White Rabbit is intended to be the next-generation deterministic network based on synchronous Ethernet, allowing for low-latency deterministic packet routing and transparent, high precision timing transmission. The network consists of master node (only one active at a time) which provides time and frequency reference, switches interconnected through fibers or twisted-pair copper in star topology and slave nodes.

White Rabbit Switch (WRS) is the key component in new CERN timing system, allowing for multiplexing of high-precision timing and control data in single fiber connection.

Moreover of the White Rabbit Protocol requirements, one of the most important is to allow the possibility of two WRS configurations:

- using a uTCA rack.
- as a standalone switch equipment inside a 19" case.

This document describes design guidelines and technical details for White Rabbit switch uTCA MCH, AMC modules and miniBackplane board.

2 White Rabbit Switch Functional Requirements.

Although all the requirements that the White Rabbit Switch must accomplish are described in the "White Rabbit Switch Functional Specification" document (ref. [3]), it has been designed according to the following baselines:

- minimizing cost in order to get the lowest WRS price.
- using most of the components and architecture, if is possible, used in WRS V2.
- improving design flexibility in order to facilitate compatibility between uTCA and miniBackplane configurations.

3 Switching Core Board.

The main element of the WRS will be the ***"Switching Core Board".*** Fundamental functionality of White Rabbit Protocol will be implemented inside this board. Standalone and uTCA WRS configurations, both will use SCB as their main component.

3.1 SCB Hardware

In this section, SCB hardware is described. Figure 1 shows its basic functional block diagram.

3.1.1 Interfaces

SCB provides the following interfaces:

- **Front panel I/O connectors:**
  - 5 SMC coaxial connectors:
    - 125 MHz reference clock input (EXTREFIN125M). Used to provide external 125 MHz clock as a reference for downlink ports instead of clock recovered from uplinks.
    - 10 MHz reference clock input (EXTREFIN10M). Allows for using 10 MHz reference clock from cesium/GPS.
    - 125 MHz reference clock output (EXTREFOUT125) with 125 MHz network reference clock, recovered from uplink or provided by external input.
    - PPS input (EXTPPSIN). Receives PPS pulses from external source. Along with EXTREFIN can be used to synchronize the switch to external timing source.
    - PPS output (EXTPPSOUT). Outputs PPS signal.
  - 100Mbit Ethernet port (ETH_MNG). Twisted-pair ethernet port connected to management CPU. Allows for using external, non-WR network to manage WRS.
  - 4 LEDs for informational/debugging purposes.
- **Connectors 1:** these connectors (we will use 2 connectors in order to provide a better mechanics connection) will be the main connectors of the SCB board. They allow connection to uTCA board 1 (in uTCA configuration) and miniBackplane (in standalone configuration):
  - 3.3V main power supply.
  - 40 general purpose IOs (for controlling dual color LEDs or SFPs...).
  - 20 GTX connections for uplink and downlink ports.
  - Serial Communication Port: will provide FPGA serial communication (SPI, IOSERDES...).
  - INT: connected to ARM9 IRQ0 interrupt signal.
  - RESET: connected to NRST signal of ARM9 to reset board.
  - Configurable RS232 port (RS232_MNG). RS232 port, which can be used either by local serial management console or as input of NMEA timecode from GPS/cesium.
  - SPI: connected to SPI1 port of ARM9 will allow serial communication with this CPU.
  - BOOT_SEL: this signal could be used to select the boot process of ARM9.

- **Connector 2:** allows connection to uTCA board 2:
  - Clock 1: differential LVDS 125 MHz low-jitter reference clock phase-locked with System Timing Master clock.
  - Clock 2 configurable frequency, differential LVDS, low-jitter reference clock phase-locked with System Timing Master clock.
- **Connector 3:** PCB kind according to tongue 3 of uTCA standard:
  - 12 differential LVDS pairs for SMI_LINK[1:6].
- **Connector 4:** allows connection to uTCA board 4:
  - 12 differential LVDS pairs for SMI_LINK[7:12].

3.1.2 Main FPGA

We've chosen a Xilinx Virtex-6 device (LX130T) in 1156-pin BGA package because it has both, the enough block RAM and GTX transceivers required in this application.

Inside the FPGA there are:

- WR switch core
- CPU interface- EBI slave for accessing various configuration registers. It uses 32-bit asynchronous bus with wait states, additional interrupt lines and DMA request signal.
- Clocks management
- Switch management interface SMI_LINK endpoints
- Up to 40 GPIO ports for driving LEDs and SFP detection
- 32 bit wide interfaces with 3 ZBT synchronous SRAM memory chips each with 512KB of capacity used to store routing tables which need fast access.
- Up to 20 GTX transceivers for uplink and downlink ports.
- Power supply monitoring by means of System Monitor inside FPGA.
- Temperature sensors control.

3.1.3 CPU

For the CPU we have decided to choose Atmel AT91SAM9263 chip (ARM926E core running at 200 MHz) because of its popularity, well-known architecture (ARM) and OS support (Linux).

