White Rabbit Switch gateware

status and plans

Grzegorz Daniluk
Maciej Lipinski

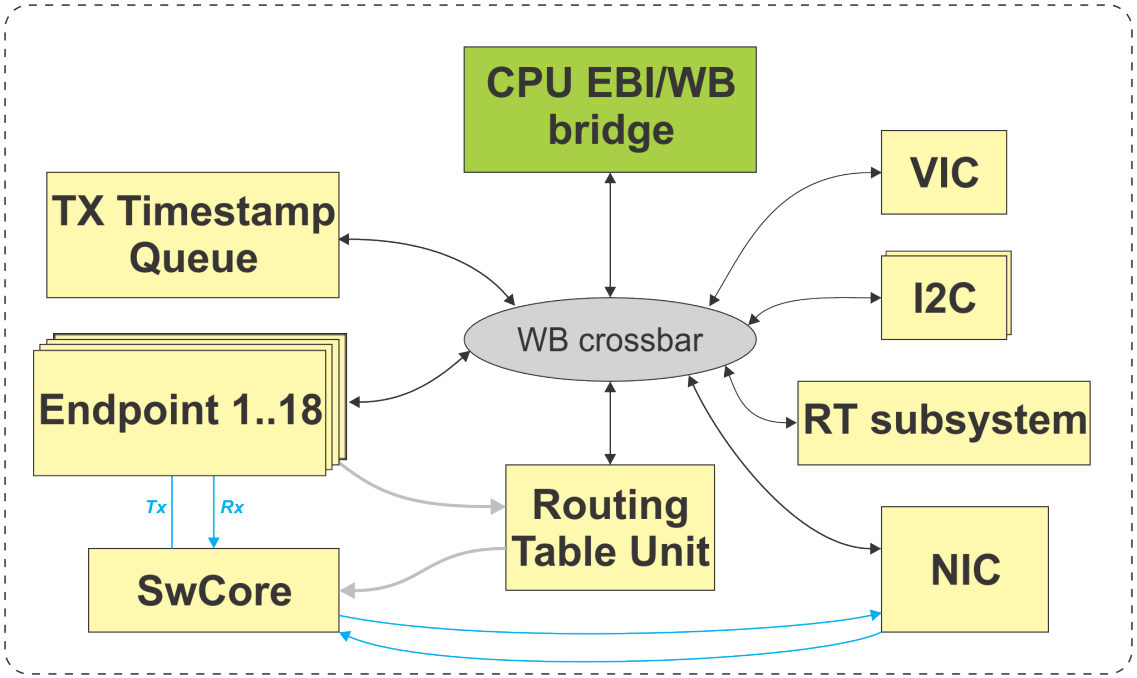
CERN BE-CO-HT

November 27, 2012

Agenda:

1. Current release
   - HDL architecture
   - Timing modules
   - Ethernet switching
   - WR Endpoint
2. Work in progress
   - RMON counters
   - Endpoint modifications
   - Switching Core
   - Routing Table Unit
   - Topology Resolution Unit
   - Resource Utilization

