The Development of a General Purpose ARM-based Processing Unit for the ATLAS TileCal sROD

Mitchell A. Cox*, Robert Reed and Bruce Mellado

*University of the Witwatersrand, Johannesburg,
1 Jan Smuts Avenue, Braamfontein, Johannesburg, South Africa
E-mail: mitchell.cox@cern.ch*

ABSTRACT: The Large Hadron Collider at CERN generates enormous amounts of raw data which present a serious computing challenge. After Phase-II upgrades in 2022, the data output from the ATLAS Tile Calorimeter will increase by 200 times to 41 Tb/s! ARM processors are common in mobile devices due to their low cost, low energy consumption and high performance. It is proposed that a cost-effective, high data throughput Processing Unit (PU) can be developed by using several consumer ARM processors in a cluster configuration to allow aggregated processing performance and data throughput while maintaining minimal software design difficulty for the end-user. This PU could be used for a variety of high-level functions on the high-throughput raw data such as spectral analysis and histograms to detect possible issues in the detector at a low level. High-throughput I/O interfaces are not typical in consumer ARM System on Chips but high data throughput capabilities are feasible via the novel use of PCI-Express as the I/O interface to the ARM processors. An overview of the PU is given and the results for performance and throughput testing of four different ARM Cortex System on Chips are presented.

KEYWORDS: Computing; Data processing methods.

*Corresponding author. On behalf of the ATLAS Tile Calorimeter System.

Contents

1. Introduction	1
2. TileCal Read Out Architecture	2
3. General Purpose Processing Unit	2
3.1 CPU Performance Testing	3
3.2 Data Throughput Testing	3
4. Future Research	4
5. Conclusion	5

1. Introduction

Projects such as the Large Hadron Collider (LHC) generate enormous amounts of raw data which presents a serious computing challenge. After planned upgrades in 2022, the data output from the ATLAS Tile Calorimeter (TileCal) will increase by 200 times to over 40 Tb/s (Terabits/s) [1]. It is infeasible to store this data for offline computation.

A paradigm shift is necessary to deal with these future workloads and the cost, energy efficiency, processing performance and I/O throughput of the computing system to achieve this task are vitally important to the success of future big science projects. Current x86-based microprocessors such as those commonly found in personal computers and servers are biased towards processing performance and not I/O throughput and are therefore less-suitable for high data throughput applications otherwise known as High Volume throughput Computing (HVC) [2].

ARM System on Chips (SoCs) are found in almost all mobile devices due to their low energy consumption, high performance and low cost [3]. One of the first steps to a true HVC system is a high data throughput Processing Unit (PU). The authors are developing an ARM-based PU for potential use by ATLAS TileCal as a high data throughput, general purpose co-processor to the read-out system Super Read Out Driver prototype (sROD) which can be used to combat the issue of pile-up. A general purpose co-processor is able to easily run more sophisticated and memory intensive algorithms than FPGA-based devices such as the sROD, although the jitter is inferior which is why FPGAs are typically used in the data path.

A brief discussion of the ATLAS TileCal read out architecture and the sROD prototype is given in Section 2. The Processing Unit (PU) is described in Section 3. Future research is briefly described in section 4 and section 5 concludes.

Figure 1. ATLAS TileCal current read out architecture [1].

Figure 2. ATLAS TileCal upgraded read out architecture showing the sROD [1].

2. TileCal Read Out Architecture

The current ATLAS Trigger and Data Acquisition System, shown in figure 1, must be upgraded in order to select interesting events at much higher data rates as described in section 1. For the upgrade, the triggering logic will be moved off the detector and the front-end will be replaced by fast analog to fibre-optic digitisers. A diagram of the upgraded system is shown in figure 2.

The sROD is located in the back-end, off the detector to avoid the requirement for expensive radiation-hard electronics as well as easing maintenance. The sROD prototype will be located in an industry standard AdvancedTCA (ATCA) chassis with Advanced Mezzanine Card (AMC) form-factor which enables comprehensive redundancy and monitoring to ensure maximum uptime.

