Fermilab Summer Student Program 2015

Final Report

CMS L1 Silicon based tracking trigger

Evaluation of Virtex-7 and Kintex UltraScale FPGAs

Intern:
Irene Degl'Innocenti

Supervisor:
Ted Liu

Introduction

High luminosity conditions at the LHC pose many unique challenges for potential silicon based track trigger systems. CMS L1 Tracking Trigger will need to reconstruct charged particle trajectories for every beam crossing (one every 25 ns, or 40 Million beam crossings per second), from an ocean of input data (bandwidth required to transfer up to ˜ 100 Tb/s).

One of the major challenges is data formatting, where hits from thousands of silicon modules must first be shared and organized into overlapping eta-phi trigger towers. Communication between nodes requires high bandwidth, low latency, and flexible real time data sharing. The system not only requires a very large number of computing elements, but also high-speed transfer of data among processing elements and high-speed manipulation of very large quantities of data. So it's essential to every processing element to be performant in communication and data storage.

Field Programmable Gate Array (FPGA) devices, which are abundant in local cells, memory, and high speed serial transceivers, were selected for the core processing element on each Data Formatter board.

The high performances asked lead us to look for the most technology advanced devices and it is important to have a deep knowledge of their features and how they can be used at their best.

Indice

1 The Pulsar IIb board

The Pulsar IIb boards will be used in the ATLAS FTK Data Formatter system starting in 2015. The Pulsar IIb design forms the basic building block of a high performance scalable architecture, which may find applications beyond tracking triggers, and may serve as a starting point for future Level-1 silicon-based tracking trigger for CMS, where the full mesh backplane is used most effectively for sophisticated time multiplexing data transfer schemes.

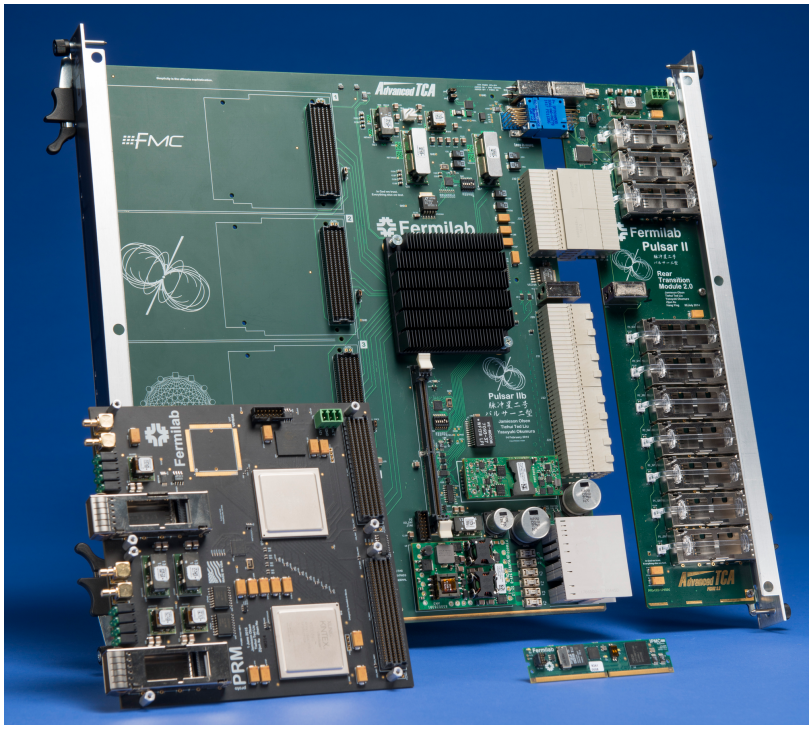

Figura 1: Prototype of Pulsar IIb

A prototype of a Pulsar IIb board has been realized and it is now under test. The main features of the boards are:

- Designed around a single **Virtex-7 FPGA**. Supported parts are: XC7VX415T, XC7VX485T, XC7VX550T or XC7VX690T. The prototype is provided with a XC7VX690T;
- 80 High speed SERDES transceivers (GTX/GTH)
  - 40 for Rear Transition Module (up to 400Gb/s)
  - 28 for the Full Mesh Fabric Interface (Ch 1 is 40Gb/s, Ch 2-13 are 20 Gb/s)
  - 12 for the FMC cards (30Gb/s per FMC)
- Four FMC mezzanine cards, each has:
  - 34 LVDS user-defined signals (LA00-LA34) + 2 LVDS output clocks
  - 3 GTH/GTX bidirectional SERDES channels
  - I2C bus