Following CPU peripherals are used:

- External bus interface 0 (EBI0) connected to 64 MBytes of SDRAM (32-bit bus chip) and 256 MB NAND flash chip containing main flash firmware
- External bus interface 1 (EBI1) operating in Static Memory mode, connected to CPU i/f in main FPGA
- SPI interface (SPI) connected to DataFlash memory containing failsafe boot image
- Ethernet MAC (EMAC) connected via appropriate magnetics to 100Base-T RJ45 port on front panel of MCH (local ethernet management)
- UART local serial port routed to Connectors 1.

- GPIO pins connected to FPGA configuration pins/JTAG port, used to boot the FPGA.
- SPI interface to provide serial communication
- Serial interface to configure FPGA

CPU is interfaced with FPGA using 32-bit asynchronous bus (with external wait state support) and few independent interrupt lines.

Main tasks of CPU are:

- PTPv2 daemon
- WRP protocol daemon for WR node discovery and sub-nanosecond phase sync
- Switch management through SNMP/remote console.
- Handling of high-layer IEEE802.1x protocols (multicasting, spanning tree, GMRP/-GARP)
- Configuration of special WR features

3.1.4 CPU/FPGA boot process

To prevent bricking the switch by faulty remote firmware update, special boot procedure was developed.

The idea is to have two separate flash memories and two versions of firmware:

- Main firmware - currently working version, which can be updated at any time.
- Failsafe firmware - minimal system and FPGA bitstreams required for device operation and main firmware update. This flash is not remotely writable.

Across of BOOT_SEL signal at Connectors 1 we can select between both versions:

- BOOT_SEL asserted (HI or '1') → Failsafe firmware
- BOOT_SEL deasserted (LOW, '0' or not connected) → Main firmware.

3.1.5 Power Supply

SCB will be supplied with 3.3 V incoming by Connectors 1. All others needed voltages on board will be generated inside SCB from this 3.3V main power supply. In order to get the best efficiency, powerful supplies will be done with DCDC in buck topology and external discrete MOSFET switches.

