Parallel Patterns using Heterogeneous Computing

Mr Andreas Vermeulen
University of St Andrews
Saint Andrews, Fife KY16 9AJ
University of Dundee
Nethergate,Dundee DD1 4HN
a.f.vermeulen@dundee.ac.uk

Dr Vladimir Janjic
University of St Andrews
Saint Andrews, Fife KY16 9AJ
vj32@st-andrews.ac.uk

Mr Andy Cobley
University of Dundee
Nethergate, Dundee DD1 4HN
acobley@computing.dundee.ac.uk

ABSTRACT
An enhancement of a Research Information Factory using heterogeneous computing and parallel knowledge-extraction patterns.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous

General Terms
Theory, Framework, Application, Research, Hardware

Keywords
knowledge-extraction, patterns, information factory, RIF, RIFF, RIFC, heterogeneous computing, parallel patterns, cassandra, spark, opencl, fastflow, cuda, 3D torus network

1. INTRODUCTION
The increasing demand for data into knowledge conversion requires more volume, variety, velocity and veracity [8] in processing solutions with energy and natural resources requirements that are undesirable.

Research goal is to develop effective processing patterns with less overall energy cost.

2. BACKGROUND
Heterogeneous computing systems [8] uses central processing units (*CPU*), graphical processing units (*GPU*) and field programmable gate arrays (*FPGA*) to enable low energy processing.

Parallel Patterns are common libraries like (CUDA [11], OpenCL [9, 5, 13, 14], FastFlow [2] and ZeroMQ [6]) to guide the processing.

Efficiency and Energy-awareness is controlling efficiency of processing [15] in Floating-point Operations Per Second per Watt (*FLOP/S/W*) to achive energy requirements adviced by new euopean energy laws.

3. PROPOSED SOLUTION
The research uses parallel patterns for knowledge extraction, mechanisms for storing and extracting data while using minimum amounts of energy. It covers three basic stages:

3.1 Heterogeneous systems.
The research will study the fundamental behavior of heterogeneous computing components using a nVidia Jetson TK1 development kit. [12] and Tilera TILE-Mx100 processor [10].

3.2 Research Information Factory Framework
The framework (*RIFF*) [1] uses a parallel processing pattern via **R**etrieve-**A**ssess-**P**rocess-**T**ransform-**O**rganise-**R**eport rules.

3.3 Research Information Factory Cluster
The cluster (*RIFC*) is a 3D torus appliance [1] using Cassandra database [3, 4] and Spark Engine [4, 7] for data processing.

4. CONCLUSION
Heterogeneous systems with parallel patterns is the optimum option to achieve the goal. The Research Information Framework is a new set of guidelines to achieve the research goal is to process with less energy.

5. REFERENCES
[1] Yuichiro Ajima, Shinji Sumimoto, and Toshiyuki Shimizu. Tofu: A 6d mesh/torus interconnect for exascale computers. *Computer*, (11):36–40, 2009.
[2] Marco Aldinucci, Marco Danelutto, Peter Kilpatrick, and Massimo Torquati. Fastflow: high-level and efficient streaming on multi-core.(a fastflow short tutorial). *Programming multi-core and many-core computing systems, parallel and distributed computing*, 2011.
[3] Cassandra. Apache cassandra.
[4] Datastax. Getting started with apache spark and cassandra.
[5] Khronos OpenCL Working Group et al. The opencl specification, version 1.2, 15 november 2011. *Cited on pages*, 18(7):30.
[6] Pieter Hintjens. Ømq-the guide. *Online: http://zguide. zeromq. org/page: all, Accessed on*, 23, 2011.
[7] Apache Incubator. Spark: Lightning-fast cluster computing, 2013.
[8] Ashfaq A Khokhar, Viktor K Prasanna, Muhammad E Shaaban, and Cho-Li Wang. Heterogeneous computing: Challenges and opportunities. *Computer*, 26(6):18–27, 1993.
[9] Khronos. Opencl.
[10] Timothy G Mattson, Rob Van der Wijngaart, and Michael Frumkin. Programming the intel 80-core network-on-a-chip terascale processor. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, page 38. IEEE Press, 2008.
[11] nVidia. Cuda toolkit.
[12] nVidia. The world's first embedded supercomputer.
[13] Kavya Subraya Shagrithaya. Enabling development of opencl applications on fpga platforms. 2012.
[14] John E Stone, David Gohara, and Guochun Shi. Opencl: A parallel programming standard for heterogeneous computing systems. *Computing in science & engineering*, 12(1-3):66–73, 2010.
[15] Qiang Wu, Yajun Ha, Akash Kumar, Shaobo Luo, Ang Li, and Shihab Mohamed. A heterogeneous platform with gpu and fpga for power efficient high performance computing. In *Integrated Circuits (ISIC), 2014 14th International Symposium on*, pages 220–223. IEEE, 2014.