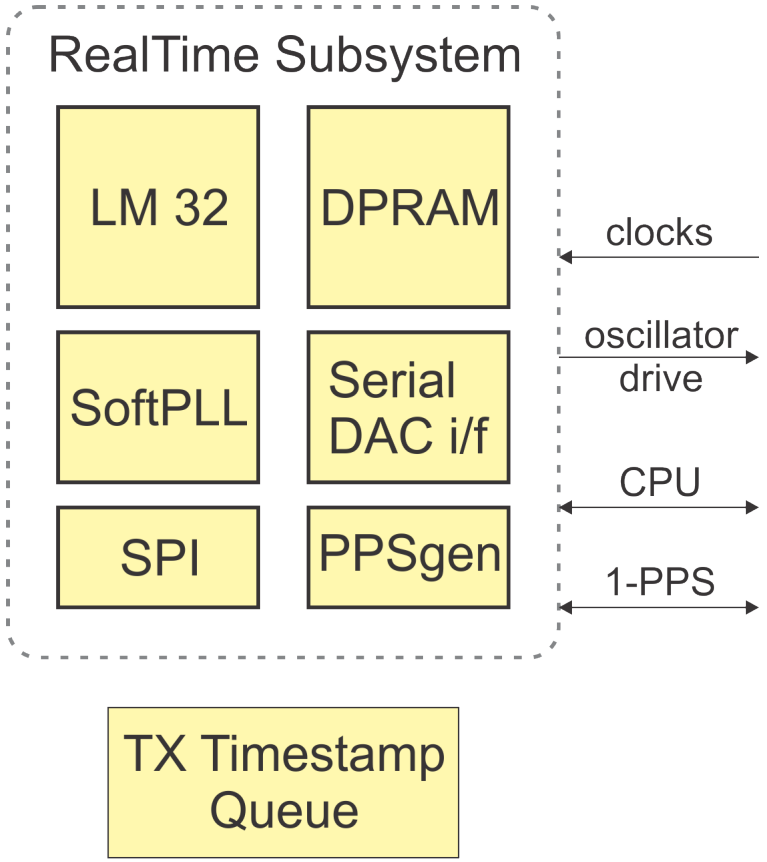
HDL architecture



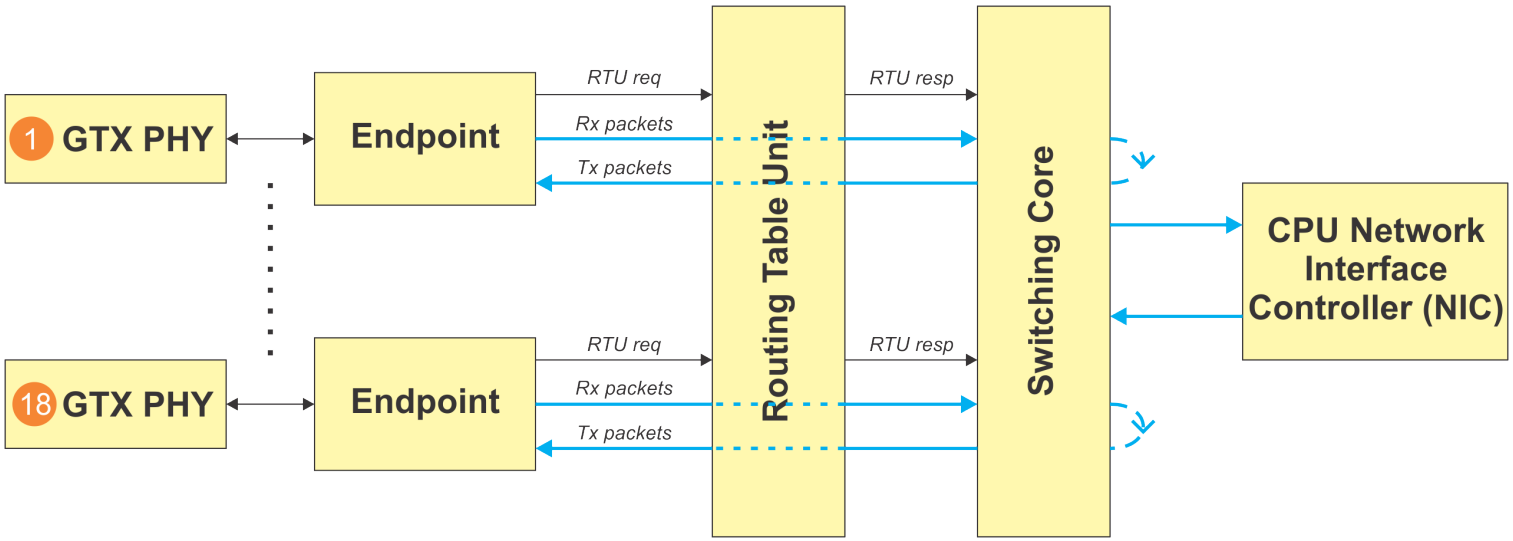
- first stable release for v3 of the hardware (tag: *wr-switch-sw-v3.0*);
- runs on Xilinx Virtex 6 (XC6VLX130T) but is fully generic, no vendor specific IP-cores required;
- Wishbone bus between all HDL modules;
- internal packets flow using pipelined Wishbone.

Timing modules

- software PLL implementation (LM32 + SoftPLL);
- SoftPLL the same as in WR PTP Core;
- current time stored in PPSgen block;
- SPI, Serial DAC interface for clock tuning;
- TX timestamps, frame IDs, port IDs stored in FIFO.