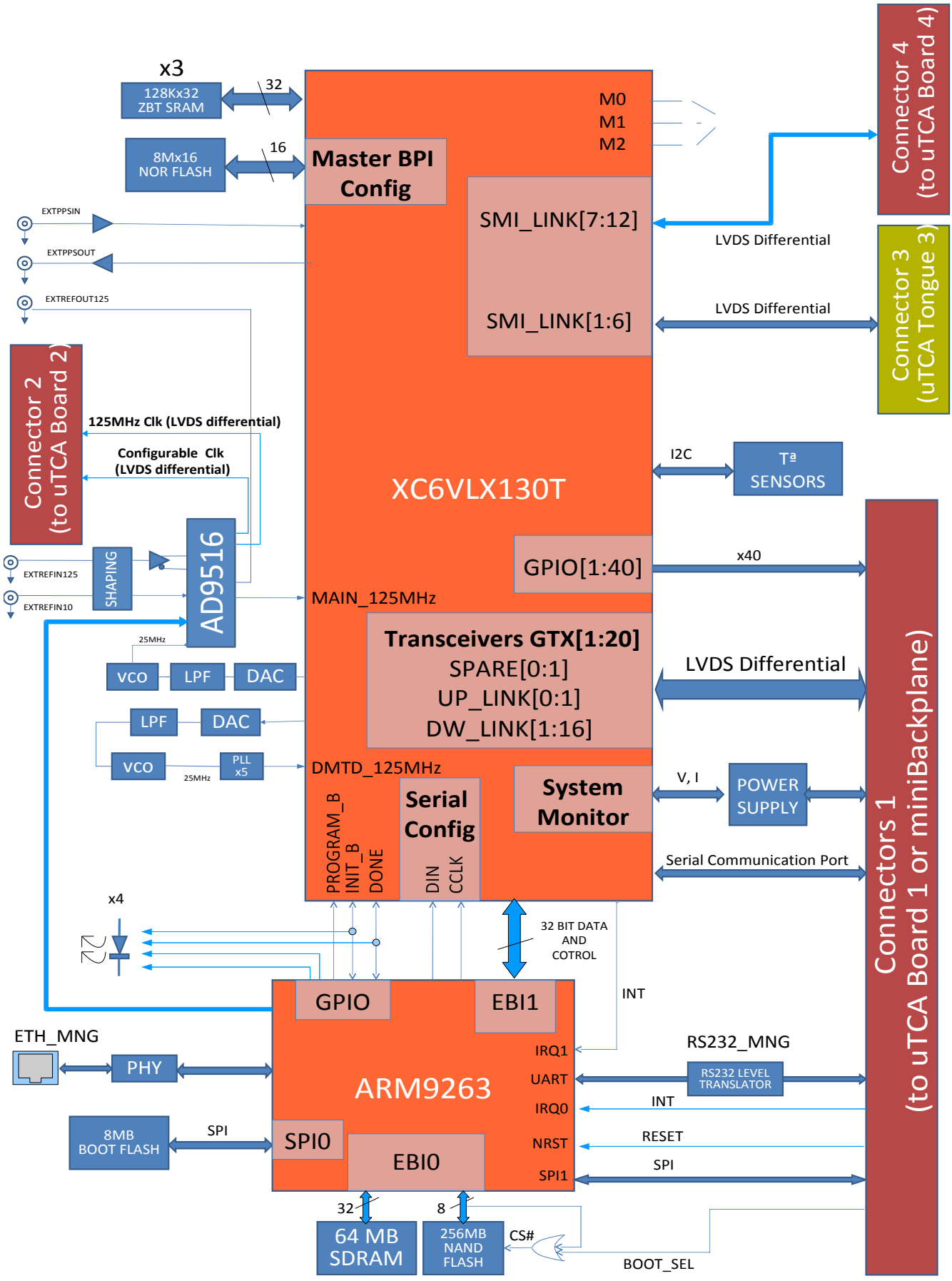


Figure 1: SCB Functional Block Diagram.

4 uTCA White Rabbit Switch Configuration.

The complete uTCA White Rabbit Switch consists of two different modules (described in following chapters) which are working into a uTCA crate with appropriate backplane:

- WR MCH module with two uplink ports acting as master card for WR AMCs and responsible for high-level switch features like switch/crate management, PTP, etc. Its main component is SCB board placed at uTCA board 3 (see Figure 4).
- WR AMC cards containing "subswitches" with 12 downlink ports each uplinked through backplane to WR MCH.

According to uTCA standard, this White Rabbit switch only will be able to provide 12 downlink gigabit ports through FabricA lane. System uses following uTCA backplane lanes:

- Fabric A (from WR MCH to AMC slots port 0/1) - main gigabit Ethernet link (LVDS), interconnecting WR MCH with subswitch modules in AMCs or other cards supporting GbE.
- Fabric D (from WR MCH to AMC slots port 4 - fat pipe 1) - additional synchronous serial link (LVDS) running at 125 Mbps. Used for management and conguration of WR AMCs and for transmitting data between WR MCH and WR AMCs which cannot be multiplexed into main Ethernet link without making it nondeterministic (like timestamps). It is also used to transmit PPS/timecodes to slave cards with simplied timing receivers (not requiring PTP).
- Clock 1 (from WR MCH CLK1 to AMC port CLK1) - 125 MHz low-jitter reference clock phase-locked with System Timing Master clock. These clocks are generated independently for each card in the crate.
- Clock 2 (from WR MCH CLK2 to AMC port CLK2) – configurable frequency, low-jitter reference clock phase-locked with System Timing Master clock. These clocks are generated independently for each card in the crate.

4.1 Inputs and outputs

- **Front panel I/O**: Moreover the interfaces provided by SCB, MCH will provide this signals:
  - Two SFP module sockets for uplink ports (UP0, UP1). Uplink port 0 is the primary uplink port, used in normal conditions. Uplink port 1 is used as timing and HP source when primary uplink connection is not functioning properly.
  - Configurable RS232 port (RS232_MNG). RS232 port, which can be used either by local serial management console or as input of NMEA timecode from GPS/cesium.
  - LEDs:
    - 14 dual-color LEDs to inform about the state of the Uplink and Downlink ports.
    - 4 for informational/debugging purposes (uTCA LEDs).
- **microTCA Backplane I/O:**
  - 12 LVDS Gigabit Ethernet lanes (from WR MCH to each AMC slot port 0/1)
  - 12 LVDS management i/f lanes (from WR MCH to each AMC slot port 4)
  - 12 LVDS 125 MHz TCLKs (from WR MCH to each AMC slot port CLK1)
  - 12 LVDS frequency configurable TCLKs (from WR MCH to each AMC slot port CLK2)
  - IPMB-A/B to PSU and fans, IPMB-L lines to rst 12 AMC slots.
  - Power and auxiliary uTCA control signals

