D7.2

Dissemination and Exploitation Activities report 1

Daniel Strimpel

Dissemination and Exploitation Activities report 1

DOCUMENT IDENTIFIER:	D7.2
DUE DATE:	31/12/2018
DELIVERY DATE:	05/02/2019
CLASSIFICATION:	Confidential
EDITORS:	Daniel Strimpel
DOCUMENT VERSION:	1.0

CONTRACT START DATE:	01/01/2018
CONTRACT DURATION:	36 months

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

CONTRIBUTORS	
Name	**Organisation**
Daniel Strimpel	iProov
Kostas Akasoglou	EXUS
Yannis Katsaros	EXUS
Charalampos Psaropoulos	EXUS
Ioannis Komnios	EXUS
Georgios Mylonas	CTI
Christos Kotselidis	UNIMAN
Juan Fumero	UNIMAN
Foivos Zakkak	UNIMAN
Vassilis Spitadakis	NEUROCOM
Ioannis Konstantinou	ICCS
Katerina (Aikaterini) Doka	ICCS
Nikos Kanakis	SPARKWORKS
Viktor Rosenfeld	DFKI
Sebastian Breß	DFKI
Hazeef Mohammed	KALEAO
Monica Vaterroni	KALEAO

PEER REVIEWERS	
Name	**Organisation**
Ioannis Komnios	EXUS
Giorgos Mylonas	CTI

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

REVISION HISTORY		
Version	**Date**	**Modifications**
0.1	23/01/2019	Template
0.1	24/01/2019	Full draft sent to reviewers
1.0	05/02/2019	Final version for submission

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

Executive Summary

This report summarises the diverse dissemination activities carried out by the E2Data partners during the first 12 months of the project. These activities included:

a) Publication and presentation of research papers at conferences and workshops;

b) Poster presentations at conferences;

c) Publication of journal papers;

b) Presentations of the E2Data project;

d) Organisation of an academic workshop;

e) Development and maintenance of an online presence;

f) Participation in the Big Data Value Public Private Partnership (BDV PPP);

g) Release of the open source Tornado code.

All partners have been active in the dissemination of the project's aims and intermediate results. Our approach is in line with the dissemination strategy outlined in Work Package 7 of the E2Data Grant Agreement.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Table of Contents

Executive Summary 5
1. Introduction 8
2. Publication and Presentation of Research Papers at Conferences and Workshops 10
2.1 MoreVMs Workshop 10
2.2 14th ACM International Workshop on Data Management on New Hardware (DaMoN'18), colocated with SIGMOD/PODS 11
2.3 21st International Conference on Extending Database Technology (EDBT'18) 12
2.4 38th IEEE International Conference on Distributed Computing Systems (ICDCS 2018) 13
2.5 HIPEAC ACACES 2018 14
2.6 HIPEAC ACACES 2018 15
2.7 Managed Runtime Languages (Manlang 2018) 16
2.8 Computer Systems Week (HIPEAC#2018) 17
2.9 Virtual Machines and Language Implementation (VMIL#18) 18
2.10 Xtreme Cloud International Workshop 2018 19
2.11 1ST International Workshop On Next Generation Clouds For Extreme Data Analytics (XtremeCLOUD 2018) 20
2.12 10th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2018) 21
3. Poster Presentations 22
3.1 Poster at HIPEAC ACACES 2018 22
3.2 Poster at HIPEAC ACACES 2018 23
3.3 Poster Presentation at SPLASH#2018 Conference 24
3.4 Poster Presentation at Google Compiler Summit 2018 25
3.5 Poster Presentation at Google Compiler Summit 2018 25
4. Journal Publications 26
4.1 Very Large Databases (VLDB) 26
4.2 Datenbanken Spektrum 27
5. Presentations of the E2Data project 28
5.1 Heterogeneity Alliance at Hipeac '18, Manchester, UK 28
5.2 3rd BDV PPP Steering Committee Meeting, Brussels, Belgium 29
5.3 Elec.Tec 2018 exhibition, Athens, Greece 30

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

5.4 HiPEAC CSW Autumn 2018 31
6. Conference Support and Organisation 32
6.1 Workshop on Next Generation Clouds for Extreme Data Analytics 32
7. Online Presence 35
7.1 Website 35
7.2 Tornado Github Public Repository 35
8. Big Data Value Public-Private Partnership 36
9. Conclusions 37

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

1. Introduction

In the scope of Work Package 7 “Dissemination, Communication and Exploitation”, the E2Data Consortium is conducting specific dissemination activities, following the rationale presented in E2Data Grant Agreement. The overall goal of these dissemination activities is to ensure the visibility of the project and maximise its impact in the European ecosystem (researchers, industry, societies etc.) and beyond. The project dissemination activities are organised according to the following areas:

- Publication and Presentation of Research Papers at Conferences and Workshops;
- Poster Presentations;
- Journal Publications;
- Presentations of the E2Data project;
- Organisation of a workshop;
- Online Presence and
- Releasing open source code.

In addition, the E2Data project joined the Big Data Value Public-Private Partnership (BDV PPP). The BDV PPP aims at creating a functional Data Market and Data Economy in Europe, in order to allow Europe to play a leading role in Big Data in the global market. Ioannis Komnios, the E2Data project coordinator, and Christos Kotselidis, the E2Data technical coordinator, joined the BDV PPP steering committee.

Overall, during course of the first 12 months of the project, the consortium focused especially on dissemination via presentations, whether given to small groups in poster sessions, or to large numbers of conference attendees. These presentations were usually based on original research by the presenter, which had also been published in scholarly journals.