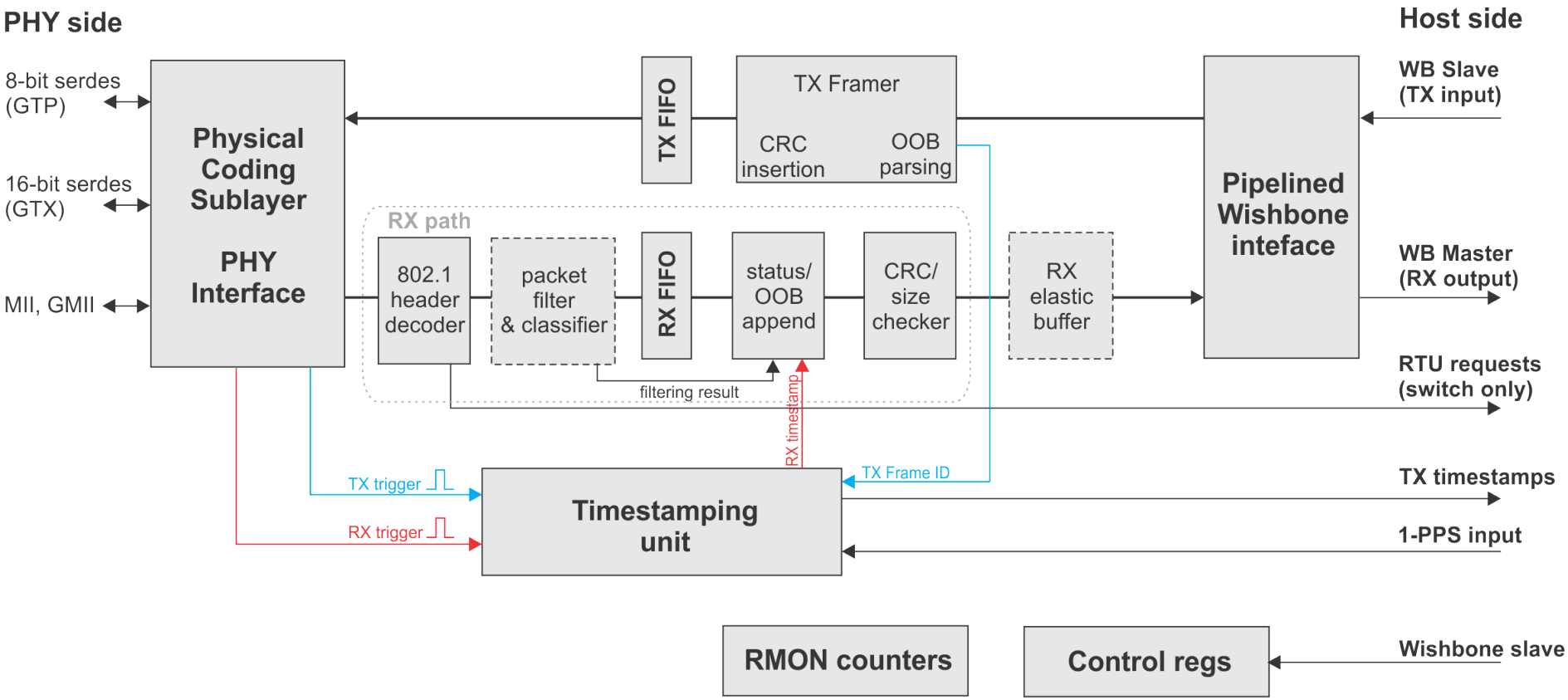


Ethernet switching

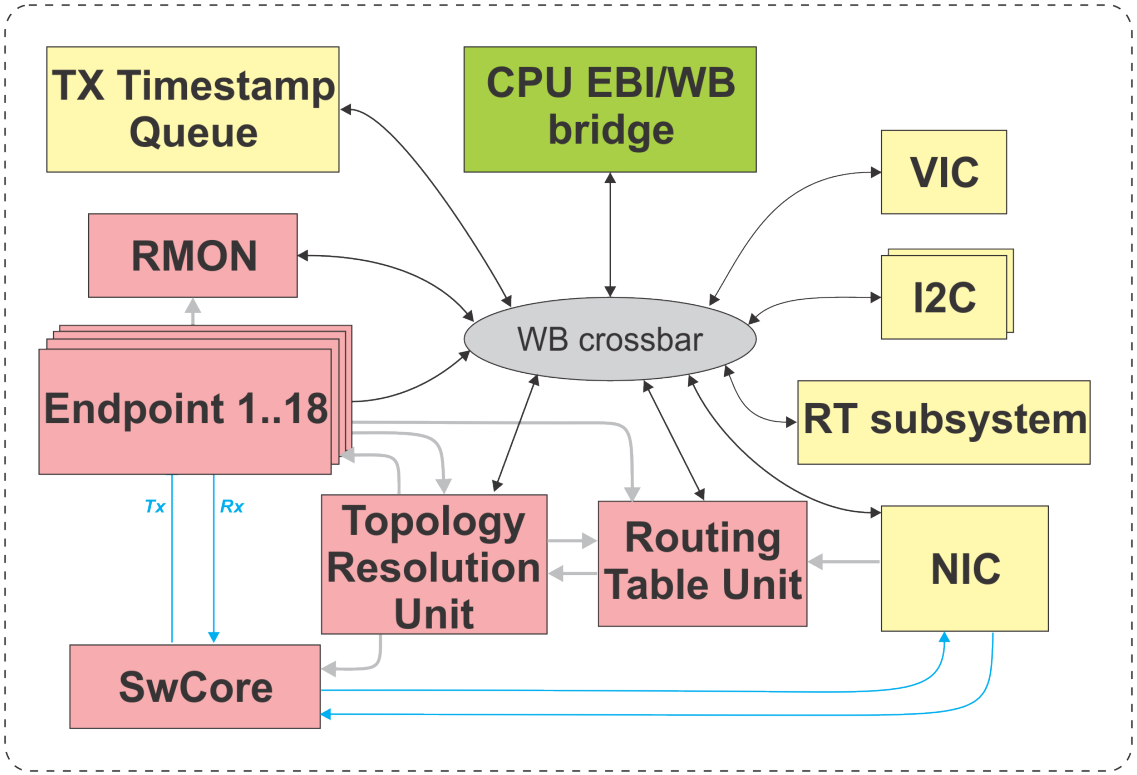


- Endpoint: Ethernet MAC implementation with precise timestamping;
- RTU: switching decisions for 18 WR Ports and internal NIC;
- SwCore: actual Ethernet frames switching;
- NIC: packets from/to main CPU.

