

Institution:	Imperial College London
Unit of Assessment:	13A – Electrical & Electronic Engineering
Title of case study:	Case 4 - Reconfigurable Computing for High Performance Applications
<p>1. Summary of the impact</p> <p>In the last 20 years, reconfigurable technology has transformed High-Performance Computing and Embedded Systems Design. Research of the Custom Computing and Reconfigurable Systems teams at Imperial made pivotal contributions to this transformation, targeting particularly Field-Programmable Gate Array (FPGA) technology. Since 2008, the impact of this research has been to</p> <ol style="list-style-type: none"> 1) underpin design flow for partial run-time reconfigurable designs for Xilinx FPGA devices; 2) contribute to the start-up company Maxeler, pioneering reconfigurable computing systems and cloud services for high-performance computing in the financial and other sectors; 3) enable near real-time risk analysis for JP Morgan's global portfolio to analyse and manage risk much faster than previously possible; 4) achieve about 250 times speedup for Chevron's seismic modelling for oil and gas exploration, compared to the alternative use of CPU-based machines; 5) accelerate a financial market integrity platform for BlueBee and HL Steam in hardware. 	
<p>2. Underpinning research</p> <p>Field-Programmable Gate Arrays (FPGAs) are a technology introduced in the 80's. They provide an affordable vehicle for hardware acceleration of sophisticated software. The emergence of FPGAs has enabled novel computing engines for solving complex scientific and commercial problems. Research groups in Department of Electrical & Electronic Engineering (EEE) and the Department of Computing (DoC) conducted the underpinning research for the above impacts (I1-I5). The group in EEE, consisting of Prof. Peter Cheung (lead), Prof. George Constantinides and Dr. David Thomas, focused mainly on hardware architecture, synthesis algorithms and applications in financial modelling and digital signal processing. The DoC group, consisting of Prof. Wayne Luk (lead) and Dr. Oskar Mencer, focused mainly on application-specific programming languages, compilers, run-time reconfigurability and system aspects. These research groups and their collaborations have produced significant research advances that underpin the above impacts, clustered in five topics (R1-R5) and mapped to references and the specific impacts:</p> <p>R1 Design of FPGA-based reconfigurable computing systems [1,2; I1]: FPGA-based computing systems provide the flexibility of run-time reconfigurability. However, this advantage also adds a new dimension to the complexity of system design, especially when the FPGA device supports partial run-time reconfiguration. Since 1995, our teams have developed new automatic design tools that efficiently exploit the dynamic reconfigurable nature of FPGAs in producing novel implementations of computing systems. For specific classes of devices, the techniques we developed reduce reconfiguration time from time linear in their size to constant time at best and logarithmic time at worst. This work has resulted in two granted patents (US6369610, US7543283).</p> <p>R2 FPGA-specific compilation and synthesis [3,4; I2-I4]: This stream compiler provides a software-like programming interface to hardware design; it targets FPGAs but maintains the performance of hand-designed circuits. This compiler and interface improve productivity by letting programmers optimize implementations at <i>multiple</i> levels – algorithms, architecture, arithmetic, and gates – and all within the <i>same</i> C++ program. This increased productivity is demonstrated in hardware acceleration of a range of applications, e.g. encryption with Kasumi and IDEA, function evaluation, and Wavelet and LZ-like compression.</p>	

- R3 Optimizing word-length and data for area, power, performance, and accuracy trade-offs [5; I2-I5]:** In conventional computers, the word-length (e.g. in a 32-bit or 64-bit ALU) and data format (e.g. fixed or floating point) are fixed by the architecture. FPGAs provide the freedom of optimally determining word-length and data format according to the needs of specific applications. Based on an approach that combined analytical and heuristic methods, our teams pioneered – since 1999 – a number of techniques for determining word-lengths and data formats that will both optimize area, performance and power consumption for a given system-level specification (e.g. worst-case accuracy and signal-to-noise ratio). These techniques also allow us to trade off the accuracy of computations with other desired performance characteristics (e.g. performance).
- R4 Transferring reconfigurable FPGA-based technology into applications [2-5; I3,I4]:** Since 1995, our teams have researched how the aforementioned new techniques can be applied to solve industrial and commercial problems. Our research has demonstrated the applicability of these techniques in various industrial applications domains such as financial analysis, seismic modelling and digital signal processing. In particular, we are responsible for some of the earliest research on accelerating financial applications, which produced a powerful mathematical framework for rapid execution of a wide class of Monte-Carlo simulations. In the framework, Monte-Carlo applications can be written in a high-level language, which a streaming compiler such as ours can nonetheless convert into high-performance data pipelines. Our research teams also identified generic design templates that can help reduce design efforts of future systems in these application domains.
- R5 Novel tools for heterogeneous multiprocessor systems [6; I5]:** We are among the first (starting this research in 2006) to develop a compilation tool chain for high-level programs targeting heterogeneous systems with different types of processing elements such as general-purpose processors, GPUs, and FPGAs. The core of these tools includes a task transformation engine, a mapping selector, an optimizer for data representations, and a hardware synthesizer. The tool chain uses as intermediate representation C programs enriched with source annotations, thus making it easy for users to comprehend and control the compilation process.