As part of E2Data’s online presence, the Consortium has developed and maintains a website presenting the rationale of the project, as well as news and activities. It maintains visibility in social media such as Twitter, aiming to raise awareness not only of E2Data research but also of the real-world possibilities that flow out of that research.

Table 1 below summarises the dissemination activities of the E2Data project during the first 12 months of the project.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Type of Activity	Extent/Details of Involvement
Publication with Presentation at Events	12 conference publications
Poster Sessions	5 poster presentations
Journal Publications	2 journal articles published
Presentations of the E2Data Project	4 presentations
Conference Support and Organisation	Cloud Computing workshop with over 40 participants
Online Presence	Complete redesign of the E2Data website
Big Data Value Public-Private Partnership	Participation at three separate events in Belgium and Bulgaria

TABLE 1: SUMMARY OF ACTIVITIES

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2. Publication and Presentation of Research Papers at Conferences and Workshops

2.1 MoreVMs Workshop

Event Website: https://2018.programming-conference.org/track/MoreVMs-2018

Dates: 09/04/2018

Presenter: Juan Fumero, UNIMAN

Presentation Title: Towards Practical Heterogeneous Virtual Machines

Description:

Heterogeneous computing emerged as a means to achieve higher performance and energy efficiency. However, this trend has been accompanied by changes in software development norms that do not necessarily favour programmers. A prime example is the two most popular heterogeneous programming languages, CUDA and OpenCL, which expose several low-level features to the API making them difficult to use by non-expert users.

Instead of using low-level programming languages, developers tend to prefer more high-level, object-oriented languages typically executed on managed runtime environments. Although many programmers might expect that such languages would have already been adapted for execution on heterogeneous hardware, the reality is that their support is either very limited or totally absent. This paper highlights the main reasons and complexities of enabling heterogeneous managed runtime systems and proposes a number of directions to address those challenges.

Link to Resources:

Paper: https://dl.acm.org/citation.cfm?id=3191730

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.2 14th ACM International Workshop on Data Management on New Hardware (DaMoN'18), co-located with SIGMOD/PODS

Event Website: https://sites.google.com/view/damon2018/

Dates: 11/06/2018

Presenter: Clemens Lutz, DFKI

Presentation Title: Efficient k-Means on GPUs

Description:

k-Means is a versatile clustering algorithm widely-used in practice. To cluster large data sets, state-of-the-art implementations use GPUs to shorten the data to knowledge time. These implementations commonly assign points on a GPU and update centroids on a CPU.

We show that this approach has two main drawbacks. First, it separates the two algorithm phases over different processors, which requires an expensive data exchange between devices. Second, even when both phases are computed on the GPU, the same data are read twice per iteration, leading to inefficient use of memory bandwidth.

In this paper, we describe a new approach that executes k-means in a single data pass per iteration. We propose a new algorithm to updates centroids that allows us to perform both phases efficiently on GPUs. Thereby, we remove data transfers within each iteration. We fuse both phases to eliminate artificial synchronization barriers, and thus compute k-means in a single data pass. Overall, we achieve up to 20x higher throughput compared to the state-of-the-art approach.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.3 21st International Conference on Extending Database Technology (EDBT'18)

Event Website: http://edbticdt2018.at/

Dates: 11/06/2018

Presenter: Tobias Behrens, DFKI

Presentation Title: Efficient SIMD Vectorization for Hashing in OpenCL

Description:

Hashing is at the core of many efficient database operators such as hash-based joins and aggregations. Vectorization is a technique that uses Single Instruction Multiple Data (SIMD) instructions to process multiple data elements at once. Applying vectorization to hash tables results in promising speedups for build and probe operations. However, vectorization typically requires intrinsics – low-level APIs in which functions map to processor-specific SIMD instructions. Intrinsics are specific to a processor architecture and result in complex and difficult to maintain code.

OpenCL is a parallel programming framework which provides a higher abstraction level than intrinsics and is portable to different processors. Thus, OpenCL avoids processor dependencies, which results in improved code maintainability. In this paper, we add efficient, vectorized hashing primitives to OpenCL. Our results show that OpenCL-based vectorization is competitive to intrinsics on CPUs but not on Xeon Phi coprocessors

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2.4 38th IEEE International Conference on Distributed Computing Systems (ICDCS 2018)

Event Website: https://icdcs2018.ocg.at

Dates: 02-05/07/2018

Presenter: Katerina Doka, ICCS

Presentation Title: Docker-sec: A Fully Automated Container Security Enhancement Mechanism

Description:

The popularity of containers is constantly rising in the virtualization landscape, since they incur significantly less overhead than Virtual Machines, the traditional hypervisor-based counterparts, while enjoying better performance. However, containers pose significant security challenges due to their direct communication with the host kernel, allowing attackers to break into the host system and co-located containers more easily than Virtual Machines. Existing security hardening mechanisms are based on the enforcement of Mandatory Access Control rules, which exclusively allow specified, desired operations. However, these mechanisms entail explicit knowledge of the container functionality and behavior and require manual intervention and setup.