WR Endpoint

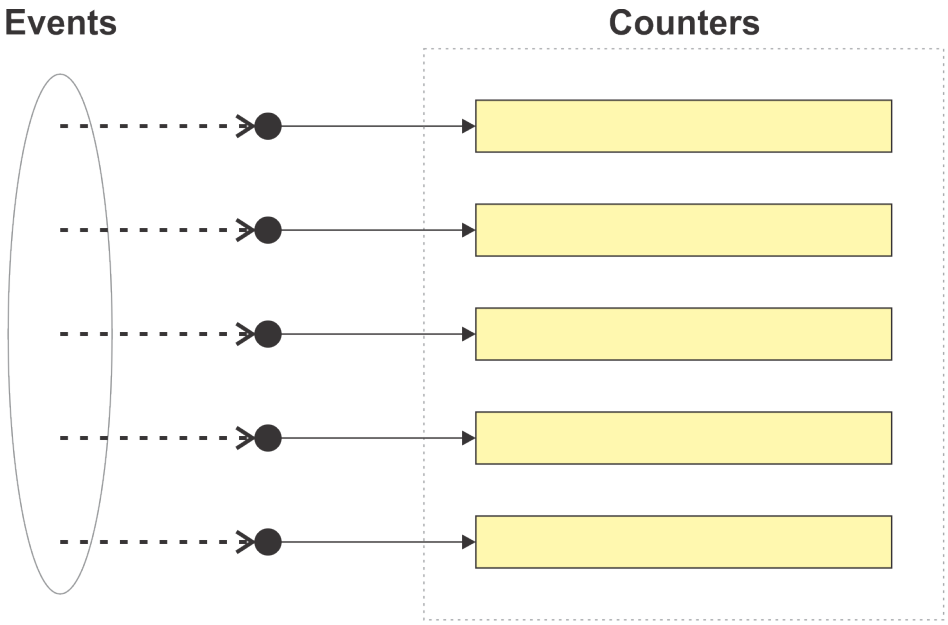

PHY side
Host side
8-bit serdes (GTP)
16-bit serdes (GTX)
MII, GMII
Physical Coding Sublayer
PHY Interface
TX FIFO
TX Framer
CRC insertion
OOB parsing
Pipelined Wishbone inteface
WB Slave (TX input)
WB Master (RX output)
RTU requests (switch only)
RX path
802.1 header decoder
packet filter & classifier
RX FIFO
status/ OOB append
CRC/ size checker
RX elastic buffer
filtering result
RX timestamp
TX trigger
RX trigger
TX Frame ID
Timestamping unit
TX timestamps
1-PPS input
RMON counters
Control regs
Wishbone slave

Work in progress



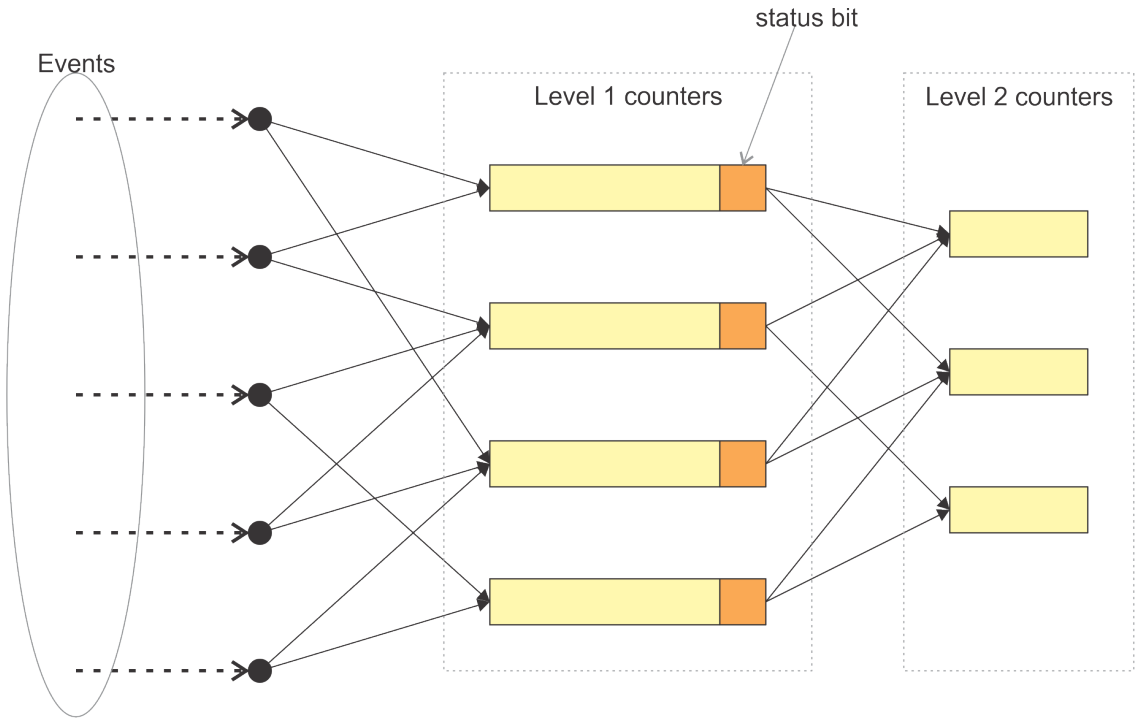
- RMON counters;
- Endpoint modifications;
- Switching Core - new functionality;
- Routing Table Unit - optimized;
- Topology Resolution Unit.