3. References to the research (*References that best indicate quality of underpinning research.)

- 1.* **W. Luk**, N. Shirazi and **P.Y.K. Cheung**, “*Compilation tools for run-time reconfigurable designs*”, IEEE Symposium on FPGAs for Custom Computing Machines (FCCM), (1997), pp.56-65, doi: 10.1109/FPGA.1997.624605.
- 2.* S.D. Haynes, J. Stone, **P.Y.K. Cheung**, **W. Luk**, “*Video image processing with the Sonic architecture*”, IEEE Computer, (2000), Vol. 33 pp.50-57, doi: 10.1109/2.839321.
3. **O. Mencer**, D.J. Pearce, L.W. Howes and **W. Luk**, “*Design space exploration with A Stream Compiler*”, IEEE International Conference on Field Programmable Technology (FPT), (2003), pp. 270-277, doi: 10.1109/FPT.2003.1275757.
4. J.A. Bower, **D.B. Thomas**, **W. Luk** and **O. Mencer**, “*A reconfigurable simulation framework for financial computation*”, International Conference on Reconfigurable Computing and FPGAs (ReConFig), (2006), pp.1-9, doi: 10.1109/RECONF.2006.307750.
- 5.* **G. Constantinides**, **P.Y.K. Cheung** and **W. Luk**, “*Optimum and heuristic synthesis of multiple word-length architectures*”, IEEE Transactions on VLSI Systems, 13 (1), 2005, pp.39-57, doi: 10.1109/TVLSI.2004.840398.
6. B. Cope, **P.Y.K. Cheung**, **W. Luk**, “*Bridging the Gap between FPGAs and Multi-Processor Architectures: A Video Processing Perspective*”, IEEE Int. conf. on Application-specific Systems, Architectures and Processors, (2007), pp.308-313, doi: 10.1109/ASAP.2007.4429998.

4. Details of the impact

We now provide details of the five aforementioned impacts and their link to underpinning research:

- 11) Xilinx Corp. is a US\$2.2 billion company (2012) with nearly a 50% share of the FPGA devices market. The impact of our research in FPGA technology is evidenced in their testimony (Dr. G. Brebner, Distinguished Engineer from Xilinx [E1]):

“The research group led by you (Prof. Luk) in Department of Computing and Professor Peter Cheung in Department of EEE ... is second-to-none internationally, both in terms of quality and quantity. Your team has pioneered many key technologies which have resulted ... in significant industrial impact.”

For example, Dr. Brebner refers specifically to our work on partial run-time reconfiguration [1], which “... has had a major impact on multiple generations of Xilinx devices that support run-time reconfiguration”, and “..... underpins the design flow for the latest Xilinx device supporting partial run-time reconfiguration”. The specific devices introduced since 2008 were Virtex-5 (65nm), Virtex-6 (40nm) and Virtex-7 (28nm) devices as well as their latest Zynq device family, which embeds ARM processor cores within the FPGA fabric. Xilinx’s document UG702, entitled “Partial Reconfiguration User Guide” and published in October 2012, demonstrates that our work [R1], first published in 1997, has stood the test of time: its impact is still found in design flow for Xilinx Zynq devices introduced in 2011 [E1].

- 12) Our underpinning research into acceleration of compute-intensive algorithms using FPGA technology has led to its successful commercial exploitation in the start-up company, Maxeler. The company, with Head Office in Hammersmith, employs 70 members of staff in the UK and US. Many of the key technologies it employs can be traced directly to the research conducted by the team at Imperial College [3,4]. In particular, the MaxCompiler and MaxGen tools are rooted in Imperial tools. These tools enable MaxCloud, the industry’s first FPGA-based cloud computing service, to provide high performance dataflow computation capability through the cloud [E2]. Underpinning research by Cheung and Luk also led to several US patents (US6369610, US7543283, US12/747650) assigned to Maxeler on 2 Nov 2012. Moreover, Maxeler US patent US20130139122 A1 cited their work on word-length optimisation (R5).
- 13) The FPGA-based technology from Maxeler is used by JP Morgan [E3] and by other finance companies such as Scottish Widows [E4] for accelerating a variety of financial modelling calculations including risk analysis. This is therefore direct impact of our research in [2-5]. In one example, the compute time was reduced from 8 hours to less than 4 minutes [E5]. The effectiveness of this technology resulted in JP Morgan purchasing a 20% stake in Maxeler [E6]. According to JP Morgan’s Managing Director and Global Head of their Applied Analytics Group, the total investment made by JP Morgan in deploying our reconfigurable technology amounts to US\$30m with an annual estimated saving in running cost of US\$6m, in addition to compute time reduction. Furthermore, the reconfigurable computing systems “support our business units that manage the risk on positions in the trillions of dollars” [E7].
- 14) In addition to financial applications, our reconfigurable computing technology [2-5] has had a strong impact on oil and gas exploration. The oil and gas industry is a major user of high-performance computing. In geoscience, computational cycles are dominated by relatively few and well defined kernels. Using Maxeler’s FPGA-based hardware platform and optimising the algorithm implemented with the Maxeler tool flow, speedups of almost 250 times compared to the use of a single CPU core have been reported by Chevron [E8]. Faster seismic analysis and imaging allows more application runs or modelling with higher fidelity, or both. Quicker and more accurate results enable oil and gas companies to make more informed bids on parcels and subsequent drilling. Shorter turnaround time on these seismic applications is critical to the company’s profit, especially when they are in a bidding process for drilling rights.