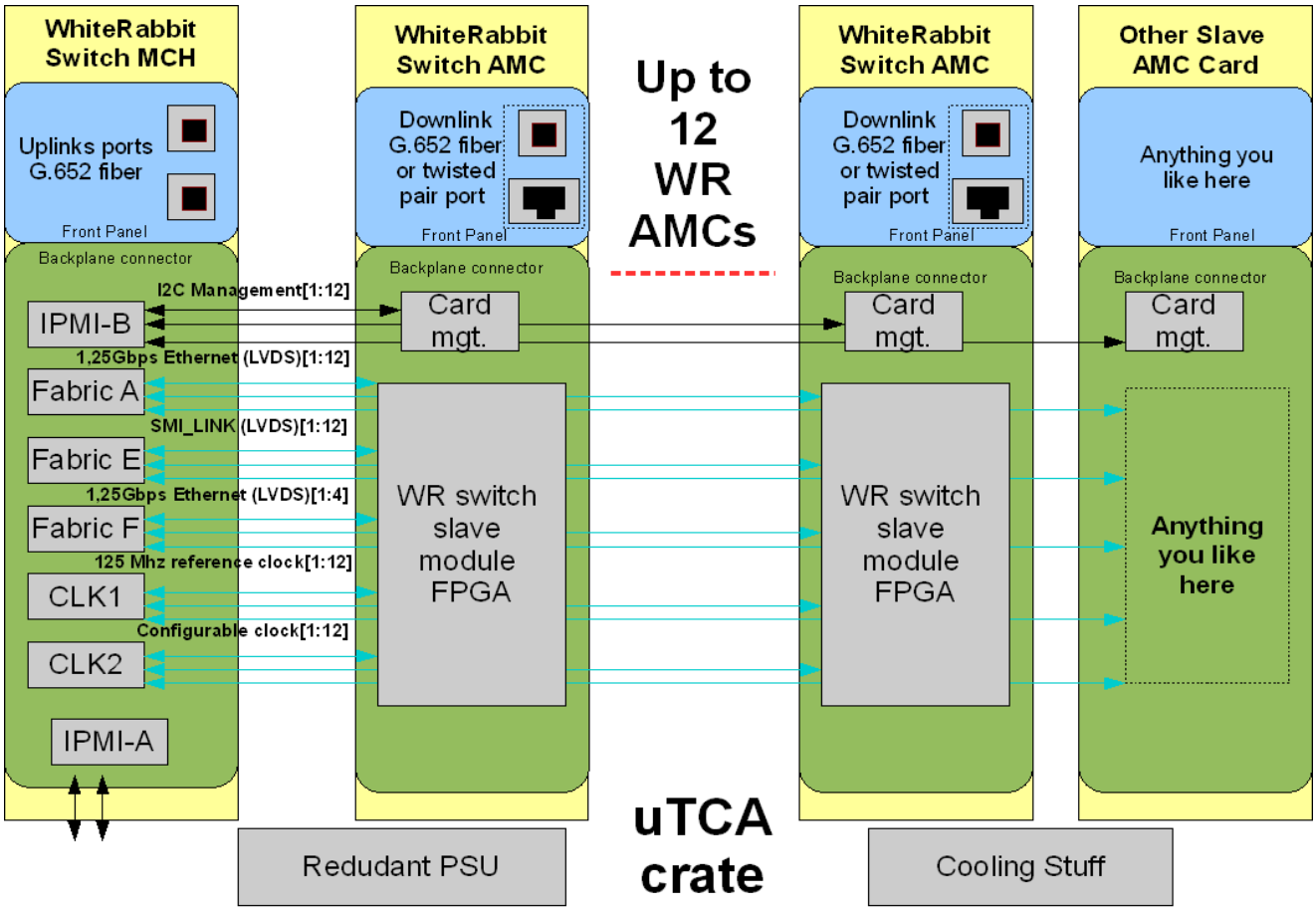


Figure 2:General block diagram of White Rabbit Switch.

4.2 Main digital section

SCB is the main part of WR MCH, containing the FPGA which takes care of packet routing and CPU which handles switch/microTCA crate management and PTP tasks. For more information read Paragraph 3. The ARM7 CPU provides all uTCA control and watchdog functionality. The block diagram is shown on Figure 3.