RMON counters



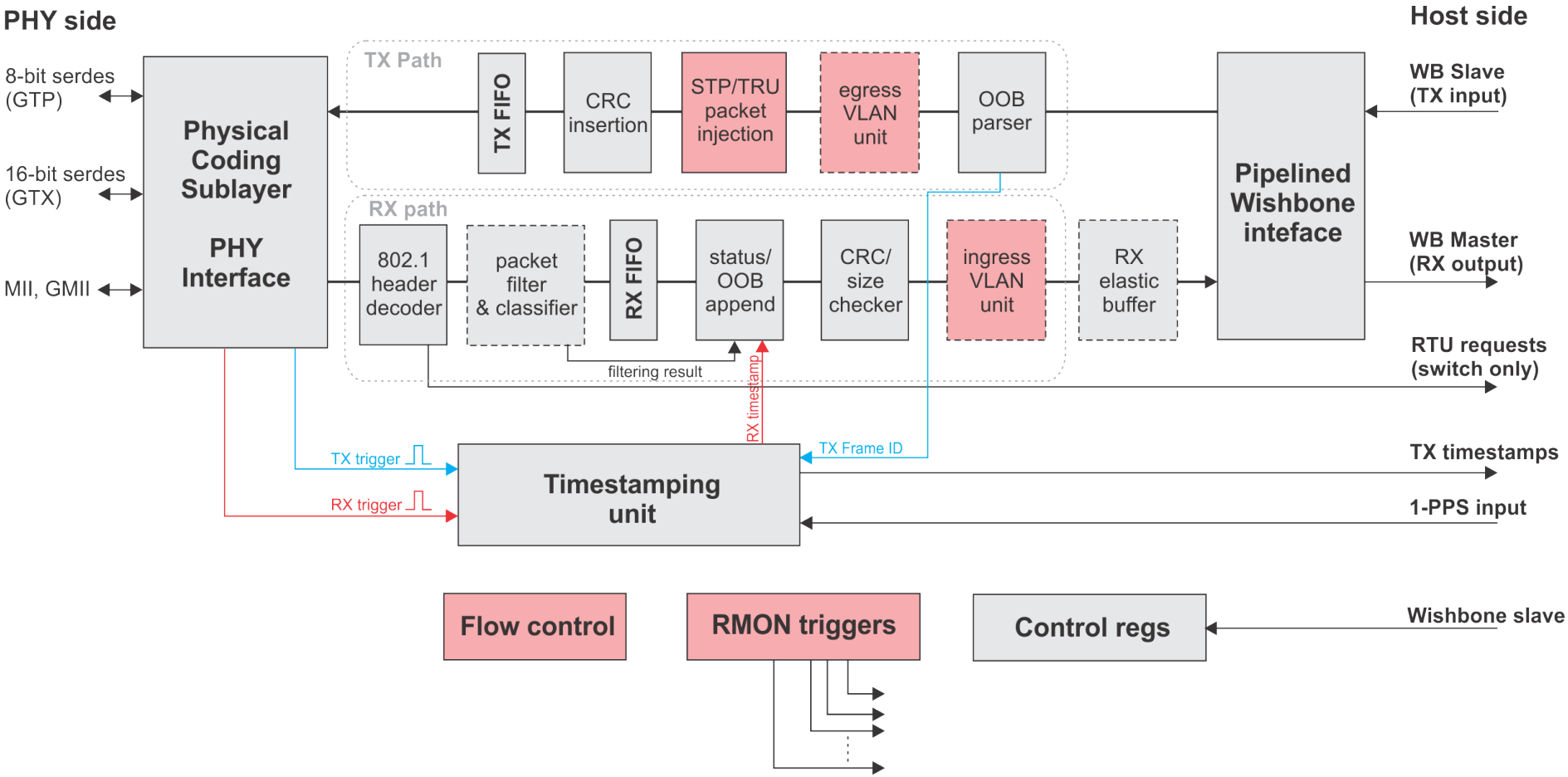
- direct counting consumes lots of resources;
- algorithms for approximate counting - use very little resources but cannot produce accurate statistics;
- Counter Braids seems to be the perfect solution.