In both the existing and the upgraded systems, a pipeline is used to store events until the level one trigger provides an accept signal. This short delay is required while the level one trigger performs computations. In the upgraded system the sROD will perform some calculations such as Optimal Filtering, before sending data to the rest of the triggering and data acquisition system [1].

A general purpose Processing Unit can be used to enhance as well as provide new functionality that is difficult to implement on FPGA.

3. General Purpose Processing Unit

ARM System on Chips (SoCs) are low cost, energy efficient and high performance which has led to their extensive use in mobile devices. Testing has been done and a summary of ARM performance and energy efficiency results is in section 3.1.

The completed PU will be located in the ATCA chassis on an AMC next to the sROD prototype or as a separate board connected to the back-plane. The PU will be able to process at least 40 Gb/s raw data, fed through the ATCA carrier board from the sROD prototype. A PCI-Express I/O interface will be used to link the FPGA on the sROD to a cluster of ARM SoCs on the PU. Figure 3 illustrates this connection. The PU will be flexible in that data can be fed via XAUI or bonded SFP+ connectors if PCI-Express is not suitable. This allows for generic operation in other environments.

Figure 3. Schematic of the PU and SROD prototype connection.

PCI-Express is one of the few viable alternatives for high data throughput I/O on a commodity SoC. Parallel Gigabit Ethernet connections would require too many SoCs and would defeat the power efficiency and cost effectiveness of the solution. USB 3.0 is an option that provides up to 5 Gb/s of I/O but there is significant protocol overhead and the power efficiency is not as good as PCI-Express due to the fact that it is intended to work over long cables. Section 3.2 provides further details of PCI-Express testing that has been done.

3.1 CPU Performance Testing

Four different ARM SoCs have been thoroughly tested to ascertain CPU and memory performance characteristics. Table 1 summarises the specifications and results of an ARM Cortex-A7, A9 and two A15 SoCs.

Although no comparison to Intel x86 CPUs is presented, the ARM Cortex-A9 has similar performance to an Intel Atom N2800 but double the energy efficiency [4].

Each generation of ARM SoC sees a significant increase in performance and energy efficiency making their future use a good candidate for a compact and energy efficient system.

3.2 Data Throughput Testing

PCI-Express throughput tests have been performed on a pair of Freescale i.MX6 quad-core ARM Cortex-A9 SoCs clocked at 1 GHz, located on Wandboard development boards [5]. The results are presented in table 2 and a photo of the custom test setup designed by the author is in figure 4.

Three tests were run to ascertain the maximum data throughput that can be obtained from the i.MX6 SoC: a simple CPU based memcpy command and two Direct Memory Access (DMA) transfers, initiated by the Endpoint (EP) and the Root Complex (RC).

The theoretical maximum throughput for the PCI-Express Gen 2 x1 link that was used is 500 MB/s. The best result is using DMA initiated by the RC but it is only 72% of the theoretical maximum. The RC-mode drivers are more optimized than the EP-mode drivers due to limited manufacturer support for EP-mode. The read results are lower than write because of overheads to initiate the read. The PU architecture will take these differences into account and use a data push rather than a pull based approach.

Table 1. Summary of performance and specifications for four modern ARM SoCs.

ARM Cortex	-A7	-A9	-A15	-A15
Core Revision	r0p4	r2p2	r3p2	r3p3
SoC	AllWinner A20	Freescall i.MX6Q	Samsung 5410	NVIDIA TK1
Clock (MHz)	1008	996	1000	1090
Cores	2	4	4	4
Feature Size (nm)	40	40	28	28
SP GFLOPS	1.76	5.12	10.56	11.70
DP GFLOPS	0.70	2.40	6.04	6.26
CoreMark	4858	11327	14994	16689
GPGPU GFLOPS	-	-	-	350
Load Power (W)	2.85	5.03	7.48	6.11
DP GFLOPS/W	0.42	0.80	1.10	1.63
RAM [MB/s]	1996	1448	6114	6878
Ethernet (Mb/s)	100	1000	100	1000
PCIe (Gb/s)	-	5	-	40

Table 2. PCI-Express throughput results of an i.MX6 pair.

