PROCEEDINGS OF SCIENCE

Serenity: An ATCA prototyping platform for CMS Phase-2

A. Rose[1], D. Parker, G. Iles
Imperial College London
Prince Consort Road, London, SW7 2BW, United Kingdom
E-mail: awr01@imperial.ac.uk

Ö. Sahin, P.-A. Bausson
CEA Saclay

A. Tsirou
CERN

G. Fedi, P. G. Verdini
INFN Pisa

L. Ardila, M. Balzer, T. Schuh
Karlsruhe Institute of Technology

T. Williams, A. Thea, K. Harder
STFC Rutherford Appleton Laboratory

S. Dugad, R. Shukla, I. Mirza
Tata Institute of Fundamental Research

Serenity is an ATCA prototyping platform designed to explore alternative, novel design choices for CMS Phase-2. It uses a newly available interconnect technology from Samtec (Z-RAY) to mount a removable processing unit (FPGA) that should mitigate risk and provides significant flexibility in processing unit choice and connectivity. We explore the pros and cons of using an industry-standard Computer-On-Module running standard Centos Linux and a small service FPGA for low level control. Specially designed Kapton heaters have been used to validate the thermal design of the card and broader considerations of ATCA systems.

Topical Workshop on Electronics for Particle Physics (TWEPP2018)
17-21 September 2018
Antwerp, Belgium

[1]Speaker

PoS(TWEPP2018)115

1. Overview

Serenity is an Advanced Telecommunications Computing Architecture (ATCA) prototyping platform consisting of three elements: An ATCA Carrier-Card which provides the common board services, that is, power, clocking, optical interfaces, electrical interconnections between daughter-cards, the intelligent platform management controller (IPMC) functionality and an on-board CPU to control and manage the board; daughter-cards which host the data-processing elements themselves (in our case FPGAs) and a framework of generic, flexible firmware and software.

The aim of Serenity is to be as simple and flexible as possible, and to that end, uses Commercial Off-The-Shelf (COTS) components wherever possible. Specifically, the IPMC functionality required by ATCA blades is provided by the CERN IPMC module [1], which itself runs the commercial PigeonPoint software [2]. Likewise, the board-level control is provided by a ComExpress type-10 computer-on-module which, being an industry standardized form-factor [3], does not restrict us to a single vendor. A single lane of PCIexpress (Gen-2) connects the CPU to each interposer site, and also to a very small Xilinx Artix-7 FPGA which provides the protocol and voltage conversions necessary to interface to the JTAG and I2C chains which control the board: such an arrangement provides a very clean separation of hardware, firmware and software, simplifying the parallel development of the three. An additional advantage of the ComExpress form-factor over system-on-chip solutions is that, because it is utilizes a standard CPU, it can also run a standard operating-system, simplifying long-term maintenance, the primary consideration of the experimental system-administrators.

Serenity originated out of the UK CMS collaboration, which is involved in the development of back-end electronics for many subsystems for the phase-II upgrade of the CMS experiment, including the trigger and data-acquisition systems for the high-granularity endcap calorimeters [4], the outer-tracker readout [5], and the upgrade of the level-1 trigger towards particle-flow type algorithms (the so-called level-1 correlator) [6]. It was, however, always intended to be a generic, open processing platform developed by the entire community, building upon and pushing further the successful common hardware model first implemented by its predecessor, the MP7 [7]. The systems for which Serenity was originally targeted all have very different requirements in terms of logic resources, cost considerations and connectivity. To meet such disparate requirements, as well as those of potential future users, the chosen solution was to provide a standardized interface to a very large number of connectivity options, which could be used if required and simply ignored if not. This was done by defining a footprint for a standardized daughter-card, onto which an FPGA in any package, from any family or generation, and possibly even vendor, could be mounted: the design of the daughter-card selects to which carrier resources the FPGA is connected, and provides the specialization of the generic board to a specific application. The locating of the FPGAs on daughter-cards has a second advantage in that, by placing the FPGA on the simple daughter-card, financial risk is reduced by isolating the FPGAs, which constitute the bulk of the cost, from the carrier, which carries the bulk of the potential failure modes.