To overcome these limitations, we present Docker-sec, a user-friendly mechanism for the protection of Docker containers throughout their lifetime via the enforcement of access policies that correspond to the anticipated (and legitimate) activity of the applications they enclose. Docker-sec employs two mechanisms: (a) Upon container creation, it constructs an initial, static set of access rules based on container configuration parameters; (b) During container runtime, the initial set is enhanced with additional rules that further restrict the container's capabilities, reflecting the actual application operations. Through a rich interaction with our system the audience will experience first-hand how Docker-sec can successfully protect containers from zero-day vulnerabilities in an automatic manner, with minimal overhead on the application performance.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2.5 HIPEAC ACACES 2018

Event Website: http://acaces.hipeac.net/2018/

Dates: 04/07/2018

Presenter: Maria Xekalaki, UNIMAN

Presentation Title: Dynamic Acceleration of Big Data Applications on Heterogeneous Hardware Resources

Description:

The rapid development of data intensive applications, such as social networks, IoT, etc., has caused an explosion in Big Data processing. Soon, homogeneous CPU clusters reached their limits, as they could no longer satisfy their clients' requirements, which were to execute their applications within a strict time and budget limit. Consequently, cloud providers shifted to diverse hardware architectures, like GPUs and FPGAs, which can greatly accelerate certain operations. However, exploiting such architectures is not trivial; it requires developers to adapt their current software for each target device, a process that is both inefficient and error prone.

In this paper we present our approach in accelerating Big Data Applications on heterogeneous hardware resources. The framework we propose utilizes Apache Flink, an open source stream and batch processing system, and Tornado, a heterogeneous Virtual Machine, which allows Big Data practitioners to exploit the diverse hardware resources in a transparent, dynamic and efficient way, without requiring any code modifications. Our framework is innovative as it does not rely on precompiled kernels or depend on priori knowledge of the hardware, unlike its equivalent state-of-the-art frameworks.

Links to Resources:

Paper

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.6 HIPEAC ACACES 2018

Event Website: http://acaces.hipeac.net/2018/

Dates: 04/07/2018

Presenter: Michail Papadimitriou, UNIMAN

Presentation Title: Exploiting Programmability of FPGAs Through Managed Runtime Systems

Description:

The slowdown of Moore's law along with the end of Dennard's scaling and the ever-increasing demand for computing power have shown the performance limitations of homogeneous systems. To address this issue, computer architects have taken advantage of the recent technological advancements in order to come up with heterogeneous solutions where the extensive use of special purpose processing units brings the promise of further improvements both in terms of performance and power efficiency. However, this new reality imposes new challenges both in terms of manageability and programmability. System providers face difficulties to integrate the growing number of accelerators into their existing systems, while application developers cannot harness the full potential of the hardware due to the variety of programming models. For example, GPU hardware is SIMD, while FPGAs are non-Von Neumman, while OpenCL attempts to close this gap and provide a unified development environment. However, this is not yet feasible as deep understanding of the hardware is required to develop optimized solutions.

In this work, we tackle the aforementioned issue related to the programmability of FPGAs. We propose an extension to a virtual machine framework (TornadoVM), which transparently compiles and executes Java programs onto GPUs, to also support FPGAs. In more detail, we generate FPGA executable code with the use of OpenCL, high-level-synthesis (HLS) software and by applying compiler techniques in managed runtime.

Link to Resources:

Paper

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.7 Managed Runtime Languages (Manlang 2018)

Event Website: http://ssw.jku.at/manlang18/

Dates: 12/09/2018

Presenter: Juan Fumero, UNIMAN

Presentation Title: Exploiting high-performance heterogeneous hardware for Java programs using graal.

Description :

The proliferation of heterogeneous hardware in recent years means that every system we program is likely to include a mix of compute elements; each with different characteristics. By utilizing these available hardware resources, developers can improve the performance and energy efficiency of their applications. However, existing tools for heterogeneous programming neglect developers who do not have the time or inclination to switch programming languages or learn the intricacies of a specific piece of hardware.

This paper presents a framework that enables Java applications to be deployed across a variety of heterogeneous systems while exploiting any available multi- or many-core processor. The novel aspect of our approach is that it does not require any a priori knowledge of the hardware, or for the developer to worry about managing disparate memory spaces. Java applications are transparently compiled and optimized for the hardware at run-time.

We also present a performance evaluation of our just-in-time (JIT) compiler using a framework to accelerate SLAM, a complex computer vision application entirely written in Java. We show that we can accelerate SLAM up to 150x compared to the Java reference implementation, rendering 107 frames per second (FPS).

Links to Resources:

Paper: https://dl.acm.org/citation.cfm?id=3237016

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2.8 Computer Systems Week (HIPEAC#2018)

Event Website: on hipeac.net

Dates: 30/10/2018

Presenter: Juan Fumero, UNIMAN

Presentation Title: On the Road to Heterogeneous VMs for Distributed Systems

Description:

With the proliferation of the big data analysis and machine learning computation for many of the industries and researchers, there is a need for processing and analyzing data as efficient and fast as possible. To make big data analysis feasible, researchers and industry make use of big data centers and computing clusters that enable parallel processing. During the recent years, data centers started including servers with different types of processors such as GPUs and FPGAs for accelerating and increasing the power efficiency of these big data computations. However, many programming frameworks such as Apache Flink and Hadoop, cannot use directly heterogeneous resources without using wrappers or additional low-level libraries, requiring deep-knowledge and understanding of parallel architectures and programming models. In this talk, we present an alternative to execute heterogeneous big data applications on heterogeneous clusters by enabling automatic GPU and FPGA just-in-time compilation of high-level programming languages.