	CPU memcpy	DMA (EP)	DMA (RC)
Read (MB/s)	$94.8 \pm 1.1\%$	$174.1 \pm 0.3\%$	$236.4 \pm 0.2\%$
Write (MB/s)	$283.3 \pm 0.3\%$	$352.2 \pm 0.3\%$	$357.9 \pm 0.4\%$

Figure 4. PCI-Express test setup for a pair of i.MX6 SoCs situated on Wandboards.

4. Future Research

A Wandboard cluster via PCI-Express is under development. Figure 5 shows two adapters that have been designed and manufactured to enable Wandboards to connect to a standard PCI-Express x1 connector. A PCIe switch development board will be used to connect up to 8 of these devices to each other. A custom Linux device driver is being implemented to encapsulate standard Ethernet packets and transmit them over PCI-Express. This will allow existing applications to use the high data throughput PCIe-based cluster without modification.

There are several new ARMv8 cores (Cortex-A5x) being released with 64-bit capabilities which enables higher performance. It is important to benchmark and characterise these for the purposes of scientific computing.

Figure 5. Two Wandboard to PCI-Express x1 adapters.

Other low power and cost-effective platforms will be considered in future, such as Intel Atom. Previous generations of Atom are inferior to ARM but new, as yet unavailable, Atom SoCs will be tested when they become available.

ARM SoCs also have potential application in higher level triggering and reconstruction systems. The synthetic performance benchmarks indicate that ARM SoC performance is similar to the Intel and AMD CPUs currently in use in the Event Builder, for example [6]. In order to pursue this line of investigation, the ATLAS software would have to be recompiled for the ARM instruction set which will be a significant and challenging undertaking.

5. Conclusion

A massively parallel computing system is required to handle the digital data processing requirements of future big science projects such as the Large Hadron Collider Phase-II upgrades at CERN. The ATLAS Trigger and Data Acquisition system has several levels of data processing beginning with the front-end which is dominated by FPGA technology. At higher levels where the volume of data has been reduced, conventional server-grade hardware is typically used.

Commodity ARM SoCs have good performance and excellent power efficiency as shown by synthetic benchmarks. By using PCI-Express as a high data throughput I/O interconnect, the otherwise low I/O bandwidth of commodity ARM SoCs is addressed. This combination of performance, power efficiency and I/O throughput makes ARM SoCs a candidate for future high data throughput workloads.

An ARM-based Processing Unit high level design is proposed which can be installed and used nearer the detector front-end than conventional servers could. Since the PU is programmable in programming languages such as C or C++, complex algorithms for additional processing or supervisory are made possible.

The high-level design presented includes a method to connect to the TileCal sROD prototype for testing on raw data from the detector. PCI-Express testing of a pair of ARM Cortex-A9 SoCs (Freescale i.MX6Q) has been done and data throughputs of 357 MB/s have been achieved which is superior to gigabit Ethernet. A PCIe-based cluster is being built as a PU proof of concept: the hardware is complete and the Linux device drivers are in progress.

Acknowledgments

The financial assistance of the South African National Research Foundation (NRF) towards this research is hereby acknowledged. We would also like to acknowledge the School of Physics, the Faculty of Science and the Research Office at the University of the Witwatersrand, Johannesburg.

References

- [1] F. Carrió et al., *The sROD module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator*, *Journal of Instrumentation* **9** (Feb., 2014) C02019–C02019.
- [2] J. Zhan et al., *High Volume Throughput Computing: Identifying and Characterizing Throughput Oriented Workloads in Data Centers*, in *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*, pp. 1712–1721, IEEE, May, 2012.
- [3] T. Krazit, *ARMed for the living room*, 2006.
- [4] EEMBC, *CoreMark Online Database*, 2014.
- [5] Wandboard.org, *Wandboard - Freescale i.MX6 ARM Cortex-A9 Opensource Community Development Board*, 2012.
- [6] H. P. Beck et al., *Performance of the Final Event Builder for the ATLAS Experiment*, *IEEE Transactions on Nuclear Science* **55** (2008), no. 1 176–181.