The connectivity options for the daughter-cards are 144 channels, nominally 72 Tx + 72 Rx, routed to Samtec Firefly connectors [8], 64 differential pairs routed between the two daughter-card sites (the so-called inter-interposer bus) and 2 Tx + 2 Rx channels to a Quad Small Form-factor Pluggable (QSFP) optical module. Although the Firefly signals were given a nominal direction assuming the use of 12-channel unidirectional optical Firefly modules, bidirectional and

passive electrical (non-directional) modules are also available and compatible with the Serenity design. The inter-interposer bus is passive and makes no assumption of directionality. The ATCA carrier includes provision for an additional 48 channels of Firefly connectivity (nominally 24 Tx + 24 Rx) for each daughter-card which could be accessed via a minor revision of the carrier to reroute part of the inter-interposer bus to eight "spare" Firefly connectors, increasing the optical bandwidth by up to 33% at the cost of reducing the inter-site connectivity by up to 75%. These features are shown in Figure 1. The connectivity from the carrier to the daughter-cards is provided via a factory-customized 1990 pin Samtec Z-RAY connector [9].

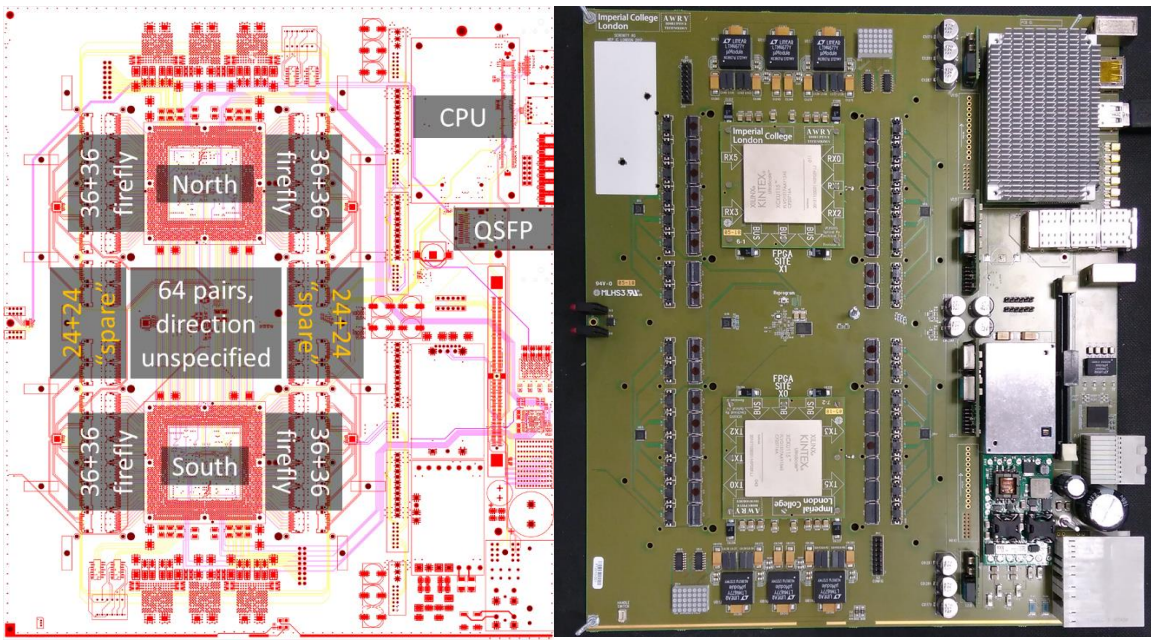

Figure 1: Left – The layout of the top surface of the Serenity ATCA carrier highlighting the two daughter-card sites (labelled North and South although they are functionally identical), the 64 inter-interposer pairs, the positions of the Firefly optical channels (including the optional "spare" channels), the QSFP for DAQ applications, and the on-board CPU. Right – The Serenity ATCA carrier hosting two different design of daughter cards each utilizing Xilinx KU115 FPGAs, the"North" receiving data optically and transmitting across the inter-interposer bus and the "South" receiving data on the inter-interposer bus and transmitting optically.