Links to Resources: Talk

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2.9 Virtual Machines and Language Implementation (VMIL#18)

Event Website: https://2018.splashcon.org/track/vmil-2018

Dates: 04/11/2018

Presenter: Juan Fumero, UNIMAN

Presentation Title: Using Compiler Snippets to Exploit Parallelism on Heterogeneous Hardware: A Java Reduction Case Study

Description:

Parallel skeletons are essential structured design patterns for efficient heterogeneous and parallel programming. They allow programmers to express common algorithms in such a way that it is much easier to read, maintain, debug and implement for different parallel programming models and parallel architectures. Reductions are one of the most common parallel skeletons. Many programming frameworks have been proposed for accelerating reduction operations on heterogeneous hardware. However, for the Java programming language, little work has been done for automatically compiling and exploiting reductions in Java applications on GPUs.

In this paper we present our work in progress in utilizing compiler snippets to express parallelism on heterogeneous hardware. In detail, we demonstrate the usage of Graal's snippets, in the context of the Tornado compiler, to express a set of Java reduction operations for GPU acceleration. The snippets are expressed in pure Java with OpenCL semantics, simplifying the JIT compiler optimizations and code generation. We showcase that with our technique we are able to execute a predefined set of reductions on GPUs within 85% of the performance of the native code and reach up to 20x over the Java sequential execution.

Links to Resources:

Paper: https://dl.acm.org/citation.cfm?id=3281292

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.10 Xtreme Cloud International Workshop 2018

Event Website: http://xtremecloud2018.cslab.ece.ntua.gr

Dates: 10/12/2018

Presenter: Maria Xekalaki, UNIMAN with Daniel Strimpel, IPROOV, in attendance

Presentation Title: Challenges and proposals for enabling dynamic heterogeneous execution of Big Data frameworks

Description:

The efficient execution of Big Data applications requires a large quantity of compute and memory resources. Typically, these resources are in the form of data centres with numerous processing elements connected through a computer network. Although initially the majority of data centers were utilizing only CPU resources, nowadays we can find heterogeneous accelerators such as GPUs and FPGAs. Ideally, Big Data frameworks and applications should exploit those diverse hardware resources in order to push their performance boundaries or increase resource utilization. Despite ongoing work to enable such functionality, the majority of the solutions revolve around external libraries that provide pre-compiled kernels for heterogeneous accelerators.

This fact imposes programmability and code fragmentation challenges that can only be addressed by enabling Big Data platforms to dynamically compile and execute their code on such devices.

In this paper, we analyze and discuss the major challenges for programming and executing Big Data processing applications on distributed systems with heterogeneous hardware. In addition, we present our work-in-progress towards providing a heterogeneous programming framework for running Big Data applications on systems that include diverse hardware resources including CPUs, GPUs, and FPGAs. In contrast to existing approaches, our envisioned solution employs JIT compilation and runtime support, integrated in the data flow engine, enabling the automatic acceleration of Big Data platforms completely transparently to the user and without sacrificing programmability.

Link to Resources: Paper

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

2.11 1ST International Workshop On Next Generation Clouds For Extreme Data Analytics (XtremeCLOUD 2018)

Event Website: http://xtremecloud2018.cslab.ece.ntua.gr/

Dates: 10/12/2018

Presenter: Konstantinos Bitsakos, ICCS

Presentation Title: The Vision of a Heterogeneous Scheduler

Description:

Modern Big Data processing systems, scheduling platforms and cloud infrastructures employ specialized hardware accelerators such as GPUs, FPGAs, TPUs, ASICs, etc. to optimize the execution of resource intensive workloads such as Machine Learning, Artificial Intelligence or generic Data Analytics tasks. Nevertheless, this support is mostly a user-dependent, manual process that requires careful and educated decisions on both the amount and type of required resources to exploit the underlying hardware and achieve any user-defined higher-level policies. In this work we present the initial design of the Heterogeneous Scheduler (HRS), an intelligent scheduler that can make automated decisions on both how and where to map arbitrary data analytics tasks to underlying cloud hardware that may consist of a mix of hardware accelerators and clusters with general purpose CPUs. We experimentally evaluate the performance trade-offs between hardware accelerators and CPUs where we show that there are cases where one technology outperforms the other. We finally present an initial architecture of HRS where we depict its different components and their interactions with the Big Data framework and the cloud infrastructure.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

2.12 10th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2018)

Event Website: https://cyprusconferences.org/cloudcom2018/

Dates: 10-13/12/2018

Presenter: Konstantinos Bitsakos, ICCS

Presentation Title: DERP: A Deep Reinforcement learning cloud system for elastic resource provisioning

Description:

Modern large-scale computer clusters benefit significantly from elasticity. Elasticity allows a cluster to dynamically allocate computer resources, based on the user's fluctuating workload demands. Many cloud providers use threshold-based approaches, which have been proven to be difficult to configure and optimise, while others use reinforcement learning and decision-tree approaches, which struggle when having to handle large multidimensional cluster states. In this work we use Deep Reinforcement learning techniques to achieve automatic elasticity. We use three different approaches of a Deep Reinforcement learning agent, called DERP (Deep Elastic Resource Provisioning), that takes as input the current multi-dimensional state of a cluster and manages to train and converge to the optimal elasticity behaviour after a finite amount of training steps. The system automatically decides and proceeds on requesting/releasing VM resources from the provider and orchestrating them inside a NoSQL cluster according to user-defined policies/rewards. We compare our agent to state-of-the-art, Reinforcement learning and decision-tree based, approaches in demanding simulation environments and show that it gains rewards up to 1.6 times better on its lifetime. We then test our approach in a real-life cluster environment and show that the system resizes clusters in real-time and adapts its performance through a variety of demanding optimisation strategies, input and training loads.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

3. Poster Presentations

3.1 Poster at HIPEAC ACACES 2018

Event Website: http://acaces.hipeac.net/2018/

Dates: 04/07/2018

Presenter: Maria Xekalaki, UNIMAN

Presentation Title: Poster - Dynamic Acceleration of Big Data Applications on Heterogeneous Hardware Resources

Description:

The rapid development of data intensive applications, such as social networks, IoT, etc., has caused an explosion in Big Data processing. Soon, homogeneous CPU clusters reached their limits, as they could no longer satisfy their clients' requirements, which were to execute their applications within a strict time and budget limit. Consequently, cloud providers shifted to diverse hardware architectures, like GPUs and FPGAs, which can greatly accelerate certain operations. However, exploiting such architectures is not trivial; it requires developers to adapt their current software for each target device, a process that is both inefficient and error prone.