  - 12V and 3.3V power, up to 35W total

- Compatible with the LAPP IPMC module
  - 31 temperature, current, and voltage sensors
  - Board supports 10/100/1000-BASE-T Ethernet on Base Interface port 1

- 256MB DDR3 RAM
- Shelf-wide clock distribution via the backplane Synchronization Interface
- RTM is PICMG 3.8 compliant and supports hot swap
- Compatible with PRM (Pattern Recognition Memory) Mezzanine Card, provided with two **Kintex UltraScale FPGAs**.

1.1 Virtex-7 FPGA

Virtex-7 FPGAs, developed by Xilinx, are optimized for advanced systems requiring the highest performance and highest bandwidth connectivity. Between the 7-Series FPGAs, the Virtex-7 Family is optimized for highest system performance and capacity with a 2X improvement in system performance. Highest capability devices enabled by stacked silicon interconnect (SSI) technology.

Virtex FPGAs are typically programmed in hardware description languages such as VHDL or Verilog, using the Xilinx ISE or Vivado Design Suite computer software.

The Virtex series of FPGAs are based on Configurable Logic Blocks (CLBs), where each CLB is equivalent to multiple ASIC gates. Each CLB is composed of multiple slices, that differ in construction between Virtex families.

Virtex FPGAs include an I/O Block for controlling input/output pins on the Virtex chip, that support a variety of signalling standards. All pins default to "input" mode (high impedance). I/O pins are grouped into I/O Banks, where each Bank can support a different voltage.

In addition to configurable FPGA logic, Virtex FPGAs include fixed-function hardware for multipliers, memories, microprocessor cores, FIFO and ECC logic, DSP blocks, PCI Express controllers, Ethernet MAC blocks, and high-speed serial transceivers.

1.2 Kintex UltraScale

Kintex UltraScale devices, always developed by Xilinx, deliver ASIC-class system-level performance, clock management, and power management for next-generation price-performance-per-watt. These second generation devices expand the mid-range by delivering the highest throughput with lowest latency for medium-to-high volume applications, including 100G networking, wireless infrastructure, and other DSP-intensive applications. Based on the ASIC-class advantage of the UltraScale architecture, Kintex UltraScale devices are co-optimized with the Vivado Design Suite. High DSP and block RAM-to-logic ratios and next-generation transceivers, combined with low-cost packaging, enable an optimum blend of capability and cost. The family incorporates:

- ASIC-like clocking for scalability, performance and lower dynamic power;

- Next-generation routing for rapid timing closure;
- Enhanced logic infrastructure for maximum performance and device utilization.

2 Serial I/O Links Test on Virtex-7

2.1 Virtex-7 GTH Transceivers

The 7 series FPGAs GTX and GTH transceivers are power-efficient transceivers, supporting line rates from 500 Mb/s to 12.5 Gb/s for GTX transceivers and 13.1 Gb/s for GTH transceivers. The GTX/GTH transceiver is highly configurable and tightly integrated with the programmable logic resources of the FPGA.

The CORE Generator tool in the Vivado Design Suite includes a wizard to automatically configure GTX/GTH transceivers to support configurations for different protocols or perform custom configuration. The GTX/GTH transceiver offers a data rate range and features that allow physical layer support for various protocols.

The Pulsar IIb prototype FPGA XC7VX690T is provided by 80 GTH transceivers, supporting line rates up to 13.1 Gb/s.

2.1.1 VC709 Evaluation Board

To better understand the properties of the GTH Transceivers I was given a VC709 board. The VC709 evaluation board for the Virtex-7 FPGA provides a hardware environment for developing and evaluating designs targeting the Virtex-7 XC7VX690T-2FFG1761C FPGA. The VC709 board provides features common to many embedded processing systems, including dual DDR3 small outline dual-inline memory module (SODIMM) memories, an 8-lane PCI Express interface, general purpose I/O, and a UART interface.