Serenity is currently in the prototyping stage, and at the outset we had a set of concerns about the thermal and acoustic issues related to the ATCA form-factor, about 16 and 25Gbps signal integrity, and more abstractly, whether different countries could work together on different aspects of common hardware so that all might benefit from the economy of scale, a model which is somewhat at odds with the traditional idea of "system ownership". This Serenity prototype was intended to answer these questions.

2. Signal integrity testing

Given that the primary feature of Serenity is its large amount of serial connectivity, testing the signal integrity above 16Gbps was the highest priority. Infrastructure was put in place to produce simultaneous eye-diagrams from all links on both processing FPGAs. Tests were done at 16Gbps between two Xilinx XCKU115 FPGAs (Figure 2) and preliminary tests at 25Gbps between two Xilinx XCKU15P FPGAs (Figure 3). Even with the relatively closed optical eyes,

both the inter-interposer and optical link transferred 8×10^{14} bits without error. We thus conclude that the electrical performance of both the interposer and the Firefly connectors looks excellent but that it is prudent to undertake more in-depth studies into the signal integrity of the optical modules, how the eyes are affected by the FPGA transceiver and optical module settings.

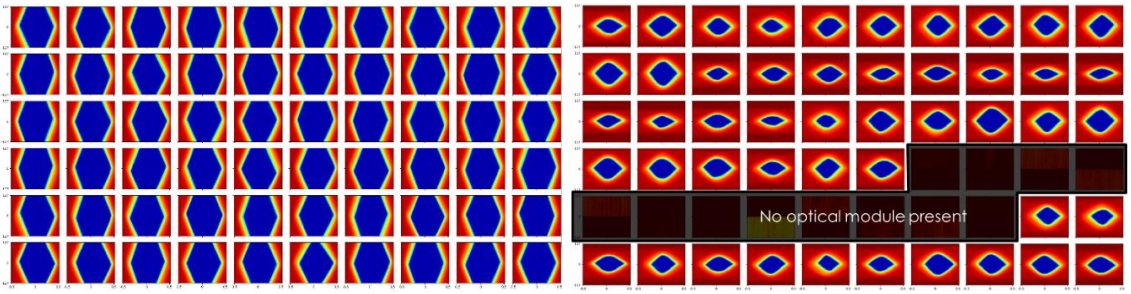

Figure 2: Left – Simultaneous eye-diagrams for the 60 available inter-interposer links. Right – Simultaneous eye-diagrams for 48 of the 60 available optical links. Both sets of measurements are between two independent FPGAs, the links configured to run at 16Gbps, PRBS-7, Decision Feedback Equalization (DFE) disabled, and no pre- or post-cursor. The optical eyes were measured over 10m of optical fibres with two MTP connectors, and the optical modules were used with their default settings.

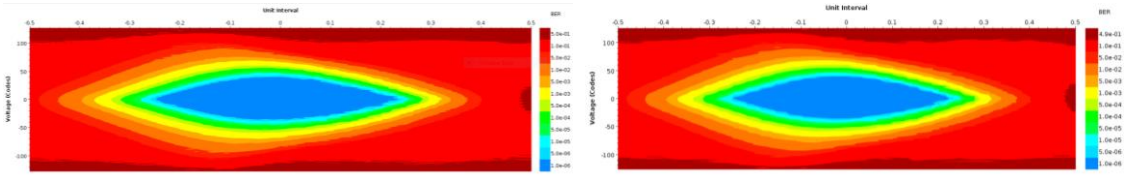

Figure 3: Preliminary eye-diagrams for a serial link running (left) over passive copper Firefly loopback-cable and (right) over the inter-interposer bus between two independent FPGAs. Both tests were configured using 25Gbps, PRBS31.