In this paper we present our approach in accelerating Big Data Applications on heterogeneous hardware resources. The framework we propose utilizes Apache Flink, an open source stream and batch processing system, and Tornado, a heterogeneous Virtual Machine, which allows Big Data practitioners to exploit the diverse hardware resources in a transparent, dynamic and efficient way, without requiring any code modifications. Our framework is innovative as it does not rely on precompiled kernels or depend on priori knowledge of the hardware, unlike its equivalent state-of-the-art frameworks.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

3.2 Poster at HIPEAC ACACES 2018

Event Website: http://acaces.hipeac.net/2018/

Dates: 04/07/2018

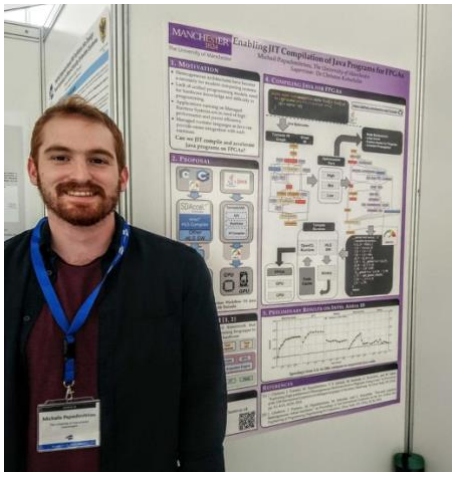
Presenter: Michail Papadimitriou, UNIMAN

Presentation Title: Exploiting Programmability of FPGAs Through Managed Runtime Systems

Description:

The slowdown of Moore's law along with the end of Dennard's scaling and the ever-increasing demand for computing power have shown the performance limitations of homogeneous systems. To address this issue, computer architects have taken advantage of the recent technological advancements in order to come up with heterogeneous solutions where the extensive use of special purpose processing units brings the promise of further improvements both in terms of performance and power efficiency. However, this new reality imposes new challenges both in terms of manageability and programmability. System providers face difficulties to integrate the growing number of accelerators into their existing systems, while application developers cannot harness the full potential of the hardware due to the variety of programming models. For example, GPU hardware is SIMD, while FPGAs are non-Von Neumman, while OpenCL attempts to close this gap and provide a unified development environment. However, this is not yet feasible as deep understanding of the hardware is required to develop optimized solutions.

In this work, we tackle the aforementioned issue related to the programmability of FPGAs. We propose an extension to a virtual machine framework (TornadoVM), which transparently compiles and executes Java programs onto GPUs, to also support FPGAs. In more detail, we generate FPGA executable code with the use of OpenCL, high-level-synthesis (HLS) software and by applying compiler techniques in managed runtime.



Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

3.3 Poster Presentation at SPLASH#2018 Conference

Event Website: https://2018.splashcon.org/track/splash-2018-Posters

Dates: 30/10/2018

Presenter: Juan Fumero, UNIMAN

Presentation Title: Using Compiler Snippets to Exploit Parallelism on Heterogeneous Hardware: A Java Reduction Case Study

Description:

Parallel skeletons are essential structured design patterns for efficient heterogeneous and parallel programming. They allow programmers to express common algorithms in such a way that it is much easier to read, maintain, debug and implement for different parallel programming models and parallel architectures. Reductions are one of the most common parallel skeletons. Many programming frameworks have been proposed for accelerating reduction operations on heterogeneous hardware. However, for the Java programming language, little work has been done for automatically compiling and exploiting reductions in Java applications on GPUs.

In this paper we present our work in progress in utilizing compiler snippets to express parallelism on heterogeneous hardware. In detail, we demonstrate the usage of Graal's snippets, in the context of the Tornado compiler, to express a set of Java reduction operations for GPU acceleration. The snippets are expressed in pure Java with OpenCL semantics, simplifying the JIT compiler optimizations and code generation. We showcase that with our technique we are able to execute a predefined set of reductions on GPUs within 85% of the performance of the native code and reach up to 20x over the Java sequential execution.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

3.4 Poster Presentation at Google Compiler Summit 2018

Dates: 04/12/2018

Presenter: Maria Xekalaki, UNIMAN

Presentation Title: Towards heterogeneous and distributed computing

Description:

The University of Manchester received two fully funded invitations to present their work on E2Data at the Google Compiler Summit of 2018. This event is geared towards PhD students. The participation this year was around 80 PhD students across Europe and Maria as well as Michalis presented their work in Heterogeneous distributing computing and JIT compilation of Java programs on FPGAs.