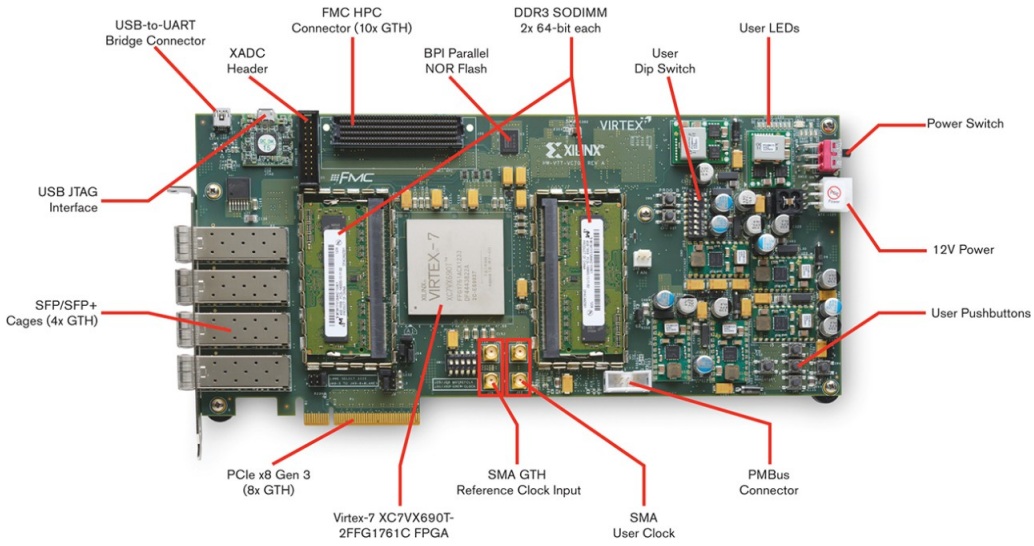

Figura 2: View of a VC709 board

The VC709 board provides access to 22 GTH transceivers:

- Eight of the GTH transceivers are wired to the PCI Express x8 endpoint edge connector (P1) fingers.
- Ten of the GTH transceivers are wired to the FMC HPC connector (J35).
- Four of the GTH transceivers are wired to the four SFP/SFP+ connectors (P2, P3, P4, P5).

The GTH transceivers in 7 series FPGAs are grouped into four channels described as Quads. The reference clock for a Quad can be sourced from the Quad above or Quad below the GTH Quad of interest. There are six GTH Quads on the VC709: Quad 113, Quad 114, Quad 115, Quad 117, Quad 118, Quad 119. Through SFP+ connectors we have access to Quad 113.

2.1.2 IBERT Test on optical cables

The customizable LogiCORE IP Integrated Bit Error Ratio Test (IBERT) core for 7 series FPGA GTX transceivers is designed for evaluating and monitoring the GTX transceivers. This core includes pattern generators and checkers that are implemented in FPGA logic, and access to ports and the dynamic reconfiguration port attributes of the GTX transceivers. Communication logic is also included to allow the design to be run time accessible through JTAG. This core can be used as a self-contained or open design, based on customer configuration.

A single eye scan measurement consists of accumulating the number of data samples (sample count) and the number of times that the offset sample disagreed with the data sample (error count). The bit error ratio (BER) at the programmed vertical and horizontal offset is the ratio of the error count to the sample count.

Transceivers can be customized for the desired line rate, reference clock rate, reference clock source, and datapath width. The test requires a system clock that can be sourced from a pin or one of the enabled GTX transceivers.

Using the Vivado Design Suite and the VC709 board I programmed the FPGA on it with the IBERT IP CORE; after connected the Quad 113 Transceivers through the SFP+ connectors of the board with some optic cables and loopbacks I was able to manage and test the communication on run time.