Counter Braids

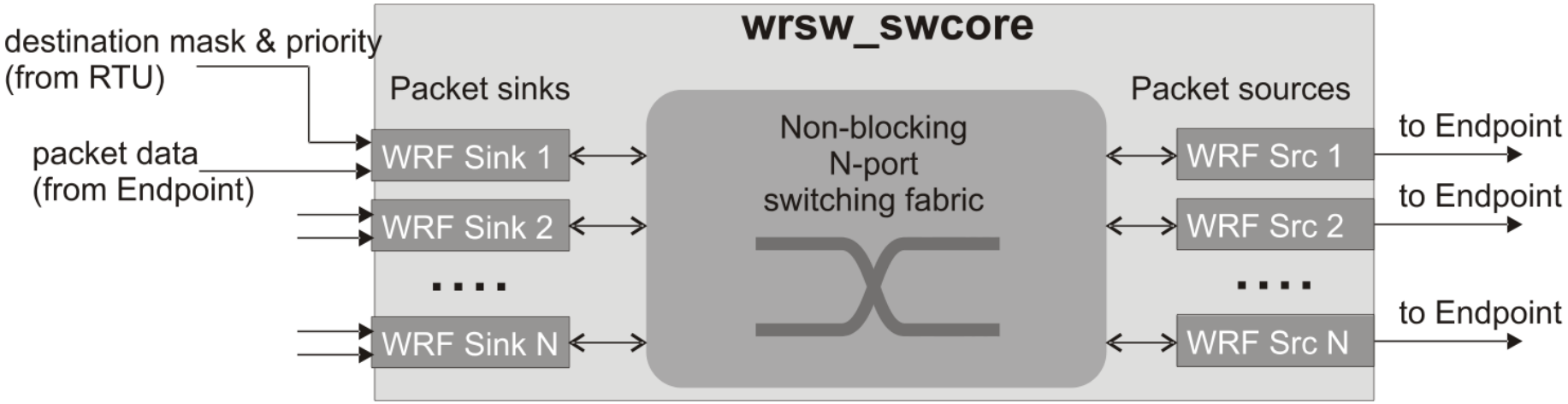


- each event increments more than one counter;
- few levels of small counters (most often 2-3 levels);
- overflowing counter on 1st level increments multiple counters on 2nd level;
- decoding done in software