3.5 Poster Presentation at Google Compiler Summit 2018

Dates: 04/12/2018

Presenter: Michail Papadimitriou, UNIMAN

Presentation Title: Enabling JIT Compilation of Java Programs for FPGAs

Description:

Joining Maria at the same event, Michalis also presented his work on Tornado and FPGA acceleration of Java programs on FPGAs. Both participants disseminated the work done in E2Data to Google, as well as received feedback from Google engineers specifically from the v8 compiler team and Google chrome.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

4. Journal Publications

4.1 Very Large Databases (VLDB)

Full Title: Very Large Databases Journal, Volume 27, Number 6 (VLDB J. 27(6)) (2018)

Journal Website: http://vldb.org/vldb_journal/

Dates: 27-31/08/2018

Author: Sebastian Breß, DFKI

Submission Title: Generating Custom Code for Efficient Query Execution on Heterogeneous Processors

Description:

Processor manufacturers build increasingly specialized processors to mitigate the effects of the power wall in order to deliver improved performance. Currently, database engines have to be manually optimized for each processor which is a costly and error prone process. In this paper, we propose concepts to adapt to and to exploit the performance enhancements of modern processors automatically. Our core idea is to create processor-specific code variants and to learn a well-performing code variant for each processor. These code variants leverage various parallelization strategies and apply both generic and processor-specific code transformations. Our experimental results show that the performance of code variants may diverge up to two orders of magnitude. In order to achieve peak performance, we generate custom code for each processor. We show that our approach finds an efficient custom code variant for multi-core CPUs, GPUs, and MICs.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

4.2 Datenbanken Spektrum

Full Title: Datenbanken Spektrum, Volume 18, Issue 3 (2018)

Journal Website: https://link.springer.com/journal/13222/18/3

Dates: 06/09/2018

Author: Clemens Lutz, DFIK

Submission Title: Efficient and Scalable k-Means on GPUs

Description:

k-Means is a versatile clustering algorithm widely used in practice. To cluster large data sets, state-of-the-art implementations use GPUs to shorten the data to knowledge time. These implementations commonly assign points on a GPU and update centroids on a CPU.

We identify two main shortcomings of this approach. First, it requires expensive data exchange between processors when switching between the two processing steps point assignment and centroid update. Second, even when processing both steps of k-means on the same processor, points still need to be read two times within an iteration, leading to inefficient use of memory bandwidth.

In this paper, we present a novel approach for centroid update that allows us to efficiently process both phases of k-means on GPUs. We fuse point assignment and centroid update to execute one iteration with a single pass over the points. Our evaluation shows that our k-means approach scales to very large data sets. Overall, we achieve up to 20 × higher throughput compared to the state-of-the-art approach.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

5. Presentations of the E2Data project

5.1 Heterogeneity Alliance at Hipeac '18, Manchester, UK

Dates: 22/01/2018

Event Website: http://heterogeneityalliance.eu/content/hetereogeneity-alliance-workshop-hipeac-conference-2018

Presenter Name: Christos Kotselidis

Presentation Title: E2Data Pitch Talk

Description:

The Heterogeneity Alliance consists of a number of industrial and academic institutions with core interest in heterogeneous execution. In the annual workshop hosted by the Hipeac conference, we had the opportunity to present the whole stack of the E2Data project in front of an audience of approximately 50-60 people.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

5.2 3rd BDV PPP Steering Committee Meeting, Brussels, Belgium

Dates: 08/02/2018

Event Website: http://www.bdva.eu/?q=node/973

Presenter Name: Ioannis Komnios

Presentation Title: E2Data Project Overview

Description:

The Steering Committee of the BDV PPP met on 08 February 2018, hosted by ATOS Brussels, and with the presence of the European Commission, the PPP projects and the BDVA.

Important focus was given to the presentations of the projects starting in 2018 (including E2Data) and present in the meeting (CLASS, BigMedilytics, BOOST 4.0, E2DATA, FABDANGO, EDI, ICARUS, TheyBuyForYou, Cross-CPP, DataBench and I-BiDaaS). The E2Data project coordinator had the opportunity to pitch about the main guidelines, objectives and strategies. BDVe project presented the main assets the project is providing to the PPP (marketplace, educational hub, and Big Data landscape). BDVA emphasized the importance of the Monitoring Report for the PPP, and it was reviewed the strategy of communication and marketing of the PPP for the year 2018. Finally, it was highlighted again the importance of these meetings in order to find synergies between the different projects, strengthen their collaboration and keep coordinated in the different activities.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

5.3 Elec.Tec 2018 exhibition, Athens, Greece

Dates: 23/02/2018

Event Website: http://www.elec-tec.gr

Presenter Name: Katerina Doka

Presentation Title: Exploiting Heterogeneity in Big Data Analytics

Description:

The 2nd Elec.Tec is an international exhibition on automation systems, telecommunications, software, power generation, lighting, electrical materials and equipment. ICCS members presented the E2Data goals and technical challenges, as part of the ongoing research activities of the Computing Systems Laboratory, School of Electrical and Computer Engineering, NTUA that relate to the exploitation of heterogeneity in data, software platforms and hardware in order to render Big Data analytics workflows faster and more cost- and energy-efficient.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

5.4 HiPEAC CSW Autumn 2018

Event Title: HiPEAC CSW Autumn 2018

Dates: 29-31/10/2018

Event Website: https://www.hipeac.net/csw/2018/heraklion/

Presenter Name: Ioannis Mytilinis

Presentation Title: E2Data: European Extreme Performing Big Data Stacks

Description:

Presentation of the E2Data software stack at a thematic session on Energy-efficient accelerator-enhanced Data Centers. The session was organized under the auspices of the HiPEAC Computing Systems Week that was held in Autumn 2018.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

6. Conference Support and Organisation

6.1 Workshop on Next Generation Clouds for Extreme Data Analytics

Participants: >40

Event Title: 1st International Workshop on Next Generation Clouds for Extreme Data Analytics, co-located with IEEE CloudCom 2018

Event Website: http://xtremecloud2018.cslab.ece.ntua.gr/

Dates: 10/12/2018

Description:

Katerina Doka (ICCS, WP4 leader of E2Data) has co-organized the 1st International Workshop on Next Generation Clouds for Extreme Data Analytics (XtremeCLOUD 2018), co-located with IEEE CloudCom 2018.