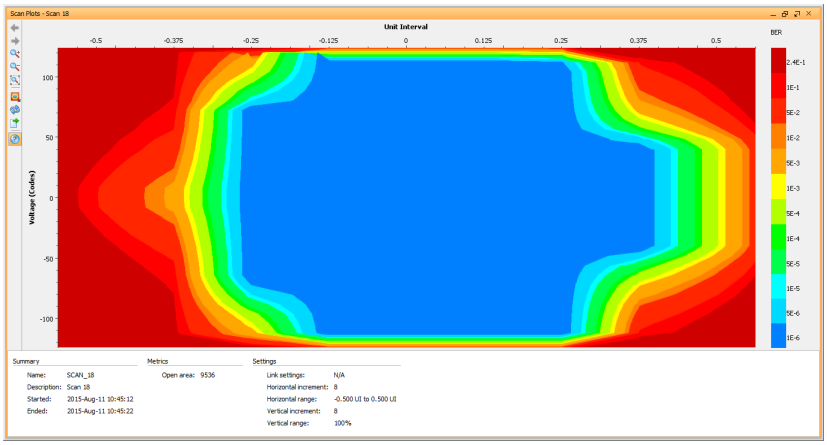

Figura 3: Example of BER map for the optical cable AMPHENOL CABLE LSZH 04/13 0942M

Vivado gives also the opportunity to change the transceivers setting on run time.

It is possible to produce a BER map, referred to as a statistical eye, where the color map represents $\log_{10}(BER)$. The optical cables I tested are:

- Prismian O.C. 2F 50/125 OM2 BB,
- AFL Telecommunications 1-800-AFL_FIBER 50/125,
- AMPHENOL CABLE LSZH 04/13 0942M.

3 Data storage: Block Ram

3.1 Block RAM in Virtex-7 FPGAs

The BRAM Block is a configurable memory module. The block RAM in Xilinx 7 series FPGAs stores up to 36 Kbits of data and can be configured as either two independent 18 Kb RAMs, or one 36 Kb RAM. Each 36 Kb block RAM can be configured as a 64K x 1 (when cascaded with an adjacent 36 Kb block RAM), 32K x 1, 16K x 2, 8K x 4, 4K x 9, 2K x 18, 1K x 36, or 512 x 72 in simple dual-port mode. Each 18 Kb block RAM can be configured as a 16K x 1, 8K x2 , 4K x 4, 2K x 9, 1K x 18 or 512 x 36 in simple dual-port mode. Write and Read are synchronous operations; the two ports are symmetrical and totally independent, sharing only the stored data. Each port can be configured in one of the available widths, independent of the other port. In addition, the read port width can be different from the write port width for each port. The memory content can be initialized or cleared by the configuration bitstream. During a write operation the memory can be set to have the data output remain unchanged, reflect the new data being written or the previous data now being overwritten. XC7VX690T FPGA has 1470 total 36 Kb Block RAM Blocks per Device. Embedded dual- or single-port RAM modules, ROM modules, synchronous FIFOs, and data width converters are implemented using the Xilinx CORE GeneratorTM block memory modules in Vivado Design Suite.

3.1.1 Possible configurations

TDP (True Dual Port) The true dual-port 36 Kb block RAM dual-port memories consist of a 36 Kb storage area and two completely independent access ports, A and B. Similarly, each 18 Kb block RAM dual-port memory consists of an 18 Kb storage area and two completely independent access ports, A and B. The structure is fully symmetrical, and both ports are interchangeable.

Data can be written to either or both ports and can be read from either or both ports. Each write operation is synchronous, each port has its own address, data in, data out, clock, clock enable, and write enable. The read and write operations are synchronous and require a clock edge.

SDP (Simple Dual Port) In this mode, the block RAM port width doubles to 36 bits for the 18 Kb block RAM and 72 bits for the 36 Kb block RAM. In simple dual-port mode, independent Read and Write operations can occur simultaneously, where port A is designated as the Read port and port B as the Write port. When the Read and Write port access the same data location at the same time, it is treated as a collision, identical to the port collision in true dual-port mode.

Cascadable Block RAM Two 32K x 1 RAMs can be combined to form one 64K x 1 RAM without using local interconnect or additional CLB logic resources. Any two adjacent block RAMs can be cascaded to generate a 64K x 1 block RAM. Increasing the depth of the block RAM by cascading two block RAMs is available only in the 64K x 1 mode.

3.1.2 Write Modes

A write operation is a single clock-edge operation. The write address is registered on the write port, and the data input is stored in memory.

Three settings of the write mode determines the behavior of the data available on the output latches after a write clock edge: WRITE_FIRST, READ_FIRST, and NO_CHANGE. Write mode selection is set by configuration. The Write mode attribute can be individually selected for each port. The default mode is WRITE_FIRST. WRITE_FIRST outputs the newly written data onto the output bus. READ_FIRST outputs the previously stored data while new data is being written. NO_CHANGE maintains the output previously generated by a read operation.