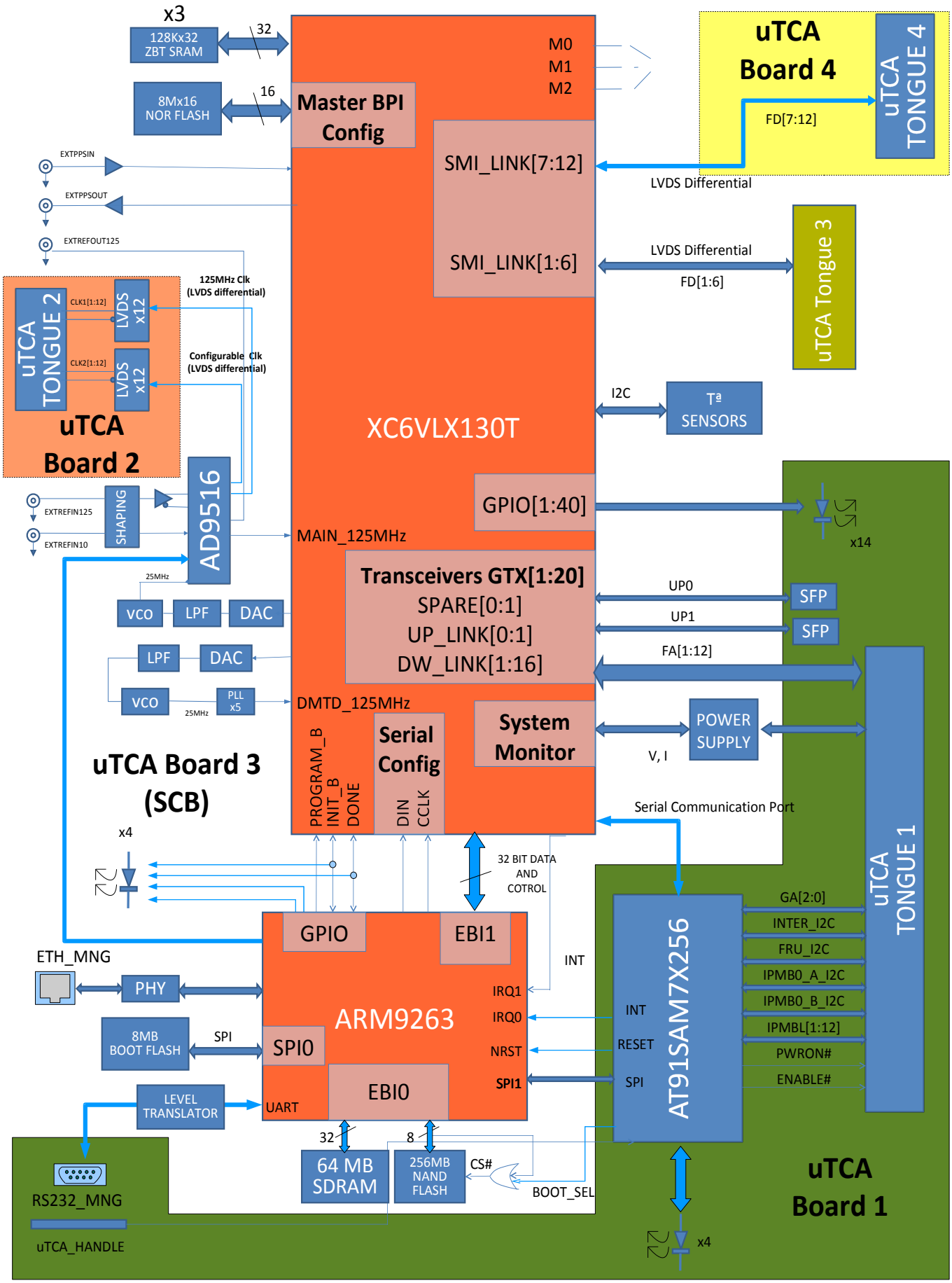


Figure 3: uTCA WRS Functional Block Diagram.

4.3 MCH Board 1 Functionality

4.3.1 Watchdog MCU

In order to achieve high system reliability and fault tolerance, there is another, small AT91SAM7X256 microcontroller implementing following functions:

- Watchdog unit, periodically polling important switch components (FPGAs, CPU), checking if they are alive and resetting them if needed
- IPMB bus I2C hub
- Power supervisor for the card - powering the card, initializing low-level uTCA stuff handling hot-swap
- Performing failsafe boot sequence after faulty firmware update
- Monitoring temperature sensors

Watchdog MCU is interfaced with main CPU via SPI link and external interrupt line. Also it has I/O lines going to both FPGAs which are used by polling mechanism. The MCU is supplied from 3.3V Management Power line.

4.3.2 MCH Power Supply

As the microTCA backplane provides only +12V supply voltage, separate power supply is required. Its block diagram can be found on figure 4.

Main digital power supplies are done in buck topology, with external discrete MOSFET switches.

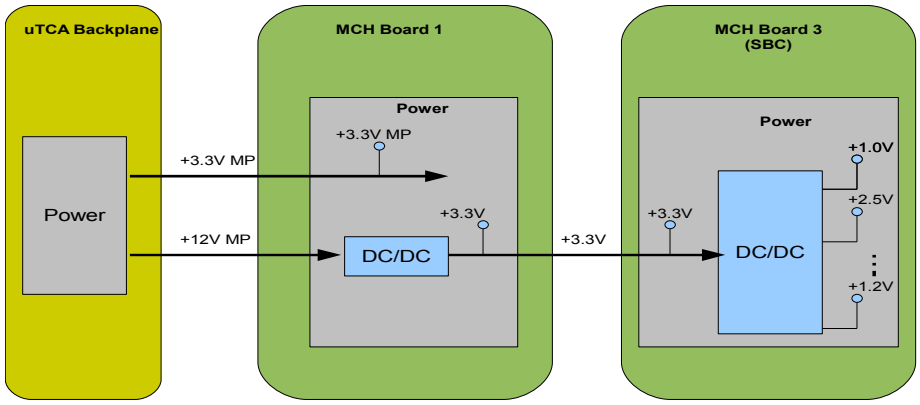


Figure 4: Power supply basic architecture.

4.3.3 CPU/FPGA boot process

Normal boot procedure is described below

1. Card gets management power (3.3V MP). Watchdog MCU powers up and enables main management power for MCU. Also it deasserts FBOOT_SEL (see Figure 3) to force CPU to boot from main flash.
2. CPU firmware is loaded, special loader is launched which boots the FPGAs.
3. Device begins to operate.

In case we wanted to perform remote firmware update, we do:

1. Tell watchdog MCU that we are going to update the rmware and set up countdown timer. When timeout is exceeded, watchdog should reset main CPU and force it to boot using failsafe firmware (it asserts FBOOT_SEL).
2. Flash the new firmware and reboot.
3. Login remotely again, do necessary checks (if new firmware works as we expected) and disable countdown timer in watchdog MCU.