Motivation:

Organizations and companies worldwide rely on Big Data analytics to extract significant value out of the enormous amounts of data that stream into their businesses. Correctly and timely harnessing such data can help them reach more insightful decisions, accelerate reactions to key challenges, improve the overall customer experience and, eventually, gain competitive advantage. Due to the enormous expansion rate of the digital Universe, modern Big Data analytics applications are driven by the requirement to process ever increasing amounts of complex batch as well as streaming data within strict time constraints, while having a limited budget for infrastructure resources.

Cloud Computing has prevailed as the computing model of choice for this class of applications, offering access to a wide pool of resources and scale-out capabilities. Despite their proliferation as a dominant computing paradigm, cloud computing systems lack effective mechanisms to manage their vast amounts of resources efficiently, leading to severe resource waste and ultimately limiting their applicability to large classes of critical applications that pose non-moderate resource demands. Moreover, current Big Data software stacks fail to efficiently exploit the heterogeneity in available architectures and hardware which are paving into mainstream cloud infrastructures, due to the monolithic way in which execution is performed. This fact hinders the performance and energy efficiency benefits that accelerators (GPUs, FPGAs, ASICs, etc.) and novel architectures could offer.

Thus, the alarming scalability concerns of extreme data applications create a significant need to lift existing technological barriers of actual fluidity and scalability of cloud resources towards the Next Generation Cloud that can act as a critical cornerstone for digital economy.

To that end, XtremeCLOUD 2018 aims to redefine the state of the art in fundamental aspects of Cloud Computing, such as heterogeneity, resource management, scalability, that enable extreme scale data processing. This workshop, inspired and co-organized by two European H2020 projects, ACTiCLOUD and

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

E2Data, aims to solicit original research work on the design, implementation and evaluation of novel tools and methods for optimizing Big Data applications and workflows.

We envision that this is only the beginning of a series of successful workshops promoting research, collaboration and innovation in the area of Extreme Data.

Workshop Topics of Interest:

- Hardware-aware Big Data frameworks;
- Benchmarking/modeling of performance/cost/energy consumption of resource demanding cloud applications on heterogeneous hardware;
- Algorithms, methods and tools to improve the utilization and scalability of Cloud infrastructures;
- Holistic and efficient management of Cloud resources;
- Novel architectures and programming models for extreme scale data processing;
- Extreme scale batch/streaming applications optimization;
- Scheduling algorithms and tools for heterogeneous execution;
- Applications and use cases of Big Data analytics over heterogeneous architectures/hardware;
- Visionary ideas on extreme data analytics and heterogeneous environments.

Programme Committee and Organizers:

- Bastian Koller, High Performance Computing Center Stuttgart, Germany;
- Christos Kotselidis, University of Manchester, UK;
- Ioannis Konstantinou, National Technical University of Athens, Greece;
- Manos Athanassoulis, Harvard University, USA;
- Nectarios Koziris, National Technical University of Athens, Greece;
- Ying Zhang, MonetDB Solutions, The Netherlands;
- Juan Fumero, University of Manchester, UK;
- Panagiotis Kokkinos, Computer Technology Institute and Press Diophantus, Greece;
- Huasong Shan, JD.com.

Organizers:

- Georgios Goumas, NTUA, Greece;
- Katerina Doka, NTUA, Greece (ICCS, WP4 leader of E2Data).

Paper Selection and Workshop Program: All submitted papers (7) have been peer reviewed by three members of the Program Committee and 6 of them have been selected for presentation during the workshop.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

The selected papers in XtremeCLOUD 2018 present the scientific results of researchers and practitioners from both academia and industry focusing of different aspects of extreme scale data processing. Scheduling and executing Big Data tasks over heterogeneous infrastructures as well as resource allocation and placement in various settings - ranging from NUMA architectures to databases and entire data centers - are in the spotlight of this year's workshop.

Moreover, XtremeCLOUD 2018 featured a keynote talk by Prof. Peter Boncz, a tenured researcher at CWI and professor at VU University Amsterdam. Peter Bonz is a highly esteemed scientist, recipient of numerous awards and currently active both in academia and industry. His expertise lies in the field of Data Management and Database architectures, and he is the creator of the MonetDB database.

His talk outlined his current understanding of analytical data management in the cloud, describing some of his experiences as scientific advisor of Databricks, the company behind Spark and hosting data science software as-a-service in both AWS and Azure. Finally, the talk focused on exploiting future, heterogeneous, hardware in the cloud. Beyond machine learning, it is currently unclear how data science pipelines, and specifically in ETL, warehousing and BI, will be able to use novel hardware elements in both storage and processing.

Attendance: ExtremeCLOUD 2018 had a remarkable attendance of more than 40 participants.