3.1.3 Waveform Simulation

To run a behavioral simulation of a block RAM on a Virtex-7 FPGA I wrote a VHDL file using the MACRO provided by Xilinx, describing the hardware of the block RAM, and a VHDL test bench to simulate possible reading and writing events.

Parameters for different simulations

- Block RAM size: 18Kb ? 36Kb;
- Optional output registers: enabled ? not enabled;
- Data Width: 16 bits ? 32 bits;
- Write Mode: WRITE_FIRST ? READ_FIRST;
- Clock frequency: 200 Mhz ? 400 Mhz.

Observations

- The clockedge-to-output delay is modeled 0.1 ns.
- When WE_A is low, port A outputs the previuosly store data while the new one is being written.
- When WE_B is low, port B outputs the newly written data onto the output bus.

3.2 Block RAM in Kintex UltraScale FPGAs

3.2.1 Main changes from 7-series

- SDP memory supports NO_CHANGE mode;

- New data cascading scheme (more than x2 block RAMs);
- Address enable added;
- Dynamic Sleep mode preserving data content.

3.2.2 Waveform Simulation

To run a behavioral simulation of a block RAM on a Kintex UltraScale FPGA I used the IP BRAM generator CORE provided by Vivado Design Suite. So I implemented that in my VHDL project and I wrote a test bench to simulate the behavior of the memory with different configurations:

- #1 36Kb block RAM TDP with totally different width and length in reading and in writing and for the two independent ports;
- #4 36Kb block RAMs in single-port mode cascaded to form a 36x4096 RAM;
- #1 36Kb block RAM SDP with ECC enabled.

ECC Hamming Code Error Correction The RAMB36E2 in simple dual-port mode can be configured as a single 512 x 64 RAM with built-in Hamming code error correction using the extra eight bits in the 72-bit wide RAM.

Eight protection bits (ECCPARITY) are generated during each write operation and stored with the 64-bit data into the memory. These ECCPARITY bits are used during each read operation to correct any single-bit error, or to detect (but not correct) any double-bit error.

During each read operation, 72 bits of data (64 bits of data and 8 bits of parity) are read from the memory and fed into the ECC decoder. The ECC decoder generates two status outputs (SBITERR and DBITERR) that are used to indicate the three possible read results: No error, single-bit error corrected, and double-bit error detected. In the standard ECC mode, the read operation does not correct the error in the memory array, it only presents corrected data on DOUT. To improve FMAX, optional registers controlled by the DO_REG attribute are available for data output (DOUT), SBITERR, and DBITERR.

The encoder and decoder can be accessed separately (independently) for external use in RAMB36E2 in simple dual-port mode and by extension in the FIFO36E2, both in 72-bit mode. To use the encoder by itself, the data needs to be sent through the DIN port, and the ECCPARITY output port can be sampled. To use the decoder by itself, the encoder is disabled, the data is written into the block RAM, and the corrected data and status bits are read out of the block RAM.

4 Conclusions

Pulsar IIb board is now under test. The group, after testing the reliability of the communications with instruments similar to the ones I used to test the optical cables, is now working on the firmware aimed to implement the passage from the local coordinates of the silicon modules to the global coordinates needed to build a global image of the detected events. The firmware is based on associative memories, which cores are LUTs (Look Up Tables). The idea is to realize the LUTs with appropriate block RAMs. My work has given a first general view of the possibilities provided by FPGAs' block RAMs. From my point of view I had the opportunity to put my hands on the most technology advanced devices available now; I also could taste how the design targets are totally different from the ones my education made used to me.

5 Bibliography

Y. Okumura, J. Olsen, T. Liu, H. Yin, *Prototype performance studies of a Full Mesh ATCA-based General Purpose Data Processing Board*,(2013)

7 Series FPGAs Overview, (05, 2015)

UltraScale Architecture and Product Overview, (08, 2015)

7 Series FPGAs GTX/GTH Transceivers User Guide, (08, 2015)

VC709 Evaluation Board for the Virtex-7 FPGA User Guide, (09, 2015)

VC709 GTH IBERT Design Creation, (06, 2014)

7 Series FPGAs Memory Resources User Guide, (11, 2014)

UltraScale Architecture Memory Resources User Guide, (02, 2015)