4.4 WR MCH implementation

This section describes important implementation details of MCH module. It also explains some ideas used in design.

4.4.1 PCB and mechanical considerations

The switch design is too complex to be t on single uTCA card. Also the required backplane layout forces us to use separate PCBs. Proposed switch physical implementation is shown on figure 5. Look to figure 3 to see functionality distribution for each board.

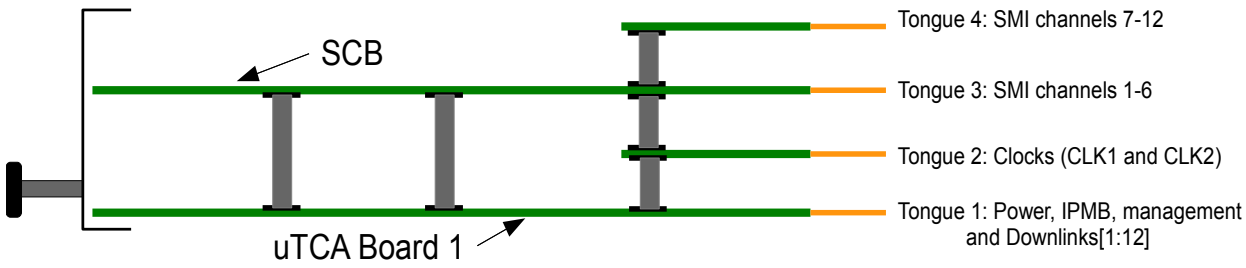


Figure 5: uTCA WRS physical implementation.

PCBs are interconnected with dense (0.8mm) SMD connectors. For mechanical stability, there are spacers binding boards to each other (4 spacers between adjacent boards). In order to preserve board space and minimize the cost, we've decided to use PCB edge connectors for backplane interface instead of special (and very expensive) uTCA snap-on connectors.

Another problem we have faced is very limited space on uTCA card front panel. Figure 6 shows proposed front panel layout.

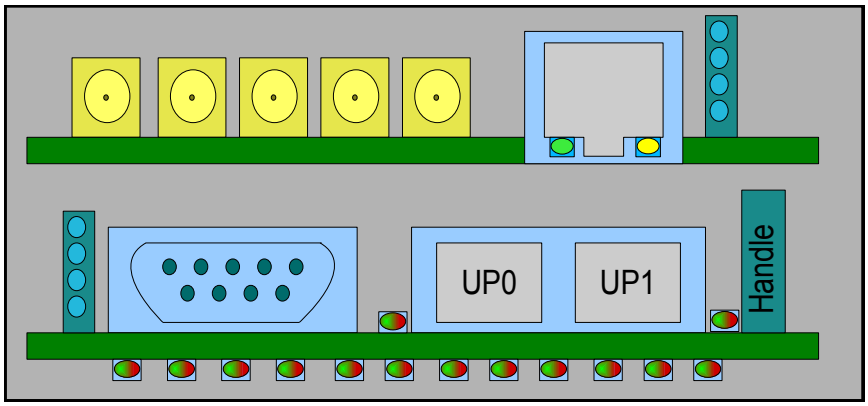


Figure 6: uTCA WRS front panel.

5 White Rabbit Switch in miniBackplane Configuration.

In order to expand up to 16 downlink ports (using LX130T FPGA we could expand up to 18), MCH board 3 can be connected to a miniBackplane board. Furthermore, there will be two version of this backplane to provide the desired physical interface (fiber or twisted pair).

5.1 Inputs and outputs

Moreover the interfaces provided by SCB, the WRS front panel will provide this signals:

- LEDs:
  - 18 dual-color LEDs to inform about the state of the communication port.
  - 4 for informational/debugging purposes.
- 2 uplink ports (UP0 and UP1) and 16 downlink ports (DP[1:16]) with two options:
  - SFP module sockets
  - 1Gbit Twisted-Pair Ethernet
- Configurable RS232 port (RS232_MNG). RS232 port, which can be used either by local serial management console or as input of NMEA timecode from GPS/cesium.

5.2 Main digital section

SCB is the main part of WRS, containing the FPGA which takes care of packet routing and CPU which handles switch management and PTP tasks. For more information read Paragraph 3. The XC6S FPGA performs SFP and PHY control and watchdog functionality. The block diagram is shown on Figure 7.