3. Thermal studies

In order to understand the power, cooling, thermal and acoustic limitations of Serenity, and ATCA systems in general, a test-stand was constructed at CERN. Custom kapton heaters were designed and mounted on aluminium blocks to represent optical modules and FPGAs, and these assemblies mounted on a blank Serenity/daughter-card PCBs, along with off-the-shelf heatsinks and twelve thermocouples to probe the air temperature at various locations. Measurements were made at various fan-speeds and thermal-loads, and the results compared against thermal simulations made with both Ansys Fluent [10] and Mentor FloTherm [11]. The Fluent simulations were found to match well the observed behaviour, and measurements indicate that, even with the naïve heatsink designs, the optics can be kept below the 50°C threshold required for them to survive for the 10-15 year lifetime of the experiment. Alternative designs for both the FPGA and optical heatsinks have been simulated in Fluent, with some designs offering significant improvement. These optimized designs are being manufactured and will be tested. The initial concerns over the noise and power-requirements of ATCA systems have only been partially allayed, however: to operate two fan-trays at maximum speed requires 2kW of power and produces 92dBA at a distance of 1m, which would cause considerable problems when scaled to systems such as those required by CMS phase-II. By lowering the fan-speed, however, a 75% reduction in fan-power resulted in only an 11°C increase in FPGA temperature, which suggests that these extreme fan-speeds have limited advantage.

4. Clocking

Serenity is intended as the clock-distribution system for at least one system requiring a high-performance (320.624MHz) LHC clock. A further advantage of having the FPGAs mounted on daughter-cards is that the performance of the carrier and interposers can be measured independent of the performance of the FPGA and firmware. A daughter-card was produced to expose the clock-signals: a low-jitter (1.3ps RMS) clock was injected through the ATCA zone-2 connector and the signals measured on the daughter-card using a 6GHz, 20GS/s oscilloscope configured to acquire 1.3 million clock-cycles with a large acquisition window (20 million samples) corresponding to a 1ms continuous acquisition, thereby scanning jitter frequencies above 1kHz. Across all 18 available clocks (9 on each interposer site), the channel-to-channel RMS jitter was measured to be 2.8ps, sufficient for Serenity to be considered "an ideal clock-distribution node".

5. Conclusion

Serenity has succeeded in its initial aims of allowing a single board to support multiple types of FPGA and user-specified connectivity. A number of different institutes have successfully produced their own daughter-cards with different types of FPGAs suggesting that the community is open to this new "common-hardware" approach. Initial signal integrity tests suggest that concerns over interposer signal integrity were unfounded, since performance at both 16 and 25Gbps looks good. The clock-distribution system has been measured and found to perform beyond expectations. Physical measurements and simulations indicate that the cooling of both the optics and FPGAs on daughter-cards looks manageable, although general concerns over power and noise of HL-LHC scale systems built using ATCA remain.

References

[1] Design of an AdvancedTCA board management controller (IPMC), J. Mendez et al. 2017 JINST 12 C03010, 10.1088/1748-0221/12/03/C03010
[2] Pigeon Point PICMG products, http://www.pigeonpoint.com/products_picmg.html
[3] ComExpress, PICMG, https://www.picmg.org/openstandards/com-express
[4] The Phase-2 Upgrade of the CMS endcap calorimeter: Technical Design Report, CMS-TDR-019, 9 Apr 2018, https://cds.cern.ch/record/2293646/files/CMS-TDR-019.pdf
[5] The Phase-2 Upgrade of the CMS tracker: Technical Design Report, CMS-TDR-014, 1 Jul 2017, https://cds.cern.ch/record/2272264/files/CMS-TDR-014.pdf
[6] The Phase-2 Upgrade of the CMS Level-1 trigger: Interim Technical Design Report, CMS-TDR-017, 12 Sep 2017, https://cds.cern.ch/record/2283192/files/CMS-TDR-017.pdf
[7] Development and testing of an upgrade to the CMS level-1 calorimeter trigger, A. Rose et al., 2013 JINST 9 C01006, 10.1088/1748-0221/9/01/C01006
[8] Firefly™ micro flyover system™, Samtec Inc., https://www.samtec.com/optics/optical-cable/mid-board/Firefly
[9] Z-RAY® ultra-low profile arrays, Samtec Inc., https://www.samtec.com/connectors/high-speed-board-to-board/high-density-arrays/zray
[10] Fluent, Ansys Inc., https://www.ansys.com/en-gb/products/fluids/ansys-fluent
[11] FloTHERM, Siemens Mentor Graphics, https://www.mentor.com/products/mechanical/flotherm/

PoS(TWEPP2018)115