Endpoint modifications

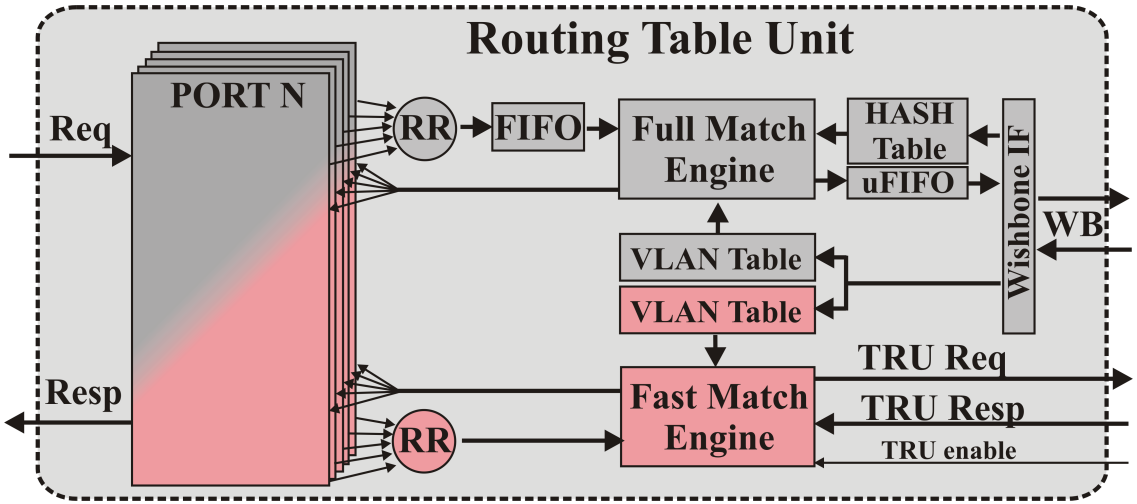


Switching Core



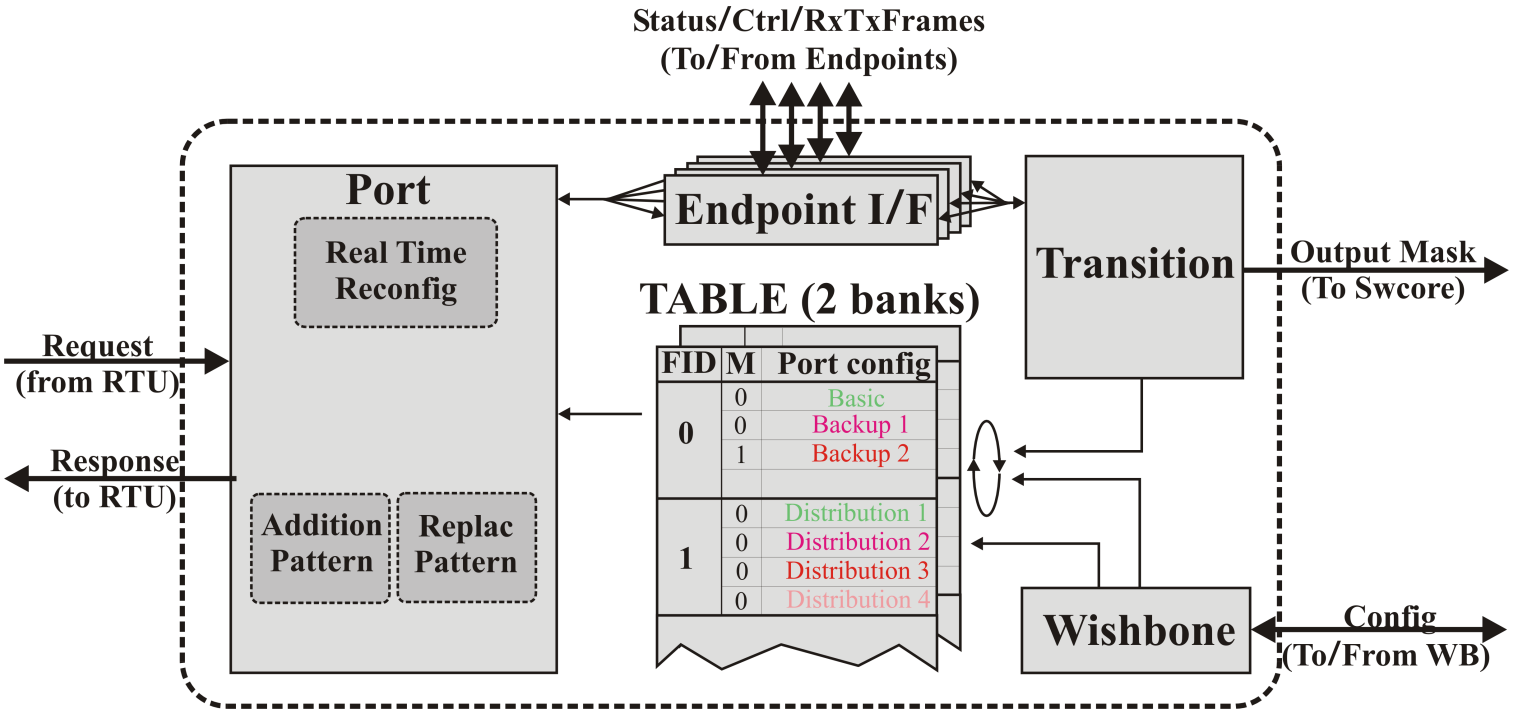
- optimization (wire speed)
- separation resources for HP traffic
- improvement output queues scheduling (decoupled)
- implementation of time-triggered output scheduling
- interface with TRU

Routing Table Unit



- fully deterministic and non-dropping for HP traffic
- Fast Match Engine
  - PTP/broadcast/Link-limited/configurable MACs traffic
  - configurable hardware support for HP traffic
  - response in max (N+5) cycles
  - interface with TRU
- port
  - port mirroring
  - drop/broadcast when full match too slow

Topology Resolution Unit



- universal unit for topology resolution protocol
  - port switch-over between redundant ports (eRSTP)
  - traffic distribution between redundant ports (eLACP)
- fully pipelined (3 cycles to answer, each cycle new request)

Resource Utilization

Feature	Resource utilization [%]								
	Bigger FPGA			Current FPGA					
	18 ports			18 ports			8 ports		
	LUTs	Slices	RAM	LUTs	Slices	RAM	LUTs	Slices	RAM
Master v3	42	57	37/9	79	96	58/15	39	56	49/7
Master v3 RTU optimization	47	61	45/9	85	98	63/15	39	55	55/7
Master v3 RTU optimization Endpoint new features (VLAN, flow control)	50	64	50/11	90	99	68/17	41	57	57/8
Master v3 RTU optimization Endpoint new features (VLAN, flow control) TRU	53	69	55/11	95	99	79/17	50	72	60/11

Thank you