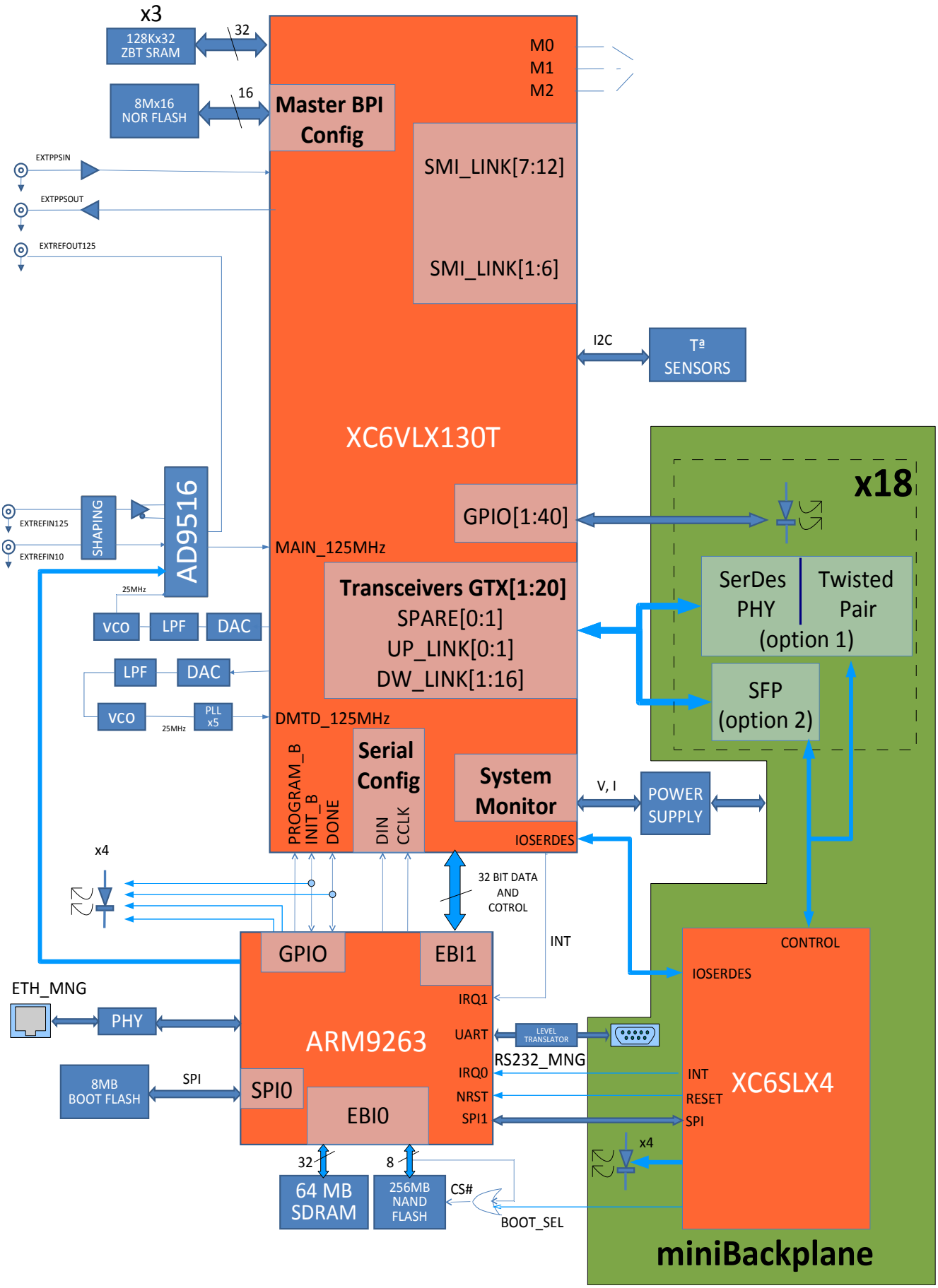


Figure 7: miniBackplane WRS Functional Block Diagram.

5.3 MiniBackplane functionality

5.3.1 XC6S FPGA Control

In order to achieve high system reliability and fault tolerance, there is another small XC6S FPGA (Spartan-6) implementing following functions:

- Watchdog unit, periodically polling important switch components (FPGA and CPU inside SCB), checking if they are alive and resetting them if needed
- Performing failsafe boot sequence after faulty firmware update

XC6S is interfaced with main CPU via SPI link and external interrupt line. Also it has I/O lines going to main FPGA which are used by polling mechanism.

5.3.2 WRS Power Supply

Its block diagram can be found on Figure 8. Main digital power supplies are done in buck topology, with external discrete MOSFET switches.

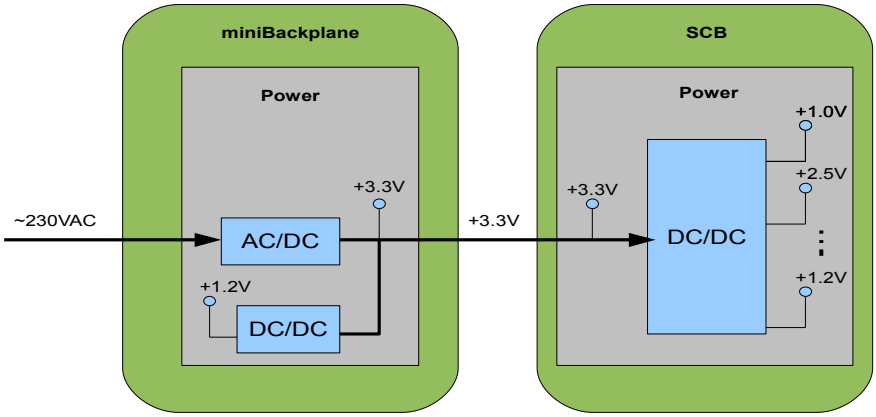


Figure 8: Power supply basic architecture.