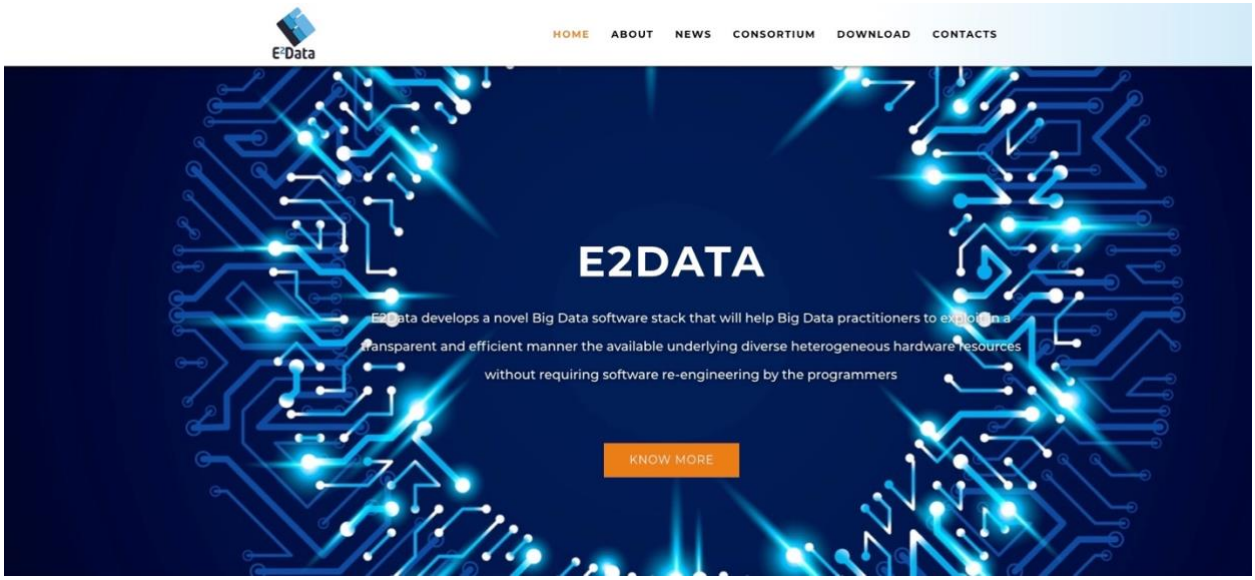
Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

7. Online Presence

7.1 Website

A first version of the E2Data website was launched during the first two months of the project. A redesigned and expanded E2Data website was launched on 12/10/2018.

The new website includes links to download papers, presentations and code: https://e2data.eu/download/publications



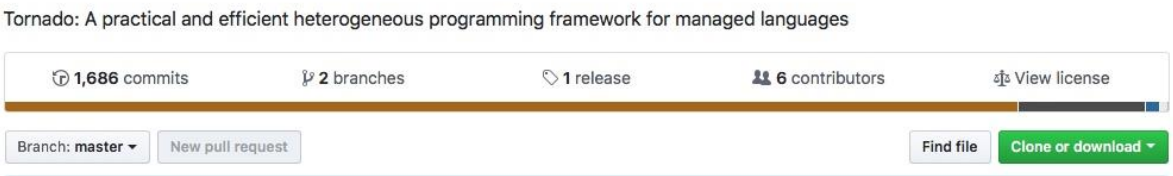
7.2 Tornado Github Public Repository

Tornado project released as Open Source on Github.

URL: https://github.com/beehive-lab/Tornado

Publish Date: September 2018

Stars: 48



Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

8. Big Data Value Public-Private Partnership

As one of the newly funded H2020 Big Data projects, the E2Data project joined the Big Data Value Public-Private Partnership (BDV PPP). Ioannis Komnios, the E2Data project coordinator, and Christos Kotselidis, the E2Data technical coordinator, have joined the BDV PPP steering committee.

The Big Data Value Public-Private Partnership aims at creating a functional Data Market and Data Economy in Europe, in order to allow Europe to play a leading role in Big Data in the global market. The Big Data Value PPP is a partnership between the European Commission and the Big Data Value Association (BDVA) . The contractual arrangement on the cPPP was signed on 13 October 2014. The BDV PPP is developing an interoperable data-driven ecosystem as a source for new businesses and innovations using Big Data. To achieve this the BDV SRIA has defined four implementation mechanisms: i-Spaces, Lighthouse projects, technical priorities and coordination and coordination projects.

E2Data participated in:

- The 3rd BDV PPP steering committee meeting in Brussels, Belgium, on 08/02/2018 where Ioannis Komnios presented an overview of the project and its main goals;
- The Big Data Value Meet-up in Sofia, Bulgaria, on 14-16/05/2018;
- The 4th BDV PPP steering and technical committee meetings in Sofia, Bulgaria, on 15/05/2018.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.

Partners
EXUS (Coordinator), UNIMAN, ICCS, DFKI, NEUROCOM, KALEAO, CTI DIOPHANTUS, SPARKS, iProov

9. Conclusions

The current report presented the diverse dissemination activities carried out by the E2Data Consortium throughout the first 12 months of the project; Table 1 at the beginning of the deliverable summarises these activities.

The entire Consortium has been very active in the dissemination of the project results, with extensive presentations in scientific conferences and workshops, journal publications, and participation in the H2020 BDV PPP initiative, as well as putting in a significant effort in redesigning and implementing a more effective Web presence.

The dissemination progress is in line with the foreseen measurable outcomes defined in the dissemination strategy which was outlined in the E2Data Grant Agreement.

The consortium will continue its intensive dissemination activities, with a special eye towards increasing social media involvement and online engagement, strengthening our existing online presence and seeking out new channels for expanding our reach. In addition, we will continue to foster our relationships with relevant conferences, journals, workshops and other H2020 initiatives.

Co-funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement nº 780245.