- 15) BlueBee Technologies is a start-up company that provides tools for heterogeneous multi-core platforms. The success of its tools was recognised by a prestigious Valorisation award from the Dutch Technology Foundation, STW, in December 2012. According to BlueBee's CEO [E9], several key components in the BlueBee tool chain – such as task transformation and mapping selection – are directly based on our research on high-level compilation tool chains for heterogeneous systems [6]; success of BlueBee can be attributed to our pioneering research which laid the foundations for many important developments in word-length optimisation [3] and high-level design transformation [6]. Recently BlueBee has teamed up with HL Steam, a company responsible for the Ancoa financial market integrity platform, to support FPGA-based acceleration for such platforms. The impact of our research is therefore also ongoing: the enabling of real-world workloads of millions of financial transactions per second to be efficiently processed at lower dollar and energy cost [E10].

5. Sources to corroborate the impact

- E1. Letter by Distinguished Engineer, Xilinx Research Labs, (21/5/13) stating the impact of our research on FPGA technology, particularly on Xilinx products.
- E2. "Maxeler launches Maximum Performance Cloud Computing with Maxcloud", 20/9/11: <http://www.maxeler.com/maxeler-launches-maximum-performance-cloud-computing-with-maxcloud/> or <https://www.imperial.ac.uk/ref/webarchive/nwf>.
- E3. S. Shah, "JP Morgan goes live with Maxeler supercomputer", 20/12/11, (access on 26/4/2013): <http://www.computing.co.uk/ctg/news/2133724/jp-morgan-goes-live-maxeler-supercomputer> or <https://www.imperial.ac.uk/ref/webarchive/pjf>.
(A news article reporting JP Morgan's deployment of Maxeler's FPGA-based high performance supercomputer for risk evaluation.)
- E4. O. Mencer, E. Vynckier, J. Spooner, S. Girdlestone and O. Charlesworth, "Finding the right level of abstraction for minimizing operational expenditure", Workshop on High Performance Computational finance, (2011), pp.13-18, <http://dl.acm.org/citation.cfm?id=2088262>.
(A publication by Scottish Widows and Maxeler on the effectiveness of their HPC on financial computing by Scottish Widows.)
- E5. S. Weston, J. Spooner, S. Racaniere and O. Mencer, "Rapid Computation of Value and Risk for Derivatives Portfolios", Concurrency and Computation: Practice and Experience, 24 (8), 10 June 2012, pp. 880-894, DOI: 10.1002/cpe.1778.
(A publication by JP Morgan and Maxeler on using FPGA-based HPC in 2009 and 2010 on risk computation with 30x performance acceleration.)
- E6. "Maxeler Technologies Sells 20 Percent Stake in Company to J.P. Morgan", March 31, 2011 (accessed on 26/4/13): <http://www.hpcwire.com/hpcwire/2011-03-31/maxeler-technologies-sells-20-percent-stake-in-company-to-j-p-morgan.html> or <https://www.imperial.ac.uk/ref/webarchive/qjf>
(A news article reporting JP Morgan taking a 20% stake in Maxeler.)
- E7. Managing Director and Global Head of Applied Analytics Group at J.P. Morgan (29 May 2012). *A letter from JP Morgan stating the impact of our reconfigurable research on J.P. Morgan's business with quantitative estimates.*
- E8. T. Nemeth, J. Stefani, W. Liu, R. Dimond, O. Pell and R. Ergas, "An Implementation of the Acoustic Wave Equation on FPGAs", 78th Society of Exploration Geophysicists (SEG) Annual Meeting, November 2008.
<http://www.maxeler.com/media/documents/MaxelerSummaryAcousticWaveEquation.pdf>.
 Archived [here](#) on 23/10/2013.
- E9. Letter by CEO of BlueBee Technologies, (27 Jan 2013) stating the impact of our research on their products. (Email: Koen.bertels@bluebee-tech.com).
- E10. News and views from the HL Steam team, "HL Steam teams up with BlueBee for hardware acceleration", 16 November 2012. <http://hlsteam.com/hl-steam-teams-up-with-bluebee-for-hardware-acceleration> or <https://www.imperial.ac.uk/ref/webarchive/vqf>.