5.3.3 SCB CPU/FPGA boot process

Normal boot procedure is described below

1. After FPGA control (XC6S) power-up it enables main management power for MCU. Also it deasserts FBOOT_SEL (see Figure 7) to force CPU to boot from main flash.
2. CPU firmware is loaded, special loader is launched which boots the FPGA.
3. Device begins to operate.

In case we wanted to perform remote firmware update, we do:

1. Tell XC6S that we are going to update the firmware and set up countdown timer. When timeout is exceeded, watchdog should reset main CPU and force it to boot using failsafe firmware (it asserts FBOOT_SEL).
2. Flash the new firmware and reboot.
3. Login remotely again, do necessary checks (if new firmware works as we expected) and disable countdown timer in watchdog XC6S.

5.4 WRS implementation

This section describes important implementation details of MCH module. It also explains some ideas used in design.

5.4.1 PCB and mechanical considerations

Proposed switch physical implementation is shown on Figure 9. WRS will be inside a 19" standard case.

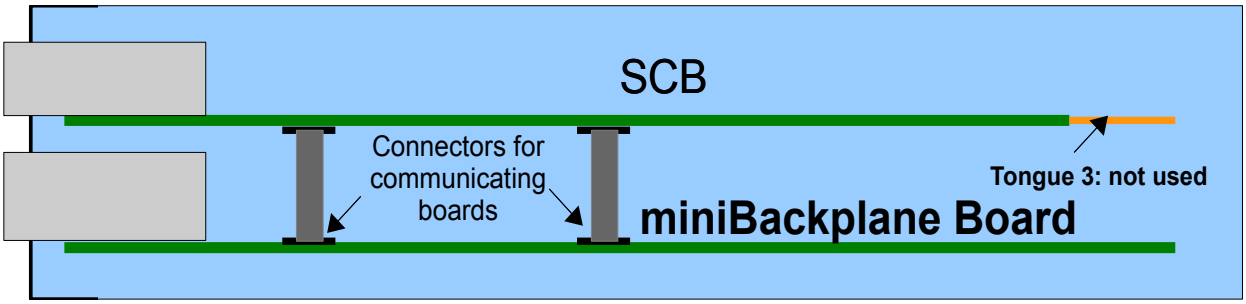


Figure 9: WRS physical implementation.

PCBs are interconnected with dense (0.8mm) SMD connectors. For mechanical stability, there are spacers binding boards to each other (4 spacers between adjacent boards). SCB uTCA tongue 3 connector is not used.

Figure 10 and 11 show proposed front panel layouts, both SFP and Twisted-pair options.

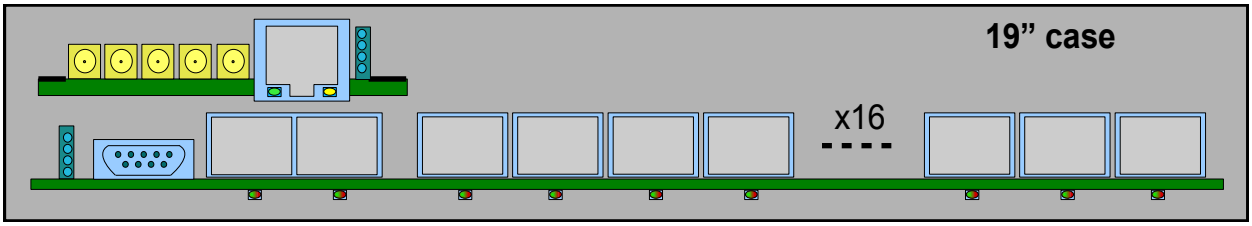


Figure 10: WRS front panel SFP option.

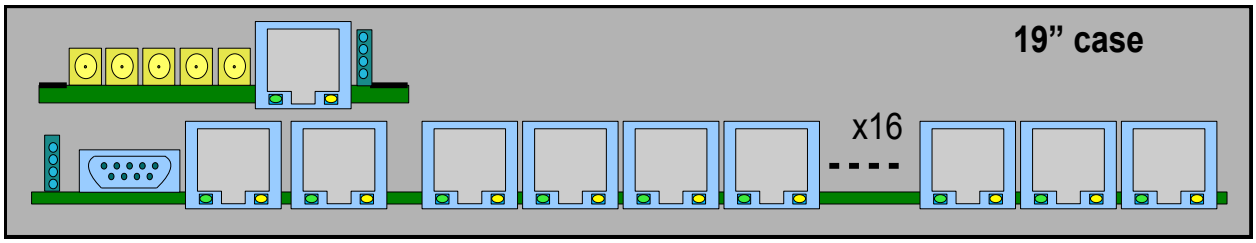


Figure 11: WRS front panel Twisted-pair option.

End of document.